Journal of Teacher Education
http://jte.sagepub.com

The Teacher Effectiveness Movement: How 80 Years of Essentialist Control Have Shaped the Teacher Education Profession

David G. Imig and Scott R. Imig

Journal of Teacher Education 2006; 57; 167
DOI: 10.1177/0022487105285672

The online version of this article can be found at:
http://jte.sagepub.com/cgi/content/abstract/57/2/167

Published by:
SAGE Publications
http://www.sagepublications.com

On behalf of:
AACTE
American Association of Colleges for Teacher Education (AACTE)

THE TEACHER EFFECTIVENESS MOVEMENT

HOW 80 YEARS OF ESSENTIALIST CONTROL HAVE SHAPED THE TEACHER EDUCATION PROFESSION

David G. Imig
University of Maryland–College Park

Scott R. Imig
University of Virginia

The locus of control in teacher education has been outside the hands of those who educate our nation's teachers for more than a century. Essentialists have long controlled the agenda for public schooling in America, and it is evident as well that their influence has prevailed in both the form and function of teacher education. The authors suggest that the contest between progressives and essentialists regarding teacher education has been repeatedly decided in favor of the essentialists. The current attempt to recast teacher education to focus singularly on effectiveness of classroom teachers in raising the test scores of their students is a not-unanticipated result of this enduring contest.

Keywords: *teacher effectiveness; essentialists; progressives*

> The results of the teacher's work [should be] measured by the growth of pupils. It is in terms of such growth that the outcomes of teaching must ultimately be evaluated, and the young teacher should be accustomed from the outset to think of his work as measured finally by this standard.
>
> —Learned & Bagley (1920, p. 219)

In 1914, the Carnegie Foundation for the Advancement of Teaching (CFAT) undertook an ambitious study of teacher education in the state of Missouri. This study was inspired by the now famous Flexner (1910) report on medical education in the United States and Canada. The Flexner report highlights the disparities in quality among schools of medicine and was embraced by policy makers and practitioners alike. Although the Flexner study led to the rapid closure of more than half of the medical schools then in existence, the Learned and Bagley (1920) study was met with indifference and hostility. At a time of ascendant progressivism in American education, their philosophy regarding student learning and the role of the teacher was decidedly "out of step." Other reasons for the indifferent response to the extraordinary study on teacher education they undertook was the authors' insistence that teachers had to learn how to make use of various measures of student learning and to have these measures guide their practice throughout their professional lives. Their appeal for a teacher education in which all courses (academic and professional) were professionalized and their insistence on single purpose collegiate-level institutions devoted exclusively to teacher preparation were controversial. More challenging was the fact that Learned and Bagley also insisted that student learning should be the guiding principle for teaching and that teachers should be judged on their effectiveness in promoting student learning.

Journal of Teacher Education, Vol. 57, No. 2, March/April 2006 167-180
DOI: 10.1177/0022487105285672

Although few policy makers or practitioners today are aware of the recommendations of the Learned and Bagley (1920) report, it would be wrong to assume that the ambitious study did not have a lasting effect on teaching and teacher education. It is impossible to read the quote above and not realize the prescient nature of the report. As the National Academy of Sciences' Education Center moves forward in 2005 to 2007 with the conduct of the congressionally mandated study of teacher preparation programs in the United States, or what has been described as "a Flexner-type study of teacher education" (National Academies, 2005), the findings and recommendations of the Flexner (1910) report on teacher education should be considered. Although the first report was met with suspicion and skepticism from the leaders of the teaching profession, state policy makers, and the college presidents who would be most affected, there is much in the report that deserves consideration. Perhaps the most important consideration is the report's repeated attempts to describe effective teachers and effective teaching and its ultimate acceptance of the need for multiple criteria to make such determinations. Unlike the wave of reform brought about by Flexner, the Learned and Bagley report fed the slow trickle of change that continues to erode teacher education down to its narrowest definition.

Essentialists have traditionally railed against progressives' influence on teaching and teacher education, but history tells a different story (Labaree, 2004). Our contention is that essentialists, often with the ear of policy makers, have long controlled the agenda for public schooling in America. Progressives, from John Dewey to John Goodlad, have always been on the outside attempting to recast the role and purpose of schooling and to expand the definitions of *quality teaching* and *teacher effectiveness*. In this article, we argue that the policy recommendations of the Carnegie study on teacher education (Learned & Bagley, 1920) and the mandates of the No Child Left Behind Act (NCLBA, 2001), signed into law 80 years later, illustrate that the locus of control in teacher education has long been outside the hands of those who educate our nation's teachers. We draw attention to selected movements within the profession to embrace a research-based conception of teaching effectiveness. Finally, we argue that essentialists or neoessentialists (a more apt term than *neoconservatives* to describe those "in charge" of contemporary American education) have prevailed in the battle to define the purpose of schooling, the structure of the school curriculum, the role of the teacher, and the measures used to evaluate student growth and teacher effectiveness. This is not an advocacy for essentialism but rather a call to be informed in the making of policy for teaching and teacher education. Subject matter knowledge and student achievement gains are the currency of the realm in which we must operate. This is the place we have to begin in the transformation of teaching and teacher education.

THE LEARNED AND BAGLEY STUDY AND THE NCLBA

William C. Bagley (1874-1946) was the father of essentialism in American education. A recent biography of Bagley (Null, 2003) does a superb job of describing his life and the experiences he gained as teacher, principal, faculty member, and education dean that led to his articulation of the essentialist agenda. He was the descendant of a long line of traditionalists who would contest the efforts of "progressives" (a term with as many definitions as there were educational philosophers in 20th century America). The contest between essentialists and progressives has been a dominant factor in the shaping of educational policy throughout the century. Essentialists contend that content matters and the focus of schooling should be on student learning. Teachers are responsible for leading whole classes of students and for the setting of high expectations and directing student learning toward measurable ends. Bagley's long tenure at Teachers College, Columbia University, where he was a contemporary of the leading progressives, including William H. Kilpatrick, George S. Counts, and Harold O. Rugg, helped to sharpen the distinctions between progressive ideas and essentialist principles that persist today. The different visions articulated for

schools and schooling produce different views of the good teacher and the means for judging whether a teacher is effective or less effective. Education historian Diane Ravitch (2002) described Bagley as the "leading dissident" in an era when Bagley's Teacher College colleagues were fashioning a progressive agenda for schools. His success should be measured by the fact that essentialism is the undergirding philosophy for the NCLBA of 2001. Bagley argued that schools should promote cultural transmission and emphasize a traditional course of studies. Although Kilpatrick and his fellow progressives called for a schooling that would "reach into the thick of life" and for teachers to be "socially minded and socially disposed," Bagley and his essentialist colleagues were appealing for a traditional curriculum and teacher-centered schools (Kilpatrick, 1932).

In an era of child-centered advocacy, Learned and Bagley (1920) used the CFAT report to emphasize their belief that the teacher was key to the learning of all children. Dissatisfied with what they observed in the schools they examined, they called for teachers with both the "external elements of skill" and the "internal elements of insight and resourcefulness." Learned and Bagley wanted teachers who possessed

- aptness and readiness in illustration,
- clearness and lucidity in explanation and exposition,
- keen sensitiveness to evidences of misunderstanding and misinterpretation on the part of students,
- dexterity and alertness in devising problems and framing questions,
- a sense of humor,
- an attitude that requires reasoned support of each point presented,
- quickness to detect inattention, and
- a sense of proportion to distinguish between the fundamental and the accessory.

Learned and Bagley (1920) deplored the level of instruction then occurring in public schools, which they attributed to teachers' being neither prepared nor treated like professionals. The Learned and Bagley report, published by the CFAT as Bulletin No. 14, insisted that the nation should no longer tolerate school systems that failed to focus on good teaching. Although the report fails to differentiate between teacher practices and teacher attributes, it does contain much discussion of what a teacher should be able to do to promote student learning. Although Bagley was an opponent of the use of intelligence testing in schools, the bulletin calls for a summative evaluation of teacher candidates as they leave the professional preparation program, and there is much attention in the study to guiding student effort through the use of various norm-referenced measures of literacy and numeracy. Learned and Bagley concluded that effective teachers should be gauged by the performance of their pupils, but the authors also sought to understand and present the qualities and practices effective teachers should demonstrate.

We contend that the NCLBA (2001) represents a set of strategies and action steps that could well be based on the recommendations of the Learned and Bagley (1920) report. It is doubtful that the authors of the NCLBA were any more aware of the Learned and Bagley report than are most teacher educators, but the appeals for action, the attention to student learning, and the assertions about good teachers and their importance could well be derived from the bulletin. As Gutek (2004) and others have argued, neoessentialism undergirds many of the premises and directions of the now controversial NCLBA. Although the appeal for educating all children is a decidedly progressive or liberal ideal, the attention to classroom actions and instructional practices is essentialist. Reliance on assessment with the careful disaggregation of scores by ethnic and racial groups (and other characteristics) drives the NCLBA. This reliance on tests and measures to guide student practice (with the appeal to scientism and evidence) are consistent with the recommendations of the Learned and Bagley report.

As David Labaree (2004) contended, the traditionalists or essentialists have long prevailed in the conduct of schooling. He argued that progressivism was the prevailing rhetoric of school systems (and of those who led those systems) but that actual classroom practices have always been "traditional." The contention that the enactment of the NCLBA (2001) was a radical

departure from current practice, therefore, is misguided because what actually occurs in classrooms and what was mandated by the Bush administration are depressingly similar. For all of the interventions of teacher educators (and policy initiatives of reform-minded groups such as the National Board for Professional Teaching Standards and the Interstate New Teacher Assessment and Support Consortium [INTASC]) to transform schools, what actually occurs in most classrooms would be strikingly familiar to Learned and Bagley (1920). Although historians contend that Learned and Bagley had little impact on the curriculum or practices of the then teacher preparation institutions, we believe their report should be read as a road map for realizing a corps of teachers capable of teaching in traditional schools; it is the template for measuring effectiveness in contemporary schools (Frazier, in press). If the purpose of such schools is academic (i.e., the acquisition of facts and knowledge of core subjects), then the use of tests to assess the attainment of such facts and knowledge is understandable.

Those who drafted Bush administration policy proposals for education embraced the belief that highly effective teachers are those who realize student achievement gains. This is evident in the NCLBA (2001) and in subsequent federal authorizations for education. Although teacher performance (and the traits and characteristics of effective teachers) was once the subject of much consideration, during the past decade attention has shifted to a narrow definition of *effectiveness* based on the ability of teachers to realize student achievement gains on various measures of student learning. Other measures of teacher effectiveness (that have been considered during the course of the past century) are either ignored or dismissed in the assertion that student score gains on tests of subject matter knowledge are what matters (Mitzel, 1960). As a result of the NCLBA, *high quality* has become synonymous with *effective*. The promise of high-quality teachers who are "effective" in promoting student learning by all children is at the heart of Bush administration efforts to transform schools and schools of education. They have relied on only two measures to describe the characteristics of highly qualified teachers—teacher candidate scores on standardized tests of subject matter knowledge and degree attainment in a particular core or academic subject. In the drafting of the signature education policy of the Bush administration, the NCLBA, the Bush administration officials employed these two characteristics or indicators despite the lack of solid scientifically based research evidence that either matters. The law defines a highly qualified teacher as

> [a person who] has obtained full State certification as a teacher (including certification obtained through alternative routes to certification) or passed the State teacher licensing examination, and holds a license to teach in such State [and who has] not had certification or licensure requirements waived on an emergency, temporary, or provisional basis. (NCLBA, 2001, § 9101)

Today, 4 years after passage of this monumental law, efforts to encourage states to define a highly qualified teacher as one with the prerequisite academic degree and passage of the state licensure examination persist. Both the Education Department and various interest groups seek better ways to document the current qualifications of teachers and to measure local school district compliance with the law. There is also significant encouragement for states to align state licensure provisions with the NCLBA and to focus on the verbal ability of candidates for teaching.

The Bush administration has also sought to influence state boards of education and other state agencies that control teacher licensure to adopt policies more conducive to the provisions of the NCLBA (2001). The ability to affect student learning is the measure of quality that best satisfies this definition. Announced in the education secretary's first annual report on teacher quality (Paige, 2002), those intentions represent a bold reach by this administration into a policy arena that has traditionally been controlled by the states. Using funding authorized in Title II, Part A of the NCLBA, the Bush administration encouraged states to dramatically reshape state policy for teacher education. Specifically, Paige's (2002) report calls on states to (a) end the "exclusive franchise" of schools of education and to curtail the "*shocking number of . . . mandated* [our emphasis] education courses to qual-

ify for certification" (p. 33), (b) assist state efforts to uncouple education school courses from state licensure and make "attendance at schools of education . . . optional," (c) "streamline" licensure requirements to place a premium on verbal ability and content knowledge, (d) develop new and "challenging assessments" for teacher candidates, and (e) require "content area majors for prospective teachers." Teacher licensure and teacher certification are the focus of these efforts to transform teaching and teacher education.

ATTEMPTS TO DEFINE EFFECTIVE TEACHING

Although the neoessentialists have hammered the theme that good teaching matters, their measure of good teaching is the gains of students on standardized measures of student achievement. For at least the past 80 years, researchers and others have sought to identify measures of teacher effectiveness. From the CFAT's issuance of the Learned and Bagley (1920) report on teacher education to the Carnegie Corporation's establishment of Teachers for a New Era, it has been a remarkable 80 years of such efforts with profound implications for teacher education. A repeated inquiry during the course of those 80 years was the matter of an effective teacher—What are the traits and characteristics of effective teachers? What must an effective teacher carry to the classroom in terms of requisite skills and knowledge, background, and experience? What do effective teachers do in a classroom? How would one know an effective teacher? Does demonstrating the knowledge and skills called for by the National Board for Professional Teaching Standards meet the effective teacher standard? Does meeting state teacher standards (or INTASC standards) or passing a PRAXIS assessment or National Evaluation Systems test serve as an indicator of high quality? To the satisfaction of no one, educational researchers have been unable to identify those characteristics that do make a difference.

Teacher Characteristics

For decades teacher educators have struggled to define the characteristics of good teachers. They have attended to such considerations in their admission policies and in the exit practices. Many have talked about student learning as the ultimate measure of teacher success but struggled to define *learning* and to find agreement. Is it multidimensional or can it really be defined so narrowly as to include only score gains on particular measures of student achievement. The difficulties of measuring student performance and attributing gains and loses to individual teachers (particularly with the passage of time) prohibited much consideration of such measures of teacher effectiveness. There was always the promise that such connections could be established. Gage's (1972) classic *Teacher Effectiveness and Teacher Education* offers the hope that one could show relationships between teacher practices and student performances. Encyclopedias of research findings were produced and research syntheses prepared that examine aspects of teaching strategies and the impact these methods had on classrooms and individual students' performance (Berliner, 1979; Evertson, 1982; Good & Brophy, 1986; Medley, 1978; Stallings, 1985). An enormous amount of time was spent in examining good teaching and observing what good teachers did in their classrooms. Faculty careers were devoted to constructing observation tools and knowing what teachers did in different types of classrooms with different groups of students under certain kinds of conditions. Yet the difficulties associated with using standardized achievement tests to determine teacher effectiveness precluded such use in the policy arena until new technologies were available and new research methodologies were accepted. Using economic research methodologies, a generation of researchers seized on the availability of large databases of student scores and began to use such measures in determining the effectiveness of particular instructional or curricular interventions (Abell Foundation, 2001; Ballou & Podgursky, 2000; Ferguson & Womack, 1993; Goldhaber & Brewer, 2000; Hanushek, 1997).

Until that work appeared, those who make policy for teacher education largely depended on a fairly robust set of research findings to define the characteristics of effective teachers that had been reported in the several hand-

books of research on teaching and in various journals and research syntheses. Teacher educators (and researchers) contributed to examining the characteristics of teachers identified by parents and principals as effective and then determined through peer and supervisor evaluations whether those characteristics really mattered. In such ways, it was decided that one or more of the following mattered in the determination of high-quality teaching: (a) years of teaching experience—experienced teachers were believed more effective than novices or beginners; (b) possession of an advanced degree—advanced study was seen as a necessity for successful teaching; (c) the teaching assignment—teaching subjects that one had been prepared to teach was determined to be necessary; (d) licensure by the state—meeting the state's expectations for entry to teaching was thought to facilitate their success in teaching the state's PK-12 curriculum; (e) the selectivity of the preparing institution—novice teachers who were "smart" were thought to be much more effective than those who attended less selective colleges; (f) possession of an academic degree in the subject to be taught—passion and compassion are important but they are secondary to subject matter knowledge; (g) the availability of ongoing professional development, particularly development closely related to current lessons and particular curriculum; (h) the accreditation status of the institution—graduates of accredited colleges and universities were considered higher quality than those who had not; and (i) candidate scores on various teacher tests and measures of verbal ability—with the presumption that high scores correlate with satisfactory student performance in the classroom (Allen, 2003; Rice, 2003; Wilson, Floden, & Ferrini-Mundy, 2001).

During the past decade, research scholars and conservative politicians have challenged each of these contentions regarding effective teachers. Their overall claim is that these are "input" measures and that there is little or no compelling evidence that any of them matter in producing student score gains. They attack the research methods used and the claims made and assert that student achievement is the only measure that has validity. With claims that the research evidence regarding effectiveness was insufficient or inadequate, the teacher education community was left without an adequate response to the question of what is an effective teacher and how does one prepare one. The recent efforts of the research community to produce meta-analyses of the literature on teacher preparation and teacher performance have generally raised more questions than answers and resulted in much uncertainty regarding effective teaching and teacher education (Cochran-Smith, 2005; Darling-Hammond & Bransford, 2005).

In large measure, because of the lack of confirmation regarding particular strategies to realize certain outcomes, teacher educators are left to defend current practice without "scientifically based evidence." Those who control the policy agenda now frame the questions, define the way that researchers answer those questions, and determine ways that those answers will affect policy. The only question that is raised is whether this intervention or that action will result in student learning. We find it difficult to defend professional preparation programs when past studies are judged lacking in "scientific rigor" and we must plead for additional time to build a defensible case for our actions. In large measure, because of the uncertainties of the research evidence available, the policy community has come to embrace a single criterion for determining who is or is not an effective teacher—the ability of the teacher to realize and maximize student achievement gains on various assessments of student achievement. Those measures have been accepted as the criterion and norm-referenced tests of student achievement now used to show that schools and students are making "adequate annual progress" as required by the NCLBA (2001).

Teacher Performance

In their chapter in the *Handbook of Research*, Medley and Mitzel (1963) made the case that researchers should turn their attention toward matching teaching behaviors with student per-

formance. The advent of process-product research and the effort to show that particular performances by teachers resulted in particular student outcomes would drive a generation of research scholars. It was always the elusive goal of the competency-based teacher education movement, and the University of Texas Research and Development Center to devote much effort to examining effective teachers and their preparation. Scholars sought to determine the most effective ways of observing teacher work, with much debate centered on whether anecdotal records or rating scales or checklists were more reliable.

Although federally enacted antipoverty programs (particularly Title I of the Elementary and Secondary Education Act or the NCLBA, 2001) and the Education of All Handicapped Children Act (1975) had relied on testing to determine program effectiveness, reliance on minimal competency testing became a characteristic of all schools following the issuance of *A Nation At Risk* (National Commission on Excellence in Education, 1983). That report on the condition of schooling in America prompted virtually every state to test students on the basic skills of literacy and numeracy. The report also seemed to invite even more testing as school reformers and policy makers questioned every aspect of schooling and relied on test results to judge whether reform measures were effective. Testing in schools expanded in the 1980s with policy maker mandates of high school graduation and grade-to-grade promotion tests. Results of these various tests were reported school by school, district by district, state by state, and even internationally. Testing became the way to satisfy the demand for greater school accountability, and it was probably inevitable that score results were now associated with particular teachers.

Teacher testing evolved in ways not dissimilar from PK-12 student assessment. Although teacher testing was widely used by local districts during the 19th and early 20th centuries, it was largely abandoned prior to World War II. Although the National Teacher Examination would emerge in the 1940s, high-stakes teacher testing would not become a major factor in deciding who should teach until the 1970s. States adopted the National Teacher Examination to use in judging the readiness of candidates to teach with the intention that it would be used to identify effective teachers. These were tests of teacher knowledge—basic skills, professional knowledge, and content or subject matter. Recognizing the limitations of such paper-and-pencil tests, states, institutions, and local schools also relied on observational measures to assess the effectiveness of both candidates and experienced teachers. The development and use of candidate portfolios as an assessment tool was probably the most pervasive development of the 1980s, with teacher education programs relying on portfolios to screen candidates into programs, showcase accomplishments throughout the program, and ascertain compliance with program objectives and state licensure standards. The development of commercial products to facilitate the development and use of portfolios was part of this move to emphasize teacher performance.

It would take the establishment of the National Board for Professional Teaching Standards and the creation of the INTASC (both in 1987) to transform the nature of teacher testing and to make the tie between teacher practices and student performance. There was considerable skepticism about whether this could or should be done, however, with one of the leading researchers in the field protesting that

> using student achievement data to evaluate teachers . . . is too susceptible to intentional distortion and manipulation to engender any confidence in the data; moreover, teachers and others believe that no type of test nor any manner of statistical analysis can equate the difficulty of the teacher's task in the wide variety of circumstances in which they work. (Glass, 1990, p. 239)

Medley (1982) provided a state-of-the-art piece on teacher effectiveness in the *Encyclopedia of Educational Research*. He reviewed nine decades of efforts to define effective and ineffective teachers and commented that such research was "most difficult" and offered schemata with nine variables to guide ongoing and future research efforts. The definition offered by Medley for *teacher effectiveness* was "the results

the teacher gets" and the "pupil make towards some specified goal of education." Medley added that the first half of the 20th century had been spent "fumbling" for an answer to the question of what is an effective teacher and warned that policy makers "would act on the best and most recent evidence they have available" (p. 1350). He raised a set of concerns, including the fact that many of the factors that contributed to student learning, "particularly when using student achievement data," were situations not under the teacher's control.

Writing at the time the Reagan administration was gaining influence over education policy, Medley (1982) sensed both the impatience of the policy community and their determination to act with the best information they had on matters of teacher quality and teacher effectiveness. Obviously, his predictions proved to be prescient in the actions taken by policy makers to frame policies pertaining to teaching and teacher education.

Expert Consensus Building

Twenty-five years ago, John Goodlad led a group of experts in the identification of postulates or belief statements about high-quality teacher education. Compiled in Goodlad's (1990) *Teachers for Our Nation's Schools*, the experts reached consensus on the identification of 19 postulates. An example of a so-called Goodlad postulate is that

> the responsible group of academic and clinical faculty members must seek out and select for a predetermined number of student places in the program those candidates who reveal an initial commitment to the moral, ethical, and enculturation responsibilities to be assumed, and make clear to them that preparing for these responsibilities is central to the program. (p. 57)

Goodlad and a carefully selected group of policy makers and professionals identified these belief statements and then studied some 24 colleges to produce "evidence" for their postulates. Goodlad and his colleagues "read selectively and quite a lot . . . studied the histories of education in other professions . . . talked with knowledgeable others . . . probed into the question of current agreement on existing good teacher education . . . and exchanged and discussed various position papers" (pp. 34-35) to arrive at a set of "presuppositions."

Another approach to expert consensus building, taken by Linda Darling-Hammond (2000) in *Studies of Excellence in Teacher Education*, establishes a set of quality indicators by examining teacher preparation programs at a preselected list of 12 colleges and universities that exhibit certain characteristics. A team of scholars visited and then wrote case studies of the approaches to teacher education taken at these institutions; Darling-Hammond then analyzed their findings and prepared a summary. She asserted that high quality in teacher preparation was evident when (a) there was a shared and clear understanding of good teaching; (b) the faculty had practice and performance standards for themselves and their program; (c) the curriculum focused on child and adolescent development and learning theory, including theories about cognition and motivation; (d) the curriculum focused on a context of practice; (e) the curriculum included extensive clinical practice; (f) the institution exhibited common agreements and shared beliefs between university faculty and school practitioners; and (g) the institution made use of multiple instructional strategies to inform candidates for teaching—including the use of modern technologies.

Although Goodlad (1990) began with postulates and Darling-Hammond (2000) concluded with them, there was a common commitment to using quality determinants to judge high-quality teacher preparation. Expert knowledge was used to establish a set of necessary conditions for high-quality teacher education, and teacher educators were expected to make use of these characteristics in the design and conduct of their programs.

Educational Research Meta-Analyses

A second way that academics or scholars have built consensus regarding good teaching and/or teacher education is through the educational research meta-analysis. This can be described as a research-based approach to defining high-quality teaching and teacher education. Essentially it is an effort to unpack the

hundreds of studies conducted during the past 25 years to see if they provide research-based answers. Two major efforts have recently concluded their work and published their results. The first of these works was undertaken by the American Educational Research Association and seen by many as an attempt to provide research evidence to support claims about teacher education. Marilyn Cochran-Smith and Kenneth Zeichner (2005) led this extraordinary effort that generated debates about consensus, appropriate research methodologies, and scientific rigor in the definition of what constitutes effective teacher education. The second of these was an effort of the National Academy of Education Committee on Teacher Education. Linda Darling-Hammond and John Bransford (2005) are leading this continuing work along with a leading group of educational researchers. This group is examining research evidence in nine domains or areas to arrive at consensus about high-quality teacher education. The efforts at defining high quality through research-based approaches should have considerable impact on policy making. Both consensus panels assembled the best American scholars and educational researchers on teaching and teacher education but then had difficulty in arriving at a consensus about what research tells us regarding the effectiveness of teacher education.

Professional Consensus Model

This is an approach that draws on the wisdom of practice and relies on a system of standards and criteria to render judgments about the quality of particular approaches to teacher preparation. Not dissimilar from Goodlad's (1990) identification of a set of postulates or Darling-Hammond's (2000) use of quality indicators, this approach involves professional educators drawn from a broad range of professional societies and organizations to arrive at a consensus about what constitutes good practice. Embraced in the accreditation standards of the National Council for Accreditation of Teacher Education (2001), this approach asserts that "knowledge of the subject matter" is important, that the teacher candidate must be able "to provide multiple explanations and instructional strategies" (pedagogical content knowledge), and that the "candidate work with students, families and communities in ways that reflect the dispositions expected of professional educators" (pp. 14-16). Professional educators—practitioners and scholars—have reached agreements that these standards matter. The professional consensus model has also been used to set licensing requirements for teacher candidates. Definition of the desired skills, knowledge, and dispositions of beginning teachers, with the expectation that teacher education programs will set compatible standards and expectations, is represented here.

The professional consensus model has produced both certification standards and standards for national accreditation. Perhaps the best example of its use is the nearly 15-year effort to define a set of core principles for the licensure of teachers. The work undertaken by the INTASC was done through a coalition of professional groups that put forth a set of standards for the licensure of beginning teachers. Examples of such INTASC standards are that "the teacher understands how students differ in their approaches to learning and creates instructional opportunities that are adapted to diverse learners" or "the teacher understands how children learn and develop and provides learning experiences that support their intellectual, social and personal development" (Council of Chief State School Officers, 2005, pp. 16-18).

Value-Added Modeling

Today these approaches are threatened and professional consensus is dismissed. Teacher educators continue to insist on a broad array of skills, knowledge, and dispositions to judge teachers and an even wider array of standards to judge student performance, whereas neoconservatives or essentialists have captured the policy discourse and insist on a single measure to judge both student success and teacher performance—the ability to show measurable gains on tests of student achievement. Student learning is the mantra of policy makers, and

teacher educators are at peril if they fail to attend to this demand. Our argument is that all of these approaches to defining effective practice stand threatened by a new technology that could remake quality assurance and quality control. Although professionals embraced these approaches and asserted that they represented the means for identifying good practice and effective teaching, it took an agricultural statistician to stand these traditional ways of judging good teaching on their head and to recast the conversation to focus on teacher effectiveness. The work of William Sanders (Sanders & Rivers, 1996) transforms the debates regarding teacher quality by offering a quantifiable methodology for measuring teacher quality. Sanders's work has withstood the scrutiny of other scholars and elevated value-added modeling to a significant place in the policy discourse.

A professor and director of the University of Tennessee Value-Added Research and Assessment Center for nearly 35 years, Sanders and his colleagues (Sanders & Rivers, 1996) developed a statistical model that educational statisticians have reluctantly come to accept and that policy makers assert accurately measures the effects of teachers, schools, and districts on pupil achievement. Using each student's test performance history to determine expected gain scores, Sanders's value-added model reveals the effects of individual teachers on their pupils' test performance. Specifically, Sanders's work highlights the detrimental impact of poor teachers on pupil test performance. Plotted test scores of pupils exposed to 3 consecutive years of "ineffective" teachers dart precipitously downward, whereas those showing pupils lucky enough to have three consecutive effective teachers arch skyward. Skeptics have long pointed to the fact that Sanders's model and the other value-added methods rely on the wholehearted acceptance of multiple-choice tests as the indicator of teacher effectiveness (McCaffrey, Lockwood, Koretz, & Hamilton, 2004). This format, many claim, does not lend itself to assessing the critical thinking skills we agree schooling must instill or enable schools to embrace the democratic or social justice agenda that is an essential part of the progressive vision for schooling. Regardless of its methodological shortcomings or narrow conception of schooling and student achievement, in the 1990s, value-added modeling of teacher effects came to dominate policy making.

This research has had a profound influence on policy makers. Citations of Sanders's work are found throughout policy reports prepared by business-related groups, such as the National Alliance of Small Business report (Koppich, 2001) and the report of the Teaching Commission (2004), whereas his appearance before policy groups and groups of business and academic leaders to describe the Tennessee Value-Added Assessment System greatly influenced the policy discourse on teacher quality. Sanders used data from Tennessee's system of standardized tests, the Tennessee Comprehensive Assessment Program, given annually to students in Grades 3 through 8, to assess teacher performance based on their students' test score gains. His findings about effective and ineffective teachers and their influence on student learning have shaped a series of policy interventions regarding teacher quality and brought new attention to the importance of new forms of professional development for school personnel. The adoption by states of the so-called Sanders's model to determine the effectiveness of their teacher education programs took the effort into the realm of categorizing highly effective and less effective teacher education programs and now shapes the policy agendas of state agencies and higher education coordinating boards in a half a dozen states.

Given the dearth of solid research evidence about what contributes to high-quality teaching that has come from the profession, it is not surprising that policy makers have embraced such measures as represented in value-added modeling. They believe techniques such as value-added modeling offer a tool to make judgments about who is a high-quality teacher. The setting of proficiency targets and creating expectations for adequate yearly progress are part of a mindset that has its roots in essentialism. Today, in particular, policy makers see the possibility of moving beyond simple "end-of-year" achievement scores to measure school or classroom

accomplishments to the possibility of tracking student achievement and linking the results to particular teachers. Teachers who raise test scores will be judged to be high performing and determined to be highly effective. The promise, according to a recent policy brief released by the centrist Progressive Policy Institute, is that "by measuring test score gains from one year to the next, researchers and [school] administrators can better determine the characteristics and conditions that lead to effective teaching" (Leigh & Mead, 2005, p. 1). This statement is prescient for somewhere, lost in the value-added euphoria, is the bigger question—What is it that effective teachers do? How are they prepared? What knowledge, skills, and dispositions do they bring to teaching that make them effective? What are the conditions that make their practice effective? Sanders and others have created a means of identifying those classrooms where the answers may lie. It is the task of today's researchers to enter those classrooms and separate those teachers who produce standardized test takers from those teachers who produce high-performing thinkers.

USING THE ESSENTIALIST ROAD MAP FOR CHANGE

Today we are witnessing the almost paralyzing impact of the so-called teacher wars with sides pitted against sides in an ideological struggle for the future of teaching and teacher preparation. One side emphasizes achievement as opposed to learning, offers an essentialist curriculum and a constrained set of school purposes, seeks the "good enough" teacher, and promotes subject matter knowledge over pedagogy; the other side urges a philosophy of progressivism and a psychology of constructivism and argues that the centerpiece of a democratic society is its public schools. One side promotes teacher centeredness, whereas the other contends that the child has to be at the center of good schooling. Different pedagogies and epistemologies undergird these separate conceptions of teaching and teacher education as the essentialist philosophy of Bagley (Learned & Bagley, 1920) bumps up against the progressive ideology of John Dewey (1933) or William Kilpatrick (1932). Today essentialism is the ascendant philosophy, and it has gained enormous political clout as its advocates have gained political power, and they have used that power to lay the groundwork for changing every aspect of teaching and teacher education.

Legislative bodies at the state and federal level as well as executives at all levels of government are contesting the way we identify and recruit prospective teachers, the way we prepare those teachers for the rigors of today's classrooms and communities, and the way that we induct them and socialize them to the realities of today's schools. The meaning and definition of a highly qualified teacher is the subject of debates by school officials and policy makers. For years, teacher educators wanted to be at the center of the political debates and now that we are, we are often found wanting in our ability to provide solid evidence for the effectiveness of the approaches we use. Unfortunately, we lack valid evidence for many of our assertions and get caught because we often cannot show that what we do makes a positive difference in the lives and learning of all PK-12 students. Justifiably, this increases the intensity of the debates.

During the course of the past century, researchers have sought to define *teacher quality* by looking at teachers' characteristics, measures of teacher competence, and teacher performance. Today, the emphasis is on teacher effectiveness, and an effective teacher is one who can raise student achievement scores on a variety of norm-referenced and commercially produced tests of student achievement. Although standards-based schooling has produced a demand for tests aligned with course objectives and curricular goals, the public is far less demanding of the tests used than that the tests should produce evidence that PK-12 students are learning the content prescribed in the district or state course of study. The use of both norm-referenced and criterion-referenced achievement tests, mandated by the NCLBA (2001), with the fixation on "score gains" for all students, is now a part of the policy discourse. It is perhaps indicative of the times that the Teachers for a New Era program, which bills itself as an initiative to improve the quality of teaching,

places evidence of pupil learning front and center as an indicator of quality teaching programs. The design principles of the Teachers for a New Era program call on all teacher education programs to be driven by evidence, including attention to "pupil learning that has occurred under the tutelage of teachers who are graduates of the program" (Carnegie Corporation of New York, 2001, The Role of Pupil Learning section, para. 1).

As much as we are concerned about effective schools, we seem even more concerned about the efficaciousness of teacher education. Does it make a difference? Defining *teacher quality* by determining the effects of teaching on student learning is the goal of contemporary efforts to measure the effectiveness of teaching. Matters of teacher effectiveness and teacher quality have been greatly influenced by the efforts to measure student performance. Almost a century after the commissioning of the CFAT study of teacher education (Learned & Bagley, 1920), there remains much uncertainty about the appropriateness or efficacy of using student scores on various measures of attainment. Advocates of such efforts are criticized for their narrow vision of school (as learning institutions) and urged to consider the full range of student needs and the larger purposes of schools.

The traditional abhorrence with which progressives view standardized testing and the use of narrowly described measures of student learning to judge the effectiveness of teachers and schools is difficult to overcome. Antipathy to such measures by teacher education is noted in the original Carnegie bulletin (Learned & Bagley, 1920) on teacher education and has persisted throughout the intervening 85 years. Measures used to judge student performance are criticized for their flaws in conception and design and for the adverse effects they have on particular populations. Regardless of the inappropriateness and/or limitations of such assessments, the neoessentialists have prevailed on this front and student learning is largely defined as represented by these score gains with disaggregation by race and ethnicity, gender, handicapping condition, migratory status, or socioeconomic status. Despite the efforts of progressives to have policy makers rely on other measures (rather than relying on single test scores or one-time testing programs) and much concern regarding schools as "test factories" or "test prep" sites, with teachers "teaching to the test" or being "testers" rather than "teachers," the truth is that in the quest to make schools more accountable, test scores are about the only measure consistently embraced by the public (Johnson & Duffett, 2005).

Measures of student attrition and indiscipline, course-taking patterns and advanced placement test successes, successful transition to college or work, time to completion of study, and parent and/or employer "satisfaction" are measures commonly offered by educators. Policy makers have accepted such measures but want means for knowing that students have learned a prescribed course of study and met learning objectives. Composite scores for classes and schools, districts, and states have become the measure used. So-called progressives can choose to criticize policy making that narrows the purposes of schooling to this degree, but they do so at their own peril until they are able to convince the public of the worthiness of other measures (Cochran-Smith, 2004).

It is right and appropriate to decry this reality, but most contemporary policy makers have embraced this narrower definition of school purpose and school curriculum. They want singular measures to be used for accountability purposes. Borrowing from Bagley and the essentialists, they have accepted a philosophy that places a premium on student learning in a narrow range of academic subjects. Policy makers and others have put this kind of student learning at the center of the current debate regarding school reform and renewal. After a century of efforts to embrace a more expansive set of goals and purposes for American schools, policy makers have succumbed to the need for readily definable measures of success and the temptation to measure "our" success against the success of other nations and schools. Using student assessment and a well-defined curriculum to guide the practice of teachers has been difficult to do; more difficult has been using the results of student achievement tests to assess

the performance of teachers. There has been an aversion to using such measures and much resistance to such approaches by teachers who contend that nonschool factors are at least as important as in-school variables in determining the success of students. Thus far, teacher organizations have been successful in withstanding this movement and teacher educators have protested against such efforts, but it seems unlikely that they will forestall reliance on pupil scores in determining either their own efficaciousness or their tenure. At the same time, teacher educators will have little success in claiming professional consensus or seeking expert opinion until they can show that what they do matters in the learning of pupils. PK-12 student performance is now not only the condition that guides the definition of teacher effectiveness but also the measure that ultimately will decide whether one form or another of teacher preparation or professional development will prevail.

REFERENCES

Abell Foundation. (2001). *Teacher certification reconsidered: Stumbling for quality*. Baltimore: Author.

Allen, M. (2003). *Eight questions on teacher preparation: What does the research say?* Denver, CO: Education Commission of the States.

Ballou, D., & Podgursky, B. M. (2000). Reforming teacher preparation and licensing: Continuing the debate. *Teachers College Record*, *102*(1), 5-27.

Berliner, D. C. (1979). Tempus educare. In P. Peterson & H. Walberg (Eds.), *Research in teaching* (pp. 120-135). Berkeley, CA: McCutchan.

Carnegie Corporation of New York. (2001). *Teachers for a New Era prospectus*. New York: Author. Retrieved June 8, 2005, from http://carnegie.org/sub/program/teachers_prospectus.html

Cochran-Smith, M. (2004). Taking stock in 2004: Teacher education in dangerous times. *Journal of Teacher Education*, *55*(1), 3-5.

Cochran-Smith, M. (2005). Teacher education and the outcomes trap. *Journal of Teacher Education*, *56*(5), 411-417.

Cochran-Smith, M., & Zeichner, K. (2005). *Studying teacher education: The report of the AERA Panel on Research and Teacher Education*. Mahwah, NJ: Lawrence Erlbaum.

Council of Chief State School Officers. (2005). *INTASC standards development*. Retrieved December 26, 2005, from http://www.ccsso.org/Projects/interstate_new_teacher_assessment_and_support_consortium/projects/ standards_development/791.cfm

Darling-Hammond, L. (Ed.). (2000). *Studies of excellence in teacher education*. Washington, DC: American Association of Colleges of Teacher Education.

Darling-Hammond, L., & Bransford, J. (Eds.). (2005). *Preparing teachers for a changing world: What teachers should learn and be able to do*. San Francisco: Jossey-Bass.

Dewey, J. (1933). *How we think*. Lexington, MA: D. C. Heath

Education of All Handicapped Children Act of 1975, Pub. L. No. 94-142.

Evertson, C. M. (1982). Differences in instructional activities in higher and lower achieving junior high English and math classes. *Elementary School Journal*, *82*(4), 329-351.

Ferguson, P., & Womack, S. T. (1993). The impact of subject matter and education coursework on teaching performance. *Journal of Teacher Education*, *44*(1), 55-63.

Flexner, A. (1910). *Medical education in the United States and Canada: A report to the Carnegie Foundation for the Advancement of Teaching*. New York: Carnegie Foundation for the Advancement of Teaching.

Frazier, J. (in press). *Preparing the American teacher: Teacher education in the United States, 1776-2006*. New York: Teachers College Press.

Gage, N. (1972). *Teacher effectiveness and teacher education*. Palo Alto, CA: Pacific Books.

Glass, G. V. (1990). Using student test scores to evaluate teachers. In J. Millman & L. Darling-Hammond (Eds.), *The new handbook of teacher evaluation* (pp. 229-240). Newbury Park, CA: Sage.

Goldhaber, D. D., & Brewer, D. J. (2000). Does teacher certification matter? High school teacher certification status and student achievement. *Educational Evaluation and Policy Analysis*, *22*(2), 129-146.

Good, T. L., & Brophy, J. E. (1986). School effects. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 570-602). New York: Macmillan.

Goodlad, J. I. (1990). *Teachers for our nation's schools*. San Francisco: Jossey-Bass.

Gutek, G. L. (2004). *Philosophical and ideological voices in education*. Boston: Allyn Bacon.

Hanushek, E. A. (1997). Assessing the effects of school resources on student performance: An update. *Education Evaluation and Policy Analysis*, *19*(2), 141-164.

Johnson, J., & Duffett, A. (2005). *Attitudes about teaching: Including the views of parents, administrators, teachers and the general pubic*. New York: Public Agenda.

Kilpatrick, W. H. (1932). *Education and the social crisis: A proposed program*. New York: Liveright.

Koppich, J. (2001). *Investing in teaching*. Washington, DC: National Alliance of Small Business.

Labaree, D. F. (2004). *The trouble with ed schools*. New Haven, CT: Yale University Press.

Learned, W. S., & Bagley, W. C. (with McMurry, C. A., Strayer, G. D., Dearborn, W. F., Kandel, I. L., & Josselyn, H. W.). (1920). *The professional preparation of teachers for American public schools: A study based upon an examination of tax-supported normal schools in the state of Missouri* (Bulletin No. 14). New York: Carnegie Foundation for the Advancement of Teaching.

Leigh, A., & Mead, S. (2005). *Lifting teacher performance policy report*. Washington, DC: Progressive Policy Institute.

McCaffrey, D., Lockwood, J. R., Koretz, D. M., & Hamilton, L. S. (2004). *Evaluating value-added models for teacher accountability*. Santa Monica, CA: RAND.
Medley, D. (1978). *Teacher competence and teacher effectiveness: A review of the process-product research*. Washington, DC: American Association of Colleges for Teacher Education.
Medley, D. (1982). Teacher effectiveness. In H. E. Mitzel (Ed.), *Encyclopedia of educational research* (pp. 1345-1352). New York: Free Press.
Medley, D., & Mitzel, H. (1963). Measuring classroom behavior by systematic observation. In N. L. Gage (Ed.), *Handbook of research on teaching* (pp. 247-328). Chicago: Rand McNally.
Mitzel, H. E. (1960). Teacher effectiveness. In C. W. Harris (Ed.), *Encyclopedia of educational research* (pp. 1481-1485). New York: Macmillan.
National Academies. (2005). *Study of U.S. teacher education programs (STEP)*. Retrieved December 5, 2005, from http://www.nationalacademies.org/cfe/teacherprep.html
National Commission on Excellence in Education. (1983). *A nation at risk*. Washington, DC: Government Printing Office.
National Council for Accreditation of Teacher Education. (2001). *Professional standards for the accreditation of schools, colleges and departments of education*. Washington, DC: Author.
No Child Left Behind Act of 2001, Pub. L. No. 107-110, 115 Stat. 1425 (2002).
Null, J. W. (2003). *A disciplined progressive educator: The life and career of William Chandler Bagley*. New York: Peter Lang.
Paige, R. (2002). *Meeting the highly qualified teacher challenge: The secretary's annual report on teacher quality*. Washington, DC: U.S. Department of Education.
Ravitch, D. (2000). *Left back: A century of battles over school reform*. New York: Simon & Schuster.
Rice, J. K. (2003). *Teacher quality, understanding the effectiveness of teacher attributes*. Washington, DC: Economic Policy Institute.
Sanders, W. L., & Rivers, J. C. (1996). *Cumulative and residual effects of teachers on future academic achievement*. Knoxville, TN: University of Tennessee Value Research and Assessment Center.
Stallings, J. (1985). Effective elementary classroom practices. In M. J. Kyle (Ed.), *Reaching for excellence: An effective schools sourcebook*. Washington, DC: National Institute of Education.
Teaching Commission. (2004). *Teaching at risk: A call to action*. New York: Author.
Wilson, S., Floden, R., & Ferrini-Mundy, J. (2001). *Teacher preparation research: Current knowledge, gaps, and recommendations*. Seattle: University of Washington, Center for Teaching Policy.

David G. Imig *was the president and chief executive officer of the American Association of Colleges for Teacher Education for 25 years. Beginning in January 2006, he began serving as a professor of practice in the College of Education at University of Maryland–College Park.*

Scott R. Imig *is an assistant professor in the Curry School of Education at University of Virginia. He also serves as director of the Teaching Assessment Initiative—the research component of the University of Virginia's Teachers for a New Era grant.*