

Assessment as a Critical Element in Clinical Experiences for Teacher Preparation

April 29, 2010

With emendations, September 29, 2010

Emerson J. Elliott, *NCATE*¹

INTRODUCTION

This paper examines three complementary dimensions of assessment that teachers will encounter in their careers. These three dimensions involve creation and/ or constructive use of data—data from assessments as well as other sources—for effective practice. They all require teacher preparation settings that offer candidates opportunities to work with students in a variety of contexts, over time and through several iterations. All can be mastered most effectively by candidates when they are exposed to environments that feature close working relationships among the various school and preparation program partners, where collaboration and mutual responsibility are expected, and where engagement with problem-solving data is nurtured. The three dimensions are:

Using assessment in teaching to advance student learning (see page 6)—The first dimension is teacher understanding of the powerful role assessment plays as a means to advance student learning and improve teaching. Research on learning emphasizes the way knowledge is represented, organized, and processed in the mind, and good assessment practices need to make use of those findings by focusing not just on component skills and discrete bits of knowledge but by addressing more complex aspects of student achievement³. Clear goals for instruction help focus teaching and inform students what they are expected to learn. Knowledge of varied forms of assessment, and ways to fit them to particular learning goals, make for effective teaching and provide information for teachers about what their students have learned. Appropriate assessments can require students to solve problems, demonstrate different forms of writing they have mastered, undertake systematic research or scientific studies, develop products, or make oral presentations. Constructive feedback on these varied demonstrations of student work can provide strong incentives for learning. Collaborative experiences should equip new teachers to develop a set of shared standards with their peers and colleagues by which student work is reviewed and scored in relation to structured rubrics.

Using assessment in teacher preparation to monitor candidate growth and proficiency (see page 12)—The second dimension of assessment is measuring teacher candidate growth and proficiency. Whatever candidates are “taught” about good assessment practice can be undermined by poor assessment practice in teacher preparation; thus, modeling the best of what is known about effective use of assessment is a critical aspect of teacher preparation, not just in coursework, but also in exposure to student work and all of the field and clinical experiences of candidates. Portfolio

assessments, such as the California “PACT” are culminating measures of candidates’ readiness for employment, but need consistent, program-spanning support.

Using assessment and other data in longitudinal record systems to enhance teacher practice (see page 15)—The third dimension of assessment illustrated in this paper is information contained in the rapidly growing state longitudinal data systems monitoring individual students and teachers. These systems are repositories for student performance results on state tests and are intended as resources, especially for state oversight and policy purposes.

One argument advanced in their support is sometimes controversial: policymakers can accumulate performance information about the students assigned to particular teachers for use in measuring “value added” by individual teachers. However, in addition to serving as potential sources of information to be used when judging teachers, these systems will also be official records of results of state accountability testing. Prospective teachers need to understand accountability assessments, how the results will be used, and how to teach effectively in states and districts that have aligned their assessments with explicit standards.

The longitudinal systems also contain information of value for programs to share with school districts about the preparation of candidates they might hire, and for districts/states to share with programs about the career progression of former candidates, including the performances of their students and other indicators of teacher’s professional work.

CONTEXT

Around the turn of the millennium, the National Academy of Sciences published a consensus report on learning research and measurement that speaks to an essential aspect of teacher preparation. This report sketched “a vision for the future of assessment” in teaching⁴:

In the future envisioned by the committee, educational assessments will be viewed as a facilitator of high levels of student achievement. They will help students learn and succeed in school by making as clear as possible to them, their teachers, and other education stakeholders the nature of their accomplishments and the progress of their learning.

Teachers will assess students’ understanding frequently in the classroom to provide them with feedback and determine next steps for instruction. Their classroom practices will be grounded in principles of how students think and learn in content domains and of assessment as a process of

reasoning from evidence. Teachers will use this knowledge to design assessments that provide students with feedback about particular qualities of their work and what they can do to improve.

Students will provide evidence of their understanding and thinking in a variety of ways—by responding to teachers’ questions, writing or producing projects, working with computerized tutoring systems, or attempting to explain concepts to other students. Teachers, in turn, will use this information to modify instruction for the class and for individuals on the basis of their understanding and thinking patterns.

The Academy report grounds assessment in accumulating research on learning. It references the mind’s cognitive architecture, including both short-term memory (which is limited) and long-term memory (described as a “virtually limitless store of knowledge”). Much of what an individual knows is domain- and task-specific and organized into structures. Experts use meta-cognitive strategies, reflecting on and directing their own thinking, for monitoring understanding during problem solving and for performing self-correction. Learning entails the transformation of naïve understanding into more complete and accurate comprehension. Practice and feedback are critical aspects of the development of skill and expertise.

The report connects these findings with their implications for assessment. Assessment of how people organize information in long-term memory is critical for determining what they know, how they know it, and how they are able to use that knowledge to answer questions, solve problems, and engage in additional learning. Assessments should evaluate how people organize acquired information, encompassing both strategies for problem solving and ways of chunking relevant information into manageable units. Assessors should attempt to determine whether an individual has good meta-cognitive skills by identifying specific strategies children are using for problem solving, giving particular consideration to where those strategies fall on a developmental continuum of efficiency and appropriateness for a particular knowledge and skill domain. Classroom instruction assessments should focus on making students’ thinking visible to both their teachers and themselves so that instructional strategies can be selected to support an appropriate course for future learning. Finally, one of the most important roles for assessment is the provision of timely and informative feedback to students during instruction and learning.

On the basis of its findings about learning and their implications for assessment, the Academy committee included one recommendation explicitly concerning teacher preparation and professional development:

Instruction in how students learn and how learning can be assessed should be a major component of teacher pre-service and professional development programs. This training should be linked to actual experience in classrooms in assessing and interpreting the development of student competence. To ensure that this occurs, state and national standards for teacher licensure and program accreditation should include specific requirements focused on the proper integration of learning and assessment in teachers' educational experience⁵.

There is overwhelming evidence that this recommendation has generally not been followed. Anecdotal reports and statistical surveys suggest that teacher preparation programs typically provide little instruction about assessments and their positive use as tools of learning, and few opportunities exist for candidates to practice their use in field and clinical experiences. Nor is there evidence that preparation programs, in general, effectively encourage candidates to search for, analyze, and act on other data about students, families, their community, or their own practice.

The current national policy discussion lends urgency to this gap in preparation. As this is written, in March 2010, test experts are trying to make sense of the moves of reformers, and, especially, to take early steps in response to a forthcoming "next generation assessment" competition from the U. S. Department of Education⁶. This new round of assessment development would include state accountability measures for learning that would evaluate the effectiveness of the (also) new Common Core State Standards as they are adopted by states and elaborated in curricular decisions.

For the first time there also seems to be recognition at the Federal level that single end-of-year accountability testing — as it has been practiced in the United States — has little power to enhance learning, as described in the Academy report noted above. New assessments, however, could involve teachers in developmental roles and, perhaps, in scoring. Some of the ideas would restructure accountability assessments to include teacher administered classroom tasks in addition to external, on-demand assessment, using the cumulative descriptor, "assessment system." The Department competition predicts that more diverse forms of assessment will be employed that are explicitly linked to a curriculum, such as oral presentations, written papers of significant length, end-of-course exams, individual projects, and group projects.

Currently, then, the national education reform debate is beginning to recognize — as it had not previously — that teacher-created assessments administered during the course of instruction are an important aid to

learning. Such formative assessments can be conceptually aligned with the same learning progressions that underlie accountability assessments. A kind of hybrid, referred to as “interim assessments,” is now also on the table for discussion. These are conceptually aligned with the learning progressions on which accountability assessments are based, may be developed externally, are standardized tasks, are administered upon completion of a unit of instruction, and are scored with reliable rubrics. The intent is that such interim assessments will accumulate as part of the summative accountability “system,” and will acquire some of the attributes of formative assessments, such as relevance to teachers and relatively timely reporting⁷. The resolution of this current national conversation about accountability assessments and how they might be comprised and used is beyond the timeline and scope of this paper. The key point for the paper is that *teacher use of assessment for learning* and *teacher understanding of assessment practices for accountability* purposes must both be embedded in preparation.

1. ASSESSMENT FOR LEARNING

The Council of Chief State School Officers (CCSSO) has recently released a report on the qualities of assessments that “support high-quality learning.”⁸ The paper is cast as a description of “what a student assessment system could look like if built from the principles and best practices found in current educational research and effective education systems in the U.S. and high-achieving nations around the world.” Such a system would have the following characteristics:

All students have a clear idea of how learning progresses and what they can do to improve. Next generation learners are encouraged to demonstrate their learning as a continuous process.

Parents understand the expectations for their children’s learning as well as the information they receive from school, district, and state assessments. They can work with educators to support their children’s growth and progress.

Teachers are skilled at developing and using a range of assessments based on standards, learners’ needs, and their professional judgment. Scoring student work based on shared learning targets is common classroom practice for teachers. Teachers are well educated and supported in these new expectations.

Supportive educators, including school principals, administrative staff, and leaders at the school and district levels, understand the standards and assessment elements and create conditions for successful learning.

Student achievement information generated at all levels of the assessment system becomes part of the longitudinal state data system and contributes to a rich profile of accomplishment for every student.

Assessments “embedded in the curriculum” influence the day-to-day work of teaching and learning and focus on the use of knowledge to solve problems. The CCSSO paper notes that curricula created to align with common core standards will include:

“the abilities to find and organize information to solve problems, frame and conduct investigations, analyze and synthesize data, self-monitor and improve one’s own performance, communicate well in multiple forms, work in teams, and learn independently.”

Assessments are not only for external accountability. Instead, they are a means for teachers to be explicit about what they are teaching and for students to know what they are expected to learn. They will make use of open-ended items such as essays and problem solutions that require students to analyze, apply knowledge, and write extensively. Teachers assign research projects, science investigations, development of products, and presentations as school-based tasks — tasks that blur the line between what has been thought of as “instruction” and some definitions of assessment. Teachers also work with others to develop, review, score, and use the results of assessments.

If these are qualities that the “next generation” assessments could exhibit in elementary and secondary classrooms, then an unavoidable question is how teachers can be prepared to teach in an environment where they are used. In the decade since NCATE adopted a “performance-based” approach to evidence used in accreditation, it has written, published, and posted on its website several documents intended as technical assistance for institutions preparing for accreditation. One of these addressed what NCATE means when it refers to K-12 “student learning” in its accreditation standards and the implications of focusing on student learning (in contrast with teacher actions) for preparation programs. Preparation programs that prepare teacher candidates for a standards- and assessments-based curriculum meet these conditions⁹:

1a. Course-based and clinical experiences prepare candidates to be “assessment literate” — that is, to know of, select, and create appropriate and effective assessments in teaching that will provide dependable information about student achievement.

Certainly, candidates must be well acquainted with different forms of assessment, know when to use them, understand how to create them, be familiar with ways to interpret results, and know how interpretations can be used to make constructive changes in teaching practices. While some of this candidate knowledge might be generic, some must be specific to the candidate’s specialty field.

Richard J. Stiggins, the president of the Assessment Training Institute (now a part of Pearson professional development) in Portland, Oregon, described “classroom assessment competencies” for educators that flesh out this assessment literacy concept¹⁰:

What teachers need to know:

Professional development in classroom assessment must build a deep understanding of the difference between sound and unsound assessment and a complete understanding of how to use assessment as a teaching tool. . . . (Teachers) must be crystal clear about the achievement targets that they want their students to hit. Different forms of achievement (i.e., mastery of content knowledge, reasoning proficiency, performance skills and product creation proficiencies) require the application of different modes of assessment. Assessment literates know that the first question they must answer is, “What do I want to assess?” Only with that answer in mind can they determine how best to assess it. . . . (T)hey must be prepared to use the full range of assessments to track student achievement, including selected response, essay, performance, and personal communication-based assessment formats. Teachers must know how to select an appropriate assessment method for their particular context and understand when to use it and why it fits best.

Teacher preparation faculty should examine their program curricula and experiences to determine whether and how well candidates are provided opportunities to learn about, select, and create assessments — that is, to become “assessment literate.”

1b. Course-based and clinical experiences assure that candidates have adequate opportunities to develop and practice their skills in teaching so that students learn.

Dr. Stiggins has written a description of assessment “for learning” — that is, assessment of student achievement to advance, not merely check on, student learning¹¹. Teachers assess “for learning” by:

- Understanding and articulating *prior to teaching* the achievement targets that their students are to hit;
- Informing their students about those learning goals, *in terms that students understand*, from the very beginning of the teaching and learning process;
- Becoming assessment literate and thus able to transform their expectations into assessment exercises and scoring procedures that *accurately reflect student achievement*;
- Using classroom assessments to *build students’ confidence* in themselves as learners and help them take responsibility for their own learning, so as to lay a foundation for lifelong learning;
- Translating classroom assessment results into frequent *descriptive feedback* (versus judgmental feedback) for students, providing them with specific insights as to how to improve;
- Continuously *adjusting instruction* based on the results of classroom assessments;
- Engaging students in *regular self-assessment*, with standards held constant so that students can watch themselves grow over time and thus feel responsible for their own success; and
- Actively involving students in *communicating* with their teacher and their families about their achievement status and improvement.

Readers should note the obvious congruence of concepts appearing in Dr. Stiggin’s list with the findings from research on learning and the implications of those findings on assessment that were described above, on page 4, from the National Academy of Sciences report, *Knowing What Students Know*.

The clinical experiences offered in the course of teacher preparation must afford candidates opportunities to bring together what they have learned and experience applying it in realistic settings. The clinical experiences may need to be more frequent and structured developmentally; the final, capstone student teaching experience may need to provide extended time, and offer greater independence for candidate work with students. For example, many teacher preparation

institutions have created multiple opportunities for candidates to work in clinical settings as they progress through a program. Some have made their culminating student teaching arrangements longer in duration. Others have structured their student teaching to allow a second experience with a class of students following a candidate's reflections on the first. A number of education units are partnering with professional development schools. These schools are designed to create settings in which candidates and teachers, themselves, are active learners and candidates participate along with experienced teachers in teams that make student learning the central purpose¹². Whatever the arrangements, the teacher preparation unit faculty, and the classroom supervisors of candidates in student teaching, must have shared perspectives and goals about what constitutes appropriate instruction, the role of curriculum and standards, the relationships between teachers and students, and appropriate uses of assessment for learning.

One dean observed that the State of Washington has embedded expectations like the ones that describe student-learning-as-the-focus-of-teaching already enumerated here. He has found that the biggest change comes in the *focus of conversation* about candidate portfolios. Candidates and faculty *reflect on results in terms of student learning, rather than on specific activities in which the candidate participated* while teaching.

NCATE's interest in accrediting high-quality "alternative" providers from non-Institutions of Higher Education (IHE) will necessitate thinking about the differences that divergent pathways present for candidates to use outcome data and how these differences inform accreditation expectations and processes. Residency programs and other pathways that place candidates in a classroom while they are completing preparation provide candidates with broad and extensive opportunities to apply assessment strategies and use student outcome data that differentiate instruction and improve teaching practices.

Partnerships between preparation programs and schools that focus on school improvement and transformation of school and classroom environments to improve student learning can provide rich clinical settings for candidates. Strong partnerships with school districts also afford opportunities for candidates to work with school instructional teams, possibly as cohorts, to analyze various kinds of assessment data, develop appropriate strategies to improve student learning, and possibly apply those skills in the classroom. The use of "instructional rounds" in schools is "a network approach to improving teaching and learning" that can provide

opportunities for candidates to *observe* teachers use of assessment strategies and to debrief on what has been observed.¹³

While partnerships of the sort described here already exist in some places, it is also true that they do not exist in others. Moreover, geography, population characteristics, and other circumstances may impede an evolutionary move toward these “ideal” arrangements. Among other means to address such impediments, technological tools, including video, distance learning, and simulations, may be necessary “partners” as well.

1c. Course-based faculty and clinical supervisors assess effects that teacher candidates have on P-12 student learning within their unit assessment systems and assessments for each program.

All institutions will find that information about P-12 student learning effects is a critical focal component in monitoring each candidate’s clinical experience progress. NCATE’s Unit Standard 2 calls on preparation programs to conduct assessments that examine:

candidates’ attainment of content knowledge and demonstration of teaching that leads to student learning.

The word “content” here would include assessment knowledge, of course. Many, probably most, institutions will place special weight on data gathered during candidates’ culminating student teaching experiences. Some may rely on evaluations that are routinely conducted through their affiliated professional development school. A growing number of graduating candidates are employed in states with formal induction or mentoring programs, of which one feature is a comprehensive evaluation of new teacher performance. In those situations, faculty may choose to use the state or employer-gathered information as a complement to their unit and program assessments if it includes candidates’ effects on P-12 student learning.

A particularly important source of information for preparation programs is the National Academy of Education book *Preparing Teachers for a Changing World: What Teachers Should Learn and be Able to Do*. This volume describes some pedagogical approaches that “seem particularly promising in helping new teachers develop an understanding of the core ideas” about assessment¹⁴. Four are described:

Analysis of student work and learning—Teacher educators may engage student teachers in examining samples of students' work (for example, writing, projects, lab reports, mathematical equations and explanations) and of student learning (video of student discussions, for instance) in order to evaluate and assess the learning of the student or of a group of students. The purpose of these kinds of analyses is to help new teachers develop an understanding of how such evaluations of learning can inform their instructional choices. There is evidence that teachers feel they can improve their practice when they collectively review student work to analyze what has been learned by different students, uncover misconceptions or difficulties, and reflect on curriculum or teaching adaptations that may be needed to produce stronger student understanding.

Engagement in assessment design—Another pedagogical approach teacher educators are finding useful in helping students understand some of the central ideas about assessment is engaging students in assessment design as part of their learning about instructional design. Rather than planning activities for lessons or units without considering assessment, candidates design an assessment plan as part of a unit or curriculum plan. They learn to "map backward" (Wiggins and McTighe, 1998) from their teaching objective to the design of culminating assessments. The idea is that students' initial and ongoing knowledge and progress will be monitored through a series of periodic formal and informal assessments, leading to the concluding one.

Examining motivation and learning and how they relate to assessment—Teacher educators have found that it is extremely important to help prospective teachers understand core principles of learning that relate to assessment, such as motivation and meta-cognition. For instance, in a course on learning, faculty at one university have found it useful to help student teachers understand the concepts of intrinsic and extrinsic motivation as they relate to the design of feedback and rubrics. They focus upon the concept of "clear criteria" and how that aids in developing understanding. During the course, faculty share key principles of feedback. They discuss, for example, the ways in which the principle of "providing specific positive feedback first before constructive critique" is important not only because it helps fuel students' motivation (by helping them feel a sense of accomplishment in their current work) but also because it contributes to students' meta-cognitive awareness of the progress of their current learning (students become more able to identify their strengths, current levels of understanding, and then the gaps and areas needing development).

Working with standards to design and evaluate assessments for accountability—Finally, teacher educators have developed some pedagogical approaches that seem particularly effective in helping

new teachers understand the relationship between their own curricula and standards and standardized testing, as well as big ideas like norm-and-criterion-referenced testing. Some teacher education programs engage student teachers in a technique called “domain mapping.” Starting with the state’s curriculum frameworks or national content standards, they draw a Venn diagram or construct a table to illustrate what subpart of the desired curriculum is covered by the test and what is not. This activity helps student teachers recognize what parts of each content strand are represented by state- or commercially developed tests and what parts are not. Candidates can then see, for example, that if a test may cover the easiest-to-measure part of each content strand in a domain, this does not mean that the domain has been adequately represented. Saying what has been left out helps to make clear the limitations of the test as a curriculum guide. Based on this explicit analysis, pre-service teachers can then consciously plan units of study and allocation of instructional time in ways that keep attention to tested content in its proportional place.

Above all, faculty and classroom teachers involved in preparing new candidates should, themselves, serve as models for their teacher candidates by exhibiting those instructional and dispositional behaviors that focus teaching on student learning and that employ appropriate assessments “for learning.”

2. ASSESSMENT TO MONITOR CANDIDATE GROWTH AND PROFICIENCY

“Performance assessment” in higher education has evolved through several swings in interest among both institutions and states over the past three or more decades¹⁵. While states have fostered or sponsored assessment of student learning outcomes since the mid 1980s, “their involvement has been fitful and complex¹⁶..” Currently, some thirteen states use a common standardized placement battery to determine if entering students are prepared for college in terms of reading, writing, and mathematics¹⁷. Sometimes the assessment interest has focused on improving learning, and sometimes on institutional accountability. The 2006-2008 Spellings Commission reignited widespread interest in assessment of learning outcomes for accountability purposes, with two-thirds of 32 responding states reporting that they were “planning or implementing [new] state/system assessment policies, so there may be renewed interest at the state level¹⁸.” These efforts to gauge college performance or learning achievement cross all institutional programs.

One institution that has persisted in developing the concept of “assessment as learning” both across the institution and specifically in teacher education is Alverno College. This Wisconsin private liberal arts institution has pioneered the use of assessment for “ability-based education” that does not rely on

standardized tests or traditional exams¹⁹. The specific abilities that Alverno faculty have identified as central to liberal arts and professional education are: communication, analysis, problem solving, valuing in decision making, social interaction, developing a global perspective, effective citizenship, and aesthetic engagement. For both course-based and cross-course expected learning outcomes, Alverno faculty have constructed multidimensional performance assessments to judge candidates “in action.”

The Alverno interest is not limited to *possession* of knowledge, but is especially focused on a candidate’s *use* of knowledge²⁰. Faculty and trained assessors observe and judge a candidate’s performance based on explicit criteria. Their feedback, as well as the reflective practice of self assessment by each candidate, helps to create a continuous process that improves learning and integrates it with assessment. Faculty have adopted these “foundational assumptions:”

- Education goes beyond knowing,, encompassing the ability to do what one knows.
- Educators are responsible for making learning more available to the learner by articulating outcomes and making them public.
- Abilities must be carefully identified in relation to the requirements of contemporary life.
- Assessment is integral to learning.

Alverno’s approach creates a developmental record, and its long history — stretching back to the early 1970s — underscores how much the elementary and secondary education system lags behind in performance measures and the use of assessment *for* learning, since it is only now appearing in the national education “reform” conversation.

The concept of measuring the effects of teaching on student learning actually has its own story. Means for measuring the *act* of teaching and its consequences have been created over the past three decades²¹. One example is the National Board for Professional Teaching Standards in its assessments for National Board Certified Teachers. A second is the Council of Chief State School Officers cross-state effort known as “InTASC,” the Interstate Teacher Assessment and Support Consortium, in its portfolio assessments of new teachers administered during induction.

Other examples occur in teacher education. The idea of “teacher work samples” was devised at Western Oregon University in the 1970s and was more recently extended to eleven Renaissance Partnership teacher preparation colleges and universities. Connecticut, California, and other states initiated the evolution of these measures. Oregon has played a leading role in integrating P-12 standards with teacher education standards, and in demonstrating proficiencies to be demonstrated by teachers through teacher

work samples. More recently, Kansas adapted portfolio performance measures, including student learning components, as an induction assessment for new teachers. Florida has extended its assessment system, incorporating student learning features into its training and certification of teachers through alternative routes as well as through traditional preparation programs.

Perhaps the assessment that is becoming most prominent in teacher education is one created in the State of California by Stanford University in collaboration with 30 California state universities and other institutions known as the Performance Assessment for California Teachers (PACT). In a current association with AACTE and CCSSO, the basic features of PACT will be the foundation for development of a nationally available instrument for evaluating teachers across the continuum of their careers²². Data that come from the assessment are to guide pre-service and in-service training.

PACT is comprised of three assessments²³. The first, “embedded signature assessments,” differ from one institution to another and include such items as case studies, lesson plans, analyses of student work, and observations and reflections on student teaching. The second, the PACT “teaching event,” is designed to capture four categories of teaching: planning, instruction, assessment and reflection. Each candidate plans a learning segment of three to five lessons, videotapes that learning segment, analyzes student learning, and reflects on his or her practice. The teaching events are subject-specific, with separate forms for elementary and secondary credential areas, and include evidence from artifacts of teaching (such as lesson plans, video clips of instruction, student work samples, teacher assignments, daily reflections, etc.) and reflective commentaries which explain the professional judgments underlying the teaching and learning artifacts. The third part of PACT is “subject-specific tasks,” with separate versions for elementary and for secondary candidates., addressing at least one category of teaching (planning, instruction, and/or assessment) in the core content areas not explicitly the focus of the teaching event — i.e., math, science, and history/social science in contrast with language literacy.

The goals of the AACTE/Stanford/CCSSO partnership are to:

- Allow school districts to analyze teachers’ ability to support and advance student achievement;
- Develop a performance assessment that can be a key part of a system of state assessments, beginning with educator preparation, that will be used in the professional development of in-service educators throughout their careers; and
- Contribute in an important way to the development of a more coherent national policy environment for teacher licensure, recruitment and in-service evaluation, and to a more effective national agenda for improvement of teacher quality.

To date the partnership has commitments from 20 participating pilot states.

To reinforce the connection between teaching and student learning, the PACT assessment provides information about P-12 student learning as the outcome of teaching in a classroom situation. That information is gathered in the context of a performance of teaching skills in the classroom.

Finally, preparation programs should carefully consider how their assessments of candidate growth and proficiency actually model ways that teachers should employ assessment in their own practice. If teachers are expected to encourage the development of student meta-cognition that is reinforced by their approach to assessment, preparation programs should also embrace this same approach to foster *candidate learning*. Creating opportunities for candidates to practice during their preparation programs also should be accompanied by systematic support for these candidates in interpreting their experiences and expanding their repertoires. Preparation programs that include this focus on candidate meta-cognition have been shown to strengthen their candidates' performance in several dimensions.²⁴

3. USING ASSESSMENT AND OTHER DATA IN LONGITUDINAL SYSTEMS

In the early 2000s, new Federal legislation establishing the Institute for Education Sciences included an authorization for grants to states to develop “longitudinal data systems.” The term describes an electronic arrangement to collect and retain data about each individual student, retrievable through a unique tracking number, and containing wide ranging information about the student’s demographics, record of progress, success on assessments, course taking, and the like.

While the rhetoric regarding the capabilities of these systems has often outpaced their actual performance on the ground, large amounts of Federal money are being invested (more than \$500 million to date) in the K-12 versions, and is now used to build seamless connections with college data systems. The Gates Foundation is strongly supportive of these systems also as a means to make administrative and policy decisions better informed by data. To that end it has sponsored the Data Quality Campaign (DQC) as an advocacy group to encourage state participation.

The most recent DQC summary of the implementation of these data systems was issued in January 2010. It reported near universal adoption of some features of these systems — e.g., 50 states²⁵ had unique student identifiers, 50 include test data, and 51 were able to provide graduation and dropout data.

However, only 33 states had direct connections to match student level P-12 and higher education data, and just 24 reported that teacher identifiers can be matched with student records.

The point is that these systems are being created virtually everywhere in the United States, that their growth trend has been sharply upward over the five years that the Data Quality Campaign has tracked their progress, and that state policymakers will have access to far higher quality and more analytic information than previously. Three features of these burgeoning systems need to be taken into account during preparation programs, especially in the field and clinical experiences of such programs: their use in evaluation of teachers, their use in compiling accountability information, and their use for exchanges of information between teacher employers and teacher preparation programs. Each of these is described in more detail below.

3a. Use in accumulating information for teacher evaluation

A selling point for longitudinal data systems to state policy leaders is their ability to link information about students with their teachers. That means that performance data on students in a classroom can be associated with data that identify and describe their teacher; the result can be used, at least in part, to judge the teacher's performance. This aspect of longitudinal data has been controversial both among teachers and their representatives as it influences evaluations, and among members of the measurement and statistics profession with regard to the validity of the measurements. Especially strong cautions have been stated about "value-added" judgments in which student learning gains are the sole measure — or a predominant one — particularly when the goal is to evaluate individual teachers as opposed to groups of teachers in a school. Value-added measures are criticized for failing to account for extraneous factors, such as the lingering effects of previous instruction or other events in children's lives that could influence their performance. However, recent reports are providing more guidance describing how "value-added" measures can be used appropriately²⁶. Further, following years of conversation, it seems that those who advocate "multiple measures" of teachers' work (including structured observations, videos and other approaches) are being heard by policymakers as federal legislation and grant announcements are considered. In any event, teachers should be knowledgeable enough about the principal features of this issue to participate in the continuing debate²⁷.

3b. Use in accumulating accountability information

Longitudinal systems, by design, will include information about student performance on state accountability tests. Preparation programs need to equip teachers with an understanding of the implications of testing in a standards- and assessment-based accountability system; high school teachers, for example, need to know more about the SAT, ACT, Advanced Placement tests and other standard course-based assessments, as well as state required tests, end-of-course tests such as those sponsored by Achieve, and state graduation tests. Teachers in earlier grades need to know about accountability tests in their state or district. All teachers need a sophisticated understanding of how test-prep and test coaching can undermine real learning and how to teach with integrity in the face of high-stakes pressure. This does not mean ignoring the tests, but rather analyzing the learning goals and staying focused on them instead of practicing on narrow item formats. Grade level teams of teachers also need to know how to do once-per-year formative program evaluations using state test results from the previous year to revise curricula in preparation for the coming year²⁸.

3c. Use as a source of information for reciprocal exchange between employers and preparers of teachers

The longitudinal data systems will, or can, include information about preparation programs that would inform district hiring decisions, and, reciprocally, data about teacher career paths and their students' learning that can help preparation programs know their own strengths and weaknesses²⁹. Information about the teacher's preparation courses and experiences, licensure, certification, academic record, and — perhaps eventually — PACT-like performance assessments could be retained in the higher education extension of longitudinal data systems. These systems could record information about candidates' understanding and practice in applying: knowledge of child growth and development, standards- and assessments-based instruction, subject content knowledge, and content pedagogy that involves extensive experience with students and student "work." Such information could be invaluable for prospective employers.

Similarly, information about a teacher's career progress, responsibilities, school setting, collaborative and professional development activities, and multiple measures of students' performance would all be appropriate to share with preparing institutions. Some examples of useful information are:

- **teacher career paths** (such as hiring, type of school, whether employed in STEM or other hard-to-fill subjects, tenure, and contextual data on the characteristics of students in the school and in the teacher's classroom);

- **teacher induction experience** (such as teacher satisfaction with their preparation, assistance provided by the school or district, principal satisfaction with quality of the teachers, and licensure status); and
- **teacher performance measures** (such as value added student learning, performance evaluation results or other measures of teacher impact, observational measures of classroom performance, evidence of leadership roles, and merit pay status).

Such information—especially when compared with similar information from comparable institutions and with the preparing institution’s own program goals—would help faculty determine points of particular strength or weakness in their courses and program experiences. There is value for both districts and preparation programs in such reciprocal exchanges of data. Finally, there is value for states as well, since the longitudinal systems would permit them to match college preparation information with on-the-job teacher performance to see what patterns emerge among new teachers from different institutions or from different types of preparation.

The Administration published (in March 2010) *A Blueprint for Reform: The Reauthorization of the Elementary and Secondary Education Act* that outlines concepts that will be pursued in Congress. The “Blueprint” has gathered many of the currently discussed ideas into one package. We find, for example: support for teacher career ladders; a call for stronger traditional and alternative pathways into teaching; making grants available for “recruitment, preparation, placement, and induction of promising teacher candidates for high-need schools, subjects, areas, and fields;” recognition that principals have important leadership roles in supporting teachers; asking for state plans to ensure equitable distribution of teachers and principals who receive at least an “effective” rating, “fair and meaningful teacher and principal evaluation systems, working in collaboration with teachers, principals and other stakeholders”; requesting that districts build instructional teams of teachers; and requiring states and districts to publish “report cards” on key indicators such as teacher qualifications and on the “performance of teacher and principal preparation programs by their graduates’ impact on student growth and other measures such as job placement and retention.”

If such policies become the law of the land then they will necessitate changes in the management practices of districts and schools. Longitudinal data systems should make it possible to merge data that describe these practices; of particular importance, by design they make it possible to associate

information about students with information about teachers. As the systems are fully implemented, the information they contain will be extensive and both highly useful for analytical purposes and highly sensitive with regard to protecting confidentiality.

Teachers in preparation need to know that these systems are under construction and understand how they, themselves, can tap into them during their teaching career. There are privacy issues, of course; frequently some intermediary at the state level is a transfer point for data exchange. Teachers can, however, combine information on their own students with, for example, information about other classrooms with similar students as one means to extend their own professional knowledge. Exposure to the qualities of these systems is especially urgent now if, as the Administration proposes, the conversation about teacher evaluation is to move beyond the restricted measurements that have been available of student achievement growth to a more collaborative one seeking a range of valid measures that can make up a teacher evaluation system.

SUMMARY AND RECOMMENDATIONS

This paper began with a lament that there is ample evidence that the following recommendation from a National Academy of Sciences study on assessment and learning has not been implemented:

Instruction in how students learn and how learning can be assessed should be a major component of teacher pre-service and professional development programs. This training should be linked to actual experience in classrooms in assessing and interpreting the development of student competence. To ensure that this occurs, state and national standards for teacher licensure and program accreditation should include specific requirements focused on the proper integration of learning and assessment in teachers' educational experience.

The text of the paper is intended to outline key features of preparation that could help to put the Academy recommendation into widespread use. Its three sections:

- illustrate the kind of knowledge of and practice in assessment that teachers need in order to capture its powerful effects on learning (section 1);
- summarize and bring up-to-date efforts to create valid assessments of the *act* of teaching that are appropriate for use during teacher pre-service preparation (section 2); and
- describe the fast-developing longitudinal data systems that states are creating — and that will contain valuable information about individual students with both demographic and

instructional aspects of their progress — which teachers need to be able to draw from as well (section 3).

All of these points emphasize the *use* of data — assessment data and other data — *as means to improve teacher effectiveness*. All emphasize the instrumental role that pre-service preparation must have through courses, candidate work, and field experiences under the caring eyes of mentors and peers.

Implementation of the Academy recommendation should be one of the critical goals of the Blue Ribbon Panel. This calls for courses and field and clinical experiences so that candidates can study, practice, and master use of assessment and other data that will advance student learning. The specific recommendations are:

That candidates be presented with multiple and rich course material in their preparation that will enable them to become *assessment-literate* and *data-wise* — There should be explicit instruction about different modes of assessment, when and how to use these differing modes, the need for clarity about achievement targets for students, how to create assessments that are appropriate for the target, interpretation of results, and providing constructive and relevant feedback to students to reinforce and motivate learning. There should be explicit instruction in applications of data to understand situations and solve problems, together with opportunities to practice its use. As longitudinal data systems are rapidly built across the states, candidates need to know how to negotiate these systems for their own professional purposes and where there are points they can influence about the content, access to, and proper use of the wide array of information these systems will include.

That candidates participate in multiple and rich experiences with students so they can develop use of assessments in their teaching practice — Preparation programs should provide opportunities through candidates' own developing stages in use of assessment for learning. Beginning early in their preparation candidates should have exposure to diverse examples of student work and collaborative experiences to develop their proficiencies in knowing what student work says about achievement, level of student development, student preconceptions, and student performance in a standards- and assessments-based instructional system. Candidates should have opportunities to observe assessments used well and to participate in the administration and interpretation of assessments; they should experience the attention of caring mentors who observe their fledging assessment practices, and they should experience full practice as professionals in the design and execution of their own assessments

for learning. There should be frequent clinical opportunities to practice use of assessment to judge whether classroom teaching goals are met and provide feedback to improve student learning.

That candidates' preparation experiences showcase faculty modeling that consistently demonstrates the application of good assessment practices for learning — It is often said, with strong justification, that candidates for teaching have experienced 14 or 15 years of education in which assessment is used poorly. “Poor” assessments are not focused on specific learning goals, are not constructed in a mode appropriate to specific goals, do not provide a basis for constructive and descriptive feedback, are not used to motivate student learning, and do not function as a diagnosis of teaching strengths and weaknesses. Far too often the main effect of assessment is to displace the regular curriculum in order to prepare for coming state required standardized tests. Moving candidates to a new understanding about learning and use of assessment to measure what is being taught is difficult. Faculty in teacher preparation must model good assessment practices, and the partnering preparation program and school and district teachers need to share consistent views about what those “good assessment practices” are. The partners — faculty and classroom teachers — should also provide models for candidates to make their practice more effective by searching out and using data from sources other than assessment, including, for example, data about student demographics and families in the community, data about community resources, and data about attendance.

CODA on research in teacher preparation:

Other Blue Ribbon Panel papers and principles make reference to one of the important conditions of teacher preparation, namely its limited research base. However, it is appropriate to mention the topic in this paper on assessment as well because the main propositions advocated in the paper — that (1) teacher candidates acquire assessment literacy, that (2) candidates experience a wide variety of forms of assessment, used effectively, during their preparation, and that (3) candidates understand and have occasions to apply information on student assessment, teacher evaluation, student course taking and other contextual information from longitudinal data systems—all contribute to a more data-based, data-driven profession, and also to systematic research about preparation.

The rapidly expanding longitudinal databases will provide, and already are providing, far stronger underpinnings for conclusions about state policy and administrative actions than we have had available previously. Half a dozen or so of these systems are the sources of information used by the Urban Institute Center for Analysis of Longitudinal Data for Education Research (CALDER) studies which

associate teacher and student data from longitudinal sources, and frequently supplement them with other data or surveys for a more complete research data base.

But the emphases on student learning and multiple forms of assessment described in this paper underscore the role that partnerships for effective clinical and field experiences can play in *research and effective use of data about teacher preparation itself*. Data can be collected about particular pedagogies, interaction situations with students, collaborations among school and institution partners, recruitment and selectivity choices, and other aspects of these partnerships. Assessments of wide variety can be created to help judge the effects of the different interventions on student learning. In *using* data to examine its own performance, each partnership can understand what it has done that works, re-engineer those activities that are deficient, and identify others that have negative attributes. The accumulation of these individual partnership experiences, supplemented by the systematic research studies that will be conducted through some of the partnerships, will gradually help to fill in the blanks in our knowledge base on effective teacher preparation. The clinically-based partnerships will be both exhibitors of systematic use of data to improve their own performance, and sites to collaborate in structured research.

¹ The author thanks reviewers who made constructive, insightful and important comments on earlier versions of this paper, especially Lorrie Shepard and Jim Cibulka. I remain responsible for how their suggestions are reflected in this final version.

³ The summarized excerpts about learning research that underlie what are thought of as good assessment practices are condensed and adapted from pages 3-5 of *Knowing What Students Know: The Science and Design of Educational Assessment*, National Research Council Committee on the Foundations of Assessment. Pellegrino, 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, 2001.

⁴ *Knowing What Students Know: The Science and Design of Educational Assessment*, p. 292

⁵ *Knowing What Students Know: The Science and Design of Educational Assessment*, p. 309

⁶ On September 2, 2010, Secretary Duncan announced two grants, totaling \$330 million, that had been awarded to consortia of states, each of which would undertake development of new assessments for accountability purposes and include formative or “interim” assessments in some form. The assessments are to be aligned with the Common Core State Standards in mathematics and English language arts.

⁷ The assessment distinctions in this paragraph are based on comments from Lorrie Shepard on April 17, April 27, and April 28.

⁸ *Performance Counts: Assessment Systems that Support High-Quality Learning*, Linda Darling-Hammond, Council of Chief State School Officers, Washington, DC, 2010

⁹ Adapted from *Student Learning in NCATE Accreditation*, NCATE, Washington, DC, 2005

¹⁰ *Learning Teams for Assessment Literacy: A Concept Paper*, Rick Stiggins, Journal of Staff Development, volume 20

¹¹ *Assessment Crisis: The Absence of Assessment FOR Learning*, Richard J. Stiggins, Phi Delta Kappan, online article, V. 83, No. 10, pages 758-765, June 2002.

¹² See *Standards for Professional Development Schools*, Marsha Levine, Project Director, National Council for Accreditation of Teacher Education, Washington, DC, 2001.

¹³ *Instructional Rounds in Education: A Network approach to Improving Teaching and Learning*, Elizabeth A. City, Richard F. Elmore, Sarah E. Fiarman, and Lee Taitel. Cambridge, MA: Harvard Education Press, 2009.

¹⁴ *Preparing Teachers for a Changing World: What Teachers Should Learn and be Able to Do*, Linda Darling-Hammond, John Bransford editors, National Academy of Education, Jossey-Bass, 2005; See Chapter 8 on *Assessment*, Lorrie Shepard, Karen Hammerness, Linda Darling-Hammond, Frances Rust, with Joan Baratz Snowden, Edmund Gordon, Cris Gutierrez, Arthro Pacheco, pp. 275-326.

¹⁵ One report on this history and its application in different disciplines is contained in *Assessing Student Competence in Accredited Disciplines: Pioneering Approaches to Assessment in Higher Education*, edited by Catherine A. Palomba and Trudy W. Banta, Stylus Publishing, Sterling, Virginia, 2001. A more recent report is *Assessing Student Learning Outcomes in College: The Role of the States*, Peter T. Ewell at the National Center for Higher Education Management Systems (in draft, no date)

¹⁶ Ewell, *Assessing Student Learning Outcomes in College: The Role of the States*, p. 1

¹⁷ Ewell, *Assessing Student Learning Outcomes in College: The Role of the States*, p. 5

¹⁸ Ewell, *Assessing Student Learning Outcomes in College: The Role of the States*, p. 16, 17

¹⁹ From online description of Alverno's Ability-Based curriculum, www.alverno.edu

²⁰ *Student Assessment-as-Learning at Alverno College*, the Alverno College Faculty, Alverno College, Wisconsin, 1994

²¹ From *Student Learning in NCATE Accreditation*, op. cit.

²² Online overview, Teacher Performance Assessment Consortium, AACTE

²³ Performance Assessment for California Teachers STEP Elementary, 2009-2010, from the online site

²⁴ *Preparing Teachers for a Changing World*, op. cit. See Chapter 10 on *How Teachers Learn and Develop*, Karen Hammerness, Linda Darling-Hammond, John Bransford, with David Berliner, Marilyn Cochran-Smith, Morva McDonald, and Kenneth Zeichner, pp. 375-378.

²⁵ The term "states" includes DC and Puerto Rico. Two states are not included in the "50" that have unique student identifiers.

²⁶ See, for example, *Getting Value Out of Value-Added*, the report from a workshop jointly sponsored by the National Research Council and the National Academy of Education, committee on Value-Added Methodology for Instructional Improvement, Program Evaluation, and Educational Accountability, Henry Braun, Naomi Chudowsky, and Judith Koenig, Editors, Division of Behavior and Social Sciences and Education, National Academies Press, Washington, D.C., 2010. See also, *Problems with the Use of Student Test Scores to Evaluate Teachers*, EPI Briefing Paper # 278, August 29, 2010, Economic Policy Institute. Co-authors are: Eva Baker, Paul Barton, Linda Darling-Hammond, Edward Haertel, Helen Ladd, Robert Linn, Diane Ravitch, Richard Rothstein, Richart Shavelson, and Lorrie Shepard. Note that the Gates-supported teacher effectiveness research project, Measuring Effective Teaching (MET), is gathering information from several different resources for its study: student achievement gains on assessments; classroom observations and teacher reflections (including use of video taping as a data source); teachers' pedagogical content knowledge (that is, a test); student perceptions of the classroom instructional environment (as ways to measure engagement); and teachers' perceptions of working conditions and instructional support at their schools.

²⁷ On April 17, Lorrie wrote: "As yet, very few value-added analyses give results back teacher-by-teacher. Kane and others have acknowledged that individual teacher results are highly variable from one year to the next. Therefore, teachers should not necessarily be responding by changing their teaching in response to one year's results. Bill Sanders has addressed the volatility issue by requiring 3 years of data and using shrinkage formulae. As a consequence, only a very few teachers get a result that is good or bad compared to the average teacher. I don't think we are ready to say what it is that every teacher should know about interpreting value added results."

²⁸ This paragraph is adapted from comments made on an earlier draft of the paper by Lorrie Shepard.

²⁹ The Data Quality Campaign has circulated a "template" among states as a means to encourage discussions between states and their preparation institutions about data in their student longitudinal data systems that would inform preparation programs of the strengths and weaknesses exhibited by teachers they had prepared. The URL is: http://www.dataqualitycampaign.org/files/DQC_Teacher_Template_8-16.pdf