1-2007

From Teacher Education to P-12 Learning Outcomes: The New Burden of Proof

Marcy Singer-Gabe/la

Chris Iddings,

Kim Paulsen

Margaret Smithey

Marie Hardenbrook

See next page for additional authors

[How does access to this work benefit you? Let us know!](#)

Follow this and additional works at: https://commons.und.edu/tl-nirp-journal

Part of the Scholarship of Teaching and Learning Commons

From Teacher Education to P-12 Leaming Outcomes: The New Burden of Proof

Authors

Marcy Singer-Gabe/la; Chris Iddings,; Kim Paulsen; Margaret Smithey; Marie Hardenbrook; Amy Palmeri; Leona Schauble; and Camilla Benbow

From Teacher Education to P-12 Learning Outcomes: The New Burden of Proof

Marcy Singer-Gabella, Chris Iddings, Kim Paulsen, Margaret Smithey, Marie Hardenbrook, Amy Palmeri, Leona Schauble, and Camilla Benbow
Vanderbilt University

Introduction

The federal "No Child Left Behind" (NCLB) Act (2002) has created a new "burden of proof" for colleges of education. The legislation's expansive accountability reforms for elementary and middle schools, combined with a spotlight on teacher quality, have redefined what are to be considered key indicators of teacher candidate performance in teacher preparation programs. Traditionally, schools of education have tracked dispositions, beliefs about teaching and students, content knowledge, and teaching skills. Now, however, the focal performance indicator is P-12 student learning. Hence, colleges of education have increasingly become concerned about what constitutes compelling evidence that graduates indeed have a significant and positive impact on the achievement of their students.[1]

Providing such evidence is not a trivial undertaking. This article offers a view of the requirements and challenges of establishing an approach to assessment that is tethered to P-12 student learning outcomes. We are concerned with two phases of program assessment. The first focuses on evidence of teacher candidates' progress toward beginning teaching competence *prior* to full time teaching. The second focuses on evidence of the effectiveness of program graduates in supporting the learning of their students, using data on P-12 student learning outcomes *after* graduates enter full-time teaching.

A strong assessment strategy is driven by a clear idea of the outcome. Therefore, we begin with a brief accounting of goals for student learning and achievement that permeate both policy documents and current research on student learning across content areas and grade levels. We then turn to the kind of teaching required to support such learning. In describing "good teaching," we draw on the consensus

vision that has emerged among the communities of teacher education and professional practice, discuss current thinking about the kinds of knowledge and skills required for this kind of teaching, and indicate potential challenges to this vision. Next, we pursue two facets of the assessment challenge:

1. What are the pathways to competence (i.e., good teaching) and how can we assess preservice teachers' progress along these developmental trajectories; and
2. What are the conceptual, methodological, and logistical challenges of hooking assessments of teacher practice and teacher education program quality to student learning outcomes?

To ground this discussion, we provide examples of our own efforts to document the preparation and education experiences of undergraduate and master's level teacher candidates in an education program at a private research university. We conclude with an outline of an agenda for research and development that we are pursuing at Peabody. In sharing our initial planning, we aim to create energy and direction among schools of education to move ahead with sound evaluation of programs and practices that are linked to P-12 student learning and achievement.

The Goal: Student Learning and Achievement

International standards for student learning have moved well beyond basic comprehension and computation (Programme for International Student Assessment, 2004). Professional organizations representing most curricular domains across the elementary and secondary curriculum have articulated expectations for student learning that emphasize the ability to investigate, analyze, and evaluate information and ideas, to recognize their relevance, and to apply concepts appropriately in novel contexts.[2] "Successful" schools, then, are those that effectively foster students' abilities to think critically and creatively; enable students to achieve principled understanding of subject matter that is scaffolded by relevant conceptual and factual knowledge; inspire students' civic awareness, sense of responsibility, and competence; and deepen students' social awareness, empathy, and

compassion. While curriculum and pedagogy vary among such schools, it is clear that students must have access to rich and challenging content and opportunities to apply their thinking in authentic situations. Carpenter and Lehrer (1999) have delineated the features of classrooms that support student understanding. They characterize them as places where students are encouraged to construct relationships between related disciplinary ideas, to elaborate and extend upon their knowledge, to articulate their thinking, and to develop their identities as learners.[3]

From Learning to Teaching

Review of research and standards documents of the past 15 years reveals remarkable consistency in notions of teaching that support the kinds of learning articulated above. As described in the Standards of the Interstate New Teacher Assessment and Support Consortium (INTASC, 1992), even beginning teachers must be able to support the intellectual, social, emotional, moral, and physical development of students; respond with flexibility and professional judgment to their different needs; and actively engage them in their own learning so that they can use and generate knowledge in effective and powerful ways (p. 12). Good teaching in action is profoundly complicated. Summarizing research on teaching, Linda Darling-Hammond offers an account:

> Studies of teaching (Jackson, 1968; Lortie, 1975; Clandinin, 1986) describe it as complex work characterized by simultaneity, multidimensionality, and unpredictability. In classrooms competing goals and multiple tasks are negotiated at a breakneck pace, trade-offs are continually made, unanticipated obstacles and opportunities arise. Each hour of every day teachers must juggle the need to create a secure supportive environment for learning with the press for academic achievement, the need to attend to individual students and the demands of the group, and the challenges of pursuing multiple strands of work so that students at varying places in their learning move ahead and none are left behind. (Darling-Hammond, 1997, p. 69)

Navigating these demands *well* requires that teachers possess extensive subject matter knowledge, the ability to establish respectful working relationships with children and their families, the vision and ability to create rich learning environments, as well as the commitment to continued professional learning.

Although this view of teaching and what teachers must know and be able to do is widely accepted by communities of educational research, professional practice, and accrediting agencies, there is a widening gap in views about what knowledge and skills are essential for the *beginning* teacher, and thus what teacher preparation programs should emphasize. On the one hand, university faculty generally view teaching as a learned profession and aim to prepare beginning teachers to be caring, committed, and autonomous decision-makers who possess solid knowledge of content and a beginning repertoire of pedagogical strategies to support diverse learners (Feiman-Nemser, 2001). At Peabody, faculty work deliberately to help teacher candidates develop not only subject matter expertise, but also habits of mind (dispositions) that promote teaching as an intellectual process of inquiry, discovery, and engagement with learners and learning.

In contrast, recent federal regulations emphasize technical aspects of teaching and suggest that teachers should be prepared to deliver content knowledge and to handle classroom management (Imig, 2004). According to this perspective, judgment and discretion develop later. The gap between these conceptions of what is essential for a beginning teacher has been exacerbated by an approach to accountability defined by standardized testing of student achievement in reading, language arts, and mathematics. Given a policy environment in which the main outcome of interest is test performance, rather narrowly defined, there is increasing interest in the use of highly scripted, "teacher proof" materials, designed to ensure that even teachers with little experience or content knowledge can enable students to acquire basic skills that are tested on state assessments. Implied here is the notion that "good enough" teaching is achieved by acquiring a static base of technical knowledge and skills rather than learning to negotiate a complex and dynamic enterprise.

We reject this narrower vision of "good enough" teaching. A growing body of evidence supports the conclusion that there is no one script that can support the learning of all students (National Research

Council, 2000). Further, while these scripted instructional methods are appropriately used with certain special education populations and may effectively support the development of lower level skills for some general education students, they represent one end of a continuum of strategies and target a particular subset of learning needs. A more sophisticated and flexible teaching repertoire, and the understanding and ability to judge when particular approaches are appropriate, is required to help *all* students develop the foundation for more complex understanding essential to success in secondary and post-secondary content.

Yet another perspective is that beginning teachers simply need strong subject matter preparation, combined with basic classroom management skills. While there is anecdotal evidence of individuals who succeed starting with a strong liberal arts background and basic classroom management skills, as a broad policy matter we find this perspective wanting for many of the same reasons that we reject the view of teaching as scriptable. As discussed later in this paper, teachers need not only content knowledge, but also the pedagogical under-standing and flexible repertoire to transform content into learning experiences appropriate to a wide range of learners.

Assessing the Progress of Teacher Candidates

Pathways to Competence

Part of the difficulty in determining what is essential for beginning teaching results from the absence of an empirically validated model of the path (or paths) to teaching competence. Although there is wide agreement across colleges and universities about the general experiences prospective teachers should be provided (e.g., coursework in content and methods, field experiences, etc.), most teacher education programs operate from an intuitive sense of what increasingly sophisticated understandings of teaching, learning, learners, and content look like. These intuitions—which may or may not be shared across program faculty—are expressed as teacher education faculty make judgments about candidates' readiness to proceed from one stage to the next in preservice education (e.g., formal admission to teacher

education, entry into student teaching, graduation, and certification). Some have begun to describe these levels of understanding and skill using course or program-specific rubrics (cf. Elliott, 2003). However, to date, most understandings about candidates' development of teaching competence remain locally articulated, have not been validated through empirical research, and are not tied to P-12 student learning.

The absence of an empirically validated model of teacher candidate development is due in part to the fact that most studies occur within the context of local programs or courses—focusing on the impact of specific program features on preservice teacher understanding and development. Many rely on self-report, survey of program stakeholders, or judgments by individual course instructors. The resulting developmental frameworks thus are tied to particular programs and experiences. Syntheses of these microresearch efforts do not yet exist (Cochran-Smith & Fries, 2005).

Articulation of a model is further complicated by the wide variation in experiences that prospective teachers bring to their programs. Given the impossibility and undesirability of reducing variation in teachers' prior experiences, it seems clear that any model of teacher development must recognize diverse starting points and pathways. For this reason, a developmental trajectory for preservice teachers might be anchored at one end of development to the strands of understanding and skills that beginning teachers must possess, and then describe within each strand a continuum from naïve to more sophisticated understanding. In the following pages, we sketch a conceptual model that describes teacher development in these terms and we offer examples of the kinds of assessments that might provide evidence of preservice teachers' progress.

Learning Strands and Evidence of Progress

Our thinking is informed by the work of Sharon Feiman-Nemser on teacher learning across the teaching career span. Feiman-Nemser (2001) suggests that there are five central tasks to be accomplished in the early stages of learning to teach (i.e., preservice preparation):

- analyzing beliefs and forming new visions;
- developing subject matter knowledge for teaching;

- developing understandings of learners and learning;
- developing a beginning repertoire of curriculum, instruction, and assessment; and
- developing tools to study teaching.

In the following pages, we illustrate the first four of these tasks within the context of the teacher education program at Peabody—providing examples of activities designed to engage teacher candidates in these developmental tasks and articulating the kinds of evidence we look for as we assess candidates' progress in accomplishing them.[4]

Analyzing Beliefs and Forming New Visions. By the time they enter teacher education programs, teacher candidates possess deep-rooted beliefs about teaching, learning, and schooling—beliefs forged from their years of experience as elementary and secondary students (Lortie, 1975). Although they are well-elaborated, these beliefs are naïve and self-centered—construing the work of teaching from a perspective of child or adolescent. More often than not, these beliefs about learning and teaching also are inconsistent with aspirations for student achievement and the vision of teaching we have described. Finally, as Feiman-Nemser (2001) writes, although these beliefs are naïve, they can be tenacious—filtering and, at times, "limiting the ideas that teacher education students are able and willing to entertain" (p. 1016).

The challenge for teacher educators, then, is to help candidates identify and probe their beliefs so that they may form more complex, warranted, and productive understandings of teaching:

Teacher candidates must also form visions of what is possible and desirable in teaching to inspire and guide their professional learning and practice. Such visions connect important values and goals to concrete classroom practices. Unless teacher educators engage prospective teachers in a critical examination of their entering beliefs in light of compelling alternatives and help them develop powerful images of good teaching and strong professional commitments, these entering beliefs will continue to shape their ideas and practices. (Feiman-Nemser, 2001, p. 1017)

At Peabody, teacher candidates first take up this task during introductory courses, in which they begin to discuss their ideas about schools, what it means to be a teacher, the learning process, equity and diversity, working with parents, etc. As they progress through the four-year program, teacher candidates are asked to examine their ideas more systematically.

Across three years of field experiences in the undergraduate secondary education program, teacher candidates are asked to examine and make explicit their beliefs about teaching, learning, and equity by explaining *how they know their students are learning*. Teacher candidates record their thinking in video narratives as sophomores, juniors, and seniors. The recordings serve as evidence of candidates' development of increasingly more complex and research-based understandings of what constitutes effective teaching. For example, the following students' comments made early in their program reflect a "banking model" of teaching and learning, wherein students receive and deposit information: "Teaching is offering a commodity and making people want to learn" and "Learning is being receptive to the ideas of others." A comment made subsequent to participating in a practicum in middle and high school classrooms demonstrates emerging awareness of teaching and learning as social practices: "Teaching and learning involve fostering relationships with people, with ideas, and with subject matter."

Developing Subject Matter Knowledge. As discussed above, effective teachers possess a solid command of their subject matter. They not only possess solid factual and conceptual knowledge, but also understand the disciplines as dynamic fields of inquiry, each with particular structures for investigation, standards of argumentation, and forms of notation and communication. Further, effective teachers understand how to organize subject matter and design learning experiences that make the content meaningful and coherent. They must have an understanding of appropriate learning outcomes at different levels (i.e., knowing what mastery of concepts and skills looks like) and must also have the ability to help scaffold student learning toward these outcomes. Along the way, they must be able to anticipate student misconceptions and know how to address them. Finally, they must be familiar with and be able to assess curricular resources available to

schools. To foster this pedagogical content knowledge (Shulman, 1986; Shulman, 1987), teacher preparation programs operate within the space between content and pedagogy, enabling teacher candidates to translate the knowledge gained in their disciplinary coursework into rich and developmentally appropriate learning experiences for students.

One effective tool with which we gauge candidates' evolving pedagogical content knowledge is a task that requires candidates to critique and modify available instructional resources to serve key learning outcomes in particular disciplines. Table 1 presents an example.

Table 1. *Textbook Modifications (assignment excerpt)*

Throughout the semester, you have been given opportunities in class to critique textbook lessons based on effective ways to teach certain mathematical concepts. This assignment is a compilation of these activities in a more formal format.

Describe how you will modify lessons taken from a textbook. Select three different lessons in sequence.

- Provide an overview of the general goals and objectives for the lessons as you see them.
- Critique each component on the page. This includes your assessment of its worth and usefulness.
- Reflect back on each lesson and describe generally how you would teach the lesson, including which aspects of the materials you would use, which you would delete, which you would modify, how you would modify them, **and your rationale for each decision.**

Your objective is not to write lesson plans. You are simply critiquing available information and deciding how you might best use information from textbook publishers.

Successful performance of this task requires that candidates not only understand core concepts in mathematics, but also be able to anticipate students' challenges in learning these concepts. Hence, the task provides evidence of candidates' knowledge of a content area, their understanding

of the development of conceptual understanding within the content area (and the misconceptions that typically arise), and their ability to anticipate the needs of a range of learners. As implied in this assignment, understanding of this bridge from disciplinary content to student learning is essential, whether teachers are generating new instructional materials or working with resources provided by their schools or districts.

Developing Understandings of Learners and Learning. The transformation of disciplinary content into rich, engaging, and effective learning experiences requires knowledge not only of content, but also of learners and their development. "Informed perspectives on development and learning provide necessary frameworks for understanding students, designing appropriate learning activities, justifying pedagogical decisions and actions, and communicating with parents, students, administrators, and colleagues" (Feiman-Nemser, 2001, p. 1018). As Feiman-Nemser further argues, given the fact that many teachers work with "students whose racial, cultural, and socioeconomic backgrounds differ markedly from their own," it is particularly important that candidates "cultivate the tools and dispositions to learn about students, their families, and communities and to build on this knowledge in teaching and learning" (p. 1018).

To promote understanding about learning and learners, during student teaching each teacher candidate in the elementary education program is expected to collect data on one focal student (e.g., artifacts such as samples of student work; assessments; video- and audiotapes; interviews with the classroom teacher, parents, other community members). In this process, teacher candidates are asked to closely examine social, historical, and cultural aspects of the classroom, school, home, and community that may influence their focal students' learning. After discussions with mentors and peers, each candidate is expected to analyze student work in the areas of literacy and math in light of these examinations and make appropriate accommodations in her/his practice to address the specific needs of her/his focal student. Each candidate then writes a collaborative paper within peer groups of about three or four student teachers that considers how their practices affected student achievement over time. The collaborative nature of this assignment offers opportunities for student teachers to engage in professional dialogue focused on solving problem situations they

encounter on a day-to-day and long-term basis. In addition, this assignment provides evidence of candidates' understanding of learning as a cultural practice, the relationship between culture and institutional contexts, and the inseparability of learning from the communities that support it.

Secondary education candidates have opportunities to examine their beliefs, values, and assumptions related to diversity as they pursue field experiences in a variety of urban school settings and in community settings serving children and youth. During their junior year practicum, candidates engage in a 30-hour service-learning program with agencies serving teens who are homeless, dropouts from traditional high schools, and in recovery from substance abuse. Structured, written reflections recorded prior to, during, and following the placement, as well as pre/post conferences and performance improvement plans, provide important evidence of students' evolving dispositions towards students, families, communities, and diversity.

In the special education program at Peabody, candidates' changing understanding of learners and their development is captured in the repeated assignments to create Individualized Educational Plans (IEPs). In schools, an Individualized Educational Plan (IEP) is developed for each student who receives special education services. Writing an effective IEP demonstrates knowledge and skills in several areas: (a) special education federal law, (b) formal and informal assessment techniques, (c) student's current academic and behavioral strengths and weaknesses, (d) effective curriculum and strategies, (e) the generation of realistic annual goals and short term objectives, (f) appropriate programs and services, and (g) evaluation methods. Beginning in their sophomore year, undergraduate teacher education candidates write IEPs based on mock student data. During their student teaching experience, candidates are required to write four IEPs for students in their classrooms. These IEPs are evaluated via a rubric that measures each of the seven elements listed. Among sophomores and juniors, we look for evidence of growth in the areas of assessment, identifying current strengths and weaknesses, and writing goals and objectives. During student teaching experience, when teacher candidates have opportunities to work with their students over an extended period of time and evaluate their progress, we especially look for growth in identifying proper curriculum and progress monitoring strategies.

Another core task in Special Education draws on the learning outcomes of P-12 students as evidence of candidates' ability to design appropriate learning experiences for their students. Teacher candidates in the undergraduate program complete a 30-hour field-based experience tied to a mathematics methods course. Students are required to select a concept and develop 20 one-hour lessons to teach, including detailed lesson plans for each lesson. These lesson plans are evaluated by faculty in the areas of: (a) appropriate objectives; (b) learner characteristics; (c) cultural awareness; (d) instructional modifications; (e) appropriate pacing and chunking; (f) appropriate instructional methods; (g) assessment of background knowledge; (h) use of modeling, guided practice, and independent practice; (i) student evaluation tools; and (j) teacher evaluation tools. Prior to beginning the lessons, the P-12 students are given a pre-test of the skills to be covered in the 20 lessons. After teaching each lesson, students are required to write a reflection of their lesson, focusing both on their students' learning and their teaching. At the end of the 20 lessons P-12 students are given a post-test to assess growth over the 10-week time period.

Developing a Beginning Repertoire. Wasley, Hampel, and Clark (1997) define a teaching repertoire as "a variety of techniques, skills, and approaches in all dimensions of education—curriculum, instruction and assessment—that teachers have at their fingertips to stimulate the growth of the children with whom they work" (cited in Feiman-Nemser, 2001, p. 1018). As Feiman-Nemser argues, a beginning repertoire includes a limited number of quality curricular materials, models of teaching, and assessment techniques, *augmented by a solid understanding of how to choose and use any particular approach.*

In our masters-level internship program, where interns share two mentor teachers' classrooms for an entire school year (one semester in each), two action research projects provide useful evidence of both repertoire and candidates' ability to adjust approaches based on student learning outcomes. In these projects candidates collect data to study the effectiveness of a chosen teaching strategy on student learning. The teacher intern collects pre-data (student scores/grades) before implementing the new or revised strategy, expedites the change, and collects post-data to determine if there is improved student learning. The interns conduct the project and then present it to peers and a faculty committee for review and feedback. Key evaluation criteria focus on

the degree to which interns are able to make sense of and communicate student learning data and to use these data to plan appropriate next steps for instruction.

Although these assessment tools serve distinct purposes and programs, they have a common focus on authentic and complex performance tasks. This shared emphasis reflects our belief that the most compelling evidence of candidates' teaching abilities comes from assessment tasks that mirror the actual tasks of teaching. The criteria or markers of progress that guide our assessments of these performance tasks are based on research into teacher and student learning that spans three decades. We currently are working to fine-tune both the tasks and the evaluation criteria so that we can better (more systematically and precisely) analyze the evidence that these assessments provide.

As described in the final section, we further seek to mine these performance assessments—and to establish new ones—toward the goal of articulating an empirically validated developmental trajectory of learning to teach. Given the intricacy of the understandings and skills to be assessed, this research and development project will require the construction, test, and use of complex measures and will involve both longitudinal and "snapshot" views of candidates' achievements. To add empirical strength, we are currently seeking partners among other teacher education institutions: Carefully designed contrasts and comparisons can enhance the knowledge of the field as a whole about the trade-offs of various program choices.

Tying Teacher Performance to Student Learning Outcomes

The assessment tools described in the preceding section focus on a crucial but intermediate outcome in teacher education: teacher candidates' knowledge and skills. However, knowledge and skills do not necessarily or straightforwardly translate to improved teaching performance in classrooms. Moreover, improved teaching performance, although unquestionably valuable, is not an end in itself. Ultimately, the most compelling measure of teacher effectiveness (and thus the effectiveness of teacher education programs) is P-12 student learning and understanding. Yet, using such a distal measure as P-12 student learning to assess the effectiveness of teacher preparation entails

conceptual, methodological, and logistical obstacles. In this section, we sketch these obstacles and also highlight efforts underway or needed to address them.

Conceptual Challenges

If the ultimate goal is the sort of student learning and achievement described at the outset of this paper, then evaluating teacher education programs requires instruments and measures that can tap student learning and understanding. Currently, instruments and measures that are acceptable, efficient, and affordable are lacking. As a recent National Research Council report concludes, "Much hard work remains to focus psychometric model building on the critical features of models of cognition and learning and on observations that reveal meaningful cognitive processes in a particular domain" (National Research Council, 2001, p. 6). Many assessments sample widely among the knowledge and skills considered important for students at a particular grade level, but in doing so fail to tap a sufficiently wide range of competencies to fairly characterize student understanding within any one area of content. An assessment may be an excellent choice for a certain form of accountability (e.g., to find out at a statewide level whether students are being taught the knowledge and skills delineated in the state standards) but a very poor choice for other purposes (for example, to characterize the level of understanding achieved by students in a particular teacher's class), and yet the same assessment may be used to serve both purposes. Often, curriculum, instruction, and assessment are poorly aligned so that what is tested bears little relation to what has been taught (Webb, 1997). If so, student test scores would clearly be poor choices to serve as indices of the quality of teaching they had received.

Even with the best of assessments, the road to the outcome of student learning is long and indirect, with unknown amounts of "slippage" between each of the links. Candidates in a teacher education program presumably acquire knowledge, skills, and values. These are expected, in time, to affect the teaching practices of graduates once they are working in their own classrooms. Yet, little is known about the relationship between what preservice teachers know and understand and what those same individuals can do later on as inservice teachers.

(Indeed, links between beliefs and knowledge, on the one hand, and teaching are notoriously tricky to establish even at one point in time. What teachers articulate often fails to match what they put into practice.) A second important link is between teaching and student learning. Teachers are one influence among many on their students' learning, and inadequate measures of student learning make that link even more difficult to establish and characterize.

Methodological Challenges

Assuming, for the moment, that adequate measures of student learning are in place, we confront a number of methodological challenges of linking P-12 student learning with specific teachers at a given grade level. In the last decade, there has been increasing interest in value-added modeling based on the work of William Sanders and implemented in Tennessee. The Sanders model used in Tennessee, TVAAS (Tennessee Value-Added Assessment System), attributes student gains to teacher effectiveness and regards teacher effectiveness as both additive and cumulative. A common theme in interpreting TVAAS is "that teachers, not students, are responsible for learning and that teachers hold the responsibility to produce measurable progress in learning outcomes" (Kupermintz, 2002). Although the Sanders model makes exciting promises, the requirements of such models are extensive, and methodology has yet to overcome significant concerns (McCaffrey et al., 2003; Kupermintz, 2002).

One concern is the degree to which changes in student test scores can be isolated to the effects of having a particular teacher—rather than, for example, tutoring, summer school, and other supports beyond the classroom. A second relates to the tautological definition of "teacher effectiveness." Sanders and Horn (1998) define teacher effectiveness as student academic gain. However, their model posits student academic gain as a separate and dependent variable of teacher effectiveness (Kupermintz, 2002). Finally, there are significant unknowns regarding Sanders' handling of inevitable gaps and errors in data—decisions about how to handle errors can significantly affect the results (McCaffrey et al., 2003; Ludlow, 2005).

Research is currently underway that seeks to evaluate the strengths and weaknesses of various models, test potential sources of errors, and determine consequential validity (Ludlow, 2005). Essential to the success of this research will be the broad establishment of nonproprietary databases that contain longitudinal data on student achievement in the form of individual student records and that are accessible for purposes of research and validation of emerging models. Data will need to be formatted and coded to allow appropriate disaggregation—not only by teacher and school, but also by school context, student socioeconomic and sociocultural attributes, etc. Moreover, these databases will need to be designed to handle the more sophisticated tools and measures required to capture student learning of more complex content.

Logistical Challenges

Perhaps the most immediate logistical challenge is mobility among teachers and their students. Smith and Ingersoll (2004) reported that in 1999-2000, 29% of new teachers "either changed schools at the end of the year (15%) or left teaching altogether (14%)" (p. 693). In urban systems, student mobility rates within a school can reach 80% over one academic year. Tracking impact on student learning is nearly impossible when new teachers move frequently and the student groupings are not sustained over a period of at least a year. Yet another logistical concern lies in difficulty of gaining access to relevant data; privacy concerns may complicate the collection of scores for students of individual teachers, and school systems may be hesitant to share data with schools of higher education. Overcoming these obstacles will require broad collaboration among school districts, teacher unions, state education agencies, and teacher education programs.

An Agenda for Research and Development: Toward Evidenced-Based Assessment in Teacher Education

In an effort to address these challenges, at Peabody we are embarking on a program of research and development that will enable us to design and implement an evidence-based approach to teacher education assessment. By systematically collecting and analyzing data

on teacher candidates' program achievements (including work with students) and the achievement outcomes of our graduates' P-12 students and simultaneously working to establish a series of robust measures for learning outcomes, we seek to create a system for quality assessment of our programs. Perhaps more important for the field in the long run, we also seek to provide evidence on the design trade-offs of particular teacher education program components in relation to P-12 student learning. Our efforts to collect and analyze data on our students' abilities to improve P-12 student learning are organized around four essential tasks. In the following section we briefly explain these tasks and the action steps to pursue them.

TASK 1: Identifying and Refining Current Tools and Developing New Ones

What evidence will inform our understanding of our students' developing ability to support student learning? How can we improve the quality of the evidence we get?

Action Steps:
- Gather and review course syllabi and core assignment descriptions for teacher education courses and select required courses from Arts and Sciences. Interview faculty to get detailed understanding of assignment goals and parameters and their alignment with program competencies.
- Target particularly promising core assignments at various levels across programs. Work with faculty to collect student work samples for these assignments.
- Analyze candidate work samples to understand the nature and quality of evidence these offer about teacher education candidates' learning and understanding of their students' learning. Based on analysis of work samples, identify core assignments that effectively provide (or have the potential to provide) rich/robust evidence of candidates' and P-12 students' learning and articulate features of these assignments. Consider feasibility of focusing on courses related to mathematics and English/language arts.

- Work with course instructors to refine undergraduate and masters' core assignments to increase their potential for providing evidence of learning.
- "Implement" revised core assignments in undergraduate and masters level courses. Provide coaching support in these classrooms to collect strong work samples.
- Work with recent graduates to identify work samples— including unit plans and artifacts and samples of *their students'* work—and pursue similar analysis.

TASK II: Establishing Measures

How can we differentiate levels of understanding and achievement as reflected in both preservice teachers' work and practice *and* the work of their students?

Action Steps:

- Research/pursue resources and practices at other institutions.
- Develop assessment measures for complex, open-ended, student generated tasks (for teacher education and ultimately P-12 students) related to specific literacies (e.g., mathematics and language arts). These measures should be informed by and should inform our work in articulating a tentative trajectory for preservice teacher development.
- Examine Terra Nova, Stanford 9, and other commonly used P-12 standardized assessments to see whether there are potentially fruitful strands (e.g., problem-solving, interpretation of data/graphs) that we might focus on in linking preparation and practice of our graduates to the achievement of their (elementary and secondary) students.

TASK III: Establishing Infrastructure to Support Our Learning and Systematic Analysis

What support structures/resources are needed to enable ongoing systematic inquiry and assessment of student progress and program quality?

Action Steps:
- Review organization of faculty roles and responsibilities with regard to teacher education and ongoing assessment.
- Develop a data collection and analysis plan that identifies critical assessment points, establishes a structure to support the timely collection and monitoring of tasks and data, and articulates a feedback mechanism to improve both the quality of the assessments and the data collection process itself.
- Establish an external data review team that meets periodically to score candidates and P-12 student performance data, providing a check of reliability and validity of internal processes as well as feedback on the assessment tasks.
- Seek funding to augment resources as necessary.

TASK IV: Developing and Sustaining an Evidenced-Based Culture

What kinds of structures and routines must we establish for ongoing review of evidence to inform program development and quality assessment?

Action Steps:
- Identify team(s) and timeline for ongoing review and analysis (who are the players?).
- Meet at regular intervals to review evidence reflected in student work/practice in order to articulate:
  - *What have we learned?*
  - *What do we need to know?*
  - *What programmatic changes do we need to make?*
- Develop strategies for communicating review and analysis of evidence to stakeholders. These strategies should ultimately support writing and publication.

In pursuing this agenda, we hope not only to develop a robust, but workable strategy for assessment of teacher education candidates and programs, but also to contribute to the "reconnection" of policy

and teacher education research. Despite some profound differences in views of teaching and learning, federal policymakers and university-based educators share a view of students' "right to learn" (Darling-Hammond, 1997) and of the urgency of preparing teachers who can honor that right. In reviewing the current status (and limits) of assessment in teacher preparation, our intent is both to affirm the goal of valid and meaningful assessment that links teaching with learning and to advance dialogue and research toward this goal.

References

American Association for the Advancement of Science. (1993). *Benchmarks for science literacy*. Washington, DC: Author.

Carpenter, T. P., & Lehrer, R. (1999). Teaching and learning mathematics with understanding. In E. Fennema & T. A. Romberg (Eds.), *Mathematics classrooms that promote understanding* (pp. 319-320). Mahwah, NJ: Erlbaum.

Center for Civic Education. (1994). *National standards for civics and government*. Calabasas, CA: CCE.

Cochran-Smith, M. (2003). Assessing assessment in teacher education. *Journal of Teacher Education, 103*(6), 187-191.

Cochran-Smith, M., & Fries, K. (2005). Researching teacher education in changing times: Politics and paradigms. *Studying teacher education: The report of the AERA panel on research and teacher education*.

Consortium of National Arts Education Associations. (1994). *The national standards for arts education*. Reston, VA: National Association for Music Education.

Darling-Hammond, L. (1997). *The right to learn: A blueprint for creating schools that work*. San Francisco: Jossey-Bass.

Elliott, E. J. (2003). *Assessing education candidate performance: A look at changing practices*. Washington, DC: National Council for Accreditation of Teacher Education.

Feiman-Nemser, S. (2001). From preparation to practice: Designing a continuum to strengthen and sustain teaching. *Teachers College Record, 103*(6), 1013-1055.

Gamoran, A., Anderson, C. W., Quiroz, P. A., Secada, W. G., Williams, T., & Ashman, S. (Eds.). (2003). *Transforming teaching in math and science.* New York: Teachers College Press.

Geography Education Standards Project. (1994). *Geography for life: National geography standards.* Washington, DC: National Geographic Research and Exploration.

Imig, D. (2004). *Contextual scan.* Unpublished manuscript. Washington, DC: American Association of Colleges for Teacher Education.

Interstate New Teacher Assessment and Support Consortium. (1992). *Model standards for beginning teacher licensing and development: A resource for state dialogue.* Washington, DC: Council for Chief State School Officers.

Kupermintz, H. (2002). Value-added assessment of teachers. In A. Molnar (Ed.), *School reform proposals: The research evidence* (Chapter 11). Retrieved February 2, 2005, from http://www.asu.edu/educ/epsl/EPRU/documents/EPRU%202002-101/Chapter%2011-Kupermintz-Final.htm#_ednref31

Lortie, D. (1975). *Schoolteacher: A sociological study.* Chicago: University of Chicago Press.

Ludlow, L. (2005, January 2). *Value added modeling: Where did it come from, where is it going?* Presentation to the AACTE Winter Institute on Assessment in Teacher Education, Cancun, Mexico.

McCaffrey, D., Lockwood, J. R., Koretz, D., & Hamilton, L. (2003). *Evaluating value-added models for teacher accountability.* New York: Rand Corporation.

National Center on Education and the Economy. (1997). *New standards performance standards.* Orlando, FL: Harcourt.

National Center for History in Schools. (1996). *National standards for history.* Los Angeles: NCHS.

National Council on Economics Education. (1997). *Voluntary national content standards in economics.* New York: NCEE.

National Council for the Social Studies. (1994). *Curriculum standards for social studies: Expectations of excellence.* Washington, DC: NCSS.

National Council of Teachers of English. (1996). *Standards for the English language arts.* Newark, DE: International Reading Association.

National Research Council. (2001). *Knowing what students know: The science and design of educational assessment.* Washington, DC: National Academy Press.

National Research Council. (1996). *National science education standards.* Washington, DC: National Academy Press.

National Research Council. (2000). *How people learn: Brain, mind, experience, and school.* Washington, DC: National Academy Press.

No Child Left Behind Act: Reauthorization of the Elementary and Secondary Act, Pub. L. No. 107-110. (2002, January 8). Retrieved June 10, 2002, from http://www.ed.gov

Programme for International Student Assessment. (2004). *Learning for tomorrow's world: First results from PISA 2003.* Paris: OECD.

Sanders, W. L., & Horn, S. P. (1998). Research findings from the Tennessee Value Added Assessment System (TVAAS) database: Implications for educational evaluation and research. *Journal of Personnel Evaluation in Education, 12*(3), 247-256.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4-14.

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review, 57,* 1-22.

Smith, T. M., & Ingersoll, R. M. (2004). What are the effects of induction and mentoring on beginning teacher turnover. *American Educational Research Journal, 41*(3), 681-714.

Wasley, P., Hampel, R., & Clark, R. (1997). The puzzle of whole-school change. *Phi Delta Kappan, 78*(9), 690-697.

Webb, N. L. (1997). *Criteria for alignment of expectations and assessments in mathematics and science education.* National Institute for Science Education and Council of Chief State School Officers Research Monograph No. 6. Washington, DC: Council of Chief State School Officers.

Marcy Singer-Gabella is Research Assistant Professor of Education in the Department of Teaching and Learning in the Peabody College of Education at Vanderbilt University.

Chris Iddings is Assistant Clinical Professor of Elementary Education and Director of Elementary Education Programs in the Peabody College of Education at Vanderbilt University.

Kim Paulsen is Assistant Professor of Special Education in the Peabody College of Education at Vanderbilt University.

Margaret Smithey is a Senior Lecturer in the Department of Teaching and Learning in the Peabody College of Education at Vanderbilt University.

Marie Hardenbrook is Assistant Professor in the Practice of Education in the Peabody College of Education at Vanderbilt University.

Amy Palmeri is Assistant Professor in the Practice of Education in the Peabody College of Education at Vanderbilt University.

Leona Schauble is Professor of Education and Chair of the Department of Teaching and Learning in the Peabody College of Education at Vanderbilt University.

Camilla Benbow is the Patricia and Rodes Hart Dean of Education and Human Development at Vanderbilt University.

Endnotes

1   This policy context has stimulated efforts such as that of the Teachers for a New Era (TNE) consortium to develop "value-added tracking systems to assess the impact of program graduates on pupils' learning" (Cochran-Smith, 2003).

2   See: for reading and language arts, National Council of Teachers of English (1996); for science, American Association for the Advancement of Science (1993) and National Research Council (1996); for social studies, National Council for the Social Studies (1994); for civics and government, Center for Civic Education (1994); for economics, National Council on Economic Education (1997); for geography, Geography Education Standards Project (1994); for history, National Center for History in Schools (1996); for the arts, Consortium of National Arts Education Associations (1994). See also the National Center for Education and the Economy (1997) in English/language arts, mathematics, science, and applied learning.

3   We must note that teaching for understanding is a challenging goal, one that has in the past been reserved for students who are considered "advanced" (Gamoran et al., 2003).

4   Currently we are looking at ways in which Feiman-Nemser's fifth task might emerge and evolve within the context of other developmental tasks. For example, developing a view of teaching that encompasses ongoing professional inquiry and learning and acquiring tools with which to enact such a vision are intertwined. While at some level all five tasks are interconnected, we suspect that the essential features of this one are different in kind. Therefore, in the section entitled "Analyzing beliefs and forming new visions" (pp. 121-122) and in our ongoing investigation (see final section of paper, pp. 130-134) we do not treat this task separately.