# SCALING UP RESEARCH IN TEACHER EDUCATION NEW DEMANDS ON THEORY, MEASUREMENT, AND DESIGN

H. Del Schalock Mark D. Schalock Robert Ayres Western Oregon University

> A recent report of the American Educational Research Association Panel on Research and Teacher Education confirms beyond question earlier findings exposing the limited utility of our research base in answering questions pertaining to policy or practice concerning preparation and licensing of teachers. Conditions accounting for this perplexing circumstance are described in detail by the panel, as are recommendations provided for overcoming them. A recent research project anticipating many of the recommendations led this article's authors to the view that several of the recommendations need added detail to be immediately helpful to the research community and that further recommendations are needed for "scaled-up" research called for by the panel. Accordingly, suggested additions and refinements, with accompanying rationale and examples, are proposed. A central theme of this article is the need to add explanatory power to teacher education research, with the accompanying caution that doing so brings added complexity to theory, measurement, and design.

# *Keywords:* teacher education research; teacher education theory; teaching and learning; teacher professional development

On completing a critique of more than 500 peerreviewed studies of preservice teacher education in the United States published largely between 1990 and 2003, and vetted carefully against explicit standards of quality, the American Educational Research Association (AERA) Panel on Research and Teacher Education (Cochran-Smith & Zeichner, 2005) crafted numerous recommendations for overcoming the limitations repeatedly found in the studies reviewed. The 1990 to 2003 time period corresponds to what Cochran-Smith and Fries (2005) referred to in the introductory chapter of the report as "the teacher education as a policy problem stage" in the evolution of teacher education research.

Conclusions and recommendations reported by the AERA Panel are framed against the background of findings from several previous reviews of teacher education research from the perspective of policy (Allen, 2003; Lauer, 2001; Rice, 2003; Wilson & Floden, 2002; Wilson, Floden, & Ferrini-Mundy, 2001). Cochran-Smith and Fries (2005) summarized findings from these reviews as follows:

Although there was some evidence that teacher preparation and certification had a positive impact on educational outcomes in some content areas and at certain school levels, the research base related to teacher education as policy was neither deep nor robust. Results were mixed in some areas, and there was virtually no reliable research in many other areas. . . [These syntheses] also were consistent in their conclusions about methodological and design problems and thus in their conclusions about what was needed—uniform measures of "impact" and "effectiveness" (Allen, 2003), more stable measures

Journal of Teacher Education, Vol. 57, No. 2, March/April 2006 102-119 DOI: 10.1177/0022487105285615

 $<sup>\</sup>ensuremath{\mathbb C}$  2006 by the American Association of Colleges for Teacher Education

of teacher knowledge and behavior (Wilson et al., 2001), designs that capture the interactions among various aspects of teacher quality (Rice, 2003), designs that directly link aspects of teacher preparation to pupil achievement (Wilson et al.), direct attention to elementary and middle school teaching and in areas beyond mathematics (Rice, 2003), and data linking information about individual teachers to actual performance rather than aggregating data at the school or district level (Allen, 2003). (p. 96)

Findings from each of the nine topical reviews conducted by the AERA Panel broadly confirm the findings of these preceding reviews, both with respect to relationships found between preparation and subsequent performance as a facilitator of learning in a classroom and in the focus and quality of research conducted.

The dual purpose of this article is to outline a study recently completed that confronts many of the methodological and design weaknesses pointed to in the previous quotation and to report lessons learned from undertaking research that attempts to do so. We anchor our discussion in specific recommendations made by the AERA Panel for "scaling up" research in teacher education to overcome the many weaknesses cited in the Cochran-Smith and Fries (2005) summary.

#### PANEL RECOMMENDATIONS

Recommendations made by the AERA Panel for strengthening research on teacher preparation and its effects carry considerably greater detail than those coming from previous reviews. Its recommendations fall into three broad groupings: research design and methodology, important topics to pursue that "are amenable to productive research," and infrastructure needed to develop a body of research that informs policy and practice.

We focus here on the AERA Panel's recommendations concerning theory, methodology, and design. These include

- 1. Clear and consistent definition of terms;
- Full description of data collection and analysis methods and the context in which research is conducted;
- 3. Research situated in relation to relevant theoretical frameworks;
- 4. Development of more programs of research;

- More attention to the impact of teacher education on teacher learning and teacher practices;
- 6. Research that connects teacher education to student learning;
- Total portfolio of studies that includes multi-disciplinary and multi-methodological approaches to studying the complexities of teacher education;
- Development of better measures of teacher knowledge and performance;
- Research that examines teacher preparation in different subjects in addition to mathematics and science and takes the subject into account when examining the effects of teacher education components and programs;
- More systematic analysis of clearly identifiable alternatives in teacher education using matching controls or random trials as separate studies or in conjunction with in-depth case studies;
- More in-depth multi-institutional case studies of teacher education programs and their components. (Zeichner, 2005, p. 740)

From the perspective of having engaged in research for several decades that focuses on the connections among teaching, teacher preparation, and K-12 learning (M. Schalock, 1987; M. Schalock & Schalock, 2004; D. Schalock, Schalock, & Girod, 1997), we view these recommendations as being solid and comprehensive. If implemented, they will go a long way toward overcoming the pervasive limitations now existing in the knowledge base informing teacher education policy and practice. From the perspective of designing and implementing such research, however, we find two shortcomings in the recommendations and discussion provided. First, interdependencies among the recommendations are not fully acknowledged and addressed. Clear and consistent definition of terms, for example, will help-though not much-unless they also are situated within relevant theoretical frameworks, translated into appropriate data systems (measures), and treated appropriately in related data analyses. Second, the recommendations and elaborating discussions remain broad and general. They point the way toward a "research agenda," which was an explicit aim of the AERA Panel (and is the focus of the final chapter in its report), but they do not carry the detail that provide guidelines to a researcher engaged in the actual design and implementation of a study

that is intended to reflect the recommendations advanced.

In this article, we attempt to move six of the recommendations made (1, 3, 5, 6, 7, and 8) toward the level of detail and connectedness we feel is needed for them to be concretely helpful in designing and implementing the kind of research that is called for. We also suggest several additional recommendations that have grown out of our experience with such research.

# A Longitudinal Study of Teacher Preparation Effects as a Point of Reference

The research that has shaped our reaction to the AERA Panel's recommendations was a 3year longitudinal study designed to trace the effects of several policy-designated features of teacher preparation programs on the thinking and classroom performance of beginning elementary teachers and the learning of their students. The study highlights many of the policy issues addressed by the AERA Panel in investigating the following: (a) Do teacher preparation programs have a demonstrable influence on the thinking and practice of early career teachers and the learning of their students? (b) If so, do differing emphases within programs have differing effects, and do these moderate or grow with teaching experience? and (c) To what extent do factors other than teacher preparation, for example, characteristics of the classroom and school context in which teaching and learning occur, influence differences observed in either the performance of early career teachers and/or the learning of their students? Although the study preceded the full impact of the No Child Left Behind Act of 2001 (data were collected from early 2000 through 2002), it was crafted in theory and method from the perspective of a standards-based, evidence-driven approach to schooling. The research is referred to here as the Teacher Effectiveness Study.

# Background and Context

The Teacher Effectiveness Study was conducted as the third phase of an 8-year investigation of teacher work sample methodology as a vehicle for meaningfully and defensibly connecting teaching and learning in a standardsbased school environment and the design of teacher preparation programs to support the methodology. The methodology was developed in Oregon for use in the initial preparation and licensing of teachers in response to the state's adoption of a "goal-based" approach to schooling in the early 1980s and subsequently a "standards-based" approach in the early 1990s. As experience was gained with these shifts in thinking about schooling, it became apparent that the work of teachers and students in such schools is considerably different and more demanding on both teachers and students than teaching and learning in the norm-referenced schools of the 20th century (Black & William, 1998; Bransford, Darling-Hammond, & LePage, 2005; Falk, 2000; Marzano, 1998; O'Shea, 2005).

These shifts highlighted the need to help prospective teachers learn to think systematically about connecting their teaching to designated standards for learning on the part of students and how to act systematically on the progress students were making toward these standards. The methodology of teacher work sampling was designed to serve both purposes, as well as provide evidence on the effectiveness of a prospective teacher in making these connections. A key feature of the methodology is providing evidence as to effectiveness in both the performance of related tasks and the learning progress made by students taught. The methodology is designed on 2- to 5-week units of instruction intended for all students in a regular classroom that are anchored to state- or district-established standards for learning. The instructional units are planned and implemented by prospective teachers, with review and approval by both college and school supervisors, with preinstruction and postinstruction analysis and reporting of student progress toward targeted learning outcomes.

By the mid-1980s, the methodology was sufficiently well developed and tested that the state's teacher licensing agency adopted it as part of a 1986 revision of standards for school personnel that reflected the new model of schooling being implemented (D. Schalock & Myton, 1989). By doing so, Oregon became the first state in the nation to require prospective teachers to demonstrate their ability to foster learning gains in students taught as a condition of initial licensure. By the mid-1990s, sufficient experience and research with the methodology within Oregon led to a decision by the provost and faculty at Western Oregon University to carry out the first phase of the cumulative 8year investigation of the merits of the methodology. This was a national invitational conference held to assess its merits as a vehicle for research connecting teaching, teacher preparation, and K-12 learning. Conference participants were cautious about the potential of the methodology as a vehicle for research but enthusiastic about its potential as a vehicle for teacher preparation and licensing. They recommended that funds be sought to carry out an in-depth investigation of its potential in this regard (D. Schalock & Schalock, 1995). The rationale for measures obtained from and findings from use of the methodology at Western Oregon University at this stage in its development are presented and critiqued in Millman's (1997) Grading Teachers, Grading Schools: Is Student Achievement a Valid **Evaluation Measure?** 

The conference and its recommendation led directly to the second phase in the investigation of the methodology. Funds were obtained to carry out a 3-year investigation of the methodology as a vehicle for teacher preparation and licensing. The investigation was guided and reviewed by a national advisory panel, chaired by Jason Millman until his death, and then by Daniel Stufflebeam. The work of the panel centered on providing an external assessment of the utility and defensibility of the methodology in serving three broad purposes: (a) helping prospective teachers learn to think and act systematically in connecting their teaching to the learning of students in a standards-based school environment; (b) providing defensible evidence on the effectiveness of prospective teachers in designing and carrying out units of instruction within a standards-based learning environment, including evidence on the learning progress made by each student in the classroom in which a unit is implemented; and (c) enhancing evidence-based decisions pertaining to teacher candidate progression through a preparation program, including recommendation for an initial license to teach. The panel gave solid support to using the methodology for all three purposes but cautioned that evidence generated through the methodology be treated as only one of multiple sources of evidence used in making a recommendation for licensure. A detailed description of the methodology in use at that time is provided in *Connecting Teaching and Learning: A Handbook for Teacher Educators on Teacher Work Sample Methodology* (Girod, 2002).

The Teacher Effectiveness Study was a direct outgrowth of Phase 2 activities. The logical next question was whether there were traceable effects from using the methodology in preparation programs designed specifically to support it. Because of licensing requirements, Oregon was a ready-made context for conducting such a study.

### Design and Implementation

The Teacher Effectiveness Study was a relatively small sample size investigation designed to determine whether teacher preparation programs having specific policy-related characteristics had any traceable and lasting effects on the thinking and practice of beginning teachers and the learning of their students. The three characteristics of teacher preparation programs of interest were

- a. a strong alignment with Oregon's design for standards-based schools,
- b. an overt focus on connecting teaching and learning, especially through teacher work sampling, and
- c. an ongoing developmental assessment of candidates against clear and public performance standards resembling those encountered in Oregon's K-12 schools.

Within this overall framework, we were primarily interested in whether differently prepared teachers thought about and practiced teaching differently as 1st-year teachers and whether there was a corresponding difference in the level and quality of their students' learning. We were also interested in determining whether any 1st-year effects were maintained in subsequent years of teaching and if not, to determine those factors contributing to this result. From a theoretical and empirical standpoint, we had some confidence that exposure to these program characteristics was a departure from traditional preparation and was likely to result in teachers who thought about and practiced teaching differently and in ways that would support student learning in a standards-based school.

To answer these questions, a longitudinal, causal-comparative design was employed to follow 76 beginning elementary teachers through their first 3 years of teaching. Participating teachers experienced a range of preparation programs both within Oregon and in other states. Through a combination of expert thirdparty ratings, extensive content analyses of program materials, and interviews with participants, participant programs were sorted by level of exposure to the three program emphases of interest. These analyses and resulting ratings resulted in 13 of the 76 first-year teachers being prepared in programs providing low exposure, 20 in programs providing moderate exposure, 16 in programs providing high exposure, and 27 in programs providing very high exposure. The study included 60 second-year teachers. Because of attrition, from both the demands of the study and the profession, we were able to follow an intact group of 45 teachers across all 3 years.

To adequately capture the full range of variables related to the thinking and practice of beginning teachers, the learning of their students and the related affective, contextual, and developmental variables of interest, we purposefully employed a mixed-method approach to data collection. We found, however, that few of the constructs and related variables we wished to investigate had available measures that we felt were adequate for our purposes and for the standards-based teaching environment in which we were working. Accordingly, instrument development became a major focus of the 1st year of the project. The development of the classroom observation instrument, for example, took more than a year for careful development and validation before it was used.

During the participants' 1st year of teaching, we conducted extensive structured interviews, administered a suite of attitudinal surveys, and administered a number of surveys and questionnaires to obtain descriptions of participants' teaching contexts, professional development activities, and perceived level of competence. We also conducted focused classroom observations and asked that a modified teacher work sample be prepared. These data collection activities were repeated for 2nd-year teachers participating in the study, with additional classroom observations and focus group interviews conducted in lieu of individual interviews. In the 3rd year, we continued to conduct multiple observations but did not ask participants to prepare a third work sample.

This "multi-methodological" approach to investigating the impact of teacher education reflects the complexities addressed in AERA Panel Recommendation 7. To illustrate the magnitude and diversity of data collected in an effort to trace causally the connections among teacher preparation, teaching, and the learning of K-6 students, the variables we viewed as being necessary to attend to are listed in Table 1. To highlight the complexities involved in dealing with this range of data in a single study, we have grouped the measures under the labels used to order the variables meaningfully for purposes of analysis and interpretation.

Although space does not permit a full reporting of findings from the study, the major conclusion reached was that the emphases in teacher preparation investigated were positively related to the classroom performance of 1styear elementary teachers, with effect sizes in the moderate to large range. These positive relationships, however, essentially disappeared in 2nd-year teachers and in some cases reversed themselves in 3rd-year teachers. No corresponding positive relationships between measures of student learning and program emphases or observed classroom performance were found. We believe this was, in part, because of limitations in obtaining consistent, high-quality measures of student learning across partici-

TABLE I Valiables Addressed in the reacher Enectiveness Stud	TABLE 1	Variables Addressed in the Teacher Effectiveness S	tudy
--	---------	--	------

Variable	Organizing Classification	Measure
Teacher preparation program emphases	Independent	Project staff ratings
Preparation program context	Level 1 moderator	Program-supplied descriptions
Teachers' commitment to teaching	Level 1 moderator	Survey responses
Teachers' traits and beliefs	Level 1 dependent	
Accountability for student learning		Survey/interview responses
Efficacy beliefs for teaching profession		Self-report/survey
Personal teaching efficacy		Self-report/survey
Efficacy in helping students overcome learning difficulties		Self-report/survey
Teachers' responses to demands of teaching	Level II moderator	
Sense of personal accomplishment		Maslach Burnout Inventory
Emotional exhaustion		Maslach Burnout Inventory
Depersonalization of students		Maslach Burnout Inventory
Stress		Self-report/Survey
Contextual factors	Intervening	
School/district characteristics		Survey/interview responses
Classroom/collegial characteristics		Survey/interview responses
Student characteristics		Survey/interview responses
Classroom management/instruction	Level II dependent	Structured observations
Teacher thinking/reflection on practice	Level II dependent	Extended work sample
		Interviews/focus groups
Teacher assessments of student learning	Level II dependent	Extended work sample
Teacher school improvement/professional development activities	Level II dependent	Survey/interview responses
Teacher assessment, analysis, and reporting of student progress in learning	Level III dependent	
Student learning progress through units of instruction		Extended work sample
Student progress toward standards for learning		Extended work sample
Additional indicators of teacher impact on learning	Level III dependent	
Student engagement in learning		Structured observations
Taxonomic levels of intellectual work		Structured observations
Student understanding/exploration of meaning within and across subject areas		Structured observations
Student interest in content to be learned		Structured observations

pants. Furthermore, there were a number of unanticipated findings related to the strength of contextual and affective variables on the development of teachers (M. Schalock & Schalock, 2004).

# LESSONS LEARNED AND RELATED RECOMMENDATIONS

From the perspective of this article, among the strongest lessons learned from the sequence of studies discussed are the implications such research carries for the work of researchers, teacher educators, and school personnel. They are enormous. Related complexities and their interactions among the development of measures, data collection tasks, and analysis and interpretation demands seem endless, and it is this reality in undertaking such research that the AERA Panel recommendations fail to convey. We now turn to our view of the kind of additional detail needed for the panel's recommendations on theory, design, and measurement to be immediately useful to researchers engaged in designing and implementing studies reflecting the panel's intent.

It is our assumption that the elaborations and extensions that follow will be among only the first of many needed for the guiding vision of the AERA Panel to be fully realized. As we address *theory*, we focus on the panel's first and third recommendations (clear and consistent definitions, and situating research in related theoretical frameworks); as we address *design*, we focus on its fifth and sixth recommendations (more attention to the impact of teacher education on teacher learning and teacher practice, and research that connects teacher education to student learning); and as we address *measures*, we focus on the panel's seventh and eighth recommendations (studies that include multidisciplinary and multimethodological approaches to investigating the complexities of teacher education, and the development of better measures of teacher knowledge and performance). Because of the interdependencies among theory, design, and measurement in research generally, and among the six panel recommendations addressed specifically, some cross-referencing among suggested elaborations has been provided.

# Building and Using Theory

In undertaking the project, we searched existing literature for strong theoretical guidance, both for the constructs and variables to include and their measurement and likely relationships, but at the time found little that was helpful. Connecting teaching, teacher preparation, and K-12 learning in a standards-based (pre–No Child Left Behind Act of 2001) school environment involves a long chain of conceptual and procedural connections, and these simply were not to be found in the literature available in the late 1990s.

Nor are such connections available today as guides to either research or practice. This is the case even with the recommendations of the AERA Panel and even though the demands of the No Child Left Behind Act of 2001, the National Council for Accreditation of Teacher Education's 2000 standards for the accreditation of teacher preparation programs, and the requirements of many state teacher licensing policies have been crafted on the assumption that such connections do (or should and will) exist.

The absence of such connections should not be a great surprise. As Cochran-Smith (2005) acknowledged in her editorial in a recent issue of the *Journal of Teacher Education*,

Rigorous outcomes research in teacher education (and in many other complex enterprises) is difficult and expensive to do. To get from teacher education to impact on pupil's learning requires a chain of evidence with several critical links: empirical evidence demonstrating the link between teacher preparation programs and teacher candidates' learning, empirical evidence demonstrating the link between teacher candidates' learning and their practices in actual classrooms, and empirical evidence demonstrating the link between graduates' practices and what and how much their pupils learn. Individually, each of these links is complex and challenging to estimate. When they are combined, the challenges are multiplied. (p. 303)

The vision contained in the AERA Panel's recommendations centers on making these connections known and understood sufficiently well to be able to replicate their application across multiple school and preparation contexts, across a broad range of K-12 students, and across a broad range of learning outcomes and subject areas.

According to Shoemaker, Tankard, and Lasorsa (2004), the first step in theory building is concept clarification in the form of definitions and examples. The second step is to combine concepts into theoretical statements and their rationale. Accepting this to be the case, following are six elaborations of teacher education– related theory that we think researchers investigating the connections among teachers, teacher professional development, and K-12 student learning would find helpful.

First, focus conceptual and theoretical attention on the connections among teaching, the professional development continuum of teachers, and the learning of K-12 students that carry sufficient detail and coherence to explore causal connections. This will require attention to such intervening and moderator variables as those shown in Table 1, as well as the distinction made in the table between Levels I, II, and III dependent variables. Each term in this complex set of variables requires extensive definition and the connections among them can take many forms and combinations, but central to the process of theory building is making definitions and hypothesized connections sufficiently clear that they can be tested and refined empirically. Our efforts would have been facilitated by such clarity had it existed when we began.

Second, in articulating definitions and connections, differentiate clearly between (a) mastery of enabling knowledge and skills and (b) proficiency in knowledge and skill integration, alignment, and adaptation to accommodate the learning needs of students in a standards-based



FIGURE 1: Teacher Learning as Mastery of Enabling Knowledge and Skills and Developing Expertise in Their Application

learning environment. The mastery of enabling knowledge and skills in teacher preparation programs is carried out in large part through other than education courses and represents one kind of learning. Integrating and developing skills in applying these professional underpinnings in developmentally appropriate ways is accomplished largely through education courses and practicum experiences. Figure 1 represents a way of sorting these distinctions schematically and conveying their interdependencies. It also conveys our view of the appropriateness of thinking about "teaching as a clinical profession" as emphasized in the Carnegie Corporation's Teachers for a New Era initiative.<sup>1</sup>

Third, develop path diagrams to show the expected size (in terms of beta weights) and direction (positive or negative) of the relationships anticipated among variables within each network of variables to be investigated and use these diagrams as a guide to the analysis of data collected to verify or refine the relationships predicted. Reporting findings of studies conducted within such conceptual networks (theoretical frameworks) permits replications across similar or differing contexts, similar or differing subject areas, and similar or differing groups of

teachers involved. It also permits studies of the effects of adding, deleting, or substituting variables within a network and tracing effects. This gradual, step-by-step but systematic and replicable process is how productive theory is developed and how a knowledge base that has explanatory power, as well as utility for policy and practice, is assembled.

Fourth, we think the affective dimensions of teaching and the professional development of teachers needs to be addressed systematically by researchers attempting to connect teaching, teacher preparation, and K-12 learning (hereafter referred to as CTPL research). Much is made in the literature of the intellectually demanding and emotionally and physically exhausting nature of the first years of teaching (Johnson & The Project on the Next Generation of Teachers, 2004; Patterson, 2005). Although aware of this literature prior to the Teacher Effectiveness Study, we were surprised by both the range and depth of the affective responses of the teachers we studied to their work. We also were dismayed at how often school environments did little to alleviate conditions giving rise to such responses and failed to support beginning teachers in coping with them. Because of the

pervasiveness and intensity of affect-related issues we encountered, we think an additional recommendation in this regard is fully warranted. The affective variables attended to in the effectiveness study are listed in Table 1 as Level I dependent variables. The most complete and recent analysis of these variables is found in Ayres and Cuthbertson (2004).<sup>2</sup>

Fifth, within all of the above, attention has to be given to theory-anchored qualities or characteristics of students that moderate a teacher's impact on learning, for example, progress that has been made toward targeted goals for learning prior to the period of instruction studied or history of learning concerning each targeted learning goal. CTPL research also needs to attend to student interest in each goal area and motivation toward and the habits of work concerning school-based learning generally. These factors have a powerful influence on student learning and, thus, on inferences drawn about the effectiveness of a teacher or the connections found among teaching, teacher preparation, and student learning.

Finally, frame all of the above in terms of the demands on teachers and students of a standardsbased, evidence-driven approach to schooling. Although this design for schools is still in its infancy, and its translation into the No Child Left Behind Act (2001) was in many respects overreaching and misguided, we believe it represents the general direction that has been taking shape in American education since the *Nation at Risk* report was published in the early 1980s (Gordon, 2004). The review and reauthorization of the act is scheduled for 2007, and we expect it to be refined appreciably at that time, but we also expect the fundamental intent and direction of the law to be maintained.

# **Expanding Designs**

One of the major problems that has plagued teacher education research from the perspective of its utility in informing practice, and to some extent policy, is the "black box" design of most large-scale studies connecting teachers or teacher preparation to student learning. For example, the groundbreaking work of Sanders and his colleagues in Tennessee (Sanders & Horn, 1994; Sanders & Rivers, 1996) carries no information on the knowledge, skills, or classroom performance of teachers studied and little information on the classroom and school contexts in which teaching and learning occurred. The same limitations pertain to the frequently cited studies of Hanushek and Pace (1995), Monk (1994), and Goldhaber and Brewer (2000) on teacher contributions to student learning. Such studies carry little explanatory power as to why, how, or what within teaching or teacher preparation account for relationships found and, thus, have limited utility in guiding or refining policy, practice, or research.<sup>3</sup>

Two other reasons often are cited for the weakness in teacher education effects research. One is the paucity of longitudinal studies that have been conducted, with the consequence that changes in teacher characteristics and performance across time, subject areas, and classroom/school contexts are known only in broad outline. The other is simply the complexity of teaching. Cochran-Smith (2003) has used the phrase "the unforgiving complexity of teaching" to convey the realities of the multidimensional, ever-changing context that confronts a teacher in today's schools. As Cochran-Smith described it, teaching

is not simply good or bad, right or wrong, working or failing. Although absolutes and dichotomies such as these are popular in the headlines and campaign slogans, they are limited in their usefulness.... They ignore almost completely the nuances of "good" (or "bad") teaching of real students collected in actual classrooms in the context of particular time and places. They mistake reductionism for clarity, myopia for insight. And, as Elmore (2002) suggests, they "utterly fail" to appreciate the institutional realities and complexities of accountability in various schools and school districts as well as in particular states. (p. 4)

If one adds to this the different kinds of learning outcomes to be accomplished by students in different subject areas, and the enormous diversity in talents, dispositions, and histories that students bring to what is to be learned within each subject studied, the realities involved in helping each student progress toward high standards for learning make the image evoked by Cochran-Smith's phrasing totally fitting.

Designs accommodating this complexity are required by research connecting teaching, teacher preparation, and K-12 learning that seeks causal explanation. In designing such research, as much attention and care needs to be given to intervening and moderator variables as independent (experimental, treatment) and dependent (outcome) variables. The examples of such variables cited in Table 1 are illustrative of moderator and intervening variables that can have a powerful influence on teacher performance and student learning and, thus, warrant thoughtful attention. So are the affective dimensions of teaching and teacher development listed in Table 1 as Level I dependent variables. The challenge posed to researchers in bringing order and meaning to the complexities encountered in designing and implementing CTPL research is illustrated by the sorting shown in Table 1.

In addition to moderator and intervening variables, detailed attention needs to be given in CTPL research to often ignored dimensions of sampling, for example, the sample of schools and districts in which teachers are to be studied or the sample of content areas in which students and teachers are working when classrooms are being observed. Attention also should be given to the sample of instructional periods observed in each content area of interest and the distribution of instructional periods observed throughout the school year. Decisions concerning each of these variables carry implications of importance for inferences drawn about a teacher's effectiveness and/or the connections drawn between teaching, teacher preparation, and the learning of K-12 students.

Another dimension of sampling that needs to be addressed in designs for CTPL research are the specific learning goals (standards) students are working toward or are expected to accomplish within the subject areas and time blocks selected for study. Clarity on this issue, and decisions made with respect to it, can affect inferences drawn about a teacher's effectiveness or connections among teaching, teacher preparation, and K-12 learning as powerfully as decisions made concerning the context in which teaching and learning is studied.

Finally, all CTPL research should be crafted within a longitudinal design, but in doing so understand fully the complexities that are added to each of the other design considerations outlined and the implications for measurement, analysis, and interpretation that follow therefrom. Teachers change with experience, and school contexts change continuously in ways that affect both teacher work and student learning. The only way these interactions can be determined is longitudinally.

This last recommendation brings an accompanying alert for the designers of CTPL research: Both before and following initial licensure, give detailed attention to (a) stage in professional maturation and (b) support provided for professional development and enhanced performance. Although there is recognition that teacher learning and professional development is a process that continues (or should continue) across a teacher's career and that progresses from novice, to advanced beginner, to competent, to proficient, to adaptive expert (Berliner, 1994, 2001; Borko & Putnam, 1996; Hammerness, Darling-Hammond, & Bransford, 2005), the preservice and continued professional development of teachers typically are treated as two separate cultures. For example, beyond framing the work of the AERA Panel against "the professional agenda for teacher education" (Cochran-Smith & Zeichner, 2005, pp. 43-44) and acknowledging the "clear need to look more at how teachers' knowledge and practices are shaped by their preparation including after they have completed their programs" (Zeichner, 2005, p. 746), no attention is given explicitly in the panel's report to the professional development of teachers as a continuum. Even the National Academy of Education Report on Preparing Teachers for a Changing World by Hammerness et al. (2005) devotes only two brief sections within a chapter titled "Teacher Learning and Development"; these deal with "the process of teacher development" (pp. 378-381) and "theories of teacher development in communities of practice" (pp. 382-383).

The purpose of our including this alert is to add emphasis to the AERA Panel's focus on the interactions among stage in the professional development of teachers, the support provided for their professional development and enhanced performance, and the strength of the connections one is likely to find between teaching, teacher preparation, and K-12 learning. Understanding these interactions is essential to sound policy and practice, and equally so for CTPL researchers, because a teacher's impact on learning once they are in the classroom is always powerfully influenced by the context in which they work (Feiman-Nemser, 2001).

The importance of bridging the two cultures became shockingly clear in the Teacher Effectiveness Study. Slightly less than half (44%) of the 1st-year teachers in the study had districtappointed mentors, and only half of these were viewed by their protégés as helpful. Nearly as many 1st-year teachers received help offered voluntarily by one or more of their colleagues; but here again, only about half of the new teachers receiving such assistance viewed it as helpful. Of the 78 first-year teachers studied, 13 viewed themselves as being "entirely on their own" in navigating their 1st year as a full-time professional.

The AERA Panel recognized the importance of these connections in conducting research on the preservice preparation of teachers. But they are at least as important, and perhaps more so, when conducting CTPL research with early career and experienced teachers. Schools differ greatly in the supports and staff development opportunities they provide their teaching faculties; but with the press for improved learning that nearly all schools now face, they are increasingly turning to teacher development and related support systems for that improvement (Gewertz, 2005; Guskey, 2003; O'Shea, 2005). CTPL research involving early career and experienced teachers must take these developments into account.

#### Strengthening Measures

Throughout the AERA Panel's report, reference is made repeatedly to the need for either

better measures than teacher education research has employed in the past or new and different measures. Better measures of teacher knowledge and skills, for example, are called for explicitly in Recommendation 6 (Zeichner, 2005, p. 744), and better measures of K-12 student learning are called for in Recommendation 5 (Zeichner, 2005, p. 743). New and different measures are called for in Recommendation 8:

We also think that a multidisciplinary and multi-methodological approach to studying issues in teacher education offers the best hope for producing knowledge that is useful for policy and practice. Student learning in schools is affected by a number of different but interrelated factors in addition to the general type of preparation for teaching received by teachers. Among these are the individual attributes brought by prospective teachers to their teacher education programs; the specific features of these programs and their components and the institutions in which they are situated; the nature of instruction in teacher education programs; what prospective teachers learn in these programs; the school in which teachers teach before, during and after they complete their preparation; school district policies and practices; and state and federal policies. (Zeichner, 2005, p. 743)

The measurement tasks ahead for teacher education researchers interested in investigating CTPL connections are daunting.

Beyond signaling the need for better measures than those that have been used in the past, and the need for multidisciplinary and multimethodological approaches to the many other kinds of measures needed to pursue CTPL research systematically, the AERA Panel offered little by way of practical help in either creating or using such measures. From our experience in the Teacher Effectiveness Study and its predecessors, we offer three elaborations with respect to measurement that CTPL researchers should find helpful. The first is that when thinking about CTPL research, think in terms of *indicators* of teacher impact on learning that include but go beyond state- or district-administered measures of achievement. In doing so, however, take care in assuring these are defensible to the research community; sensible to parents, teachers, and the policy community; and diverse.

The framework outlined in Table 2 that emerged from the need to deal with this issue in

Proximity to Teacher Work	Potential Source of Evidence		
Instructionally embedded evidence (through class- room observation or videotaping)	Student engagement in learning Level(s) of intellectual work pursued (e.g., Bloom's taxonomy) Student understanding and exploration of meaning within/across subject areas Student interest in content to be learned Proportion of learning tasks targeted for an instructional period that are pursued		
Instructionally linked evidence (through teacher- documented student progress in learning)	<ul> <li>Teacher-documented gains in nontrivial learning through one or more units of instruction taught, with data on learning disaggregated for designated groups of students</li> <li>Samples of student work evaluated against established state or district performance standards, with student work disaggregated for designated groups of students</li> <li>Teacher-maintained records of student progress in learning disaggregated for designated for designated groups of students</li> <li>Quality of teacher explanation and interpretation of evidence presented on the progress students have made in their learning</li> <li>A continuous progress record of learning by students provided through online computer-adapted assessment</li> </ul>		
Instructionally aligned evidence (through school-, district-, or state-administered examinations clearly linked to the curriculum guiding teacher and student work in a school)	<ul> <li>Student performance on teacher-developed assessments administered prior to and following an extended period of instruction (e.g., midterm or end of term, with data disaggregated for designated groups of students)</li> <li>Student performance on district-developed examinations administered prior to and following an extended period of instruction, with results disaggregated</li> <li>Student performance on state examinations administered at the beginning and end of a school year, with results disaggregated</li> <li>Student performance on district- or state-administered examinations at the end of a school year analyzed with a value-added methodology of the kind developed by William Sanders</li> </ul>		
Instructionally tangential evidence (through district-, state-, or nationally administered examinations not clearly linked to curricula guiding teacher and student work in a school)	Student performance on externally developed examinations not connected tightly to state-adopted "standards" for learning Student performance on the National Assessment of Educational Progress Student performance on nationally normed tests of achievement Student performance on nationally administered examinations such as the SAT or ACT		

#### TABLE 2 Potential Sources of Evidence for Assessing Teacher Impact on Learning

the Teacher Effectiveness Study is offered as a point of departure in thinking about such indicators. The broad categories of indicators suggested are organized with regard to their proximity to and alignment with the teaching and learning process. Our assumption in framing the table along these lines is that the further one moves away from these two points of focus in attempting to trace a teacher's (or teacher preparation program's) impact on the learning of K-12 students, the greater the likelihood is of getting a distorted or incomplete picture of impact.

A further assumption underlying the organization of items within the table is that teacher impact on learning takes many forms and can be captured through many indicators. An accompanying assumption is that the greater the number of indicators captured, so long as each is strong and valid in its own right, the more reliable and valid will be the picture provided of a teacher's impact on learning. From our perspective, reliance on state or district examinations administered once a year, no matter how well they are constructed or the level of sophistication brought to their analysis, will always provide a picture of a teacher's impact on learning that is less comprehensive, reliable, and valid than it can and should be.

The second elaboration concerning measurement we wish to offer is that when thinking about assessing the classroom performance of teachers in CTPL research in today's schools, either as a preservice, early career, or experienced teacher, take care in assuring the observation or videotape coding system used (a) reflects the theory and demands of standardsbased teaching and learning, (b) is grade-level appropriate, and (c) provides extensive and defensible indicators of a teacher's impact on the learning of his or her students.

The last recommendation we make with respect to strengthening measures is to create an expectation within the teacher education research community that all CTPL-related research will both draw on and contribute to a catalogue of measures that carries a detailed description of each measure used, the conditions under which it was used, and evidence as to its reliability, validity, and distribution of scores under those conditions. Without an established store of theory-embedded and defensible measures to draw on, building a knowledge base that reliably informs policy and practice will be slow to come. The CTPL Coalition for collaborative research and theory building (see Note 1) is committed to starting such a catalogue.

# NOTES ON INITIATING AND SUSTAINING THE KIND OF RESEARCH PROPOSED

In concluding its report, the AERA Panel addressed the issue of developing an infrastructure for research on teacher education. After acknowledging that much of the research that had been reviewed involved individual or small groups of teacher educators studying aspects of their own courses or programs, and why this has been so (Zeichner, 2005, p. 750), a central point made throughout the report is reiterated:

that more research be conducted that looks at the nature and impact of different components of teacher education across different program, institutional, and state policy contexts.... We cannot assume that evidence about successful practices in one setting with a particular group of teacher educators or in one type of pathway into teaching is meaningful in other types of settings. (Zeichner, 2005, pp. 750-751)

The need for additional resources to conduct such research, and new programs that prepare teacher education researchers to lead it, are then pointed to as essential to infrastructure building. The final element mentioned is the organization of teacher education research through "strategic research partnerships" consisting of "researcher-practitioner networks" (Zeichner, 2005, pp. 753-755).

From our perspective, the partnership concept is probably the most important and likely the most difficult aspect of infrastructure building addressed. Given the range and complexity of variables involved in CTPL research, and the near-to-endless variety of contexts across which an established relationship among variables needs to be validated, no single group of researcher-practitioners operating within a single institutional and policy context will be able to carry out the research agenda called for. It is our view that both intrainstitutional and interinstitutional changes of unprecedented magnitude will be needed for teacher education and teacher education research to rise to the challenges faced in the context of today's emerging standards-based, evidence-driven approach to education. We sketch briefly in the paragraphs that follow what we see these intrainstitutional and interinstitutional changes needing to be.<sup>4</sup>

# Teacher Preparation Programs and Related Follow-Up Studies as Contexts for Research

In thinking about getting from where we are to where we need to be, with both theory and research having the focus, breadth, and qualities proposed, one promising avenue to pursue is treating teacher preparation programs as contexts for research. They are the contexts where connections need to be made systematically and defensibly among the many variables involved, and they need to be made routinely and repeatedly as each new cadre of prospective teachers enters and leaves a program.

Opportunities abound in such contexts for grounded theory development and refinement (Glaser & Strauss, 1967), with practiceanchored research and measurement as its crucible. By virtue of replications made possible through the continuing cycle of response to the obligations of preparing and licensing teachers to work in an evolving system of schools, teacher preparation programs and their participating schools hold within them the wherewithal for becoming ongoing laboratories for such research.

As most teacher preparation programs are now constituted, however, few can serve such a purpose. To be productive contexts for research and theory development of the kind proposed, four conditions need to exist:

- 1. A faculty actively involved in, and preparation programs reflecting, theory development pertaining to the connections among teaching, teacher preparation, and K-12 learning in standards-based schools.
- 2. Strong, defensible measures of theory-related constructs being addressed in both program design and program-anchored research, particularly in followup studies.
- 3. A data collection and management system that supports data-driven decisions concerning candidate entry to and progression through a preparation program that is coupled to the performance of program completers as early career teachers and is used systematically for program-related research, evaluation, and theory development.
- 4. Established policies and resource commitments that support all the above.

The concept of and call for teacher preparation programs to function as contexts for research are not new (H. D. Schalock, 1980, 1983; Shulman, 2004), and both the Teacher Education Accreditation Council and the National Council for Accreditation of Teacher Education's 2000 standards for teacher preparation program accreditation (Murray, 2005; Wise, 2005) are designed to move the profession in this direction. The findings and recommendations of the AERA Panel, however, confirm the imperative of doing so.

# Collaborative Theory Development Partnerships as Contexts for Research

We have argued throughout this article that theory development, testing, and refinement are central to both the advancement of teaching as a profession and to the research needed for its advancement. In keeping with this view, we believe that the kind of collaborative research partnerships called for by the AERA Panel should center their work on theory development and refinement through systematically testing its application to practice within a variety of contexts. Partnerships could be constituted to reflect the teaching or preparation contexts of interest, for example, schools in low socioeconomic status neighborhoods or with large numbers of English language learners or alternative designs for the mentoring of early career teachers; but whatever the context of interest, the central focus of research and theory building would be the connections among teaching, teacher preparation, and K-12 learning. The press for theory building, with its accompanying demands for clarity in definitions, defensible measures, and reporting in a level of detail that make both concepts and measures accessible for use and replication by others, is essential for teaching to evolve as a mature profession.

In pursuing such work, however, we think it important to distinguish between two fundamentally different types of investigations. One needs to focus on *theory-framing studies* and one on *theory-confirming studies*. Theory-framing studies have as their purpose the identification of constructs that have demonstrable predictive power and the articulation of one or more networks of constructs within which the focal construct(s) rests. In this sense, the purpose of theory-framing studies is exploratory, descriptive, and parameter defining. They also serve to validate measures and formulate principles.

By contrast, theory-confirming studies have as their purpose the targeted testing of theoretically and empirically supported "principles of application" to determine the power of their applicability in one or more designated contexts. Both kinds of studies work from stated hypotheses; both involve experimental, quasiexperimental, or planned-variation designs; and both involve the range of variables and measures discussed previously.

Where they differ is in use of randomized control groups within their designs and in the additional degree of control theory-confirming studies must exert over unwanted sources of influence on dependent variables. Theoryframing studies can proceed with greater degrees of flexibility in this regard and still be productive with regard to the purposes they serve. Theory-confirming studies do not have this degree of flexibility in determining the applicability of principled practice to varying contexts.

# **CONCLUDING COMMENTS**

The work of Sanders and his colleagues at the University of Tennessee (Sanders & Horn, 1994; Sanders & Rivers, 1996; Wright, Horn, & Sanders, 1997) and Mendro and his colleagues in the Dallas Public Schools (Mendro, 1998) demonstrate the pivotal importance of individual teachers to student learning. Both bodies of research demonstrate unequivocally the cumulative and lasting effects of a teacher on the academic achievement of students. During a multiyear period using "value-added" analyses, Sanders focused on what happened to students whose teachers produced high achievement versus low achievement results. He found that when children, beginning in third grade, were placed with three high performing teachers in a row, they scored on average at the 96th percentile on Tennessee's statewide mathematics assessment at the end of the fifth grade. When children with comparable achievement histories in third grade were placed with three low performing teachers in a row, their average score on the same mathematics assessment in fifth grade was at the 44th percentile-a 52 percentile point difference for children who presumably had comparable abilities and educational backgrounds.

In the Dallas study (Mendro, 1998), similar results were found in both mathematics and reading. When first-grade students were placed with three high performing teachers in a row, their average performance at third grade on the math section of the Iowa Tests of Basic Skills had increased from the 63rd percentile to the 87th. When their counterparts in first grade were placed with three low performing teachers in a row, their average math scores in third grade had decreased from the 58th percentile to the 40th. Similar results were found in reading. Further analyses of both the Tennessee and Dallas data indicate that the effects on achievement of both strong and weak teachers persisted: Subsequent achievement was enhanced or limited by

the effects of earlier teachers, and the negative effects of a low performing teacher were not able to be fully overcome by placement for 3 years with high performing teachers.

Although these findings have been widely discussed within the education community since their publication, and have shaped policy decisions such as the No Child Left Behind Act (2001) target of a "highly qualified teacher" in every classroom by 2006, the AERA Panel report makes painfully clear that teacher education research has little to say about how to select or prepare teacher candidates who will be the high performing teachers identified by the Sanders (Sanders & Horn, 1994; Sanders & Rivers, 1996; Wright et al., 1997) and Mendro (1998) teams. The recommendations of the panel for future research are intended to move the field in the direction of doing so, as does the report of the National Academy of Education's Committee on Teacher Education (Darling-Hammond & Bransford, 2005), but building the knowledge base needed to select, prepare, and nurture the continued professional development of such teachers is an enormous and long-term undertaking. In scope, complexity, and importance, it should be viewed as education's equivalent of biology's genome project. Theoretically and empirically connecting teaching, teacher preparation, and K-12 learning will require no less, either intellectually or scientifically. We hope the recommendations made in these pages will be helpful in charting that undertaking.

# NOTES

1. Figure 1 is one of several related "theory-framing maps" that have been reviewed, refined, and accepted as points of departure in a theory development initiative launched recently by a coalition of institutions and agencies committed to collaborative research and development concerning the connections among teaching, teacher professional development, and the learning of K-12 students (the CTPL Coalition). Information on the composition and projected work of the coalition can be found at http://www.tr .wou.edu/tep/products.html. Related work includes creating a *Thesaurus* for teacher education researchers and a *Handbook of Measurus* for research connecting teaching, teacher professional development, and the learning of K-12 students.

2. It is interesting that only one reference is made in the entire American Educational Research Association Panel report to either the affective or emotional dimensions of teaching or teacher preparation (see Cochran-Smith & Zeichner, 2005, pp. 278-279), and this was to the contribution of college generally to the affective development of students. 3. The term *design* has many meanings in the context of research. Most commonly it refers to use of experimental or control groups and method of assignment (random or nonrandom) of participants to groups. But the term also is used in referring to broad approaches taken to scientific inquiry, for example, correlational, cross-sectional, longitudinal, single-subject, and case study designs. We use the term here to refer to the manipulation and/or control of variables likely to have an effect on the connections among teaching, teacher preparation, and the learning of K-12 students.

4. The American Educational Research Association Panel acknowledged that the formation of partnerships for research in teacher education "goes against the grain of individual accomplishment that dominates the culture of higher education" (Zeichner, 2005, p. 755), but currently there is much to suggest that this is changing. The panel points to the Teachers for a New Era initiative, the New York City Pathways study, the Ohio Teacher Quality Partnership, and several others as examples of the kind of research partnerships envisioned. Others include the Renaissance Partnership for connecting teaching and learning and the recently established Coalition for Connecting Teaching, Teacher Preparation, and K-12 Learning.

#### REFERENCES

- Allen M. (2003). *Eight questions on teacher preparation: What does the research say?* Denver, CO: Education Commission of the States.
- Ayres, R., & Cuthbertson, L. (2004, July). Convergence and divergence of early career teacher's views of and responses to context in the classroom: Implications for teacher preparation? Paper presented at the Teacher Work Sample Emerging Best Practices Conference, Glenedon Beach, OR.
- Berliner, D. C. (1994). Expertise: The wonder of exemplary performances. In J. Mangieri & C. C. Block (Eds.), Creating powerful thinking in teachers and students: Diverse perspectives (pp. 161-186). Fort Worth, TX: Harcourt Brace.
- Berliner, D. C. (2001). Learning about and learning from expert teachers. *International Journal of Education Research*, 35(5), 463-483.
- Black, P., & William, D. (1998). Inside the black box: Raising standards through strong classroom assessment. *Phi Delta Kappan*, 80(2), 129-149.
- Borko, H., & Putnam, R. T. (1996). Learning to teach. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 673-708). New York: Simon & Schuster.
- Bransford, J., Darling-Hammond, L., & LePage, P. (2005). Introduction. In L. Darling-Hammond & J. Bransford (Eds.), Preparing teachers for a changing world: Report of the Committee on Teacher Education of the National Academy of Education (pp. 1-39). San Francisco: Jossey-Bass.
- Cochran-Smith, M. (2003). The unforgiving complexity of teaching: Avoiding simplicity in an age of accountability. *Journal of Teacher Education*, 54(1), 3-5.
- Cochran-Smith, M. (2005). Studying teacher education: What we know and need to know. *Journal of Teacher Education*, 56(4), 301-306.
- Cochran-Smith, M., & Fries, K. (2005). Researching teacher education in changing times: Politics, and paradigms. In M. Cochran-Smith & K. Zeichner (Eds.), *Studying*

teacher education: The report of the AERA Panel on Research and Teacher Education (pp. 69-109). Mahwah, NJ: Lawrence Erlbaum.

- Cochran-Smith, M., & Zeichner, K. M. (Eds.). (2005). Studying teacher education: The report of the AERA Panel on Research and Teacher Education. Mahwah, NJ: Lawrence Erlbaum.
- Darling-Hammond, L., & Bransford, J. (Eds.). (2005). *Preparing teachers for a changing world*. San Francisco: Jossey-Bass.
- Falk, B. (2000). *The heart of the matter: Using standards and assessment to learn*. New York: Heinemann.
- Feiman-Nemser, S. (2001). From preparation to practice: Designing a continuum to strengthen and sustain practice. *Teachers College Record*, 103(6), 1013-1055.
- Gewertz, D. (2005, August 10). Staff investment pays dividends in Md. district. *Education Week*, pp. 1, 16.
- Girod, G. R. (Ed.). (2002). Connecting teaching and learning: A handbook for teacher educators on teacher work sample methodology. Washington, DC: American Association of Colleges for Teacher Education.
- Glaser, B., & Strauss, A. (1967). *The discovery of grounded the ory: Strategies for qualitative research*. Hawthorne, NJ: Aldine.
- Goldhaber, D. D., & Brewer, D. G. (2000). Does teacher certification matter? High school teacher certification status and student achievement. *Education Evaluation and Policy Analysis*, 22, 129-145.
- Gordon, D. (2004). A nation reformed? Twenty years after A nation at risk. Cambridge, MA: Harvard University Press.
- Guskey, T. R. (2003). What makes professional development effective? *Phi Delta Kappan*, 84(10), 748-750.
- Hammerness, K., Darling-Hammond, L., & Bransford, J. (with Berliner, D., Cochran-Smith, M., McDonald, M., & Zeichner, K.) (2005). How teachers learn and develop. In L. Darling-Hammond & J. Bransford (Eds.), *Preparing teachers for a changing world* (pp. 358-389). San Francisco: Jossey-Bass.
- Hanushek, E. A., & Pace, R. R. (1995). Who chooses to teacher (and why)? *Economics of Education Review*, 14(2), 101-117.
- Johnson, S. M., & The Project on the Next Generation of Teachers. (2004). *Finders and keeper: Helping new teachers survive and thrive in our schools*. San Francisco: Jossey-Bass
- Lauer, P. (2001). A secondary analysis of a review of teacher preparation research. Denver, CO: Education Commission of the States.
- Marzano, R. (1998). *Implementing standards-based education*. Washington, DC: National Education Association.
- Mendro, R. L. (1998). Student achievement and school and teacher accountability. *Journal of Personnel Evaluation in Education*, 12, 257-267.
- Millman, J. (Ed.). (1997). *Grading teachers, grading schools: Is student achievement a valid evaluation measure?* Thousand Oaks, CA: Corwin Press.
- Monk, D. H. (1994). Subject area preparation of secondary mathematics and science teachers and student achievement. *Economics of Education Review*, 13, 125-145.

- Murray, F. B. (2005). On building a unified system of accreditation in teacher education. *Journal of Teacher Education*, 56(4), 307-317.
- No Child Left Behind Act of 2001, Pub. L. No. 107-110, 115 Stat. 1425 (2002). Available at http://www.ed.gov/
- O'Shea, M. R. (2005). *From standards to success: A guide for school leaders*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Patterson, M. (2005). Hazed! *Educational Leadership*, 62(8), 20-23.
- Rice, J. K. (2003). *Teacher quality: Understanding the effectiveness of teacher attributes*. Washington, DC: Economic Policy Institute.
- Sanders, W., & Rivers, J. C. (1996). *Cumulative and residual effects of teachers on future student academic achievement*. Knoxville: University of Tennessee Value-Added Research and Assessment Center.
- Sanders, W. L., & Horn, S. (1994). The Tennessee Value-Added Assessment System (TVAAS): Mixed-model methodology in educational assessment. *Journal of Personnel Evaluation in Education*, 8(3), 299-311.
- Schalock, D., & Myton, D. (1989). A new paradigm for teacher licensure: Oregon's demand for evidence of success in fostering learning. *Journal of Teacher Education*, 29(6), 8-16.
- Schalock, D., & Schalock, M. (1995). Proceedings of a conference: Advances in theory and research on teacher effectiveness. Monmouth: Western Oregon University.
- Schalock, D., Schalock, M., & Girod, G. (1997). Teacher work sample methodology as used at Western Oregon University. In J. Millman (Ed.), *Grading teachers, grading* schools: Is student achievement a valid evaluation measurement? (pp. 15-45). Thousand Oaks, CA: Corwin Press.
- Schalock, H. D. (1980). Eating humble pie: Notes on methodology in teacher education research. In G. E. Hall, S. M. Hord, & G. Brown (Eds.), *Exploring issues in teacher education: Questions for future research* (pp. 519-536). Austin: University of Texas–Austin, Research and Development Center for Teacher Education.
- Schalock, H. D. (1983). Methodological considerations in future research and development in teacher education. In K. Howey & W. Gardner (Eds.), *The education of teachers: A look ahead* (pp. 38-73). New York: Longman.
- Schalock, M. D. (1987). Teacher productivity: What is it? How might it be measured? Can it be warranted? *Journal of Teacher Education*, 38(5), 59-62.
- Schalock, M. D., & Schalock, H. D. (2004). The longitudinal effects of teacher preparation on the practices and beliefs of new teachers and the learning of their students: Final project report. Available at http://www.tr.wou.edu/tep/products.html
- Shoemaker, P. J., Tankard, J. W., & Lasorsa, D. L. (2004). *How to build social science theories*. Thousand Oaks, CA: Sage.
- Shulman, L. (2004). Truth and consequences: Inquiry and policy in research on teacher education. *Journal of Teacher Education*, 53(3), 248-253.

- Wilson, S., & Floden, R. (2002). Addendum to teacher preparation research: Current knowledge, gaps, and recommendations. Denver, CO: Education Commission of the States.
- Wilson, S., Floden, R., & Ferrini-Mundy, J. (2001). *Teacher* preparation research: Current knowledge gaps and recommendations. Seattle, WA: Center for the Study of Teaching and Policy.
- Wise, A. E. (2005). Establishing teaching as a profession: The essential role of professional accreditation. *Journal of Teacher Education*, *56*(4), 318-331.
- Wright, S., Horn, S., & Sanders, W. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education*, 11, 57-67.
- Zeichner, K. M. (2005). A research agenda for teacher education. In M. Cochran-Smith & K. Zeichner (Eds.), *Studying teacher education: The report of the AERA Panel on Research and Teacher Education* (pp. 737-759). Mahwah, NJ: Lawrence Erlbaum.

H. Del Schalock, a research professor at the Teaching Research Institute, has been actively engaged in educational research, development, and improvement in Oregon for more than 30 years. He has led numerous externally funded research efforts concerning teacher effectiveness as defined by student progress in learning. He also has led several school improvement efforts and teacher education program improvement efforts that anchor to student progress in learning. Throughout his career, he has worked closely with Oregon's teacher preparation and licensing agency and the Oregon Board of Education to align teacher preparation and licensure requirements with the state's evolving model of standards-based schooling. Within this context, he has continuously searched for ways of meaningfully connecting student learning to teacher work as the embodiment of sound educational policy and practice and has been the lead developer and advocate of teacher work sample methodology, an approach to performance assessment connecting teacher classroom instruction to gains in student performance. He has authored or coauthored more than 100 articles, chapters, and monographs.

Mark D. Schalock, associate research professor at the Teaching Research Institute, has been actively engaged in educational research, development, evaluation, and technical assistance for the past 20 years at the local, state, and national levels. His work spans issues related to both regular and special education, with a significant recent emphasis on the preparation, licensure, and early career development of both regular and special education teachers. He has published more than 20 articles and book chapters on program and personnel evaluation, measurement, research, accountability, and educational productivity. **Robert Ayres** is a research professor at the Teaching Research Institute. He has been actively engaged in and taught educational research methods courses in the College of Education at Western Oregon University for 14 years and has been an evaluation consultant for Western's Education Evaluation Center for the past 26 years. In the latter capacity, he assesses cognitive, behavioral, and academic abilities of persons across a wide range of ages via standardized tests, interviews, and observations. His research interests include teacher attitudes and beliefs, such as "self-efficacy" and "commitment to the profession."