Knowing What Teachers Know

Article *in* Review of Research in Education · March 2015

DOI: 10.3102/0091732X14557001

CITATIONS
23

READS
660

2 authors:

Drew Gitomer
Rutgers, The State University of New Jersey
98 PUBLICATIONS **1,696** CITATIONS

SEE PROFILE

Rob Zisk
Rutgers, The State University of New Jersey
2 PUBLICATIONS **23** CITATIONS

SEE PROFILE

Chapter 1

Knowing What Teachers Know

Drew H. Gitomer
Robert C. Zisk
Rutgers, the State University of New Jersey

That teachers need to have deep knowledge has been an article of faith among philosophers, educators, and policymakers tracing back to Socrates. But the specifics of what teachers ought to know and how they can make use of knowledge during teaching continue to evolve (Biesta & Burbules, 2003) and are related to underlying conceptions of the role of teachers and the teaching profession that have evolved in the United States across its entire history (see Kafka, in press, for a review of the history of teaching in the United States).

Coupled with these changing ideas about what teachers need to know has been the educational authority's interest in assessing and warranting teachers' and prospective teachers' knowledge. Thirty years ago, in the last explicit treatment of teacher testing in *Review of Research in Education*, Haney, Madaus, and Kreitzer (1987) provided accounts of teacher assessments that were introduced as early as colonial times. The chapter remains an important critique of teacher testing by focusing on the use and social implications of such tests and on the ways in which teacher assessment might be improved.

Although some of the concerns that Haney et al. (1987) raised 30 years ago are still germane today, there have been substantial changes as well. Not only have assessments that serve a regulatory function (e.g., licensure) evolved, but an intensive scholarly focus, during the last 30 years in particular, has produced a broad literature that attempts to conceptualize and empirically examine the knowledge base of teachers (e.g., Ball, Thames, & Phelps, 2008; Grimmett & MacKinnon, 1992; Shulman, 1987).

Review of Research in Education
March 2015, Vol. 39, pp. 1–53
DOI: 10.3102/0091732X14557001

Haney et al. (1987) note that from colonial times through the first half of the 19th century, examinations of teachers were essentially oral interviews that were conducted by local officials and that focused on moral fitness, classroom/student management, and fundamental literacies. Over the course of the 19th century, counties and states attempted to impose more control on localities, partly due to the finding that those conducting the interviews were often ill-equipped to evaluate the examinees' responses and partly because of increasing state interest in public education. Conducting oral interviews continued to be the dominant form of assessment during this transitional period of control. As the 20th century unfolded, however, written objective tests became more common as the science and practice of standardized assessment first emerged.

In the 1930s, the Cooperative Test Service, an agency of the American Council on Education (ACE), developed objective tests in more than 40 subject areas at the senior high school and junior college levels. It was these tests that largely formed the basis for the development by the ACE of the National Teacher Examination (NTE)—the first national assessment for teachers (Haney et al., 1987). Given the national scope of these assessments and the fact that these were the first tests for teachers with any significant research base and scholarly record, it is at this point in time that we begin to focus the current analysis.

The current chapter focuses on philosophical and scientific developments in the assessment of teacher knowledge. Although we give attention to assessments that serve licensing functions, our goal is not to update the implications of regulatory assessments within a policy context. Nor do we limit the use of teacher assessments to only regulatory functions. Instead we focus on how conceptions of teaching and teacher development, together with research on teacher knowledge and advances in assessment design and related technologies, have led to new ideas about what teachers need to know, how to assess that knowledge, and how such information can serve a multiplicity of purposes.

It is important that we clarify how the chapter, as a review of the assessment of teacher knowledge, is organized. The goal is to provide an overview of major developments in the assessment of teacher knowledge and to illustrate those developments with representative and significant efforts. We made evaluative judgments of importance based on their presence in practice, their contributions to policy, and/or their contributions to the research base on teacher assessment. Our judgments are subjective in some sense, and we try to provide our rationale for including these illustrations as demonstrations of major developments. However, by no means is this a comprehensive review of every assessment of teacher knowledge that has been developed.

This chapter is also not intended to be a comprehensive review of all of the work that has been carried out in two areas that have shaped these assessments: (1) research that focuses on the nature of teacher knowledge across disciplines; and (2) contexts and policy initiatives that have influenced the nature of teacher assessment. In both cases we illustrate our points with relevant research and policy efforts that have shaped the directions of assessments, but our intent is not to be exhaustive in the reviews of all relevant research and policy.

We conducted our literature review for this chapter by first searching for all relevant papers that addressed teacher knowledge. Search terms included the following: teacher knowledge, conceptions of teacher knowledge, teacher-enacted knowledge, teacher practical knowledge, assessments of teacher knowledge, content knowledge for teaching (CKT), content knowledge for teaching and CKT by subject areas, pedagogical content knowledge (PCK), pedagogical content knowledge and PCK by subject areas, assessments of content knowledge for teaching, CKT validation, assessments of PCK, PCK validation, core teaching practices, and high-leverage practices. We then selected from this very broad literature articles and examples that we judged to have had significant influence and impact on assessments of teacher knowledge and used these examples to support the conceptual argument and framework of the chapter.

Given the scope of the entire volume, we conducted a more focused search on assessments of teacher knowledge with respect to English language learners (ELLs). That search did not yield much, as the knowledge base for what teachers need to know to teach ELLs is also limited (Faltis & Valdés, in press). Therefore, we do not focus much of the chapter's attention explicitly on ELL teacher assessments but instead provide at the conclusion of the chapter some discussion about the state of the field and future directions.

We organize the chapter using the validity framework described in the 2001 landmark publication by National Academies Press, *Knowing What Students Know* (National Research Council, 2001a). Key points, applicable to assessments of teacher knowledge, include the following:

1. Educational assessments have been inadequate largely because they were based on impoverished views of cognition.
2. Assessments can be thought of as having three components—a model of cognition (what we are trying to measure), a model of observation (how we collect evidence), and a model of interpretation (how we make sense of the evidence).
3. Building off a long history, the validity of assessments must be considered in terms of how scores are interpreted and used within particular contexts for particular purposes.

This chapter uses and builds on this framework to review the past 75 years of teacher assessment by considering four different assessment foci, described in Table 1, beginning with the assessments embedded within the first national assessment system for teachers, the NTE. We review the models of cognition, observation, and interpretation that undergird each of the major teacher assessment foci. We also review the nature and evidence of the validity arguments for the assessments that underlie each of the foci.

It is important to recognize that these foci are not temporally bound or linear in their development. Although it is true that the emergence of the four foci largely developed in sequence over time, assessments of basic skills and content knowledge,

TABLE 1
Major Foci in the Assessment of Teacher Knowledge

Focus	Underlying Premise	Sample Assessments	Interpretation	Use(s)
Teacher as educated professional	Teachers should have a general set of intellectual competencies.	National Teacher Examination Core Battery (General Knowledge, Communication Skills, Professional Knowledge); Praxis I	Test scores indicate extent to which teachers have fundamental knowledge that is critical for teachers to know.	Licensure
Teacher as content knowledge professional	Teachers should understand the specific subject-matter content they will teach.	Praxis II; state-specific content licensure assessments	Test scores indicate extent to which teachers have fundamental knowledge of the subject(s) they will teach.	Licensure
Teacher as content knowledge for teaching professional	Teachers should understand both the subject-matter content and how to teach that content.	Content knowledge for teaching (CKT) assessments (Mathematics Knowledge for Teaching [MKT]); National Board for Professional Teaching Standards (NBPTS) Assessment Center	Test scores indicate extent to which teachers hold subject-matter knowledge in ways needed to successfully teach.	Research into teaching; professional development; potential uses: teacher education; licensure; advanced certification
Teacher as knowledge-rich practitioner	Teachers draw on content as they carry out teaching practices.	NBPTS portfolios; high-leverage practices; core practices	Test scores indicate extent to which teachers have the ability to carry out integrated teaching practices that call on different forms of knowledge relevant to teaching.	Research into teaching; professional development; potential uses: teacher education; licensure: advanced certification

which were the initial foci of teacher assessment, remain the dominant forms of consequential assessments for teachers. Thus, any historical trajectories represent developments in theory and assessment design but do not describe historical transitions in assessment implementation.

The four foci are characterized by assessment tasks that increasingly approximate teaching practice. The first two foci have very tenuous connections to teaching practice, as they simply hold that certain types of knowledge are prerequisite to competent teaching. Yet these foci assiduously avoid claiming that greater amounts of knowledge beyond the licensure standard imply better teaching. These tests are designed as minimum competency tests only.

The more recent CKT focus makes a much stronger claim about its relationship to teaching. Each assessment task is designed to tap knowledge that is called for in carrying out the work of teaching (Ball et al., 2008). The most recent focus goes even further and attempts to assess the enactment of teacher knowledge in the carrying out of teaching practices. Although this latter focus has not yet resulted in substantial deployment of assessments, promising work that is likely to influence the future of teacher knowledge assessment is moving forward.

We limit our discussion of assessment to structured tasks that ask test takers to provide specific types of evidence and for which there are clear expectations for what constitutes a quality response. In all of these tasks, test takers are specifically asked to demonstrate their knowledge. An incorrect or poor-quality response is taken as negative evidence of an individual's knowledge.

The implication of this stance is that we do not consider tasks as assessments of teaching that occur in more naturalistic/nonstructured settings. So, for example, we do not include teacher observations of everyday teaching as a source of evidence about teacher knowledge. This is based on the reasoning that although one may draw inferences about a teacher's knowledge based on some demonstration of skilled teaching performance, a lack of demonstration does not necessarily give information about the teacher's knowledge. For example, if a teacher is leading an effective discussion and asking productive questions, it might be reasonable to infer that a teacher has a certain command of both the content of the discussion and the ways of engaging students in the content. However, a poor discussion may be caused by a myriad of reasons (e.g., classroom management) that have nothing to do with whether or not a teacher knows the content or even principles that are associated with effective discussions. Such naturalistic observations are also not a very efficient and direct method for capturing teacher knowledge. Therefore, we examine only assessment models that ask the teacher to explicitly respond to a prompt that is designed to assess a particular aspect of teacher knowledge.

HISTORICAL PERSPECTIVES ON TEACHER KNOWLEDGE

Conceptions of teacher knowledge—what teachers should know—have developed, and continue to develop, from multiple influences including philosophical and psychological views on the nature of knowledge, philosophical and political

perspectives on the purposes and nature of teaching, sociopolitical demands, and research on teaching.

We begin this discussion by considering a long-standing dialogue that has concerned both philosophers and psychologists in terms of what constitutes knowledge. Specifically, how do we make sense of the distinction between *knowing that* and *knowing how*? Fantl (2014) summarizes much of the historical philosophical debate that harkens back to the ancient Greek philosophers' reference to *epistêmê* (knowledge) and *technê* (craft or art). Even then, the boundaries had not been clear, as Fantl references the Plato argument that to "live a good life, we ultimately need knowledge in the forms of virtue." Yet philosophers have debated whether Plato meant that a virtuous life requires only knowledge of what is virtuous or that it also requires knowledge of how to live virtuously.

The extent to which *knowing that* and *knowing how* are independent constructs has also been the subject of substantial philosophical treatment. Using an example such as riding a bicycle makes apparent the difficulty in resolving the debate. One can imagine a young person who rides a bicycle quite skillfully, yet cannot articulate accurately any of the actions that contribute to successful riding. One can also imagine an elderly, infirm engineer who can describe in exquisite detail how to ride a bicycle. Yet, because of physical limitations, the engineer cannot actually ride the bicycle. Which of these individuals knows *how* to ride a bicycle?

Ryle (1946) presents a view that is useful for considering the assessment foci described in this chapter. He argues that the distinction between *knowing that* and *knowing how* is not particularly productive. Rather, the useful distinction is between theoretical and practical knowledge. Practical knowledge is that which results in *intelligent behavior*. He makes the distinction "between the museum-possession and the workshop-possession of knowledge. A silly person can be stocked with information, yet never know how to answer the particular question" (p. 16).

Cognitive psychologists, and later researchers in artificial intelligence, reconsidered the *knowing that/knowing how* distinction in terms of declarative and procedural knowledge. Chi and Ohlsson (2005) summarize declarative knowledge as descriptive and use-independent, embodying concepts, principles, ideas, schemas, and theories. Procedural knowledge, in contrast, is prescriptive and use-specific and consists of associations among goals, situations, and actions. Substantial research in a variety of domains attempted to develop psychological models that connected and described the transformation of declarative knowledge into procedural knowledge (e.g., Anderson, 1976; Kintsch, 1998; Rumelhart & Norman, 1978). As declarative knowledge becomes proceduralized, access to the declarative knowledge may be lost. The skilled performer may have little ability to describe the actions he/she undertook in executing that performance.

The foregoing distinction is important in two ways. First, philosophical treatments of teaching have certainly focused on knowledge that can be considered practical or procedural. The focus has been dominated by what teachers do or ought to do. However, this stands in stark contrast to dominant forms of assessment that have

focused on the declarative or theoretical. The second important feature is that the assessment foci we describe in this chapter show an increasing approximation to knowledge as practice.

Considerations of what teachers should know must be made in light of conceptions of the nature and role of teaching in helping students learn. On one hand, philosophical conceptions have been far-ranging and have often viewed teaching as a set of interactions in which the teacher's role is to facilitate learning through questioning and feedback and by providing experiences that spur students' thinking and build on the knowledge they already possess. This vision is quite distant from the primarily didactic instruction that has been the hallmark of most American educational practice (e.g., Goodlad, 1984). These philosophical views are similarly inconsistent with the dominant teacher and student assessment practices during the past century. The underlying premise that has guided instructional and assessment practice has been that the teacher holds knowledge and, primarily through lecture and instructional materials, transmits content to students (e.g., Goodlad, 1984).

The perspective that teaching is both interactive and facilitating has deep roots in history. Mintz (2007) discusses Socrates's view of pedagogy as *midwifery*, in which the role of the teacher is to create the proper conditions for birth (learning) to occur. Furthermore, the importance of the relation between the midwife and the mother is analogous to the relation of student and teacher, and the match between the student and teacher is critical to the quality of the educational experience.

In *Emile*, Rousseau (1979) articulates the role of the teacher as one who guides rather than tells:

> It is a question of guidance rather than instruction. He must not give precepts, he must let the scholar find them out for himself. . . . The teacher's art consists in this: To turn the child's attention from trivial details and to guide his thoughts continually towards relations of importance which he will one day need to know, that he may judge rightly of good and evil in human society.

John Dewey, perhaps the most influential philosopher guiding progressive education, actually says little about the role of the teacher in learning. Biesta and Stengel (in press) present Dewey's teaching view as follows:

> The teacher is therefore not the "guide on the side" but the designer of "a vital and personal experiencing." The teacher cannot make growth happen but *can* "determine the medium" that makes growth possible. S/he can do that as the member of the group "having the riper and fuller experience and the greater insight into the possibilities of continuous development found in a suggested project."

Although Dewey has shifted the focus of educational attention from the act of teaching to the fact of children's lives, it does not mean that there is no authoritative task for the teacher.

> It is no longer a question of how the teacher is to instruct or how the pupil is to study. The problem is to find what conditions must be fulfilled in order that study and learning will naturally and necessarily take place, what conditions must be present so that pupils will make the responses which cannot help having

learning as their consequence. The pupil's mind is no longer to be on study or learning. It is given to doing the things that the situation calls for, while learning is the result. The method of the teacher, on the other hand, becomes a matter of finding the conditions which call out self-educative activity, or learning, and of cooperating with the activities of the pupils so that they have learning as their consequence. (Boydston, 1984, p. 267)

Interestingly, the view of interaction and facilitation has not been the dominant form of instruction. Instead, it has been much more didactic. And as we will see, until recently, teacher assessments paid little mind to these views of pedagogy and instead focused on explicit knowledge held by the teacher. One alternative theoretical position does stand out as supporting didactic instruction—Skinner's (1965) theory of positive reinforcement.

In addition to the *how* of teaching, there have been substantial consideration and tensions regarding what ought to be taught. Certainly, in the past century or so, we have seen a variety of imperatives that, at times, can be contested but can also coexist. In the United States, early views of the purposes of teaching, even through the early part of the 20th century, focused on morality and citizenship. Education's primary purpose was to acculturate students. Kafka (in press) notes, "Teachers were to instill in their students the values of hard work, respect for authority, and loyalty to God, family, and country."

The long-standing focus on cultivating citizenship, democratic, and social values through education is captured in an edited volume by Ravitch and Viteritti (2001). Noddings's (2005) work on caring addresses the role of schooling in helping students develop in ways that go beyond academic goals. However, what those values are and how they ought to be developed within a diverse community have always been contested terrain. One only has to look at the now discredited attempts to "civilize the Indians" (*Official Report of the Nineteenth Annual Conference of Charities and Correction*; Prucha, 1892/1973) to understand how terms like *citizenship* have been politicized and used for less than noble purposes. Debates about addressing the educational needs of students from the nondominant culture persist and include considerations of race, language, gender, and sexuality (e.g., August & Hakuta, 1997; Freire, 1970; Gay, 2010).

Other views of educational purposes are much more concerned with a core set of academic skills and knowledge as the basis for becoming a fully functioning member of society. Of course, the precise definition of what constitutes the core has long been a matter of debate, but various specifications have been offered by scholars and organizations. Hirsch (1988) lays out a common core of knowledge that should be held by all individuals in our society. Paideia (National Paideia Center, 2014) is an organization that has built schools and curricula around the idea that a body of core content should be integrated with critical thinking skills and activities. And, of course, the educational standards movement, including the current Common Core effort (National Governors Association Center for Best Practices & Council for Chief State School Officers, 2010), represents attempts by states and subject-matter organizations to articulate expectations for what students should learn. The definition of

core knowledge is certainly subject to evolution and includes more modern specifications described in far-ranging articulations that include works by Hirsch (1988) and National Paideia Center (2014) and, more recently, broad standards documents such as the Common Core State Standards (National Governors Association Center for Best Practices & Council for Chief State School Officers, 2010) and the Next Generation Science Standards (NGSS Lead States, 2013).

FOCUS 1—TEACHER AS EDUCATED PROFESSIONAL

Assessments built on the teacher as educated professional focus were introduced in March 1940 when 3,726 teacher candidates took the first NTE in 23 testing centers across the country. One year before, the ACE appointed a National Committee on Teacher Examinations comprised of primarily state and large district school superintendents and deans or presidents of teacher colleges. The goals were to support superintendents in selecting new teachers and to provide teacher preparation institutions with national information so that they could better gauge the selectivity of their own institutions (see A. J. Wilson, 1985).

The Cognitive Model for Teacher as Educated Professional

The National Committee viewed teaching ability within the larger frame of factor theories of intelligence and cognitive abilities that was the dominant perspective on testing at the time. Charles Spearman explored the correlations among different ability measures and identified the general factor of intelligence, *g*, along with specific factors. Louis Thurstone, using factor analytic methods, pursued the idea of multiple factors of human abilities (see Carroll, 1993). This overarching view considered human abilities, both general and specific, as sets of traits that were applied across contexts. Specific traits were not only judged to be systematically related to other traits (the general factor) but also deemed to measure something unique (the specific factor).

Within this frame, teaching ability was viewed as a complex set of intellectual factors. The ACE, under the direction of Ben Wood, also immediately recognized that all teaching could not be assessed through standardized objective tests of knowledge:

> Common sense has long told us . . . that objective examinations do not and cannot measure the total subtle complex which we call teaching ability. When we are concerned with a complex that includes such fundamentally different factors as intelligence, general culture, professional information, special subject-matter mastery, moral character, interest in children, emotional stability, personality, physical health and energy, etc., it is, or ought to be obvious that no one type of measurement, such as the examination, can be an adequate basis for selection or eligibility ranking. (ACE, 1940, p. 3)

Wood went on to argue that despite the fact that only a subset of information could be assessed using extant testing procedures, such tests could still contribute to selection processes by providing comparable measures that were not compromised by

the varying standards operating across institutions that prepared teachers. He dismissed the criticism that Messick (1989) would later label *construct under-representation* by noting:

> To neglect examinations of intelligence, general culture, and professional information because they do not also measure personality, moral character, interest in children, and other important factors that determine teaching ability, would be as illogical as to neglect the use of the clinical thermometer and stethoscope because they do not measure a thousand other important diagnostic factors…The validity of the examinations should be judged by the accuracy with which they measure, not the total complex of teaching ability, but those parts which they are designed to measure, namely intelligence (linguistic and quantitative), general and special cultures of the types judged desirable by the teacher-selecting authorities, and professional information. (ACE, 1940, p. 4)

Wood's position, in the field of teacher licensure, has remained relatively unaltered. Although there is a recognition that teaching involves a complex set of abilities, licensure tests have been designed and defended as only measuring a subset of those skills (National Research Council, 2001b). Furthermore, licensure tests are positioned as tests that should never be used alone and that should be only one part of a larger licensure/selection process. However, those who have developed and used licensure tests have rarely described the details of other elements of that process.

Thus, the first generation of these assessments was based on an assumption that every teacher has a "certain minimum of intelligence, culture, and professional preparation" (A. J. Wilson, 1985, p. 20). In fact, Wood had developed these kinds of assessments and administered them to all college graduates as a marker of general educational quality, examining differences in performance across different fields. Thus, the focus on teaching was simply an extension of this earlier development of educational accountability measures.

For Wood, intelligence was operationalized as being able to reason, comprehend, and express oneself in English. Culture was defined as knowing current social problems, history and social studies, science, fine arts, literature, and mathematics. Professional information included knowledge of education and social policy, child development and educational psychology, guidance, individual and group analysis, and either elementary or secondary school methods. There was also an assumption that there is specific content knowledge that general elementary teachers and subject-specific teachers should have. Developed in 1940, this underlying cognitive model was the foundation of teacher testing for the next 60 years. Up until the NTE transitioned over to Praxis™ during the latter half of the 1990s (S. M. Wilson & Youngs, 2005), it consisted of three primary components in its core battery: communication skills, general knowledge, and professional knowledge.

One major development that did occur during these 60 years was the introduction of basic skills tests for teachers. Amid the general concern for educational performance in the United States (The National Commission on Excellence in Education, 1983) and the perceived low standards of teacher education institutions (see Levine, 2006), states began to implement policies that mandated that teachers demonstrate

basic competency in reading, writing, and mathematics in order to be licensed. In addition to using state-specific tests, multiple states used these basic skills tests, which included the Pre-Professional Skills Test (later called Praxis I® and as of 2013 called the Praxis Core Academic Skills for Educators Test; Educational Testing Service [ETS], 2014b). Still in use today, these tests are basic literacy tests that are not tied to a theory of teaching. They were simply developed to assure the public that teachers were minimally competent in the basic skills of reading, writing, and mathematics.

The Evidence Model for Teacher as Educated Professional

In almost all cases, the evidence for assessments in this focus was developed through responses to multiple-choice items, though some writing assessments asked candidates to write short essays. The first versions of the NTE asked candidates to select the best option choice from among five choices. Several examples from the first NTE administration (Flanagan, 1941) are presented in Appendix A. The first test was extraordinarily lengthy. The common examination was 480 minutes, plus optional elementary/subject-specific tests. The time allowed to complete each part of the test was as follows: *intelligence*, 120 minutes; *general culture*, 180 minutes; *professional information*, 120 minutes; and *contemporary affairs*, 60 minutes. Ninety minutes were allowed to complete any of the optional elementary/subject-specific tests. Testing time was shortened such that the common examination was about 3 hours long by 1951 (Humphry, 1966). The NTE maintained the five-option multiple-choice items for its entire existence.

Basic skills tests generally collected the same type of evidence through multiple-choice exams. Both the California Basic Educational Skills Test™ (Pearson Education, Inc., 2014) and the Praxis I® used five-option multiple-choice items, although the more recent versions also include essays to assess writing. More recent versions of the ETS tests include a range of selected-response item formats.

Assessments in this focal category were obviously developed as paper-and-pencil tests, and some are still offered in that format. However, many tests are now delivered in computerized modes as well, though they are designed to ask the same content and are scored in the same manner. From the test developer's perspective, there is no difference between administrative modes (e.g., ETS, 2014a).

The Interpretive Model for Teacher as Educated Professional

Scores on these assessments were simply transformations of number of items correct to scaled scores. The original scaled scores on the NTE ranged from 0 to 100 with the mean set at 50 (Flanagan, 1941). The tests were designed to yield internally consistent measures. As psychometric methods advanced, processes such as test equating were used to ensure comparability across administrations. Over the years, the scales underwent some modifications and did not try to emulate the intelligence test scales that informed their original design. However, the basic transformation of items correct to scaled scores continued.

The most important interpretation for any of these assessments, as with any licensure test, is how an individual score compares with a passing standard. Because licensure testing is the province of the states, each state sets its own passing standard for a test using one of a number of common methods described in the rich literature on standard setting (e.g., Cizek & Bunch, 2007; Hambleton & Pitoniak, 2006).

The first score reports provided individuals with tabular and graphical representations of scores for each subtest. Each score was compared to several reference groups: an expected average for all individuals with an average intelligence score, the average of all candidates from that administration, and score levels designated to be superior and exceptional (scaled scores of 70 and 80, respectively; Flanagan, 1941). Percentages of individuals from the two groups selecting each response are presented along with the items in Appendix A.

Research Findings for Teacher as Educated Professional

Research on these assessments most often focused on the distribution of test scores across different population groups of interest. Even the earliest reports presented subtest scores for men and women, intended subject areas and grade levels, educational degrees, and prior teaching experience (Flanagan, 1941).

One of the most important and consistent findings has been that basic skills tests produce disparate impact. African American and Latino teacher candidates typically pass these tests at much lower rates than do White candidates. Disparate impact has been assailed as keeping otherwise qualified individuals from entering teaching (e.g., Bennett, McWhorter, & Kuykendall, 2006). Other research has contended that basic skills tests are, in fact, identifying individuals who are likely to have substantial difficulty in mastering the subject content necessary to pass licensure tests in specific content areas (Gitomer, Brown, & Bonett, 2011).

Disparate impact has led to lawsuits over the years, including one brought against California's basic skills test, California Basic Educational Skills Test, in *Association of Mexican American Educators v. State of California* (1996). In this case and others, the use of the tests has generally been upheld. Pullin (2001) provides a detailed review of legal considerations across licensure tests generally.

Another line of research has focused on attempting to establish a relationship between scores on these tests and student outcomes. Findings have been inconsistent, and when significant statistical relationships have been observed, they have generally been quite modest (National Research Council, 2001b).

FOCUS 2—TEACHER AS CONTENT KNOWLEDGE PROFESSIONAL

Whereas subject-specific tests were optional when the NTE first was introduced, during the 1980s, these content-specific tests increasingly became the key requirement for teacher certification and remain so to the present day. These tests predominantly focus on knowledge of content that is judged to be the content that is addressed in the subject and grade(s) for which the test is intended (e.g., high school chemistry, middle school social studies).

This focus was driven by a set of key policy considerations that described the importance of both knowing content and knowing how to teach that content. Knowledge of content and knowledge of how to teach content have had very different trajectories in terms of the development of assessments for teachers. It is for this reason that we treat the corresponding assessments as two separate foci.

The Cognitive Model for Teacher as Content Knowledge Professional

Concerns about educational performance in the United States and implications for future economic success were raised in a number of significant publications during the 1980s. *A Nation at Risk* (The National Commission on Excellence in Education, 1983) presented a broad range of apprehensions about education, with a particular focus on the inadequate preparation of teachers to meet ever-changing demands in a rapidly transforming global environment:

> The teacher preparation curriculum is weighted heavily with courses in "educational methods" at the expense of courses in subjects to be taught. A survey of 1,350 institutions training teachers indicated that 41 percent of the time of elementary school teacher candidates is spent in education courses, which reduces the amount of time available for subject matter courses. (The National Commission of Excellence in Education, p. 30)

A Nation at Risk led to a set of policy documents focused on a new vision of teaching and teacher preparation that concentrated on what teachers need to know and do. *A Nation Prepared* (Carnegie Task Force on Teaching as a Profession, 1986), followed by the Interstate New Teacher Assessment and Support Consortium (InTASC) standards (Council of Chief State School Officers, 1992) and the National Board for Professional Teaching Standards (1989), stressed the importance of teachers knowing their subject matter, as well as how to teach it. The importance of content knowledge was later codified into law through the 1998 amendment to Title II of the Higher Education Act (1998) requiring teacher preparation institutions to report on how well teachers do on licensure tests and then through the Elementary and Secondary Education Act (ESEA) of 2002 (No Child Left Behind; ESEA, 2002), in which teachers were required to be highly qualified by virtue of, in large part, demonstrating subject-matter knowledge through content tests. A set of research reviews that included recommendations for improving teacher preparation has continued to strengthen the argument for teachers having strong content knowledge, as well as knowing how to teach that content (e.g., Darling-Hammond & Bransford, 2007; National Research Council, 2010; S. M. Wilson, Floden, & Ferrini-Mundy, 2002).

Coinciding with these policy initiatives was the emergence of a cognitive psychology that emphasized the importance of domain-specific knowledge in the acquisition of skill (e.g., Chase & Simon, 1973; Chi, 1978; Chomsky, 1988; Norman & Rumelhart, 1975; Schank & Abelson, 1977; Wellman & Gelman, 1997). This work challenged the previously dominant view that expertise was simply a matter of intelligence or intellectual processing capacity and development. Rather, the learning and

(re)structuring of content knowledge were shown to result in more sophisticated reasoning and problem solving. For example, Wellman and Gelman, among others, challenged the traditional developmental stage model of Piaget by demonstrating that children, with sufficient familiarity in a domain, could demonstrate more advanced reasoning than a traditional Piagetian stage theory would suggest. Thus, the policy focus on content knowledge was certainly consistent with the dominant learning theories of the time.

But what specific content did teachers need to know? Here, too, *A Nation at Risk* played a substantial role. The National Education Goals Panel (National Council on Education Standards and Testing, 1992; National Education Goals Panel, 1993) called for clear and ambitious content standards for K–12 education. The response was that professional organizations and almost all states (Council of Chief State School Officers, 2000) developed detailed standards describing what students should know across grade levels and subject areas (e.g., National Council for the Social Studies, 1994; National Council of Teachers of English & the International Reading Association, 1996; National Council of Teachers of Mathematics, 1989; National Research Council, 1996).

Another body of content has also been included as part of the professional content knowledge of teachers. This body of knowledge consists of domain-general rules of pedagogy, including knowledge of child development, classroom management, teaching methods, and classroom assessment (Voss, Kunter, & Baumert, 2011). Because it is not tied to subject-matter content, we distinguish this domain-general pedagogical knowledge from the PCK that provides the foundation for the assessment of CKT discussed in the next section.

Content specifications for teacher tests are typically derived from state and disciplinary standards. Then, through a set of test validation processes, committees of educators are asked to judge the relevance of the test specifications and the tests themselves to what teachers should know (Educational Testing Service, 2005). There is limited information available about the outcomes of and research into these content validation processes (National Research Council, 2001b). Researchers have raised cautions about validation procedures by arguing that just because a test measures aspects of the intended domain, it does not imply that the full domain is represented by the test (Shepard, 1993) and that validation methods have a strong confirmation bias (Moss, 1998).

The Evidence Model for Teacher as Content Knowledge Professional

Administering multiple-choice items has been the dominant method of collecting evidence about teachers' content knowledge. However, there are a number of assessments that also include constructed-response items. An example of standards and an assessment item in English and mathematics, respectively, from the California Educator Credentialing Examinations (CSET; California Subject Examinations for Teachers®; Pearson Education, Inc., 2014) are presented in Table 2. These assessments were originally all paper-and-pencil administrations but now also are delivered via computer.

TABLE 2

California Subject Examination for Teachers (CSET): Sample Subject Matter Requirements (SMRs) and Questions

SMR	Question
English Language Arts (ELA): Language, Linguistics, and Literacy (SMR Domain 2) 0001 Human Language Structures (SMR 2.1) a. Demonstrate knowledge of the nature of human language, differences among languages, the universality of linguistic structures, and language change across time, locale, and communities b. Demonstrate knowledge of word analysis, including sound patterns (phonology) and inflection, derivation, compounding, roots and affixes (morphology) c. Demonstrate knowledge of sentence structures (syntax), word and sentence meanings (semantics), and language function in communicative context (pragmatics)	2. Which of the following statements accurately describes a significant effect that the Great Vowel Shift had on the English language? A. English vowel sounds no longer corresponded to French vowel sounds. B. Regional variations of vowel sounds emerged within England as well as in Scotland and Wales. C. Vowel sounds used by speakers in British colonies differed from vowel sounds used in Great Britain. D. The spelling of certain vowel sounds no longer corresponded to the pronunciation of those sounds.
Mathematics: Number and Quantity (SMR Domain 1) 0001 The Real and Complex Number Systems (SMR 1.1) a. Demonstrate knowledge of the properties of the real number system and of its subsets b. Perform operations and recognize equivalent expressions using various representations of real numbers (e.g., fractions, decimals, exponents) c. Solve real-world and mathematical problems using numerical and algebraic expressions and equations d. Apply proportional relationships to model and solve real-world and mathematical problems e. Reason quantitatively and use units to solve problems (i.e., dimensional analysis) f. Perform operations on complex numbers and represent complex numbers and their operations on the complex plane	3. Three numbers, x, y, and z, have a sum of 871. The ratio x:y is 4:5 and the ratio y:z is 3:8. Which of the following is the value of y? A. 134 B. 156 C. 195 D. 201

Note. To view other SMRs and questions for the CSET, see http://www.ctcexams.nesinc.com/about_CSET.asp. To view preparation materials for the Praxis® tests, see http://www.ets.org/praxis/prepare/materials?WT.ac=praxishome_prepare_121126

While the delivery mode may differ, tests are designed to include the same content and are scored in the same manner. From the test developer's perspective, there is no difference between administration modes (e.g., Educational Testing Service, 2014a).

Individual tests are typically allotted several hours for administration, and students can complete tests more quickly in a computer-based environment because they need not wait for the testing time period to be completed. Depending on licensing requirements, individual candidates may need to take multiple content tests.

The Interpretive Model for Teacher as Content Knowledge Professional

In the case of traditional content tests used for licensure, it is almost certainly true that the domain of teaching knowledge, as expressed in documents such as the InTASC and National Board for Professional Teaching Standards (NBPTS), as well as in disciplinary standards, is underrepresented (e.g., American Federation of Teachers, 2012). The deep and integrated levels of understanding that are articulated in standards documents can be underrepresented in two ways.

The first type of underrepresentation concerns the nature of the items on the tests. As described in critiques of student assessments, and equally germane to teacher assessments, traditional tests "have theoretical roots in the differential and behaviorist traditions" (National Research Council, 2001b, p. 60). The richer cognitive models described in the standards are not well assessed through the traditional testing approaches evident in teacher licensure assessments.

The second form of underrepresentation involves the requirements for passing the test as defined by the passing standard. As described in the previous section, each state sets its own passing standard for a test using one of a number of common methods described in the rich literature on standard setting (e.g., Cizek & Bunch, 2007; Hambleton & Pitoniak, 2006). Teacher licensure tests that assess content frequently have been criticized for having passing standards that are very minimal and that result in extremely high passing rates (Crowe, Allen, & Coble, 2013; U.S. Department of Education, Office of Postsecondary Education, 2011). Thus, even if a test contains a large set of items that would be considered rigorous, if only a relatively small proportion of correct items is needed to meet the passing standard, then successful candidates will not have demonstrated mastery of the tested domain.

As with other licensure tests, these content tests typically involve a transformation from a raw score to a scaled score, which is then compared with the passing standard to determine a candidate's success. Scaled scores beyond passing status may or may not be reported to users, depending on existing policies. Often subscores reflecting subdomains that the test is designed to assess will be reported as well, though the validity of reporting these scores has been questioned due to the unreliability of scores based on small numbers of items (Sinharay, 2010).

Research Findings for Teacher as Content Knowledge Professional

Three primary research directions around teacher content tests have been pursued. The first has been to examine the content skill of individuals who take the

teacher licensure tests and how that has varied across different demographic groups of individuals and over time. Such studies have been used to attempt to characterize the teaching force. For example, annual reports about licensure test results by content area are produced by the U.S. Department of Education, Office of Postsecondary Education (2011). Gitomer and Qi (2010) reported on the distribution of test scores for a number of Praxis content tests over a period of time in which there was an increase in accountability pressure from No Child Left Behind (Elementary and Secondary Education Act, 2001). They observed very high passing rates across tests, with the median score well above the passing standard across states. Additionally, they found little change in scores across years for those who passed the exams. They also observed very substantial differences between the scores of those who passed and those who did not. Finally, they concluded that the primary function of the tests is to identify a relatively small set of individuals who have very limited content knowledge. The use of the tests is not designed to make finer distinctions above these relatively low passing standards.

A second research direction focuses on how the wide use of content tests can both affect, and inform us about, the shape of the teaching pool. For example, Angrist and Guryan (2007) examined teaching employment patterns associated with increasing test demands. They found that increased use of tests led to higher wages but also dissuaded some higher ability individuals from entering teaching. They also concluded that Hispanic candidates were less likely to teach in states with increased licensure test requirements. Fuller and Ladd (2012) found that teachers in lower elementary grades had weaker licensure test scores than those in upper grades. Goldhaber, Lavery, and Theobald (2014) found that students with more educational disadvantage (e.g., free/reduced-price lunch status, underrepresented minority, low prior academic performance) have teachers with lower licensure scores. Furthermore, Gitomer, Latham, and Ziomek (1999) found that because passing rates for African American candidates were substantially lower than those for White candidates, policies that simply raised passing standards on licensure tests would likely lead to greater adverse impact. Goldhaber (2007) also argued that raising passing scores would reduce the teaching pool without having material effects on educational quality.

The third research direction has looked at relationships among licensure test scores and important educational outcomes, most especially student achievement. In general, findings have been mixed. Clotfelter, Ladd, and Vigdor (2007) found small but positive relationships between teacher test scores and student achievement in mathematics and reading, with mathematics effects substantially larger. Goldhaber (2007) found small but significant relationships between teacher scores and student achievement. However, Buddin and Zamarro (2008) found no significant relationship between teacher and student measures. As summarized in National Research Council (2001b), these studies and effect sizes must be treated cautiously. Even with more sophisticated methodologies employed in the studies reported here, there are still issues of the assignment of students to classroom teachers and the nature of both teacher and student tests that can influence the results of these studies.

FOCUS 3—TEACHER AS CONTENT KNOWLEDGE FOR TEACHING PROFESSIONAL

A new generation of teacher knowledge assessments has gradually emerged over the last two decades. These assessments are predicated on the idea that teachers' knowledge of content is situated in the practice of teaching, and knowledge is held in ways that support instruction and learning. This knowledge, however, is not separate from content knowledge; instead, it is assumed to rely heavily on the integration of content knowledge and pedagogical knowledge. This has led to both large-scale assessments and significant programs of research. Nevertheless, even as conceptions of both teaching and testing have evolved to support this focus, the dominant forms of teacher testing remain rooted in the first two foci. CKT assessments remain largely the province of research efforts. For ease of presentation, we label all assessment efforts in this category as CKT, even if they were originally discussed as PCK measures.

The Cognitive Model for Teacher as CKT Professional

During the 1980s, the same policy initiatives described in the previous section set the stage for addressing CKT. Consequently, research on teacher knowledge broadened considerably, influenced by the larger emergence of the cognitive sciences and a focus on situated cognition. From this perspective, knowledge moved from being conceived of as a stable trait of individuals to a set of understandings inextricably associated with specific contexts. For example, there were classic cross-cultural and developmental studies of individuals who demonstrated very powerful understanding of complex ideas within particular contexts even though they would not evidence such understanding on formal assessments of the underlying constructs (e.g., Cole & Scribner, 1974; Lave, 1988; Rogoff & Waddell, 1982).

The foundation of this focus is *pedagogical content knowledge*, introduced by Lee Shulman (1986, 1987). In his 1985 American Educational Research Association address, Shulman described the ever-changing movements in classifying teacher knowledge from general knowledge to content knowledge. He acknowledged that these movements identified important aspects of teaching such as classroom organization and knowledge of the content one is teaching; however, he pointed out that they miss an important aspect of knowledge—how teachers transform their content knowledge into lessons and are able to teach the content that they know to those who do not yet understand it. According to Shulman (1986), this knowledge answers the question, "How do teachers decide what to teach, how to represent it, how to question students about it, and how to deal with problems of misunderstanding?" (p. 8). In essence, this knowledge is the knowledge that enables teachers to conduct the actions necessary to teach a particular subject.

The idea of PCK built on research that was focusing on teachers' use of knowledge in the practice of teaching. For example, Elbaz (1983) studied Canadian English teachers through a series of observations and interviews in an effort to identify a

teacher's "practical knowledge," or the knowledge that teachers employ when teaching. Through this work, Elbaz developed a framework for the practical knowledge for teaching that encompassed five domains: knowledge of self, knowledge of the social content, knowledge of subject matter, knowledge of curriculum development, and knowledge of instruction. Through observation, Elbaz noted, "Practical knowledge encompasses first hand experience of students' learning styles, interests, needs, strengths and difficulties, and a repertoire of instructional techniques and classroom management skills" (p. 5).

Leinhardt and Greeno (1986) analyzed how expert and novice teachers carried out common teaching activities (e.g., homework, guided practice) in mathematics. For each, they developed action schema and planning nets, common representations used by cognitive psychologists at the time. These flow diagrams specified the conditions necessary to execute an action and the potential consequences of each action. For example, if a teacher is explaining the relationship among a set of concepts, the prerequisite condition would be that the component concepts are understood. Or, a teacher might take an action with the intent that the action engages the students and moves the lesson forward. Leinhardt and Greeno concluded that teachers use their knowledge of lesson structure and subject matter to successfully carry out these complex plans.

Shulman (1987) broke down the knowledge that informed teaching into seven distinct types of knowledge: general pedagogical knowledge; content knowledge; knowledge of the curriculum; knowledge of learners and their characteristics; knowledge of the educational context; knowledge of educational ends, purposes, and values; and, finally, PCK. It is important to note that by identifying these domains that contribute to the work of teaching, Shulman acknowledges previous movements in the identification of teacher knowledge and emphasizes the importance of general teacher knowledge and content knowledge in carrying out the work of teaching. However, it is the final domain of teacher knowledge identified by Shulman, *pedagogical content knowledge*, that is at the heart of the CKT focus in the study of teacher knowledge. Shulman (1986) conceptualized PCK as comprising the following:

> For the most regularly taught topics in one's subject area, the most useful forms of representation of those ideas, the most powerful analogies, illustrations, examples, explanations, and demonstrations—in a word, the ways of representing and formulating the subject that make it comprehensible to others . . . Pedagogical content knowledge also includes an understanding of what makes the learning of specific topics easy or difficult: the conceptions and preconceptions that students of different ages and backgrounds bring with them to the learning of those most frequently taught topics and lessons. (p. 9)

Shulman (1987) further defined PCK as representing, "The blending of content and pedagogy into an understanding of how particular topics, problems, or issues are organized, represented, and adapted to the diverse interests and abilities of learners and presented for instruction" (p. 8).

Since Shulman first developed the notion of PCK, many attempts have been made to elaborate the construct across different domains (Ball et al., 2008). For example,

Magnusson, Krajcik, and Borko (1999) defined PCK in science as a unique domain of teacher knowledge that is "a transformation of several types of knowledge for teaching, including subject matter knowledge" (p. 95). Magnusson et al. stated that teachers' PCK is influenced by their subject-matter knowledge and beliefs, their pedagogical knowledge and beliefs, and their knowledge and beliefs about the educational context.

By observing teachers in the United Kingdom teaching mathematics in the upper and lower primary grades, Rowland, Huckstep, and Thwaites (2005) identified a "knowledge quartet," or the four types of knowledge that teachers use in practice: foundational knowledge, transformative knowledge, connection knowledge, and contingency knowledge. At the heart of their framework is the foundational knowledge that teachers have gained "in the academy," such as content knowledge and knowledge of teaching methods. Building from this foundation is transformative knowledge, which is the knowledge that enables teachers to transform content knowledge and knowledge of teaching methods into a lesson in which they teach the content; connection knowledge, which enables teachers to make connections drawn by the teacher during the lesson and enables the sequencing of instructional topics both locally (in the lesson) and globally (throughout the curriculum); and contingency knowledge, which enables teachers to deviate from the planned lesson or to respond to students' ideas. What resulted was a framework of PCK that includes four components: subject matter for instructional purposes, students' understanding of the subject matter, media for instruction in the subject matter, and instructional processes for the subject matter.

Pedagogical content knowledge has been substantially elaborated over the years in domain-specific treatments. Some examples include van Driel, Verloop, and de Vos (1998), Abell (2007), Baxter and Lederman (1999), and Gess-Newsome and Lederman (1999) in science; Döhrmann, Kaiser, and Blömeke (2012) and Ma (1999) in mathematics; Grossman (1990, 1991) and Howey and Grossman (1989) in English language arts (ELA); Monte-Sano and Budano (2013), S. M. Wilson and Wineburg (1988, 1993), and Leinhardt (1993) in history; and Neiss (2005) in multiple subject areas for integrating technology into instruction.

Though definitions of PCK have varied across research efforts, there are several principles that characterize the range of PCK definitions (van Driel et al., 1998). First, PCK is centered on specific topics and domains. Well-developed PCK in mathematics does not imply well-developed PCK in other content areas. Second, although PCK relies heavily on content knowledge, it is specifically concerned with the teaching of a subject. A mathematician, for example, who has never taught may have a large subject-matter knowledge base but may not have well-developed PCK. Finally, all definitions of PCK are in terms of practices that comprise the work of teaching.

More recently, Ball et al. (2008) sought to clarify what constitutes teaching specific content knowledge and developed the construct of *content knowledge for teaching*. Although they acknowledge that the work of Shulman and others in developing the concept of PCK was critical to advancing the study of teacher knowledge, they

point out that PCK is still inadequately understood and that the domain as a whole is underdeveloped. First, they point to the many differing definitions of PCK as hindering the development of a precise definition of the construct. Second, possibly as a consequence of the lack of definition, few studies have attempted to measure PCK in teachers in an effort to validate the construct.

To better identify the content knowledge specific to teaching, Ball et al. (2008) called on their work with the Mathematics Knowledge for Teaching (MKT) project, which sought to better understand the knowledge teachers need to have and what they need to be able to do in order to effectively carry out the work of teaching mathematics (Ball, Hill, & Bass, 2005). In this study, the authors looked specifically at the work of teaching and tried to first identify what a teacher needs to do in the classroom when teaching a specific subject. Then they identified the knowledge needed to carry out the tasks of teaching. Through this work, they distinguished among the types of content knowledge used in teaching a subject: the shared or common content knowledge used in many professions (common content knowledge), the content knowledge specifically used in teaching (specialized content knowledge), and PCK. Overall, they defined CKT as the knowledge that directly links the work of teaching and the content knowledge that is required to do that work (Ball et al., 2008). For international comparisons of teacher knowledge of mathematics, Tatto et al. (2008) and Krauss, Baumert, and Blum (2008) divided the domain of teacher knowledge into content knowledge and PCK. Gitomer, Phelps, Weren, Howell, and Croft (2014) built on this work to develop a set of tasks of teaching for the domains of mathematics and ELA that are designed to be general at the highest level but then articulated at more specific levels (see Table 3). The successful execution of each task of teaching requires knowledge of content that is structured in ways particular to teaching.

Ball et al. (2008) further elaborated the construct with the idea of specialized content knowledge. This special set of knowledge is described as content knowledge beyond that of what a well-educated adult should know. It is the knowledge that enables teachers to carry out everyday tasks of teaching such as presenting ideas, responding to "why" questions, linking representations to underlying ideas, and modifying tasks to make them easier or harder (Ball et al., 2008). For example, teachers need to understand the underlying structure of a mathematical algorithm for teaching purposes (e.g., why division of fractions can be carried out with an invert and multiply algorithm), whereas the typical user of the algorithm need not have such knowledge.

Kersting, Givvin, Sotelo, and Stigler (2010) introduced the concept of *usable knowledge* by designing classroom video analysis tasks to which teachers needed to respond. They divide the domain into four subdomains: mathematical content, student thinking, suggestions for improvement, and depth of interpretation.

Although the development of the construct of *pedagogical content knowledge* by Shulman and others and the development of the construct of *content knowledge for teaching* by Ball and colleagues have led to somewhat different definitions of the

TABLE 3
Tasks of Teaching Requiring Content Knowledge for Teaching

Task of Teaching	Mathematics Examples	ELA Examples
1. Anticipating student challenges, misconceptions, partial misconceptions, alternate conceptions, strengths, interests, capabilities, and background knowledge	• Anticipating student challenges in reasoning about and doing mathematics due to the interplay of content demands and students' understanding • Anticipating likely misconceptions, partial conceptions, and alternate conceptions about particular mathematics content and practices	• Anticipating the impact of limited English language proficiency on students' comprehension of text and speech and on their written and spoken expression • Anticipating how students' background knowledge, life experiences, and cultural background can interact with new ELA concepts, texts, resources, and processes
2. Evaluating student ideas evident in work, talk, actions, and interactions	• Evaluating student work, talk, and actions in order to identify conceptions in mathematics, including incorrect or partial conceptions • Evaluating non-standard responses for evidence of mathematical understanding and in terms of efficiency, validity, and generalizability	• Evaluating student work, talk, and actions for evidence of strengths and weaknesses in reading, writing, speaking, and listening • Evaluating discussion among groups of students for evidence of understanding ELA concepts, texts, and processes
3. Explaining concepts, procedures, representations, models, examples, definitions, and hypotheses	• Explaining mathematical concepts or why a mathematical idea is "true" • Interpreting a particular representation in multiple ways to further understanding	• Explaining literary or language concepts, using definitions, examples, and analogies when appropriate • Explaining processes of reading, including why certain processes are appropriate for particular texts and/or tasks
4. Creating and adapting resources for instruction (examples, models, representations, explanations, definitions, hypotheses, procedures)	• Creating and adapting examples that support particular mathematical strategies or to address particular student questions, misconceptions, or challenges with content • Adapting student-generated conjectures to support instructional purposes	• Creating and adapting examples or model texts to introduce a concept or to demonstrate a literary technique or a reading, writing, or speaking strategy • Creating and adapting analogies to support student understanding of ELA concepts, texts, and processes

(continued)

TABLE 3 (CONTINUED)

Task of Teaching	Mathematics Examples	ELA Examples
5. Evaluating and selecting resources for instruction (examples, models, representations, explanations, definitions, hypotheses, procedures)	• Evaluating and selecting representations or models that support multiple interpretations • Evaluating and selecting explanations of mathematical concepts for potential to support mathematical learning or in terms of validity, generalizability, or explanatory power	• Evaluating and selecting examples to develop understanding of a concept, literary technique, or literacy strategy, or to address particular student questions, misconceptions, or challenges with content • Evaluating and selecting procedures for writing or working with text
6. Developing questions, activities, tasks, and problems to elicit student thinking	• Creating or adapting questions, activities, tasks, or problems that demonstrate desired mathematical characteristics • Creating or adapting classes of problems that address the same mathematical concept or that systematically vary in difficulty and complexity	• Creating or adapting prompts or questions with the potential to elicit productive student writing • Developing questions, activities, or tasks to elicit evidence that students have a particular literary understanding or skill
7. Evaluating and selecting student activities (questions, problems) to elicit student thinking	• Evaluating and selecting questions, activities, or tasks to elicit evidence that students have a particular mathematical understanding or skill • Evaluating and selecting problems that support particular mathematical strategies and practices	• Evaluating and selecting questions, activities, or tasks to elicit discussion about a specific text or literary concept • Evaluating and selecting questions, activities, or tasks to support the development of a particular literary understanding or skill
8. Doing the work of the student curriculum	• Doing the work that will be demanded of the students as part of the intended curriculum	• Doing the work that will be demanded of the students as part of the intended curriculum

Note. ELA = English language arts.
Source. Gitomer, Phelps, Weren, Howell, and Croft (2014).

knowledge that teachers use while teaching, the driving force behind the CKT focus is clear: Teachers have and employ a distinct set of content and professional knowledge when engaging in the work of teaching. This knowledge enables them to carry out the routine tasks of teaching and to help students better understand the content that is being taught. Although this knowledge may be specific to teaching, it is strongly dependent on the common content knowledge of the subject. In fact, Krauss, Baumert, et al. (2008) report a series of factor analytic studies that generally supports the argument that many of these subtler theoretical distinctions are not supported through empirical analysis.

This transformation in the characterization of teachers' content knowledge has led to major policy documents that have shaped some new types of teaching assessments. Perhaps the most significant policy development is the establishment of the NBPTS documents. Each set of disciplinary and student age-band standards elaborates on one of the NBPTS core propositions: *Teachers know the subjects they teach and how to teach those subjects to students* (National Board for Professional Teaching Standards, 1989). The InTASC standards, developed in 1992 and updated in 2011 (Council of Chief State School Officers, 1992, 2011) have also been influential in the development of a range of teacher assessments. Most of the assessments grounded in the NBPTS and InTASC standards have focused on measures of teacher practice (e.g., portfolios) and are not addressed in this chapter as knowledge measures. However, these standards have also driven the development of the NBPTS assessment center, which was a large-scale attempt at assessing CKT.

The Evidence Model for Teacher as CKT Professional

At the heart of the definitions of both PCK and CKT is that the knowledge defined by the constructs enables teachers to carry out the work of teaching. Because of this, traditional content knowledge assessments do not capture evidence of a teacher's CKT. Instead, CKT assessments must be grounded in the work of teaching so that judgments can be made regarding a teacher's professional content knowledge. The first large-scale standardized assessments in this area were those of the NBPTS. Since then, a broad range of assessment methods has been used to collect evidence of PCK across domains.

The NBPTS assessment consists of two components: the portfolio, which samples classroom practice directly through videos and artifacts, and the assessment center, which is designed to assess PCK through a set of constructed-response items. The items of a given year's assessment center are instances of an item category that has specific targets. A sample set of specifications for each of the item categories for one of the certifications (Middle Childhood Generalist [i.e., upper elementary self-contained classrooms]) is presented in Appendix B. Each response is scored on the basis of an item-level rubric, and those scores then contribute to the overall NBPTS assessment score (National Board for Professional Teaching Standards, 2013).

A relatively broad array of assessment tasks, including multiple-choice items, interviews, cognitive study tasks such as card-sorting, and constructed-response

items, has been used to collect evidence of CKT (Baxter & Lederman, 1999). In most cases, stimuli have been verbal descriptions of a teaching scenario.

However, some recent work has used video stimuli (e.g., Kersting, 2008; Kersting et al., 2010). Examinees are presented with a classroom video segment and asked to analyze, in writing, how the teacher and the student(s) interacted around the mathematical content.

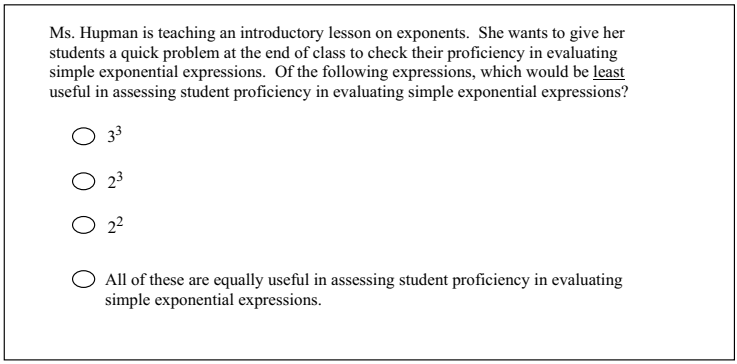
In one of the most cited attempts to measure CKT, Hill, Schilling, and Ball (2004) developed a set of assessment items that would allow them to draw inferences about teachers' CKT, specifically in the area of mathematics. These items were meant to measure common content knowledge and the specialized content knowledge that teachers use specifically when teaching. To develop the items, Hill et al. drew on the work of high-quality classroom instruction and analyzed curricular materials, student work, and personal experience in the classroom. The result of this work was a set of items meant to assess MKT.

The initial items were developed for K–6 teachers and were classified into three content areas: number concepts; operations; and patterns, functions, and algebra. For the first two content areas, separate items were developed to measure knowledge of content and knowledge of students and content. For the third content area, only items measuring knowledge of content were developed (Hill et al., 2004). After administering the assessment to more than 1,500 teachers, they found evidence of multiple factors, aligned with both content areas and type of knowledge being assessed.

Another effort to measure teachers' content-specific knowledge was the COACTIV study (Baumert et al., 2010), in which teachers of students who took the 2003–2004 Program for International Student Assessment (PISA) were given assessments of both their CKT and their common content knowledge (Krauss, Brunner, et al., 2008). The items in this study were similar to the items developed by Hill et al. (2004) in that they were paper-and-pencil items meant to address scenarios specific to teaching. However, these items differed somewhat in that they were constructed-response items and addressed specific tasks of teaching, including knowledge of multiple solution paths, knowledge of alternative student ideas, and knowledge of instructional strategies. Using the framework developed by Ball et al. (2008), the second two item types would fall under PCK (specifically, knowledge of content and students, and knowledge of content and teaching), but the first item type, knowledge of multiple solution paths, would fall into the specialized content knowledge domain. Regardless of the domains that each question type addresses, the design of the assessment is clear: to address teacher-specific knowledge by grounding the assessment items in the work of teaching. The international Teacher Education and Development Study in Mathematics (TEDS-M) also used primarily multiple-choice items to estimate content knowledge and CKT for teachers in different countries (Blömeke, Houang, & Suhl, 2011).

Related work has been done to assess CKT in the area of reading. Phelps and Schilling (2004) set out to develop an assessment of teachers' content knowledge for

FIGURE 1
Sample Mathematics Single-Selection Multiple-Choice Item



Note. The correct answer is 2^2. The relevant task of teaching is: Creating problems or questions to elicit student mathematical thinking, justifications, or explanations.

teaching elementary-level reading, resulting in a series of survey items to assess teachers' content knowledge, knowledge of content and students, and knowledge of teaching and content in reading. Further work to assess content knowledge for teaching reading includes the use of the *Language and Reading Concept* assessment that is based on required components of the Reading First legislation that was designed to improve reading skills in students enrolled in high-poverty and low-achieving schools. The assessment focuses on content knowledge and knowledge of teaching the content in five areas defined by the legislation: phonemic awareness, phonics, fluency, vocabulary, and reading comprehension (Carlisle, Correnti, Phelps, & Zeng, 2009). Kucan, Hapgood, and Palincsar (2011) also developed a constructed-response assessment designed to assess teachers' specialized knowledge for comprehension instruction. This assessment provides teachers with a passage and then asks them to state the main idea of the passage, to highlight any text features that may prove difficult for students, and then to analyze a student response about the text.

One large-scale research project (Gitomer et al., 2014) in the assessment of CKT was a part of the Measures of Effective Teaching project (Bill and Melinda Gates Foundation, 2011). In this project, the researchers set out to develop a set of CKT assessments in the areas of ELA and mathematics. The Measures of Effective Teaching study drew teachers from districts across the country and across grades 4–9. Because CKT is dependent on content knowledge and knowledge of the wide range of content taught across grade levels and districts, separate assessments were developed for grades 4–6 ELA, 7–8 ELA, 4–5 mathematics, 6–8 mathematics, and grade 9 algebra. The researchers developed assessments comprised of selected-response and constructed-response items that built on the tasks of teaching framework described in Table 3. An example from mathematics is presented in Figure 1.

FIGURE 2
Sample English Language Arts Table Item

Ms. Rice begins a unit on memoir writing by reading a passage from a literary model. She then asks students to complete a warm-up activity to help them generate ideas for their own writing.

For each assignment, indicate whether or not it will help students focus their brainstorming on generating a memoir.

	Will help focus brainstorming	**Will not help focus brainstorming**
Write a poem about the ways you have changed, using the form "I used to be…but now I am…"		
Write a sequence of sentences describing some of your experiences, beginning each sentence with the phrase "I remember."		
Write a few adjectives that describe your personality.		
Write down some of your favorite foods and describe what you like about them.		

Note. Correct answers are: Will, Will, Will not, Will not. The relevant task of teaching is: Evaluating and selecting questions to elicit productive student writing.

The mathematics item in Figure 1 does not ask the teachers to simply solve exponential expressions, which would draw on common content knowledge, but instead asks teachers to use their knowledge to identify which exponential expressions would tell them that the student knows how to work with exponents. Similarly, in the ELA item presented in Figure 2, teachers have to select the activities that will help students focus their brainstorming on a specific topic. Common content knowledge alone would not be sufficient to answer such items successfully.

The Interpretive Model for Teacher as CKT Professional

Scores on these assessments are typically reports of the number of questions the teacher answered correctly when compared to the correct answer for the given item. However, some studies use more sophisticated psychometric models, including item response theory (Blömeke et al., 2011; Hill, Ball, Blunk, Goffney, & Rowan, 2007).

For many of these assessment questions, the best answer cannot simply be justified through an appeal to the content (e.g., 2 + 2 = 4). Rather, appropriate answers rely on professional judgment of appropriate actions to take given a particular instructional context (see Hill et al., 2004; Gitomer et al., 2014).

For the NBPTS assessments and other constructed-response measures (e.g., Kersting et al., 2010), standardized rubrics are typically developed and then rated by

human judges who have been trained on the scoring criteria. Each item receives a score, and these are then weighted in computing a total certification score. Scores are provided for each of the four dimensions that define the domain. Some recent explorations in using automated techniques to interpret written responses show promising levels of reliability (Kersting, Sherin, & Stigler, 2014).

For some assessments of PCK, score interpretations based on performance on items corresponding to individual domains are provided. For example, on the MKT Assessment, there are items corresponding to common content knowledge and items corresponding to knowledge of content and students.

Research Findings for Teacher as CKT Professional

Probably because the assessment work in this focus has primarily been a research enterprise rather than a tool for teacher selection, the body of research is far more developed than for the other two foci. These assessments have been largely designed to help understand teacher knowledge and its relationship to practice at a fairly deep level.

The research in this focus can be organized into several broad categories. One body of work has attempted to establish the construct of CKT through factor analytic studies and measures of psychometric quality. A second body of work has explored the relationship of CKT scores to measures associated with the practice and effectiveness of teaching. A third body of research has examined differences in performance and reasoning on these items for individuals with differing amounts of expertise in teaching. The final research question has focused on the distribution of teacher CKT across schools differing in socioeconomic and achievement characteristics.

Examinations of the factor structure of assessments demonstrate consistently that CKT items measure something distinct from content knowledge measures, though the measures are highly correlated (Blömeke et al., 2011; Krauss, Baumert, et al., 2008; Schilling, Blunk, & Hill, 2007). Initial results from the MKT project indicated that it was possible to use a paper-and-pencil test to measure teaching-specific content knowledge. Although a general factor for overall performance on the items accounted for a large portion of the variance on individual items, common content knowledge and specialized content knowledge were distinguishable second-order factors. Similarly, in reading, Phelps (2009) found that measures of content knowledge in reading call on specialized knowledge of language, text structures, and reading processes that is different from common reading ability.

This may not be all that surprising, as it becomes clear that individuals draw on subject content knowledge when they are asked to solve CKT problems aloud (Gitomer et al., 2014; Hill, Ball, & Schilling, 2008). Hill's work (Hill, Blunk, et al., 2008; Schilling et al., 2007) illustrates the difficulty of finding factor analytic support for further distinction of constructs within the larger domain of CKT.

A number of studies have examined the relationship of CKT scores to other measures. Hill, Rowan, and Ball (2005) found that higher scores on the MKT measure were associated with greater gains in mathematics for first- and third-grade students.

Hill, Umland, Litke, and Kapitula (2012) and Hill, Blunk, et al. (2008) observed a strong relationship between MKT scores and the quality of instruction in mathematics as measured by a mathematics-specific observation protocol. The relationship to student achievement was less consistent.

Gitomer et al. (2014) studied the reasoning on assessment items in mathematics and ELA by teachers who varied in their scores on CKT assessments that were taken one year prior and included these items. They found substantial evidence that observed reasoning patterns for both correct and incorrect responses were consistent with the anticipated reasoning built into the design of the assessment items and that the quality of reasoning was associated with scores on the assessments.

Kersting's work (Kersting et al., 2010; Kersting, Givvin, Thompson, Santagata, & Stigler, 2012) examines the relationship of scores on their video analysis measures of knowledge to the MKT measures and to measures of classroom practice and student learning. The focus for all of this work was in the area of fractions, and teachers of students in grades 5–7 participated. They found that their video analysis scores were related to MKT scores and that the strongest dimension correlation was with mathematical content. They also observed a strong relationship with instructional practice as measured by an instrument the researchers developed. Finally, they found a significant relationship between video analysis scores and student learning of fractions. On further examination, only one dimension, *Suggestions for Improvement*, was associated with student learning. This is the dimension that captures teachers' suggestions for improving the instruction in the video interactions they observed.

A number of studies have compared groups with known characteristics that would be expected to show different levels of competence on CKT assessments. As part of the COACTIV study, Krauss, Baumert, et al. (2008) compared teachers who taught students who were on track to attend a university at the end of their secondary education (academic track) with teachers who taught students who were on the vocational track (nonacademic track). To teach in the academic track, teachers must essentially major in mathematics, whereas those who plan to teach in the nonacademic track have a more varied education. On each of the question types, the academic-track teachers scored significantly higher than the nonacademic-track teachers. Together with the correlation between scores on the CK items and the CKT items (.60), these findings provided additional support for the connection between content knowledge and PCK. Hill, Dean, and Goffney (2007) also found different patterns of performance and reasoning among those with different teaching and mathematics backgrounds.

In addition to testing teachers with different training, the COACTIV study tested teachers in content areas other than mathematics to develop an argument as to the validity of the construct of teacher-specific content knowledge (Krauss, Baumert, et al., 2008). University mathematics majors, high school academic-track mathematics students, and academic-track biology and chemistry teachers were all given the same PCK assessment. As expected, the biology and chemistry teachers who have neither the content knowledge nor the teaching experience in mathematics scored low on

both the CK items and the CKT items, and the high school mathematics students scored low on the CKT items. Those who taught mathematics in both academic and nonacademic tracks scored higher on both the CK and CKT items. One interesting result, however, was the high CKT scores of the university mathematics majors. Although this result supports the idea that content knowledge plays an important part in teacher-specific knowledge, it does not necessarily suggest that any mathematics major has the knowledge and the ability to teach. One possibility is that those with strong mathematics backgrounds are able to use that knowledge to reason through CKT assessment items.

Buschang, Chung, Delacruz, and Baker (2012) addressed this possibility by administering several assessments meant to measure both content knowledge and PCK for the teaching of mathematics, including the MKT Assessment and the Student Response Assessment (SRA), which asked test takers to evaluate student responses to questions. These assessments were administered to mathematics content experts (those with a PhD in mathematics and no K–12 teaching experience), expert teachers (teachers who either held a National Board certification or had experience in training teachers), and novice teachers. The SRA was specifically developed to assess the PCK needed to assess student work. On the MKT Assessment, the content experts and the expert teachers performed similarly, and both groups performed significantly better than the novice teachers. However, on the SRA, a task that is typical to the everyday work of teaching, the content experts performed poorly compared to the expert and novice teachers. This suggests that performance on the MKT Assessment is sensitive to content knowledge but that content knowledge alone is not enough to perform well on tasks that occur in the work of teaching, such as the SRA.

As with other measures of content knowledge, teachers higher in CKT tend to teach in classrooms that have fewer minority and poor students and that have higher levels of prior achievement. Hill et al. (2005) and Hill and Lubienski (2007) found that teachers of students with higher proportions of low-socioeconomic status and minority students had lower scores on the MKT Assessment.

FOCUS 4—TEACHER AS KNOWLEDGE-RICH PRACTITIONER

The last focus, teacher as knowledge-rich practitioner, is nascent. There is a small but growing literature that provides the basis for these kinds of assessments, and there are no systematic research studies that have explored the measurement characteristics of these assessments. Yet this articulation of knowledge use in practice represents a natural progression from the CKT efforts in moving from conceptions of knowledge from *knowing that* to *knowing how.*

The Cognitive Model for Teacher as Knowledge-Rich Practitioner

This focus of assessing teacher knowledge is predicated on the idea that content knowledge and CKT underlie the teacher's ability to effectively engage in critical

practices of teaching. The primary motivation behind the knowledge-rich practitioner focus is the idea that teaching is a practice-based profession and that there are core component practices (Grossman, Hammerness, & McDonald, 2009; Grossman & McDonald, 2008) that are "agnostic with respect to various models of teaching" (Grossman & McDonald, 2008, p. 186)—that is, whether a particular model of teaching is more didactic or constructivist or whether the model is in mathematics or social studies, and so on. The particular execution of those components, however, will shift with different models of teaching.

Ball and colleagues have designated these key practices as *high-leverage practices*. Ball, Sleep, Boerst, and Bass (2009) defined these high-leverage practices as follows:

> Practices in which the proficient enactment by a teacher is likely to lead to comparatively large advances in student learning . . . They include activities of teaching that are essential to the work and that are used frequently, ones that have significant power for teachers' effectiveness with pupils. (pp. 460–461)

A fundamental idea in this focus is that by focusing on practice, knowledge is studied as it is enacted (Grossman et al., 2009; Lampert, 2010). Ball and Forzani (2009) discuss a shift from knowledge to practice that is "entailed by the work" (p. 503). Teachers represent their knowledge of practice in terms of common structures or schema that they apply in situations that call for specific practices. Thus, a goal of teacher education is to help teachers develop such situated knowledge.

As the knowledge-rich practitioner focus places emphasis on developing situated knowledge during preservice teacher education, recent efforts have begun to discuss how teacher education must shift and be reconceptualized in response to the most recent practice-based focus (see Forzani, 2014; McDonald, Kazemi, & Kavanagh, 2013; Zeichner, 2012). Other efforts, such as Lampert et al. (2013), have been aimed at designing teacher education experiences that provide opportunities for preservice teachers to practice and develop the ability to carry out these practices.

Though the practices are general at an abstract level, how they develop in particular content areas is unique. A discussion around literature is different from one in physics in terms of the nature of evidence and warrants and common discourse norms, for example. Thus, the development of high-leverage or core practices is being carried out in specific domains including ELA (Hatch & Grossman, 2009), history (Fogo, 2014), science (Janssen, Westbroek, & Doyle, 2014; Windschitl, Thompson, Braaten, & Stroupe, 2012), mathematics (Boerst, Sleep, Ball, & Bass, 2011; Franke et al., 2009; TeachingWorks, 2014b), foreign language (Hlas & Hlas, 2012; Troyan, Davin, & Donato, 2013), social justice (McDonald, 2010), and culturally responsive pedagogical practices (Ladson-Billings, 1995) in mathematics (Waddell, 2014).

The Evidence Model for Teacher as Knowledge-Rich Practitioner

At this point in time, the evidence models are, at most, in a state of design. There are a number of potentially interesting methods for capturing evidence and frameworks for considering how such evidence might be developed.

One of the earliest forerunners of this kind of effort was portfolio assessments of teaching that included entries that involved evidence of practice in the form of brief classroom videos or instructional artifacts such as classroom assignments or tests. The most prominent of these efforts was associated with NBPTS. Teachers were also required to present detailed commentary around the classroom evidence. Each portfolio task, in essence, asked teachers to carry out some limited set of what would become known as core practices. For example, a portfolio might ask teachers to lead a discussion or analyze student work (National Research Council, 2008).

Parallel work led by Grossman and Ball focuses on breaking down, or decomposing (Grossman et al., 2009), the complex act of teaching into constituent parts that can be closely examined. Much of this work has referenced teacher education, but it is equally applicable to assessment and has been discussed by Moss (2010). The challenge is to create simulated task structures that are constrained in ways that focus attention on a particular practice. The structures are rich enough that they do elicit evidence that draws on knowledge and skill in executing an important teaching practice. But the tasks are constrained in ways that do not fully approximate teaching. The focus is on a particular task (e.g., leading a discussion, planning a lesson, or providing an explanation of an important concept), but the task does not require coordinating all of these components or dealing with the various competing considerations (e.g., student behavior) that teachers face in an actual classroom.

Moss (2010) describes six elements that characterize this type of assessment work in a teacher education program. First, high-leverage practices are identified and decomposed in ways that allow them to be studied and developed. Second, learning and assessment activities are designed so that they become increasingly complex and call on greater skill of the developing teacher. Third, there is a need to develop a common analytic language to interpret performance on these activities. Fourth, there is attention given to the general and immediate contexts of task design that make for effective prompts. Fifth, annotated exemplars of practice are necessary in order to give meaning to any verbal or other descriptions of the practice. Sixth, descriptions of the learning trajectories or progressions of these practices are essential.

Because these practices are enacted, the performance needs to be captured in ways that are different from traditional assessments. For certain practices, video may be an appropriate way to capture evidence. For other practices (e.g., planning), the evidence may take the form of written artifacts. One interesting possibility involves the use of simulated environments in which the preservice teacher interacts with the simulation. Dieker, Straub, Hughes, Hynes, and Hardin (2014) have developed a system that uses avatars that respond to a student in a teacher education program to develop specific competencies, including classroom management. Human actors control the avatars at this point, but in the long term, it may be possible for intelligently designed systems to drive significant parts of the interaction.

Given that assessment systems are not yet designed, there is no literature that addresses the interpretive or research dimensions of this focus. However, given what Moss (2010) identified as a shared language for guiding and analyzing practice, it

would be most natural for a central interpretive component to be some type of scoring system that relied on human judgment guided by a scoring protocol.

The knowledge-rich practitioner model has been developed in the context of teacher education. Therefore, natural places for assessment use would be in the context of developing and certifying the performance of beginning teachers. These assessments, assuming they continue to develop, would be used by teacher preparation institutions. However, there are early efforts to build out these assessments for the purpose of licensure (TeachingWorks, 2014a).

A DESIGN FRAMEWORK FOR ASSESSING TEACHER KNOWLEDGE

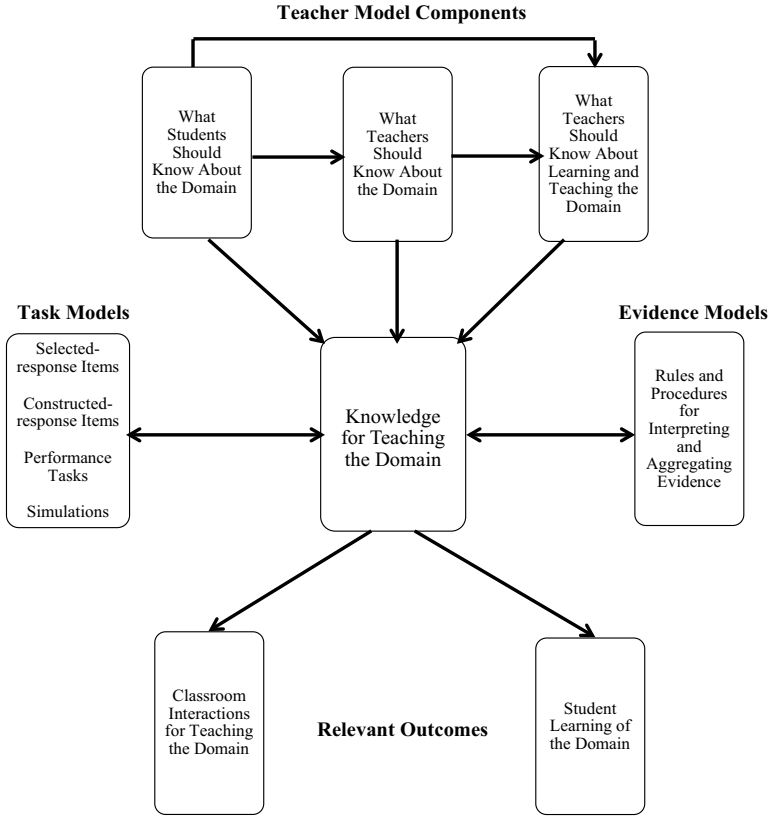
The broad foci of assessing teacher knowledge described in this chapter have been loosely coupled enterprises. Yet, underlying each focus are either explicit or implicit cognitive models for what students should know, what teachers should know, and the role of teachers in instruction. Also, whether explicit or implicit, these foci have relied on assumptions that teacher knowledge is associated with effective instruction and student outcomes. Only the CKT and knowledge-rich practitioner foci make strong theoretical linkages to those outcomes.

In this concluding section, we present a general design framework that can guide both the development and validation of assessments of teacher knowledge. We then suggest ways in which current knowledge about the teaching of ELLs can form the foundation of knowledge assessments for ELL teaching.

The design framework builds from the evidence-centered design (ECD) approach developed by Bob Mislevy and colleagues (e.g., Mislevy & Haertel, 2006). ECD is an effort to bring formal design principles to the development of assessments and begins with a conceptual assessment framework that consists of three major components. The first, the *student model*, is a detailed specification of the concepts, skills, and relationships that represent the target domain of the assessment. Though ECD uses the term *student model* generically, for the present purposes we will apply this conception to teachers. The second component, the *task model*, describes the kinds of assessment tasks and assessment environments that are used to collect evidence of understanding with respect to the concepts, skills, and relationships articulated in the student model. The third component, the *evidence model*, specifies the procedures for interpreting and scoring the evidence.

We propose a design framework in Figure 3 that supports the design of teacher knowledge assessments broadly but that is specifically intended to support the design of assessments of teacher knowledge within the CKT and knowledge-rich practitioner domains. The top of the figure represents the teacher model and includes three constituent components. The first box represents a set of learning targets for the students who are being taught. So, for example, it could represent the various concepts and skills that students in fourth grade mathematics should have. The second box represents what a teacher needs to have with regard to the content knowledge about the domain that includes all the student targets, but other content knowledge is likely to be included as well. For example, Ball et al. (2008) describe *horizon content*

FIGURE 3
Design Framework for Assessment and Validation of Teacher Knowledge



knowledge as the knowledge of content that students will encounter in later (as well as earlier) grades. As noted earlier, they also describe the construct of specialized content knowledge. The box on the right describes the knowledge needed to effectively carry out tasks of teaching, examples of which are presented in Table 3. This component of the model would include, for example, knowledge of the kinds of representations that will help students learn a concept or the ways in which students may have difficulty in grasping an idea.

Together, these three components describe the target domain of teacher knowledge, represented in the middle of the figure. At the left center of the figure are task models. Task models are evidence-capturing devices. For example, multiple-choice questions ask examinees to select from a set of given options. The choice an examinee

selects is intended to provide evidence of quality of understanding of some aspect of the teacher model. Other tasks that have been used in teacher assessment include shorter and longer written responses as well as performance tasks such as those used in the NBPTS portfolio. The work of Kersting and others described earlier provides other kinds of task models. Different task models may be more or less effective at providing evidence about different aspects of the teacher model.

We can use this framework to design assessments that address different foci of teacher knowledge, each deemed necessary to effectively carry out any task of teaching. Table 4 presents several examples of the types of assessment tasks that might be developed to assess different aspects of the teacher model. For example, if the target task of teaching is anticipating students' challenges and conceptions about a particular concept, then we can imagine one kind of assessment task that asks teachers about their direct knowledge of that concept. An assessment of CKT might pose questions that directly inquire about the kinds of conceptions and conceptual difficulties that students are likely to have. Finally, assessment tasks that are designed to call on specific types of CKT can be developed to collect evidence about a teacher's ability to enact practices, such as carrying out a discussion.

The box on the right center represents the evidence model that interprets and aggregates performance on these kinds of tasks into assessment scores. The evidence model includes such things as the answer key for multiple-choice tasks, scoring rubrics for tasks that require more qualitative interpretation, and scoring models (e.g., item response theory) that are used to develop scores that estimate a teacher's proficiency with respect to teacher knowledge.

Ultimately, it is important to validate the inferences that are made on the basis of an assessment score. The process of developing a validity argument is described by Kane (2006) and includes establishing evidence that the scores mean what they are intended to mean. Articulating a design process provides a detailed specification that can then be evaluated. Another key aspect of Kane's and other validity frameworks (e.g., Messick, 1989) is to evaluate the relationship between scores on an assessment and other outcomes that are theorized as being related to those scores.

In the case of teacher knowledge measures of CKT and practice, we establish the hypothesis that teachers who know more will engage in more effective teaching practice and also will have a greater effect on student learning. These hypotheses can be investigated directly by observing and scoring the quality of classroom interactions or evaluating measures of student learning based on some type of assessments of student knowledge and growth. Support for the validity of the assessments is achieved to the extent that there is a positive and systematic relationship between scores on these different measures.

IMPLICATIONS FOR ASSESSING KNOWLEDGE FOR TEACHERS OF ENGLISH LANGUAGE LEARNERS

How might we use such a design framework for the assessment of teacher knowledge with respect to teachers of ELLs, an area that has received minimal attention? There is a great deal of work that lays the foundation for what such assessment efforts might look like.

TABLE 4
Framework (With Examples) Connecting Different Forms of Teacher Knowledge to Tasks of Teaching

Task of Teaching	Content Knowledge	Content Knowledge for Teaching	Knowledge Through Enactment
Anticipating student challenges, misconceptions, partial misconceptions, alternate conceptions, strengths, interests, capabilities, and background knowledge	What do teachers need to know about domain content?	What do teachers need to know about the teaching of content?	What do teachers need to know or be able to do to enact content?
Mathematics example: Anticipating student challenges in proportional reasoning due to the interplay of content demands and students' understanding	Proportional reasoning Fractions	Common conceptual challenges students face in proportional reasoning Instructional resources to highlight and support resolution of challenges	Creating classroom interactions (e.g., discussions) to address challenges Using resources effectively to address challenges
ELA example: Anticipating how students' background knowledge, life experiences, and cultural background can interact with the ELA concept of character development	Fiction Literary techniques Character development	Questions that help students connect their own experiences to the writing task Exemplars that model how authors use their own backgrounds in developing their characters	Using exemplars and models that help students see potential connections between their own lives and character development Supporting classroom discussion that helps students explore strategies for character development

Note. ELA = English language arts.

The first step must be to lay out the teacher model that specifies what constellation of knowledge and skills is important for teachers to have in order to teach these students. There are very useful reports and papers that attempt to lay out what such teachers should know and be able to do, but there is currently no professional consensus on this among teacher educators (Faltis & Valdés, in press). However, important ideas about what students need to know and what teachers need to do to help students learn have been offered by the National Literacy Panel on Language-Minority Children and Youth (August & Shanahan, 2006), August and Hakuta (1997), and others. Implications for what teachers need to know and be able to do are summarized both by Faltis and Valdés (in press) as well as by the National Research Council (2010). This panel identified six areas on which there was some consensus:

- understanding of the complexity of the reading process for English-language learners;
- competence at explicit instruction in vocabulary, the development of oral proficiency;
- content instruction that focuses on learning from text, understanding and producing academic language, genre differentiation, and academic writing;
- understanding of home-school differences in interaction patterns or styles and individual differences among the wide range of English-language learners; and
- understanding of the ways language and reading interact, the skills that transfer into English, and how to facilitate that transfer; and understanding of the context in which second-language learners develop as readers. (National Research Council, 2010, p. 90)

de Jong and Harper (2005) also provide a related framework for what teachers need to know about teaching ELLs. Teachers must:

- acquire specific knowledge and skills related to the language and culture of their students;
- understand the process of sound language acquisition and acculturation;
- understand how bilingual processes are manifested in ELL oral and literacy development and how they can build on students' resources;
- understand the difficulties that ELLs may face when they are trying to learn content through a language over which they have no control (this entails a basic knowledge of the structure of the English language by the teacher);
- have a cultural awareness, particularly of the culturally-based assumptions regarding learning and literacy; and
- understand the language demands of their subject area.

Turkan, de Oliveira, Lee, and Phelps (2014) have built on this work to theorize about the specialized knowledge involved in teaching content to ELLs. They propose the idea of *disciplinary linguistic knowledge* as an understanding of the discourse of

particular disciplines and having the knowledge to engage students in such discourse. They focus on the particular linguistic challenges that ELLs face in learning academic content and the teacher knowledge needed to help students overcome those challenges.

As noted by Faltis and Valdés (in press), much of this theorizing has a very limited empirical basis. The development of assessments based on such theories, together with a validation effort, offers the opportunity to strengthen the empirical basis for the knowledge teachers need to support learning for ELLs.

The outlines of what teachers need to know, such as those presented above, begin to lay out the teacher model that would be the basis for such assessments. However, much more specificity of concepts and skills that teachers should know needs to be developed for a more complete and productive teacher model. Once the model is articulated, then corresponding task and evidence models can be designed. These models certainly would overlap with the models we have described elsewhere in the chapter.

The validation effort would then examine relationships of scores on assessments of knowledge for teachers of ELLs with measures of classroom practice and interactions, and with student outcomes including, but not limited to, academic learning. For example, it may also be useful to understand student attitudes and motivations and to examine their social interactions in and out of school. Exactly what would be considered relevant evidence depends on the student model that contributes to the teacher model (what students should learn) and the hypothesized relationship of teacher knowledge and practice to aspects of the student model.

For example, one might reasonably hypothesize that effective teachers of ELLs know a good bit about linguistic moves that will support student learning of content. We could also imagine CKT assessments that involve knowledge of these linguistic moves with respect to challenges that ELLs face. We could also imagine practice-based assessments—for example, simulations that require teachers to enact such knowledge in order to be successful. To the extent that there are positive relationships between scores on assessments like these with desired student outcomes, there is validity support not only for the assessment but also for the theory that underlies it—namely, that CKT and the enactment of linguistic moves is important for the teaching of ELLs.

CONCLUDING COMMENTS

Bringing together much more developed theories of teacher knowledge with advances in assessment design provides an opportunity to develop measures of teacher knowledge that have value far more than simply being used for certification. Such assessments, precisely because they are rooted in practice, have the potential to dramatically influence how teachers are prepared and supported throughout their careers. Obviously, substantial design work needs to be carried out to use assessments of teacher knowledge with different purposes from what they have been traditionally used for. Finally, having measures of knowledge that

are based on strong and explicit theories of practice provides rich opportunities to test and revise the assumptions that guide our theories of how teachers learn and how they help students learn.

APPENDIX A

Sample Items, First National Teacher Examination Administration

(Note: Percentages of individuals from the two groups selecting each response are presented after each response.)

English Comprehension, Part II

Slade was a matchless marksman with a navy revolver. One morning at Rocky Ridge when he was feeling comfortable he saw a man approaching who had offended him some days before. "Gentlemen," said Slade, drawing, "it is a good twenty-yard shot—I'll clip the third button on his coat!" Which he did. The bystanders admired it. And they all attended the funeral, too.

7. Slade may be best described as

    1. humorous. (3%, 0%)
    2. cold-blooded. (44%, 70%)
    3. revengeful. (34%, 19%)
    4. playful. (8%, 7%)
    5. daring. (11%, 3%)

    Omitted (0%, 1%) *not reached* (0%, 0%)

8. The passage achieves its effect chiefly by means of

    1. suspense. (13%, 10%)
    2. exaggeration. (5%, 0%)
    3. direct quotation. (20%, 3%)
    4. matter-of-fact treatment. (60%, 86%)
    5. detailed description. (2%, 0%)

    Omitted (0%, 1%) *not reached* (0%, 0%)

English Expression, Punctuation

All the childrens art classes were represented in the exhibit, and the class which had done the best work received a prize.

7.
    1. No punctuation necessary (10%, 1%)
    2. Children's. (56%, 84%)
    3. Childrens'. (27%, 15%)

    Omitted (5%, 0%) *not reached* (2%, 0%)

General Culture, Part I, Current Social Problems

60. In the past 50 years, man in the Western world has gained most in

    1. health. (52%, 97%)
    2. native intelligence. (6%, 0%)
    3. emotional stability. (4%, 0%)
    4. security of employment. (6%, 1%)
    5. stability of government. (14%, 2%)

    Omitted (0%, 0%) *not reached* (18%, 0%)

General Culture, Part III, Literature

70. The poet who "sounded his barbaric yawp over the roofs of the world," yet brought a new freedom and flexibility to poetry, was

    1. T. S. Eliot. (4%, 4%)
    2. Emerson. (6%, 0%)
    3. Edgar Lee Masters. (4%, 0%)
    4. Edwin Markham. (4%, 0%)
    5. Walt Whitman. (55%, 95%)

    Omitted (21%, 1%) *not reached* (6%, 0%)

General Culture, Part IV, Science

60. The physical discomfort produced by a warm, badly ventilated room full of people is due mainly to

    1. lack of sufficient oxygen. (24%, 13%)
    2. accumulated carbon dioxide. (24%, 15%)
    3. poisonous vapors from people's bodies. (4%, 1%)
    4. accumulated water vapor. (28%, 68%)
    5. the very low humidity of the air. (12%, 3%)

    Omitted (0%, 0%) *not reached* (8%, 0%)

General Culture, Part V, Fine Arts

76. What modern painters are largely responsible for the revival of mural painting today?

    1. English (5%, 0%)
    2. German (2%, 2%)
    3. Italian (12%, 3%)
    4. Mexican (53%, 91%)
    5. French (8%, 1%)

    Omitted (16%, 3%) *not reached* (4%, 0%)

General Culture, Part VI, Mathematics

21. The positive square root of ¼ is
    1. ½. (54%, 100%)
    2. 2. (4%, 0%)
    3. 1√2. (6%, 0%)
    4. ⅛. (4%, 0%)
    5. 1/16 (19%, 0%)

    Omitted (12%, 0%) *not reached* (1%, 0%)

Professional Information, Part I, Education and Social Policy

43. Substantial modifications in the high school curriculum have been necessary since the World War to allow for

    1. expansion of the college preparatory curriculum. (14%, 5%)
    2. the presence in high school of many persons with slight aptitude for book learning. (61%, 88%)
    3. studies on a more difficult level, necessitated by better preparation in the grades. (3%, 3%)
    4. emphasis on cultural rather than vocational subjects because of decreased employment opportunities. (5%, 2%)
    5. consolidation of rural high schools. (10%, 2%)

    Omitted (7%, 0%) *not reached* (0%, 0%)

Professional Information, Part II, Child Development and Educational Psychology

55. Six-year-old children differ most from 10-year-olds in which one of the following respects?

    1. Fine-muscle control (48%, 80%)
    2. Large-muscle control (8%, 3%)
    3. Visual acuity (3%, 2%)
    4. Auditory acuity (1%, 0%)
    5. Metabolic rate (20%, 13%)

    Omitted (14%, 2%) *not reached* (6%, 0%)

Professional Information, Part III, Guidance and Individual and Group Analysis

26. It is a basic assumption of experts in vocational guidance that

    1. interest in an occupation is the best evidence of future success in that occupation. (10%, 0%)
    2. vocational aptitudes cannot be measured objectively. (4%, 2%)
    3. pupils need assistance in estimating their own abilities, aptitudes, and interests. (84%, 98%)

4. there is one, and only one, occupation for which each individual is ideally suited. (0%, 0%)
5. vocational and educational guidance should be kept entirely separate. (1%, 0%)

Omitted (1%, 0%) *not reached* (0%, 0%)

Professional Information, Part IV, Secondary School Methods

1. The primary function of the recitation should be to

1. give individual attention to the slower pupils. (1%, 0%)
2. discover the pupils who have not assimilated previous lessons. (6%, 2%)
3. discover the extent to which the individual pupils are able to recall material read. (1%, 0%)
4. develop understanding through group discussion. (74%, 90%)
5. improve the ability of the pupils to present material orally. (17%, 7%)

Omitted (1%, 1%) *not reached* (0%, 0%)

APPENDIX B

National Board for Professional Teaching Standards (NBPTS) Middle Childhood: Generalist Assessment at a Glance (p. 4)

Assessment Center Exercises

This assessment is composed of six exercises that examine content knowledge specified in the NBPTS Standards. You are given up to 30 minutes to respond to each exercise.

Following is a description of each assessment exercise.

Exercise 1: Supporting Reading Skills	In this exercise, you demonstrate your ability to analyze and interpret a transcript of a student's oral reading of a given reading passage to identify one strength and one weakness in the student's oral reading. You are also asked to identify and justify appropriate strategies to support the identified student's ongoing reading development.
Exercise 2: Analyzing Student Work	In this exercise, you demonstrate your ability to identify mathematical misconceptions/errors in a given student's work, to identify concepts/skills necessary for student understanding of the math problem, and to provide appropriate strategies with a rationale to assist the student's understanding of the identified concepts or skills.

(continued)

APPENDIX B (CONTINUED)

Exercise 3: Knowledge of Science	In this exercise, you demonstrate your ability to identify and understand fundamental concepts and principles in science. You are asked to respond to a student's inquiry about a real-world phenomenon by identifying scientific concepts and principles that are related to the real-world phenomenon. You are also asked to describe an appropriate learning experience with a rationale that will provide student understanding of a concept/principle that relates to the real-world phenomenon.
Exercise 4: Social Studies	In this exercise, you demonstrate your ability to identify and interpret social studies/history information within a given graphic. You are asked to identify a cause-and-effect relationship based on the information in the given graphic and to identify a specific historic event related to the cause-and-effect relationship. You are also asked to describe a learning experience/activity that develops student understanding of a real-world connection related to the identified cause-and-effect relationship.
Exercise 5: Health and Wellness	In this exercise, you demonstrate your ability to identify a developmentally appropriate goal and learning activity to promote students' knowledge and skills related to a given health and wellness topic and enable them to transfer their knowledge and skills. You are also asked to explain how the learning activity would support the goal and promote students' knowledge and skills related to the health and wellness topic.
Exercise 6: Integrating the Arts	In this exercise, you demonstrate your ability to use the arts to develop student understanding of concepts in another discipline. You are asked to identify concepts in a given subject area and describe an arts-focused learning experience that will establish a connection for students' understanding of an identified concept and provide multiple paths of access for student learning of that concept. You are asked to justify how the arts-focused learning experience will enable students to develop a deeper or broader appreciation of the arts.

NOTE

This study was supported in part by the National Science Foundation, through Grant DRL-1222777. Additionally, we would like to thank the editors and reviewers for their very helpful feedback. We are also grateful to Sean Smith of Horizon Research, Inc., and our research group at Rutgers Graduate School of Education for their very useful comments. Finally, we thank Colleen McDermott for all her support in the preparation of the chapter.

REFERENCES

Abell, S. K. (2007). Research on science teacher knowledge. In S. K. Abell & N. G. Lederman (Eds.), *Handbook of research on science education* (pp. 3–28). Mahwah, NJ: Lawrence Erlbaum.

American Council on Education. (1940). *Report of the first annual administration of the National Teacher Examinations and announcement of the 1941 examinations*. New York, NY: Author.

American Federation of Teachers. (2012). *Raising the bar: Aligning and elevating teacher preparation and the teaching profession* (A report of the American Federation of Teachers Teacher Preparation Task Force). Washington, DC: Author. Retrieved from http://www.highered.nysed.gov/raisingthebar2012.pdf

Anderson, J. R. (1976). *Language, memory, and thought*. Hillsdale, NJ: Lawrence Erlbaum.

Angrist, J. D., & Guryan, J. (2007). Does teacher testing raise teacher quality? Evidence from state certification requirements. *Economics of Education Review*, *27*, 483–503. doi:10.1016/j.econedurev.2007.03.002

Association of Mexican American Educators v. State of Cal. (1996). 937 F. Supp. 1397, N.D. Cal.

August, D., & Hakuta, K. (Eds.). (1997). *Improving schooling for language-minority children: A research agenda*. Washington, DC: National Academies Press. Retrieved from http://www.nap.edu/openbook.php?isbn=0309054974

August, D., & Shanahan, T. (Eds.). (2006). *Developing literacy in second-language learners: Report of the National Literacy Panel on Language-Minority Children and Youth* (Executive summary). Mahwah, NJ: Lawrence Erlbaum. Retrieved from http://www.bilingualeducation.org/pdfs/PROP2272.pdf

Ball, D. L., & Forzani, F. M. (2009). The work of teaching and the challenge for teacher education. *Journal of Teacher Education*, *60*, 497–511. doi:10.1177/0022487109348479

Ball, D. L., Hill, H. C., & Bass, H. (2005). Knowing mathematics for teaching: Who knows mathematics well enough to teach third grade, and how can we decide? *American Educator*, *29*, 14–46. Retrieved from http://www.aft.org/pdfs/americaneducator/fall2005/BallF05.pdf

Ball, D. L., Sleep, L., Boerst, T., & Bass, H. (2009). Combining the development of practice and the practice of development in teacher education. *Elementary School Journal*, *109*, 458–476. doi:10.1086/596996

Ball, D. L., Thames, M. H., & Phelps, G. C. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education*, *59*, 389–407. doi:10.1177/0022487108324554

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., . . .Tsai, Y.-M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, *47*, 133–180. doi:10.3102/0002831209345157

Baxter, J. A., & Lederman, N. G. (1999). Assessment and measurement of pedagogical content knowledge. In J. Gess-Newsome & N. G. Lederman (Eds.), *Examining pedagogical content knowledge: The construct and its implications for science education* (pp. 147–161). Dordrecht, Netherlands: Kluwer Academic.

Bennett, C. I., McWhorter, L. M., & Kuykendall, J. A. (2006). Will I ever teach? Latino and African American students' perspectives on PRAXIS I. *American Educational Research Journal*, *43*, 531–575. doi:10.3102/00028312043003531

Biesta, G. J. J., & Burbules, N. C. (2003). *Pragmatism and educational research*. Latham, MD: Rowman & Littlefield.

Biesta, G. J. J., & Stengel, B. (in press). Thinking philosophically about teaching: Illuminating issues and (re)framing research. In D. H. Gitomer & C. A. Bell (Eds.), *Handbook of research on teaching* (5th ed.). Washington, DC: American Educational Research Association.

Bill and Melinda Gates Foundation. (2011). *Learning about teaching: Initial findings from the Measures of Effective Teaching project* (MET Project Research Paper). Seattle, WA: Author. Retrieved from https://docs.gatesfoundation.org/Documents/preliminary-findings-research-paper.pdf

Blömeke, S., Houang, R., & Suhl, U. (2011). TEDS-M: Diagnosing teacher knowledge by applying multidimensional item response theory and multi-group models. *IERI Monograph Series: Issues and Methodologies in Large-Scale Assessments*, *4*, 109–126. Retrieved from http://www.ierinstitute.org/fileadmin/Documents/IERI_Monograph/IERI_Monograph_Volume_04_Chapter_5.pdf

Boerst, T. A., Sleep, L., Ball, D. L., & Bass, H. (2011). Preparing teachers to lead mathematics discussions. *Teachers College Record*, *113*, 2844–2877. Retrieved from https://www.tcrecord.org/Content.asp?ContentID=16496

Boydston, J. A. (Ed.). (1984). *John Dewey: The later works, 1925–1953: Vol. 3. 1927–1928: Essays, reviews, miscellany, and "impressions of Soviet Russia."* Carbondale: Southern Illinois University.

Buddin, R., & Zamarro, G. (2008). *Teacher quality, teacher licensure tests, and student achievement* (RAND Working Paper No. WR-555-IES). Santa Monica, CA: RAND Corporation. Retrieved from http://www.rand.org/content/dam/rand/pubs/working_papers/2008/RAND_WR555.pdf

Buschang, R. E., Chung, G. K. W. K., Delacruz, G. C., & Baker, E. L. (2012). *Validating measures of algebra teacher subject matter knowledge and pedagogical content knowledge* (CRESST Report No. 820). Los Angeles: National Center for Research on Evaluation, Standards, & Student Testing, UCLA.

Carlisle, J. F., Correnti, R., Phelps, G., & Zeng, J. (2009). Exploration of the contribution of teachers' knowledge about reading to their students' improvement in reading. *Reading and Writing*, *22*, 457–486. doi:10.1007/s11145-009-9165-y

Carnegie Task Force on Teaching as a Profession. (1986). *A nation prepared: Teachers for the 21st century* (The report of the Task Force on Teaching as a Profession). Hyattsville, MD: Carnegie Forum on Education and the Economy.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York, NY: Cambridge University Press. Retrieved from http://steinhardtapps.es.its.nyu.edu/create/courses/2174/reading/Carroll_1.pdf

Chase, W. G., & Simon, H. A. (1973). The mind's eye in chess. In W. G. Chase (Ed.), *Visual information processing* (pp. 215–281). New York, NY: Academic Press.

Chi, M. T. H. (1978). Knowledge structure and memory development. In R. S. Siegler (Ed.), *Children's thinking: What develops?* (pp. 73–96). Hillsdale, NJ: Lawrence Erlbaum.

Chi, M. T. H., & Ohlsson, S. (2005). Complex declarative learning. In K. J. Holyoak & R. G. Morrison (Eds.), *Cambridge handbook of thinking and reasoning* (pp. 371–400). New York, NY: Cambridge University Press. Retrieved from http://chilab.asu.edu/papers/Handbook_Chapter_with_Din.pdf

Chomsky, N. (1988). *Language and problems of knowledge*. Cambridge: MIT Press.

Cizek, G. J., & Bunch, M. B. (2007). *Standard setting: A guide to establishing and evaluating performance standards on tests*. London, England: Sage.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2007). *How and why do teacher credentials matter for student achievement?* (NBER Working Paper No. 12828). Cambridge, MA: National Bureau of Economic Research. doi:10.3386/w12828

Cole, M., & Scribner, S. (1974). *Culture and thought: A psychological introduction*. New York, NY: John Wiley.

Council of Chief State School Officers. (1992). *Model standards for beginning teacher licensing and development: A resource for state dialogue*. Washington, DC: Author. Retrieved from http://thesciencenetwork.org/docs/BrainsRUs/Model%20Standards%20for%20Beg%20Teaching_Paliokas.pdf

Council of Chief State School Officers. (2000). *Key state education policies on K–12 education: 2000: Results from the 2000 policies and practices survey, state departments of education*. Washington, DC: Author. Retrieved from http://www.personal.psu.edu/dgm122/epfp/KeyState2000.PDF

Council of Chief State School Officers. (2011). *InTASC model core teaching standards: A resource for state dialogue*. Washington, DC: Author. Retrieved from http://www.ccsso.org/Documents/2011/InTASC_Model_Core_Teaching_Standards_2011.pdf

Crowe, E., Allen, M., & Coble, C. (2013). *Outcomes, measures, and data systems: A paper prepared for the CAEP Commission on Standards and Performance Reporting*. Retrieved from http://caepnet.files.wordpress.com/2012/12/caep-paper-final-1-19-2013.pdf

Darling-Hammond, L., & Bransford, J. (Eds.). (2007). *Preparing teachers for a changing world: What teachers should learn and be able to do*. San Francisco, CA: Jossey-Bass.

de Jong, E. J., & Harper, C. A. (2005). Preparing mainstream teachers for English-language learners: Is being a good teacher good enough? *Teacher Education Quarterly*, *32*, 101–124. Retrieved from http://www.teqjournal.org/backvols/2005/32_2/13dejong&harper.pdf

Dieker, L. A., Straub, C. L., Hughes, C. E., Hynes, M. C., & Hardin, S. (2014). Learning from virtual students. *Educational Leadership*, *71*(8), 54–58.

Döhrmann, M., Kaiser, G., & Blömeke, S. (2012). The conceptualisation of mathematics competencies in the international teacher education study TEDS-M. *ZDM: International Journal on Mathematics Education*, *44*, 325–340. doi:10.1007/s11858-012-0432-z

Educational Testing Service. (2005). *Validity for licensing tests: A brief orientation*. Princeton, NJ: Author. Retrieved from http://www.ets.org/s/praxis/pdf/validity.pdf

Educational Testing Service. (2014a). *Frequently asked questions about the Praxis™ series tests*. Princeton, NJ: Author. Retrieved from https://www.ets.org/praxis/faq_test_takers/

Educational Testing Service. (2014b). *Praxis® core academic skills for educators (core) test content and structure*. Princeton, NJ: Author. Retrieved from https://www.ets.org/praxis/about/core/content/

Elbaz, F. (1983). *Teacher thinking: A study of practical knowledge*. London, England: Croom Helm.

Elementary and Secondary Education Act of 2002, Pub. L. No. 107-110, 115 Stat. 1425 (2002).

Faltis, C. J., & Valdés, G. (in press). Preparing teachers for teaching in and advocating for linguistically diverse classrooms: A vade mecum for teacher educators. In D. H. Gitomer & C. A. Bell (Eds.), *Handbook of research on teaching* (5th ed.). Washington, DC: American Educational Research Association.

Fantl, J. (2014). Knowledge how. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Retrieved from http://plato.stanford.edu/archives/spr2014/entries/knowledge-how/

Flanagan, J. C. (1941). An analysis of the results from the first annual edition of the National Teacher Examinations. *Journal of Experimental Education*, *9*, 237–250. Retrieved from http://www.jstor.org/stable/20150668

Fogo, B. (2014). Core practices for teaching history: The results of a Delphi panel survey. *Theory & Research in Social Education*, *42*, 151–196. doi:10.1080/00933104.2014.902781

Forzani, F. M. (2014). Understanding core practices and practice-based teacher education: Learning from the past. *Journal of Teacher Education*, *65*, 357–368. doi:10.1177/0022487114533800

Franke, M., Webb, N., Chan, A., Ing, M., Freund, D., & Battey, D. (2009). Teacher questioning to elicit students' mathematical thinking in elementary school classrooms. *Journal of Teacher Education*, *60*, 380–392. doi:10.1177/0022487109339906

Freire, P. (1970). *Pedagogy of the oppressed*. New York, NY: Bloomsbury.

Fuller, S. C., & Ladd, H. F. (2012). *School based accountability and the distribution of teacher quality among grades in elementary school* (CALDER Working Paper No. 75). Washington, DC: Center for the Analysis of Longitudinal Data in Education Research, American Institutes for Research. Retrieved from http://www.caldercenter.org/publications/upload/wp75.pdf

Gay, G. (2010). *Culturally responsive teaching: Theory, research, and practice* (2nd ed.). New York, NY: Teachers College Press.

Gess-Newsome, J., & Lederman, N. G. (Eds.). (1999). *Examining pedagogical content knowledge: The construct and its implications for science education*. Dordrecht, Netherlands: Kluwer Academic.

Gitomer, D. H., Brown, T. L., & Bonett, J. (2011). Useful signal or unnecessary obstacle? The role of basic skills tests in teacher preparation. *Journal of Teacher Education*, *62*, 431–445. doi:10.1177/0022487111412785

Gitomer, D. H., Latham, A. S., & Ziomek, R. (1999). *The academic quality of prospective teachers: The impact of admissions and licensure testing* (ETS Teaching & Learning Report Series, ETS RR-03-25). Princeton, NJ: Educational Testing Service. Retrieved from http://www.ets.org/Media/Research/pdf/RR-03-35.pdf

Gitomer, D. H., Phelps, G., Weren, B., Howell, H., & Croft, A. J. (2014). Evidence on the validity of content knowledge for teaching assessments. In T. J. Kane, K. A. Kerr, & R. C. Pianta (Eds.), *Designing teacher evaluation systems: New guidance from the Measures of Effective Teaching project* (pp. 493–528). San Francisco, CA: Jossey-Bass.

Gitomer, D. H., & Qi, Y. (2010). *Score trends for Praxis II*. Washington, DC: U. S. Department of Education, Office of Planning, Evaluation, and Policy Development, Policy and Program Studies Service.

Goldhaber, D. (2007). Everyone's doing it, but what does teacher testing tell us about teacher effectiveness? *Journal of Human Resources*, *42*(4), 765–794. doi:10.3368/jhr.XLII.4.765

Goldhaber, D., Lavery, L., & Theobald, R. (2014). *Uneven playing field? Assessing the inequity of teacher characteristics and measured performance across students* (CEDR Working Paper No. 2014-4). Seattle: Center for Education Data & Research, University of Washington. Retrieved from http://www.cedr.us/papers/working/CEDR%20WP%202014-4.pdf

Goodlad, J. I. (1984). *A place called school*. New York, NY: McGraw-Hill.

Grimmett, P. P., & MacKinnon, A. M. (1992). Craft knowledge and the education of teachers. *Review of Research in Education*, *18*, 385–456. doi:10.3102/0091732X018001385

Grossman, P. L. (1990). *The making of a teacher: Teacher knowledge and teacher education*. New York, NY: Teachers College Press.

Grossman, P. L. (1991). What are we talking about anyhow: Subject matter knowledge for secondary English teachers. In J. Brophy (Ed.), *Advances in research on teaching: Vol. 2: Subject matter knowledge* (pp. 245–264). Greenwich, CT: JAI Press.

Grossman, P., Hammerness, K., & McDonald, M. (2009). Redefining teaching: Re-imagining teacher education. *Teachers and Teaching: Theory and Practice*, *15*, 273–290. doi:10.1080/13540600902875340

Grossman, P., & McDonald, M. (2008). Back to the future: Directions for research in teaching and teacher education. *Educational Research Journal*, *45*, 184–205. doi:10.3102/0002831207312906

Hambleton, R. K., & Pitoniak, M. J. (2006). Setting performance standards. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 433–470). Westport, CT: Praeger.

Haney, W., Madaus, G., & Kreitzer, A. (1987). Charms talismanic: Testing teachers for the improvement of American education. *Review of Research in Education*, *14*, 169–238. doi: 10.3102/0091732X014001169

Hatch, T., & Grossman, P. (2009). Learning to look beyond the boundaries of representation: Using technology to examine teaching (Overview for a digital exhibition: Learning from the practice of teaching). *Journal of Teacher Education*, *60*, 70–85. doi:10.1177/0022487108328533

Hill, H. C., Ball, D. L., Blunk, M., Goffney, I. M., & Rowan, B. (2007). Validating the ecological assumption: The relationship of measure scores to classroom teaching and student learning. *Measurement: Interdisciplinary Research and Perspectives*, *5*, 107–117. doi:10.1080/15366360701487138

Hill, H. C., Ball, D. L., & Schilling, S. G. (2008). Unpacking pedagogical content knowledge: Conceptualizing and measuring teachers' topic-specific knowledge of students. *Journal for Research in Mathematics Education*, *39*, 372–400. Retrieved from http://www.ugr.es/~pflores/2008_9/Master_Conocim/textos%20JP/[1]_Hill-Ball-Schilling-JRME2008-07.pdf

Hill, H. C., Blunk, M. L., Charalambous, C. Y., Lewis, J. M., Phelps, G. C., Sleep, L., & Ball, D. L. (2008). Mathematical knowledge for teaching and the mathematical quality of instruction: An exploratory study. *Cognition and Instruction*, *26*, 430–511. doi:10.1080/07370000802177235

Hill, H. C., Dean, C., & Goffney, I. M. (2007). Assessing elemental and structural validity: Data from teachers, non-teachers, and mathematicians. *Measurement: Interdisciplinary Research and Perspectives*, *5*(2–3), 81–92. doi:10.1080/15366360701486999

Hill, H. C., & Lubienski, S. T. (2007). Teachers' mathematics knowledge for teaching and school context: A study of California teachers. *Educational Policy*, *21*, 747–768. doi:10.1177/0895904807307061

Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal*, *42*, 371–406. doi:10.3102/00028312042002371

Hill, H. C., Schilling, S. G., & Ball, D. L. (2004). Developing measures of teachers' mathematics knowledge for teaching. *Elementary School Journal*, *105*, 11–30. doi:10.1086/428763

Hill, H. C., Umland, K., Litke, E., & Kapitula, L. R. (2012). Teacher quality and quality teaching: Examining the relationship of a teacher assessment to practice. *American Journal of Education*, *118*, 489–519. doi:10.1086/666380

Hirsch, E. D., Jr. (1988). *Cultural literacy: What every American needs to know*. New York, NY: Random House.

Hlas, A. C., & Hlas, C. S. (2012). A review of high-leverage teaching practices: Making connections between mathematics and foreign languages. *Foreign Language Annals*, *45*(Suppl. 1), s76–s97. doi:10.1111/j.1944-9720.2012.01180.x

Howey, K. R., & Grossman, P. L. (1989). A study in contrast: Sources of pedagogical content knowledge for secondary English. *Journal of Teacher Education*, *40*(5), 24–31. doi:10.1177/002248718904000504

Humphry, B. J. (1966). *Report concerning the National Teacher Examinations* (ETS Confidential). Princeton, NJ: Educational Testing Service.

Janssen, F., Westbroek, H., & Doyle, W. (2014). The practical turn in teacher education: Designing a preparation sequence for core practice frames. *Journal of Teacher Education*, *65*, 195–206. doi:10.1177/0022487113518584

Kafka, J. (in press). In search of a grand narrative on the history of teaching. In D. H. Gitomer & C. A. Bell (Eds.), *Handbook of research on teaching* (5th ed.). Washington, DC: American Educational Research Association.

Kane, M. (2006). Validation. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 17–64). New York, NY: Macmillan.

Kersting, N. (2008). Using video clips as item prompts to measure teachers' knowledge of teaching mathematics. *Educational and Psychological Measurement*, *68*, 845–861. doi:10.1177/0013164407313369

Kersting, N. B., Givvin, K., Sotelo, F., & Stigler, J. W. (2010). Teachers' analyses of classroom video predict student learning of mathematics: Further explorations of a novel measure of teacher knowledge. *Journal of Teacher Education*, *61*, 172–181. doi:10.1177/0022487109347875

Kersting, N. B., Givvin, K. B., Thompson, B., Santagata, R., & Stigler, J. (2012). Developing measures of usable knowledge: Teachers' analyses of mathematics classroom videos predict teaching quality and student learning. *American Educational Research Journal*, *49*, 568–590. doi:10.3102/0002831212437853

Kersting, N. B., Sherin, B. L., & Stigler, J. W. (2014). Automated scoring of teachers' open-ended responses to video prompts: Bringing the classroom-video-analysis assessment to scale. *Educational and Psychological Measurement*. Advance online publication. doi:10.1177/0013164414521634

Kintsch, W. (1998). *Comprehension*. Cambridge, England: Cambridge University Press.

Krauss, S., Baumert, J., & Blum, W. (2008). Secondary mathematics teachers' pedagogical content knowledge and content knowledge: Validation of the COACTIV constructs. *ZDM: International Journal of Mathematics Education*, *40*, 873–892. doi:10.1007/s11858-008-0141-9

Krauss, S., Brunner, M., Kunter, M., Baumert, J., Neubrand, M., Blum, W., & Jordan, A. (2008). Pedagogical content knowledge and content knowledge of secondary mathematics teachers. *Journal of Educational Psychology*, *100*, 716–725. doi:10.1037/0022-0663.100.3.716

Kucan, L., Hapgood, S., & Palincsar, A. S. (2011). Teachers' specialized knowledge for supporting student comprehension in text-based discussions. *Elementary School Journal*, *112*, 61–82. doi:10.1086/660689

Ladson-Billings, G. (1995). Toward a theory of culturally relevant pedagogy. *American Educational Research Journal*, *32*, 465–491. doi:10.3102/00028312032003465

Lampert, M. (2010). Learning teaching in, from, and for practice: What do we mean? *Journal of Teacher Education*, *61*(1–2), 21–34. doi:10.1177/0022487109347321

Lampert, M., Franke, M. L., Kazemi, E., Ghousseini, H., Turrou, A. C., Beasley, H., . . . Crowe, K. (2013). Keeping it complex using rehearsals to support novice teacher learning of ambitious teaching. *Journal of Teacher Education*, *64*, 226–243. doi:10.1177/0022487112473837

Lave, J. (1988). *Cognition in practice: Mind, mathematics and culture in everyday life*. Cambridge, England: Cambridge University Press.

Leinhardt, G. (1993). Weaving instructional explanations in history. *British Journal of Educational Psychology*, *63*, 46–74. doi:10.1111/j.2044-8279.1993.tb01041.x

Leinhardt, G., & Greeno, J. (1986). The cognitive skill of teaching. *Journal of Educational Psychology*, *78*, 75–95. doi:10.1037/0022-0663.78.2.75

Levine, A. (2006). *Educating school teachers*. Washington, DC: The Education Schools Project. Retrieved from http://www.edschools.org/pdf/Educating_Teachers_Report.pdf

Ma, L. (1999). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States*. Mahwah, NJ: Lawrence Erlbaum.

Magnusson, S., Krajcik, J., & Borko, H. (1999). Nature, sources and development of pedagogical content knowledge for science teaching. In J. Gess-Newsome & N. G. Lederman (Eds.), *Examining pedagogical content knowledge: The construct and its implications for science education* (pp. 95–132). Dordrecht, Netherlands: Kluwer Academic.

McDonald, M. A. (2010). Social justice teacher education and the case for enacting high-leverage practices. *Teacher Education and Practice*, *23*, 452–455.

McDonald, M., Kazemi, E., & Kavanagh, S. S. (2013). Core practices and pedagogies of teacher education: A call for a common language and collective activity. *Journal of Teacher Education*, *64*, 378–386. doi:10.1177/0022487113493807

Messick, S. (1989). Validity. In R. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13–100). Washington, DC: American Council on Education.

Mintz, A. (2007). The midwife as matchmaker: Socrates and relational pedagogy. In B. Stengel (Ed.), *Philosophy of education yearbook* (pp. 91–99). Urbana, IL: Philosophy of Education Society. Retrieved from http://www.academia.edu/171605/The_Midwife_as_Matchmaker_Socrates_and_Relational_Pedagogy

Mislevy, R. J., & Haertel, G. D. (2006). Implications of evidence-centered design for educational testing. *Educational Measurement: Issues and Practice*, *25*(4), 6–20. doi:10.1111/j.1745-3992.2006.00075.x

Monte-Sano, C., & Budano, C. (2013). Developing and enacting pedagogical content knowledge for teaching history: An exploration of two novice teachers' growth over three years. *Journal of the Learning Sciences*, *22*, 171–211. doi:10.1080/10508406.2012.742016

Moss, P. A. (1998). Rethinking validity in the assessment of teaching. In N. Lyons & G. Grant (Eds.), *With portfolio in hand: Portfolios in teaching and teacher education* (pp. 202–219). New York, NY: Teachers College Press.

Moss, P. A. (2010). Thinking systemically about assessment practice. In M. Kennedy (Ed.), *Teacher assessment and the quest for teacher quality: A handbook* (pp. 355–374). San Francisco, CA: Jossey-Bass.

National Board for Professional Teaching Standards. (1989). *What teachers should know and be able to do*. Arlington, VA: Author. Retrieved from http://www.nbpts.org/sites/default/files/documents/certificates/what_teachers_should_know.pdf

National Board for Professional Teaching Standards. (2013). *Middle childhood/generalist: Assessment at a glance*. Washington, DC: Author. Retrieved from http://www.nbpts.org/sites/default/files/documents/certificates/Aaag/MC_GEN_AssessAtaGlance_Final.pdf

The National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform* (A report to the nation and the Secretary of Education, U.S. Department of Education). Washington, DC: U.S. Department of Education. Retrieved from http://datacenter.spps.org/uploads/sotw_a_nation_at_risk_1983.pdf

National Council for the Social Studies. (1994). *Expectations of excellence: Curriculum standards for social studies*. Washington, DC: Author. Retrieved from http://files.eric.ed.gov/fulltext/ED378131.pdf

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

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: Author.

National Council on Education Standards and Testing. (1992). *Raising standards for American education: A report to Congress, the Secretary of Education, the National Education Goals Panel, and the American people*. Washington, DC: U.S. Government Printing Office. Retrieved from http://files.eric.ed.gov/fulltext/ED338721.pdf

National Education Goals Panel. (1993). *The National Education Goals Report: Building a nation of learners: Vol. 1. The national report*. Washington, DC: Author. Retrieved from http://files.eric.ed.gov/fulltext/ED360394.pdf

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common core state standards*. Washington, DC: Author.

National Paideia Center. (2014). *The Paideia principles*. Asheville, NC: Author. Retrieved from http://www.paideia.org/about-paideia/philosophy/

National Research Council. (1996). *National science education standards*. Washington, DC: National Academies Press. Retrieved from http://www.nap.edu/openbook.php?record_id=4962&page=R1

National Research Council. (2001a). *Knowing what students know: The science and design of educational assessment* (J. Pellegrino, N. Chudowsky, & R. Glaser, Eds.). Washington, DC: National Academies Press. Retrieved from http://www.nap.edu/openbook.php?isbn=0309072727

National Research Council. (2001b). *Testing teacher candidates: The role of licensure tests in improving teacher quality.* (K. J. Mitchell, D. Z. Robinson, B. S. Plake, K. T. Knowles, Eds.). Committee on Assessment and Teacher Quality, Center for Education. Board on Testing and Assessment. Washington, DC: National Academies Press. Retrieved from http://www.nap.edu/openbook.php?record_id=10090

National Research Council. (2008). *Assessing accomplished teaching: Advanced-level certification programs* (M. D. Hakel, J. A. Koenig, & S. W. Elliott, Eds.). Washington, DC: National Academies Press.

National Research Council. (2010). *Preparing teachers: Building evidence for sound policy*. Washington, DC: National Academies Press. Retrieved from http://www.nap.edu/openbook.php?record_id=12882

Neiss, M. L. (2005). Preparing teachers to teach science and mathematics with technology: Developing a technology pedagogical content knowledge. *Teaching and Teacher Education*, *21*, 509–523. doi:10.1016/j.tate.2005.03.006

NGSS Lead States. (2013). *Next generation science standards: For states, by states*. Washington, DC: National Academies Press.

Noddings, N. (2005). *The challenge to care in schools* (2nd ed.). New York, NY: Teachers College Press.

Norman, D. A., & Rumelhart, D. E. (1975). Reference and comprehension. In D. A. Norman, D. E. Rumelhart, & the LNR Research Group (Eds.), *Explorations in cognition* (pp. 65–87). San Francisco, CA: W. H. Freeman.

Pearson Education, Inc. (2014). California Educator Credentialing Examinations: California Basic Educational Skills Test™ (CBEST®). Amherst, MA: Author. Retrieved from http://www.ctcexams.nesinc.com/about_CBEST.asp

Phelps, G. (2009). Just knowing how to read isn't enough! Assessing knowledge for teaching reading. *Educational Assessment, Evaluation and Accountability*, *21*, 137–154. doi:10.1007/s11092-009-9070-6

Phelps, G., & Schilling, S. (2004). Developing measures of content knowledge for teaching reading. *Elementary School Journal*, *105*, 31–48. doi:10.1086/428764

Prucha, F. P. (Ed.). (1973). Official report of the Nineteenth Annual Conference of Charities and Correction. In *Americanizing the American Indians: Writings by the "friends of the Indian": 1880–1900* (pp. 260–271). Cambridge, MA: Harvard University Press. Retrieved from http://historymatters.gmu.edu/d/4929/ (Original work published 1892)

Pullin, D. C. (2001). Key questions in implementing teacher testing and licensing. *Journal of Law & Education*, *30*, 383–429.

Ravitch, D., & Viteritti, J. P. (Eds.). (2001). *Making good citizens: Education and civil society*. Binghamton, NY: Vail-Ballou Press.

Rogoff, B., & Waddell, K. J. (1982). Memory for information organized in a scene by children from two cultures. *Child Development*, *53*, 1224–1228. Retrieved from http://www-pmhs2.stjohns.k12.fl.us/higginj/0DBE8052-0118C716.4/Memory%20for%20Information%20Organized%20in%20a%20Scene%20by%20Children%20fro.pdf

Rousseau, J.-J. (1979). *Emile* or *on education* [The Project Gutenberg EBook version]. New York, NY: Basic Books. Retrieved from http://www.gutenberg.org/cache/epub/5427/pg5427.html

Rowland, T., Huckstep, P., & Thwaites, A. (2005). Elementary teachers' mathematics subject knowledge: The knowledge quartet and the case of Naomi. *Journal of Mathematics Teacher Education*, *8*, 255–281. doi:10.1007/s10857-005-0853-5

Rumelhart, D. E., & Norman, D. A. (1978). Accretion, tuning and restructuring: Three modes of learning. In J. W. Cotton & R. Klatzky (Eds.), *Semantic factors in cognition* (pp. 37–53). Hillsdale, NJ: Lawrence Erlbaum.

Ryle, G. (1946). Knowing how and knowing that: The presidential address. *Proceedings of the Aristotelian Society*, *46*, 1–16. Retrieved from http://www.jstor.org/stable/4544405

Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals, and understanding: An inquiry into human knowledge structures*. Hillsdale, NJ: Lawrence Erlbaum.

Schilling, S. G., Blunk, M., & Hill, H. C. (2007). Test validation and the MKT measures: Generalizations and conclusions. *Measurement: Interdisciplinary Research and Perspectives*, *5*, 118–127. doi:10.1080/15366360701487146

Shepard, L. (1993). Evaluating test validity. *Review of Research in Education*, *19*, 405–450. doi:10.3102/0091732X019001405

Shulman, L. (1986). Paradigms and research programs for the study of teaching. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 3–36). New York, NY: Macmillan.

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

Sinharay, S. (2010). How often do subscores have added value? Results from operational and simulated data. *Journal of Educational Measurement*, *47*, 150–174. doi:10.1111/j.1745-3984.2010.00106.x

Skinner, B. F. (1965, October 16). Why teachers fail. *The Saturday Review*, pp. 80–81. Retrieved from http://www.isac.psc.br/wp-content/uploads/skinner/Skinner_(1965)_Why_teacher_fail.pdf

Tatto, M., Schwille, J., Senk, S., Ingvarson, L., Peck, R., & Rowley, G. (2008). *Teacher Education and Development Study in Mathematics (TEDS-M): Policy, practice, and readiness to teach primary and secondary mathematics. Conceptual framework*. East Lansing: College of Education, Michigan State University.

TeachingWorks. (2014a). *Assessments*. Retrieved from http://www.teachingworks.org/work-of-teaching/assessments

TeachingWorks. (2014b). *High-leverage practices*. Retrieved from http://www.teachingworks.org/work-of-teaching/high-leverage-practices

Troyan, F. J., Davin, K. J., & Donato, R. (2013). Exploring a practice-based approach to foreign language teacher preparation: A work in progress. *Canadian Modern Language Review*, *69*, 154–180. doi:10.3138/cmlr.1523

Turkan, S., de Oliveira, L., Lee, O., & Phelps, G. (2014). Proposing a knowledge base for teaching academic content to English language learners: Disciplinary linguistic knowledge. *Teachers College Record*, *116*(4). Retrieved from https://www.tcrecord.org/Content.asp?ContentID=17361

U.S. Department of Education, Office of Postsecondary Education. (2011). *Preparing and credentialing the nation's teachers: The Secretary's eighth report on teacher quality based on data provided for 2008, 2009 and 2010*. Washington, DC: Author. Retrieved from http://title2.ed.gov/Public/TitleIIReport11.pdf

van Driel, J. H., Verloop, N., & de Vos, W. (1998). Developing science teachers' pedagogical content knowledge. *Journal of Research in Science Teaching*, *35*, 673–695. doi:10.1002/(SICI)1098-2736(199808)35:6<673::AID-TEA5>3.0.CO;2-J

Voss, T., Kunter, M., & Baumert, J. (2011). Assessing teacher candidates' general pedagogical/psychological knowledge: Test construction and validation. *Journal of Educational Psychology*, *103*, 952–969. doi:10.1037/a0025125

Waddell, L. R. (2014). Using culturally ambitious teaching practices to support urban mathematics teaching and learning. *Journal of Praxis in Multicultural Education*, *8*(2). doi:10.9741/2161-2978.1069

Wellman, H. M., & Gelman, S. A. (1997). Knowledge acquisition in foundational domains. In D. Kuhn & R. S. Siegler (Eds.), *Handbook of child psychology* (Vol. 2, pp. 523–573). New York, NY: John Wiley. doi:10.1002/9780470147658

Wilson, A. J. (1985, April). *Knowledge for teachers: The origin of the National Teachers Examinations Program*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL. Retrieved from http://files.eric.ed.gov/fulltext/ED262049.pdf

Wilson, S. M., Floden, R. E., & Ferrini-Mundy, J. (2002). Teacher preparation research: An insider's view from the outside. *Journal of Teacher Education*, *53*, 190–204. doi:10.1177/0022487102053003002

Wilson, S. M., & Wineburg, S. S. (1988). Peering at history through different lenses: The role of disciplinary perspectives in teaching history. *Teachers College Record*, *89*, 525–539. Retrieved from https://www.tcrecord.org/Content.asp?ContentID=540

Wilson, S. M., & Wineburg, S. S. (1993). Wrinkles in time and place: Using performance-based exercises to assess the knowledge of history teachers. *American Educational Research Journal*, *30*(4), 729–769. doi:10.3102/00028312030004729

Wilson, S. M., & Youngs, P. (2005). Research on accountability processes in teacher education. In M. Cochran-Smith & K. M. Zeichner (Eds.), *Studying teacher education: The report of the AERA panel on research and teacher education* (pp. 591–644). Mahwah, NJ: Lawrence Erlbaum.

Windschitl, M., Thompson, J., Braaten, M., & Stroupe, D. (2012). Proposing a core set of instructional practices and tools for teachers of science. *Science Education*, *96*, 878–903. doi:10.1002/sce.21027

Zeichner, K. (2012). The turn once again toward practice-based teacher education. *Journal of Teacher Education*, *63*, 376–382. doi:10.1177/0022487112445789