

Chapter 3

Developing Teachers' Knowledge and Skills at the Intersection of English Language Learners and Language Assessment

KIP TÉLLEZ

EDUARDO MOSQUEDA

University of California, Santa Cruz

The growth of teachers' professional knowledge and skills has been the topic of policy, research, and even philosophy for many decades. The assessment of English Learners (ELs), a more specific concern, has become an interest of the educational community in just the past 40 years (e.g., Harris, 1969). Our task in this chapter is to combine these two topics and consider their relations from empirical, practical, and historical perspectives (listed here in what we consider to be the rank order of importance).

At the intersection of any pedagogical practices, we are drawn into the complicated mix of generalized and specialized knowledge required for expert teaching. And although the theory/practice split might be a false dualism, we agree with Salvatori's (2003) characterization of the discipline:

That historically pedagogy has been alternately and repeatedly "elevated" to *theory* or "reduced" to *practice* . . . can be construed as an implicit but dramatic demonstration of its fundamental complexity. Indeed, I would suggest that the reasons for pedagogy's various simplifications and reductions might be found in pedagogy's complexity, rather than in its inadequacy as a discipline. (p. 67)

It will be up to others to decide whether knowledge of the proper assessment of English learners is more complex than other interrelated constructs teachers must understand; nevertheless, we are confident that oversimplifying EL assessment will diminish achievement for students already at-risk. Because this chapter, like all those in the volume, is putatively designed to offer readers a comprehensive review of the available research and theory on the issue we have identified, we begin by

Review of Research in Education

March 2015, Vol. 39, pp. 87–121

DOI: 10.3102/0091732X14554552

© 2015 AERA. <http://rre.aera.net>

circumscribing our task: The scant existing research on this topic probably does not warrant a full chapter,² but we argue that (a) the dramatic growth of the EL population in the United States and worldwide, (b) the chronic underachievement among ELs, (c) the lack of confidence and preparation in teaching English as a new language reported by many teachers and teacher candidates, (d) the limited generalized assessment knowledge and skills among teachers, and (e) the consequences of the admixture of each of these conspire to reproduce the dismal EL academic achievement patterns and negative experiences. When teachers of ELs fail to understand the nuances of general language assessment and the intersection of language and content assessment (Abedi, 2004), the specialized assessment strategies required for ELs, and the assessment of bi- or multilingual learners (Duran, 2008; Klingner & Solano-Flores, 2007; Solano-Flores & Trumbull, 2003), classroom misplacement (Artiles & Ortiz, 2002), lowered expectations (Flores, 2007), inappropriate curriculum (Gándara & Rumberger, 2009; Rumberger & Gándara, 2004), and deleterious tracking (Mosqueda, 2010; Mosqueda & Maldonado, 2013), to name just a few, interact to diminish the academic performance of ELs. With these consequences in mind, we review the existing literature on the topic but also add our own accounts of teacher knowledge and skills as they relate to EL assessment, and devote a portion of the chapter to a discussion of policy changes and future research directions that promote increased educational achievement of ELs.³ This is our overarching goal.

The chapter is organized into seven sections. In this section, we provide an overview of critical challenges in assessing ELs in their nondominant language. In the second section, we review state-of-the-art approaches for assessing the academic achievement of ELs. In the section “A Review of the Research on Preparing ‘General Education’ Teachers,” we review research and policy on the preparation of teachers with respect to the assessment of ELs. In the section “Assessment Standards for Teachers of EL,” we examine the development of disciplinary standards and the challenges of implementing such standards. In the section “An Overview of the Research on Effective Policies and Practices,” we focus on the lack of teacher preparation for assessing ELs and provide recommendations for future research. In the sixth section, we examine the implications for policy and practice regarding the preparation of teachers for EL assessment. The final section concludes with a discussion of the consequences of our findings for future research and practice.

Our chapter expands on Duran’s (2008) review of the literature delineating the challenges in assessing ELs in their nondominant language, which also provides a new direction for linking assessment and instructional practices to improve the educational experiences and outcomes of these students. In this chapter, however, we primarily review the issues relevant to teaching ELs in the U.S. context while recognizing that it is likely that a substantial proportion of English teachers worldwide now work outside of the United States, although this is a supposition on our part (the data are not clear). However, the *growth* in English teaching is almost certainly rising faster in countries outside the United States (Hu, 2003); we therefore include a few relevant international studies that mirror issues in the U.S. context. Teachers of native

Spanish-speaking ELs are the subjects in the vast majority of the research we review and proposals we advance. We admit to this bias but would point out that over 85% of ELs in the United States speak Spanish as their native language, and although limited, the majority of studies on ELs focus on Spanish-speaking students; therefore, their teachers are also far more numerous (see Téllez, 2010, for an overview). However, we want to acknowledge the diversity of languages (e.g., Mixteco, Zapoteco, Triqui, and Maya) spoken by an increasing number of immigrant students from Mexico and Central America, especially in our region of California. Last, native Spanish-speaking ELs are generally those whose academic achievement is of greatest concern and therefore the focus of teacher knowledge and skills. But before moving on to addressing these specific issues, we need to determine who is included in the definition of EL. We presently have no common national criteria for what determines EL status. Duran (2008) and others (Abedi, 2004; Celedón-Pattichis, 2004; Mahoney & MacSwan, 2005) point out that the classification of ELs is vague due to the lack of consistency and agreement for identifying such students among districts and states. Duran (2008) describes the problem:

ELLs participating in state large-scale assessments are in effect a policy construction, a category of students established by individual states to satisfy their education laws to deal with a growing group of students from non-English backgrounds who show some evidence of limited familiarity with English, patterns of low school achievement, low assessment scores in English, and propensity to drop out of school. (p. 300)

In short, ELs are in urgent need of a reformed school experience and improved opportunities to learn, but how can we create such enhancements when we do not even know who they are? Solving this problem begins with accurate and agreed-on assessment tools and systems.

As the research literature (Rivkin, Hanushek, & Kain, 2005) and popular press (Brooks, 2012) draw attention to the critical role of the teacher in academic achievement, the education community finds itself in a search for the right mix of teacher knowledge, skills, and attributes that, together, contribute to increased student achievement. At root, some teachers are more effective than others. No thoughtful educator, researcher, or policymaker would disagree on this point. The pervasive measuring and comparing of teachers (both preservice and in-service) as well as the designing of methods to improve the teaching profession through professional development offer evidence that we believe that good teachers can be made (or at least selected) and that even good teachers can significantly improve their instruction.

In fact, it is becoming increasingly clear that there are but two fundamental ways for schools to help students who are not meeting grade-level standards (a category that includes many ELs) to “catch up.” The first way, which is growing in popularity, is simply to add to the time that students spend with teachers learning school objectives. In this model, teachers simply work more. For instance, schools that demand students spend 8 hours per day or more in school and attend on weekends will obviously have differential and mostly positive academic outcomes simply because they put more teacher time to meeting educational objectives (Gándara & Rumberger,

2009; Tuttle, Teh, Nichols-Barrer, Gill, & Gleason, 2010). Adding time in the form of an extended school day, after-school programs, or weekend classes is not new, and its “dimensions” are varied and complex, as Gándara (2000) and authors in her book point out (Anderson, 2000; Minicucci, 2000). We want to underscore this strategy because it appears that an increasing number of Latino learners (who comprise the largest group of ELs) are attending such schools. In these contexts, teachers’ experience, knowledge, and skills appear to be irrelevant. In fact, some schools in this category specifically point out their teachers’ inexperience and suggest that the school’s purported success is owed primarily to the extra time given to teaching (Angrist, Dynarski, Kane, Pathak, & Walters, 2010).

The other way to “teach more” is make the time students spend with teachers more efficient, that is, use strategies that help students learn more in the same amount of time. This is the goal of teacher professional development—to expand the capacity of teachers to improve their instructional skills and, in turn, raise achievement. The relationship between time and efficiency can be expressed in the following equation: $E/T = L$, where E is efficiency, T is time, and L is learning (or what an economist would call productivity). If you lack efficiency, you must add time to learning. If you have a fixed amount of time, you need to be more efficient in order to gain the same amount of learning.

We believe this equation is particularly relevant to our topic. For instance, teachers who lack sufficient knowledge of EL assessment (i.e., efficiency) are likely to have their EL students doing work that is either too difficult or too easy and thus inefficient. On the other hand, a teacher who holds expert knowledge and skills with respect to EL assessment will know students’ language levels and have them work at their instructional capacity,⁴ which results in efficient teaching and learning.

We argue that both increased time and instructional efficiency are needed to enhance the academic success of ELs. But adding instructional time is a matter of policy decision making and resource commitment, and it is by far the easier lever to pull, although we would point out that time, in the end, is clearly fixed and limited.

Our task is to consider how we can improve the assessment of ELs, by enhancing the skills of teachers, and thus improve academic achievement for one of our most vulnerable populations. We are addressing an admittedly narrow slice of a much larger teacher development project that will be required to improve the education experiences of ELs, but for those of us who are deeply concerned about the education of our nation’s roughly 7 million ELs, this topic is crucial. Imagine an EL student whose teacher does not know her/his native language *and* who lacks skills in EL assessment. Such a teacher has no reliable way of knowing what an EL student knows or does not know in either her/his native language or in English. The work assigned to ELs may be entirely inappropriate, but ELs cannot object because they often lack the language capacity or the social capital⁵ (or both) to say so. This predicament, corroborated in research by Celedón-Pattichis (2004) and others (e.g., Rodriguez, 2009), is the fate of too many ELs, who consequently become disengaged and marginalized from healthy, productive school experiences.

For many of us in language education in the United States, the protests and eventual takeover of Crystal City, Texas, schools (Trujillo, 1998, 2005; Valenzuela, 2000) remain a watershed event, equal to Little Rock in importance.⁶ But the heroic effort of the students and citizens in this remote Texas town, who initiated La Raza Unida Party, needs wider recognition for their insistence that students' native language be respected and that the teaching of new languages be compassionate. Their legacy, as well as other communities who protested cruel language teaching practices (e.g., San Miguel, 2004), motivates our work, and we are thankful that the editors of this volume share our urgency and invited us to write this chapter.⁷

As we begin the sections that form our review, it might be useful to share some of the questions that teachers of ELs should be asking based on our review of the available research on EL assessment:

- How much English do ELs know?
- Do ELs have relative strengths within their English capacities (i.e., are they better at reading and writing than speaking and listening)?
- Can ELs read in their native language? If so, how well?
- How much content knowledge do ELs have and understand in their native language?
- How is their growing knowledge of English influencing ELs' learning of content?
- Are ELs making "average" growth in learning English? If not, is it a result of a generalized language learning challenge or a challenge in learning *English specifically*?
- If I don't know the answers to these and other questions, how can I find out?

These questions will guide the content of the sections below. As mentioned, our task is to review the relevant literature, but we are also hoping to provide a few examples of pedagogical strategies, from both the literature and our own experiences. A final caveat: Although we need to consider what the field considers strong EL assessment, our task in this work is not to conduct an exhaustive review of appropriate assessment for EL, and refer readers to other sources (e.g., Abedi, 2004; Abedi & Lord, 2001; Duran, 2008; Solano-Flores, 2011; Solano-Flores & Trumbull, 2003; Téllez, Moschkovich, & Civil, 2011). We address only the specific assessment issues that intersect with our task of considering how teachers can better gain the necessary knowledge and skills required for sound assessment of ELs.

THE STATE-OF-THE-ART IN ASSESSING THE ACADEMIC ACHIEVEMENT AND ACADEMIC GROWTH OF ELS

All ELs are better described as emerging (Téllez, 1998) or emergent (O. García, Kleifgen, & Falchi, 2008) bilinguals, whose language and/or content capacities cannot be fully assessed in any single language.⁸ As Valdes and Figueroa (1994) pointed out nearly 20 years ago, assessing bilingual learners presents distinct challenges, and

ignoring the unique constellation of language skills in any EL may result in systematic bias in the interpretation of score results. The lack of proper assessment can have serious consequences on the academic preparation of ELs, particularly for those who are misdiagnosed and placed in special education. Such practices often result in the disproportionate placement of ELs in special education when it is unclear if ELs struggle to learn because of a disability or because of language acquisition issues (Artiles & Klingner, 2006; Artiles, Rueda, Salazar, & Higuera, 2005).

Although it may not be obvious to monolingual speakers, it is quite logical that EL students may know a word in their native language that they have yet to learn in English. For instance, a native Spanish-speaking student may know the word *harina* (flour in English) as a result of family cooking experiences but have no knowledge of its English counterpart. This example points out that knowledge of terms or linguistic structures in a student's native language can exist independent of target language knowledge, especially when such knowledge is separated by the terms and structures one might learn in school versus those more commonly learned at home. The ongoing debate on the "balance" of languages in the mind is beyond our goals for this chapter (see Grosjean, 1989), but the knowledge of this concept and its implications for assessment are crucial knowledge for the teacher of ELs.

Teachers and candidates must have some knowledge of what a student knows in the native versus target language. By way of exploring this knowledge and what EL professionals might be asked to understand, we review the efforts to assess bilingual verbal capacities and highlight the work of Muñoz-Sandoval, Cummins, Alvarado, and Ruef (1998), whose Bilingual Verbal Ability Test (BVAT), can be used to determine the linguistic capacity or linguistic aptitude of bilinguals. Briefly described, the test relies on the general structure of the Woodcock–Johnson test of linguistic aptitude (Woodcock, McGrew, & Mather, 2001). Students first take the English version and then retake the missed items in a parallel test given in their native language. The scores are combined using a correction factor and a composite linguistic capacity score is computed. The score is an estimate of a bilingual's linguistic capacity, regardless of the language in which it is "housed." Before going further, we want to suggest that the interest in measuring a student's linguistic aptitude has been unfairly cast as an attempt to, for instance, label students as learning disabled or to track them into low-level classes. Although one can cite many cases, mostly historical, in which aptitude tests have been used to further subjugate an already disempowered minority group (Gould, 1996; Hilliard, 2000), we argue that teachers of ELs must understand the difference and usage between tests of linguistic achievement and linguistic aptitude, especially when considering the constellation of a bilingual student's capacities. We would also point out that the BVAT has been used to identify gifted ELs when a linguistic capacity test in English and nonverbal assessments failed to recognize EL students' extraordinary intellectual capacity (Breedlove, 2007).

We are not necessarily arguing that all teachers gain the skills needed to administer a test such as the BVAT (although they would learn much from the training), but they must understand the concept behind the test. Simply put, how can a teacher

understand proper assessment of ELs without also understanding the knowledge that a student has in the native language and the balance of the target and native languages? The BVAT results offer clear evidence that both a student's native language and English play a role in all assessment processes (Páez, 2008). We are also not recommending that schools take on the role of administering the BVAT to all its ELs. This effort would be prohibitively expensive and unnecessary. Rather, we are suggesting that an understanding of the purpose and practice of the BVAT will assist educators in knowing whether struggling ELs are challenged in learning English or learning language itself. This is a crucial distinction for educators: Mistakes can result in ELs failing to receive adequate special education services or, worse, in incorrect assumptions being made about a learner's effort and/or content knowledge understanding.

Although the BVAT and similar tests that claim to measure something called linguistic aptitude have been rightly questioned for their bias toward the experiences of middle-class children and youth (Valencia & Suzuki, 2000), the recent attacks on testing bias have come largely from critics of single language achievement tests when used to assess ELs. In a wide-ranging review of the issues connecting ELs to general academic assessment, Solórzano (2008) highlights the long-standing concerns regarding validity when students take a test in a language they are learning. Solórzano also reminds us that ELs are mandated to be included in the accountability scheme under No Child Left Behind (NCLB) and that most states, including California, make almost no allowance or modifications for ELs.⁹ A host of researchers have argued that the rules of NCLB are unfair or should not apply to schools with high concentrations of ELs (Abedi, 2004; Harper, de Jong, & Platt, 2008; McCarty, 2009; Menken, 2008), but such calls have been mostly ignored. Solórzano (2008) also emphasizes a point made by Gándara (2002; as cited in Solórzano, 2008): that even when the manufacturers of the standardized assessments suggest that their products are not valid measures of EL academic achievement, legislators and policymakers tend to ignore these warnings and require ELs take them anyway, and then insist that their scores be used to rank and sort schools and school systems. Disciplinary standards (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999) are just as easily ignored. Requiring emergent ELs, for instance, to take an achievement test in English (and only in English) is illogical, but this decision, at least in the case of NCLB, is enforced by policy, not by testing experts or even test development companies. Too often we find teachers, candidates, and other educators blaming the test itself when the rightful culprits are the policymakers (i.e., politicians) who should answer to the public, which includes a good many educators. Identifying who is at fault for irresponsible testing policies should be a primary goal of EL advocates, but teachers are often frustratingly apolitical and unwilling to confront officials (Bartolome, 2004). The test is what ELs struggle with—and what teachers see most clearly—but it is a poor metonymic device for misguided accountability schemes created by politicians motivated by the call for “higher standards,” regardless of the manifest psychometric concerns.

The fact that so many ELs are now forced to take tests in English when it is clearly inappropriate has, paradoxically, encouraged a host of studies and proposals exploring the ways by which the language bias in such tests can be reduced. The question is essentially this: How can we estimate an EL's true knowledge of content (e.g., math, science) in a language they are learning? Or put more simply, is English a barrier for ELs who know the construct but cannot "find" the correct answer because they cannot understand the language of the question itself? The answer is certainly yes, but the degree of bias has been the subject of several key studies in the past decade. For instance, Martiniello (2009), using differential item functioning methods, found four types of construct-irrelevant text that diminished EL performance on mathematics tests:

- Syntactic: multi-clausal complex structures with embedded adverbial and relative clauses; long phrases with embedded noun and prepositional phrases; lack of clear relationships between the syntactic units.
- Lexical: unfamiliar vocabulary, high-frequency words usually learned at home and not in school; polysemous or multiple-meaning words.
- References to mainstream American culture.
- Test or text layout. Lack of one-to-one correspondence between the syntactic boundaries of clauses and the lay out of the text in the printed test. (p. 176)

However, the results indicated that the bias resulting from syntactic and lexical complexity was reduced if the item was presented with a schematic representation of the test taker's task. The implication of this research for teacher knowledge is that graphic organizers, a long-standing, successful instructional strategy for second language learners (Tang, 1992), are as useful in assessment as in instructional contexts. Recent studies (e.g., Haag, Heppt, Stanat, Kuhl, & Pant, 2013; Wolf & Leon, 2009) have corroborated Martiniello's (2009) findings.

Given that content area assessments make demands of EL language capacities, and that these demands can influence performance, we should be directing attention to exactly these concerns using common criteria. To this end, Shaw, Bunch, and Geaney (2010) provide a useful taxonomy and analytic frame for understanding the language demands on performance assessments in science. This tool can be useful for exploring assessments that will help educators gain a clear understanding of language in assessment.

In an article designed to reconsider the paradigm of testing for ELs, Solano-Flores and Trumbull (2003) suggest that testing for ELs must be reframed to include fundamental shifts in test development, test review, and the treatment of language as an added source of measurement error. This final point, that the language of a test should be considered error, represents a new and compelling argument in psychometrics. Educators working with ELs would do well to understand their assertion, but they must first understand the basic tenet of assessment theory—that is, when we measure any human quality, we are approximating some unknown true score. Whatever score or mark we ascribe to an individual is composed of two components, an observed score and measurement error. Although there are various techniques for

estimating error, many of which lie beyond the needs of most educators, the simple calculation required to compute the standard error of measurement seems a reasonable expectation (McMillan, 2000; Popham, 2011).

The important point that Solano-Flores and Trumbull (2003) make is that ELs are, in fact, bilingual test takers and that each has a “unique set of weaknesses and strengths in English and a unique set of weaknesses and strengths in his or her native language” (p. 8). The error we find in the scores of EL is partly due to the distribution of knowledge across languages, a point we underscored in our discussion of the BVAT.

Thus far in this section we have addressed the shortcomings of what might be termed *summative tests* (i.e., manifold end-of-year achievement tests designed to assess a year’s worth or more of learning), but researchers and school leaders have identified formative assessment as perhaps the more crucial knowledge for teachers due to their potential to inform and enhance instructional practice. Although the distinction between formative and summative assessments is not always clear (see Scriven, 1991, for a review), we recognize the need for teachers to understand and use those types of assessments, whatever their terms, that both guide ELs to the proper instructional level and offer teachers a valid and reliable way to know what students know and do not know in “real time.”

In general, formative assessments involve a systematic process to continuously gather evidence and provide feedback about learning while instruction is underway (Heritage, 2007; Heritage, Kim, Vendlinski, & Herman, 2009), all the while making student thinking transparent to teachers and using such evidence to help teachers adapt their instructional strategies to help meet the desired learning goals (Ruiz-Primo, Furtak, Ayala, Yin, & Shavelson, 2010). Shepard (2006) has argued for a model of formative assessment where in addition to providing information that teachers can use to improve instruction, the information garnered should also guide student learning. Thus, formative assessments can include the following goals for students: identifying students’ strengths and weaknesses; aiding students in guiding their own learning, revising their work, and gaining self-evaluation skills; and fostering increased autonomy and responsibility for learning on the part of the student (Cizek, 2010). Research has shown that well-designed, formative assessments can increase student achievement (Black, Harrison, Lee, Marshall, & Wiliam, 2004).

Formative assessments vary widely depending on their purpose (Ruiz-Primo et al., 2010), but most educators agree that formative assessments must be embedded within an instructional unit and the results should be used to inform the learning goals of a particular lesson (Ayala et al., 2008; Cizek, 2010; Ruiz-Primo et al., 2010). To realize the potential of formative assessment to help support and enhance instruction and student learning, the implementation process of formative assessments must address at least three dimensions that are encapsulated in three questions posed by Ramprasad (1983, as cited in Ruiz-Primo et al., 2010): Where are we going? Where are we now? How will we get there? As Ruiz-Primo et al. explain, the “Where are we going?” question focuses on the teacher “setting and clarifying learning goals” and

identifying “evidence [for] achieving those learning goals,” whereas the “Where are we now?” question refers to “specific practices in which teachers seek to understand students’ current and prior knowledge” (p. 139). The third question, “How will we get there?” centers on how the teachers will modify their instruction to meet the needs of students (Ruiz-Primo et al., 2010). Moreover, another critical feature of formative assessments is the provision of effective feedback to students based on the assessments results (Sadler, 1989).

Despite the potential of formative assessments to improve teacher practice, their effective implementation requires a high degree of assessment and pedagogical knowledge. Heritage (2010) has found that teachers need to master four basic elements of teacher knowledge in order to implement formative assessments successfully: (a) domain knowledge, (b) pedagogical content knowledge, (c) knowledge of students’ previous learning, and (d) knowledge of assessment. As previously discussed, researchers have raised concerns about the lack of useful assessment training in preservice teacher preparation programs and professional development opportunities available to practicing teachers to help develop their assessment literacy; thus, the complex nature of effectively implementing formative assessments is often taken for granted.

The literature focused on formative assessment of ELs is scant (Duran, 2008; Llosa, 2011). What does it mean to have a linguistically and culturally responsive assessment? Do formative assessments have the potential to be sensitive to linguistic and cultural learning needs of ELs? The works of Solano-Flores (2006) and Solano-Flores and Trumbull (2003), which examine the linguistic and cultural sources of measurement error on summative assessments, provide a useful framework for evaluating formative assessments. The accurate assessment of ELs’ content knowledge mastery is a complex undertaking given the psychometric limitations not limited to construct-irrelevant variance, or the underestimation of subject matter understanding of ELs resulting from their low degrees of English language proficiency (Abedi, 2004; Duran, 2008; Heubert & Hauser, 1999).

A REVIEW OF THE RESEARCH ON PREPARING “GENERAL EDUCATION” TEACHERS FOR QUALITY ASSESSMENT PRACTICES

Before exploring the research and policy surrounding what teachers know and can do with respect to the assessment of ELs, we are obligated to understand what they know about general assessments of native English-speaking children in disciplinary content areas such as literacy and mathematics. It is our view that the assessment of ELs (or any second language learner for that matter) is a special and more complex case than that of native speakers. Monolingual, native speakers of English can be assessed in English, without regard for the distribution of language capacities in bi- or multilingual learners.

As we reviewed the literature on teachers’ knowledge of assessment, an area that has come to be known as “assessment literacy” (Mertler & Campbell, 2005; Stiggins,

1991), we find almost unanimous disappointment from the research and policy communities. Popham (2011) defines assessment literacy as “an individual’s *understandings of the fundamental assessment concepts and procedures deemed likely to influence educational decisions*” (p. 267). The malaise is summed up neatly in the titles of two articles by the assessment specialists Stiggins and Popham: “The Unfulfilled Promise of Classroom Assessment” (Stiggins, 2001) and “Seeking Redemption for Our Psychometric Sins” (Popham, 2003).

Stiggins’s (2001) admonition is particularly strident, arguing that current conditions can be explained only by presenting a fictional scenario: “It is as if someone somewhere in the distant past decided that teachers would teach, and they would need to know nothing about accurate assessment” (p. 5). Popham (2006, 2011), as previously mentioned, has made the same case, and he is joined by a host of other researchers and policymakers (Cizek, Fitzgerald, & Rachor, 1995; Daniel & King, 1998; McMillan, 2003; Randall & Engelhard, 2010) whose empirical work confirms the general sentiment.

If teachers lack assessment knowledge, are there specific areas of this shortcoming? Plake, Impara, and Fager (1993) measured teachers’ competencies on the seven basic assessment areas identified in the Standards for Teacher competence in Educational Assessment of Students. Overall, the results show that teachers were most knowledgeable in the areas of administering, scoring, and interpreting test results. The poorest performance came from those items measuring the teachers’ knowledge about communicating test results. Although the authors report that the teachers’ knowledge of assessment was deemed quite inadequate, the study offers some hope: It found that teachers who had training in measurement scored significantly higher than those who had not. This finding indicates that a general knowledge of assessment is not out of teachers’ reach, perhaps only that professional development has not been widespread enough.

With such an identifiable shortcoming exposed, it is reasonable to ask if candidates are ever introduced to classroom assessment in their preservice professional programs. Indeed, teacher education is often blamed for the shortcomings we find in the general teacher population, whether the teachers we reproach are in their first or 30th year. In our view, it is folly to argue that professional school programs should equip teachers with all the knowledge the wider educational community wishes they had, not to mention prepare them for their work 10 years into their profession. We admit to some bias in this regard, if for no other reason than that we work primarily in preservice teacher education. However, preservice licensing programs have always had a marginal influence, which has only been diminished in recent years as states limited the number of credits professional programs could require for a license (Cochran-Smith, 2001). In short, the lack of assessment literacy preparation of beginning teachers is clear. Veenman’s (1984) widely read review on the perceived problems of beginning teachers finds that “assessing students’ work” tied for fourth on a list of 24 shortcomings found in the research literature. A range of recent studies has shown that beginning teachers enter the profession with scant knowledge of assessment (MacLellan, 2004) and that

even after a course in assessment, candidates fail to enact their new knowledge (Campbell & Evans, 2000), suggesting that creating and using strong assessment practices could be developmental in nature. On the other hand, Weinstein (1989) found that candidates neither understand assessment nor regard it as important knowledge, so perhaps the lack of assessment literacy is more a matter of attitude than programmatic omission (see also Volante & Fazio, 2007).

The news on assessment literacy is not all bad. A recent study by Mertler (2009) demonstrated that even a 9-day course of study in assessment practices addressing general introductory topics (e.g., conducting an item analysis, developing valid grading procedures, developing a performance assessment and a scoring rubric) resulted in teachers who were more confident and who improved their assessment practices. It appears, then, that teachers and candidates may benefit from professional education on assessment practices but that current efforts are either nonexistent or underresourced.

ASSESSMENT STANDARDS FOR TEACHERS OF EL

The past 20 years have seen teacher licensing and professional growth turn to the development of comprehensive standards for the profession. Several major reports calling for the professionalization of teaching argued that teachers must take hold of professional standard setting if we are to establish confidence for our work among policymakers and the general public (e.g., The Holmes Group, 1986). The argument was that professionals must define high standards for those entering and continuing in the profession, set rigorous expectations, and then hold peers to these standards and expectations. Roth (1996) expertly documented—in some ways predicted—the early efforts of this movement, and we recommend that interested teacher educators recall his efforts to warn us of the consequences (Téllez, 2003). Darling-Hammond (1999) also offers a circumspect view of the development of disciplinary standards; she further suggests that the proliferation in standards is our own doing and that standards might not work as teacher educators envisioned. Even if a profession agrees to standards, it usually does not take long for policymakers to use the standards for political advantage (Cizek, 2000; Darling-Hammond, 2006; Téllez, in press).

More recently, and more forcefully, Taubman (2009) argues that standards in teacher education have had nothing but a corrosive influence on teaching, serving only to deprofessionalize the field. We also want to recognize our own bias against the pervasive use of standards in education and agree entirely with writers such as Faltis (1990), Pennycook (1999), Johnston (2002), and Kumaravadivelu (2005), who point out the reflexive and moral consequences when teaching ELs, both of which standards can obliterate.

A full discussion of costs and benefits of standards is beyond our task in this chapter, but with the widespread development of standards, we wanted to explore a select set of standards for those that specifically address assessment of ELs. If no government or professional entity required knowledge of EL assessment, then it would be less likely that candidates would gain such knowledge.

We begin by noting Bachman's (2000) call for increasing the training of language testing professionals and, predictably, for the development of standards of practice and mechanisms for ensuring their implementation. With the recognition the standards cannot alone bring about excellence, Bachman points out the dismal circumstances:

Most professional programs, including certificate courses, master's and doctor's degree courses in language teaching or applied linguistics still require no coursework or guided practice specifically in language testing, so that the majority of practitioners who develop and use language tests, both in language classrooms and as part of applied linguistics research, still do so with little or no professional training. (pp. 19–20)

In the decade following Bachman's (2000) admonition, it appears that the assessment of ELs has made its way into standards for initial licensing, university- or college-based professional programs, and organizations designed to recognize the work of advanced professionals. We will explore an example taken from each of these categories, by first exploring California's Teaching Performance Expectations (required for licensure) and a performance assessment designed to assess the Expectations. We next examine the Teacher of English to Speakers of Other Languages (TESOL) and the U.S.-based National Council for Accreditation of Teacher Education (NCATE) standards for professional programs, and we end this section with an analysis of the National Board for Professional Teaching Standards.

California's Teaching Performance Expectations

California's recent restructuring of teacher education included new standards for licensure. As a consequence of the now-infamous Senate Bill 2042 legislation, the state developed the Teaching Performance Expectations (TPEs; California Commission on Teacher Credentialing [CCTC], 2013). As the state with the most English learners (1.4 million English learners, representing 23% of the entire school-age population; California Department of Education, 2011), educators and policy-makers were set on making certain that all new teachers were certified to teach ELs. Consequently, all state programs must ensure that candidates meet TPE 7: *Teaching English Learners*. Specific to our concern, this standard requires, among a longer list of competencies, that candidates "draw upon information about students' backgrounds and prior learning, including students' assessed levels of literacy in English and their first languages, as well as their proficiency in English, to provide instruction differentiated to students' language abilities."¹⁰ In TPE 3: *Assessing Student Learning*, the standards specifically mention EL: "Candidates interpret assessment data to identify the level of proficiency of English language learners in English as well as in the students' primary language."

These standards represent challenging tasks for the preservice teacher candidate. Indeed, we are most struck by the emphasis on the knowledge and interpretation of assessments in both English and the students' native languages and wonder how

many programs can make such knowledge mandatory among its candidates. We recall our earlier discussion regarding the BVAT and the range of knowledge required to understand the balance of languages.

For better or worse, policymakers in California wrote into Senate Bill 2040 law that all new teachers must be evaluated on the TPEs in a performance context. Teacher preparation programs in California can select the standardized Teaching Performance Assessment (TPA) system developed by the California Commission on Teacher Credentialing (CCTC), in partnership with the Educational Testing Service, or an alternative assessment that meets with the CCTC's assessment quality standards. After reviewing the TPA developed by the state, a group of universities and professional programs¹¹ developed an alternative performance assessment, known as the Performance Assessment for California Teachers (PACT; see www.pacttpa.org for a description and other information). We do not intend to provide an exhaustive review of the PACT and direct readers to other sources (e.g., Bunch, Aguirre, & Téllez, 2009; Porter, 2010; Sandholtz & Shea, 2012), but we would like to note aspects of the PACT that encourage—but not guarantee—a deep knowledge of EL assessment. First, the PACT is divided among four tasks (Planning, Instruction, Assessment, and Reflection) and built around a 2- to 3-day instructional segment. The Assessment task requires that candidates identify three student work samples (one must be an EL) that represent “class trends” in what students did and did not understand. Candidates must also provide a commentary that examines the standards/objectives of the lesson segment, analyzes the learning of individual students represented in the work samples, outlines feedback to students, and identifies next steps in instruction.

The PACT is distinguished primarily by a rubric assessing the candidate's knowledge of and instruction in academic language. It is on this point that the PACT is innovative with regard to EL assessment. For instance, to pass the PACT, candidates must demonstrate that they have used “scaffolding or other supports to address identified gaps between students' current language abilities and the language demands of the learning tasks and assessments, including selected genres and key linguistic features” (PACT, n.d.). Scoring at higher levels requires candidates to point out the role of the textual resources of the specific tasks/materials and how they are related to students' varied levels of academic language proficiency. It is when candidates aim for scores beyond passing that they often approach the comprehensive knowledge outlined in the TPEs (see Bunch et al., 2009).

The TESOL/NCATE Standards

Over a decade ago, the international organization TESOL and the U.S.-based NCATE jointly developed standards for programs preparing teachers for ELs. It is important to note that even though these standards are written with the candidate as the subject in most sentences in the document, these standards evaluate *programs*. And unlike California's TPEs, the standards require no direct assessment of candidates; rather, evidence gathered from program documents offers supporting evidence.

With respect to assessment knowledge and skills for teachers of ELs,¹² the domain is divided into three standards: Issues of Assessment for ELs, Language Proficiency Assessment, and Classroom-Based Assessment for ELs. We have included the citation for the entire standards document and do not have enough space to review these standards thoroughly, but we do wish to point out the text from the third standard: "Candidates can assess learners' content-area achievement independently from their language ability and should be able to adapt classroom tests and tasks for ELs at varying stages of English language and literacy development" (TESOL International Association, 2003, p. 64). Once again, we find that standards for teachers of ELs are demanding that teachers understand the relation between a general linguistic capacity and academic knowledge.

Overall, the TESOL/NCATE standards are comprehensive and lengthy. Although we do not have data on how many programs worldwide are approved using the standards, accredited institutions are likely graduating candidates better prepared for assessment of ELs than those not reviewed. We nevertheless agree with Newman and Hanauer (2005) who offer a thoughtful critique of the standards. They point out the overlapping and confusing nature of the assessment standards, in particular, and wonder how one can meaningfully separate the domain into three mutually exclusive areas of knowledge and skills. They also critique the breadth of the standards, noting that one standard requires candidates to understand the strengths and weaknesses of norm-referenced assessments. We agree that this skill is well beyond what should be expected of a beginning educator.

The National Board for Professional Teaching Standards

This well-known effort is designed to recognize exemplary, experienced teachers. Teachers produce a wide-ranging portfolio of their work, and if the criteria are met, they are given the distinction of Board-Certified teacher. In some states, teachers are given bonus pay for board certification (see <http://www.k12.wa.us/certification/nbpts/teacherbonus.aspx> for an example). The two categories of certification germane to our discussion are the Teaching English-as-a-New-Language portfolios for Early and Middle Childhood and Early Adolescence Through Young Adulthood (National Board for Professional Teaching Standards, 2012a, 2012b). The assessment sections of the standards are extraordinarily thorough, comprising the eight sections listed below (we have included sample text from three of the standards, pp. 77–85):

1. Variety in Assessment Techniques: "[Teachers] create their own tools for assessment that might incorporate students' daily class work, artwork, or exhibits."
2. Initial Placement Assessment
3. Assessment to Guide instructional or formative assessment
4. Assessment in the five language domains
5. English Language Proficiency Assessment

Accomplished teachers understand the purpose of proficiency assessments with regard to current local, state, and federal guidelines for monitoring the progress of students' English language development. Teachers collect and analyze data from formal sources. They know how to examine such assessment instruments critically and understand their uses and limitations in the practice of informed teaching. Teachers are knowledgeable about the psychometric properties of standardized tests when administered to ELs; academic language proficiency assessments; reading placement tests; and formative instructional assessments.

6. Standardized Content Assessment

Accomplished teachers work collaboratively with school staff to confirm the eligibility of English language learners to participate in content-area assessments and ascertain that students are assessed fairly. Teachers understand test validity and reliability and are able to explain to colleagues how these concepts relate to the unique features of evaluating English language learners.

7. Assessment for special purposes, in particular identifying gifted ELs

8. Substantive Assessment information for families and others

Naturally, these standards far exceed most of the requirements for initial licensure (cf. California's TPEs). The range and depth of knowledge is impressive, but we would draw attention to Standard 6, which requires that teachers understand and make certain that ELs are assessed fairly in the content areas. We also are intrigued by the fact that the Teaching English-as-a-New-Language portfolios are evaluated using a rubric that assesses a teacher's knowledge and instruction of academic language. As a relatively new concept for assessing teachers' knowledge at the intersection of language and content, teacher knowledge of academic language eliminates the false distinction that language can be learned without learning about anything (Bunch, Abram, Lotan, & Valdes, 2001; Laplante, 2000; Téllez, 2010).

AN OVERVIEW OF THE RESEARCH ON EFFECTIVE POLICIES AND PRACTICES IN PREPARING TEACHERS FOR PROPER ASSESSMENT OF EL AND RECOMMENDATIONS BASED ON THIS RESEARCH

The notion that teachers of EL need specialized preparation at all is a surprisingly recent development. We point to E. García's (1990) key chapter that illustrated the pitiful state of teacher quality for ELs. Citing the results from the available research and several national reports, he found that linguistic minority education programs were staffed by professionals not directly trained for such programs and who lacked adequate knowledge of second language teaching and learning.¹³ What they learned, they learned "on the job." Teacher education researchers have discovered that a great many teachers of ELs, unprepared for conditions working with a linguistically diverse student population, fail to acquire much expertise at all (Gándara & Maxwell-Jolly, 2006; Gándara, Rumberger, Maxwell-Jolly, & Callahan, 2003). Instead, they grope for quick fix strategies, often becoming stressed at their limited options and lack of success. In a review of effective teacher professional growth practices, Knight and

Wiseman (2006) suggest that teachers of ELs need professional development focused on understanding language development that distinguishes between ELs' capacities for listening, speaking, reading, and writing. In fact, the early research conducted by Cummins (1982), who has more recently become a foremost international advocate for bilingual education, addressed this very concern; he demonstrated that when educators fail to recognize that ELs' proficiency in spoken English does not necessarily indicate full proficiency, ELs can be mislabeled as learning disabled, with disastrous consequences for the future learning opportunities of such students.

Self-reported data from teachers corroborate these findings, suggesting that teachers are confused and unsure about their capacities for effective teaching of ELs (Alexander, Heaviside, Farris, & Burns, 1998; Gándara & Maxwell-Jolly, 2006), although a recent study finds that teachers have greater confidence in their collective efficacy to provide strong EL instruction (Téllez & Manthey, in press).

E. García's (1990) admonition and other factors, such as a lack of teacher confidence as well as the growth in the EL population and their unacceptable academic performance, have motivated teacher educators and policymakers to initiate improvements in the quality of EL instruction. The decade of the 1990s saw a host of new policies and programs for the preparation of teachers of ELs. Many universities began specialized preparation for EL students, although some needed state legislation to initiate such improvements. Today we find a range of scholars and policymakers advocating for the importance of specialized preparation for teachers of ELs. Indeed, given the historical lack of attention to ELs, this burst of interest comes as somewhat of a surprise to those of us who have been working with ELs for decades (Theoharis & O'Toole, 2011).

As we consider the research that may point to effective strategies for preparing teachers of EL for better assessment strategies, we first consider what the research literature says about those competencies needed for quality EL teaching.¹⁴ A long list of competencies and stances is quite impractical to the professional school teacher educator, who has but a year or two, perhaps three, to "prepare" teachers for their first year, or for the school professional development coordinator who must squeeze teacher learning regarding ELs into calendars already crowded with professional development. One recent research-based guide for competencies comes from Lucas and Villegas (2010), whose "Framework for the Preparation of Linguistically Responsive Teachers" includes the following elements: (a) sociolinguistic consciousness; (b) value for linguistic diversity; (c) inclination to advocate for EL students; (d) learning about EL students' language backgrounds, experiences, and proficiencies; (e) identifying the language demands of classroom discourse and tasks; (f) knowing and applying key principles of second language learning; and (g) scaffolding instruction to promote EL students' learning. Lucas and Villegas are careful to point out that these elements are not mutually exclusive and that the framework is neither intended as a formula and nor does it include every single knowledge and skill required for teaching ELs; however, assessment is notably missing from the list. Other works outlining what EL teachers should know and be able to do focus less on specific

competencies and more on teacher socialization as a form of professional growth (e.g., Freeman, 1991; Freeman & Johnson, 1998), but knowledge of assessment is not specifically mentioned.

Again, we want to point out the value of these works and suggest that teacher educators would do well to heed their recommendations, but what might be the consequences of omitting assessment? Our discussion thus far has indicated that assessment may not make the top of the lists and that the research on professional development is quite sparse, offering scant guidance; nevertheless, the literature offers several excellent examples of EL teacher growth in assessment practices.

In a descriptive study that invited practicing bilingual teachers to develop concurrent tests of mathematics, one in Spanish, one in English, Solano-Flores, Trumbull, and Nelson-Barber (2002) found that teachers focused first on what might be called surface-level translation concerns. For instance, they wondered if the test should use metric values because these are the terms that native Spanish-speaking children, most Mexican American, would hear from their parents. As the teachers continued their translation work, they began to attend to what the authors termed the *deeper* concerns regarding the varying structures of the two languages. The teachers recognized that no two tests could ever be made entirely equal if they are written in different languages. The primary conclusion of the study revealed not that the teachers somehow got the translation correct and created equivalent forms, a complex task and arguably impossible feat, but rather that their collaborative work greatly enhanced their understanding of assessment. The study's authors conclude that teachers working in collaborative teams on assessment grow in their sophistication and that the collaboration was the key feature. We will find that the collaborative nature of teaching learning works in assessment literacy just as it does in other arenas of teacher development.

Our attention to the previous study, one focused on bilingual teachers and learners (in contrast to general English learners), raises an important question: Is bilingualism in educators a predictor of enhanced knowledge of language assessment? In other words, do we want all educators responsible for testing EL be bilingual themselves? If so, would we wish for them to be proficient in the native language of the students? The answer on both counts is, of course, yes. In fact, we always want people to speak more than one language. Who among us would not prefer to be multilingual if given the opportunity? But our question is more specific. Do bilingual educators have an advantage in assessment literacy? It would seem so, but we do not have any direct evidence. However, Zepeda, Castro, and Cronin (2011) argue convincingly that specialized assessment knowledge is crucial for teachers working in bilingual and dual language programs.

Artiles, Barreto, Pena, and McClafferty's (1998) longitudinal study found that two bilingual teachers' confidence in their teaching practice was related to the complexity of the bilingual programs in which they taught. In a study that calls to mind our earlier discussion of the assessment of bilingual verbal abilities, Ortiz et al. (2011) argue that bilingual education teachers should receive specific professional

development on the assessment used to identify learning disabilities. Their role in the education of bilingual special learners should help other educators understand the “influence of language, culture, and other background characteristics on student performance” (p. 330).

A final example of teacher learning describes a professional development model designed to assist teachers understand more about the language demands of fourth- and fifth-grade mathematics with special attention to formative assessment practices (Thompson, 2008). The program was part of a larger funded project that sought to improve EL teaching practices across a statewide geographic range.¹⁵ Therefore, bringing teachers together for multiple sessions was beyond the resources of the grant. As it turned out, this challenge turned into an opportunity. Ten teachers were invited to video record a mathematics lesson that they believed requires intensive instruction in mathematical language. The project sent a videographer to each teacher's classroom to assist in creating a quality video; the audio was especially important to capture. Each teacher's video was sent to another teacher in the project, one who taught the same grade. In nearly all cases, the teachers did not know each other prior to the project. Teachers were asked to watch their own video and their partner's and to make notes regarding specific academic language and formative assessment strategies they found. The teachers then came together for two face-to-face all-day meetings. The initial meeting was devoted almost entirely to a description of their classroom and school context, the establishment of common goals for the project, and the creation of trust among the group (Frederiksen & White, 1997; Halter 2006; Sherin & Han, 2004). The results of the study indicate that all of the teachers valued the ideas that emerged from teacher discussions; in particular, they suggested that contextualizing their teaching was an important part of the discussions. At the end of the project, they arrived at several key points that make mathematical language more comprehensible to ELs, as well as noting the importance of formative assessment in helping ELs reach new understandings.

The international context provides a few examples of studies on EL teacher assessment knowledge. For instance, Jones-Mackenzie's (2005) study of EL teacher knowledge of assessment among secondary English learners in Jamaica¹⁶ addresses the following questions: (a) What sort of Knowledge About Language (KAL) do teachers report having? (b) Do teachers with KAL apply this information to the testing procedures employed? (c) Which aspects of testing show KAL use? (d) What factors inhibit the use of KAL in the testing process? First, although all the teachers reported some background in applied linguistics, only 2 of the 30 had taken a course specifically about testing. This finding corroborates the general lack of assessment preparation in professional schools. Specifically, Jones-Mackenzie found that KAL and language learning were insufficient: The teachers were aware of the importance of using KAL in testing, but they were not sure how to do so.

In a narrative study of Chinese English language teachers, Xu and Liu (2009) found that the teaching context influences educators' sense of security, which, in turn, influences the effectiveness of their assessment. If teachers are unsure and lack

confidence in their pedagogy, they are less capable of making difficult evaluation decisions. They conclude that teachers' voices must be included in any reform policies, as well as repeating a common refrain of our review: the urgent need for professional development regarding quality assessment practices.

Given the general lack of assessment literacy among teachers of ELs, is teacher education (i.e., professional school preparation) to blame for failing to adequately prepare ELs' teachers? As we mentioned earlier, it makes little sense to expect that professional programs can be wholly responsible for everything we want teachers to know. But with respect to knowledge of EL assessment, we could not locate any research exploring candidate knowledge of EL assessment. Of course, this does not mean that professional school programs are not introducing candidates to the important concepts and practices. Some of the common texts used in EL methods courses address appropriate assessment (Diaz-Rico & Weed, 2009), but the research community has not studied the topic.

If the research on general knowledge of assessment is any guide, we can be confident that candidates lack an understanding of the nuances of EL assessment. Lacking research examples, we have decided to share a few of the practices used in our own program and invite readers to decide whether our methods might be useful.

Our own preservice teacher education program, for example, does not have a stand-alone course on assessment but rather addresses measurement concerns in each of several content-based methods courses. In particular, our program's course in English language development for elementary-level candidates requires assessment knowledge of reading by focusing on the released items from the California English Language Development Test (CELDT). (Note: the newly adopted California English Language Development Standards [California State Department of Education, 2012] will require the CELDT to be rewritten, a task not yet completed during the writing of this chapter.)

The general stance of our elementary EL preparation program is based largely on Halliday's (1969) language functions and tasks. Although not strictly a program or course of study based on Systemic Functional Linguistics, attention is drawn to the particular uses of language, and EL teachers must take care to develop each.

- Instrumental Language: For the work of life, to satisfy needs and wants
- Regulatory Language: For social control
- Interactional Language: For the establishment of social relationships
- Personal Language: To create a "self-text"
- Imaginative Language: To express and fantasize
- Heuristic Language: As a tool for learning about the world
- Informative Language: For the conveyance of information

Each of the functions could be assessed in some form, and we invite candidates to consider how they might evaluate their students' language skills in each function.

This task is open-ended and results in a wide-ranging discussion that often reaches into our candidates' general theories about teaching and learning. In contrast to this discussion, the hard realities of testing EL are presented in the form of released test items from the CELDT. We invite candidates to discuss why the items claim to assess reading, writing, and listening skills, as well as asking them to consider how the items might be biased against EL. Their general reaction is that language tests such as the CELDT are accurate but also tend to trivialize students' language capacities.

In addition, we ask students to learn more about receptive and expressive English usage by having them administer what we call a quick, informal language assessment.¹⁷ Using Carmen Lomas Garza's painting "Tamalada," taken from her book *Cuadros de Familia* (1990), as a tool, candidates work with an individual EL, assessing the student's oral language by asking, for example, "Point to all the people sitting in chairs in the picture." As part of the expressive section, they ask an EL student to point to the stove hood and ask, "What is this? What is it for?" We also invite candidates to administer an oral story retell task (Blank & Frank, 1971), which offers them the opportunity to focus on the meaning of the story that ELs gain from the narrative. Finally, we invite them to present a few verbal analogies representing common forms (synonymic, antonymic, functional, linearly ordered, and categorical membership) to the student, based on their chronological age instead of their language level. Our candidates are routinely surprised by their ELs' capacity to understand complex connections between individual words even when they assess beginning speakers, many of whom are still struggling to create complete, grammatically accurate sentences in English. The general reaction of our candidates is genuine surprise at how much English their ELs do know, as well as a deeper appreciation that teachers must create the conditions in which they can assess ELs' language capacities in individual settings.

If we can discern any common characteristics from the studies and cases described above, it is that teachers must collaborate to make genuine professional growth. In the context of EL teaching, this requirement seems only to grow in importance. In addition, they need to work within their own classroom walls. After reviewing all the literature on effective professional development, we find that the list of effective teacher learning compiled by Darling-Hammond and McLaughlin (1995) nearly 20 years ago remains relevant today. They recommend that effective professional development will adhere to the following principles:

- It must engage teachers in concrete tasks of teaching, *assessment* [italics added], observation, and reflection to enrich the learning and development processes.
- It must be based in inquiry, reflection, and experimentation that are participant-driven.
- It must be collaborative, involving a shared understanding among educators and a focus on teachers' communities of practice rather than on individual teachers.
- It must be connected to and derived from teachers' work with their students.
- It must be sustained, ongoing, intensive, and supported by modeling, coaching, and the collective solving of specific problems of practice.
- It must be connected to other aspects of school change. (p. 598)

The first casualty of this list is the so-called expert consultant who flies many miles to the school or district (after all, who can be a prophet in his own land?), deposits “expertise,” and leaves. The big check is written and off goes our expert. No one is accountable—not the expert, not the administration. The only winner—beside the consultant’s bank account—in this routine charade is the school leader, who can report that a professional development opportunity was presented to teachers. The note to the district office justifying the expense reads: “I had the expert tell the teachers how they should teach. If they don’t do it, it’s not my fault.” Sometimes the consultant comes with a canned curriculum, also for sale, which promises to raise student achievement. Claims that the curriculum is research-based are never questioned, but purchases are made regardless. It is fascinating—and a bit disappointing—to find that this expert model survives in spite of decades of evidence documenting its failure (e.g., Datta, 1981).

Contrast this dismal scenario with a teacher learning group who *reads* what theorists and researchers have to say about language learning and assessment. Expert knowledge is not ignored or disparaged but rather filtered through the teachers’ everyday experiences in their school, with their students, by reading and discussing what they have written. Teachers are compensated for their extra work; they are allowed to bill the district for the time they spend reading and learning. They test new ideas, new practices, in an environment that allows for experimentation and credible evaluation. The school system provides them with good data to make sound decisions, and given the right knowledge about the effectiveness of a program or strategy, teachers make the rational and pragmatic choice. Too many school leaders and politicians ignore the fact that teachers are nothing if not practical. If they have evidence that an instructional strategy does not help their students, they move on to what does. And if their trusted colleagues are those providing the evidence, *a fortiori*.

In this model, curriculum is built from the ground up, by teachers working together. It is difficult work and very time-consuming and sometimes more expensive than the consultant and the canned curriculum, but it has invested resources in enhancing teacher knowledge, which results in more efficient practices, which in turn raises student achievement.

We argue that if a school or school system is interested in enhancing the assessment knowledge of teachers of ELs, the principles outlined by Darling-Hammond and McLaughlin (1995) are even more relevant, with one important modification. When considering assessment practices, many of which are technical in nature (e.g., computing difficulty values), it makes sense to develop the expertise of a teacher (or teachers) who is (are) willing and able to understand these basic concepts and make them comprehensible to other educators at the school. The local teacher becomes the expert.

The most important point for school leaders to remember is that developing teacher knowledge takes time. The temptation to hire the expert, implement the off-the-shelf curriculum—to do it fast—must be restrained. Instead, schools leaders

must focus on providing resources to schools emphasizing professional development that support teachers' instructional capacity and assessment literacy that informs instructional improvement; all the while focusing on how such teacher supports can be sustained over time.

The foregoing discussion has emphasized the importance of the school context in the proper assessment of ELs. Given the existing literature's focus on "curricular validity" and inequities regarding opportunities to learn, Abedi and Herman (2010), citing Herman, Klein, and Abedi, (2000) remind us of the potential barrier to EL student academic success resulting from the "lack of effective opportunity to learn (OTL)—students' access to and engagement in the academic content they need to perform well on tests and achieve standards" (p. 726).

For ELs, course-taking *opportunities* are often determined by linguistic measures, such as the degree of English language proficiency of students. Research has shown that ELs are likely to be placed in low-level courses according to their English proficiency (Callahan, 2005; Mosqueda, 2010; Mosqueda & Maldonado, 2013). Such placement practices are often made without an accurate assessment of students' English proficiency (Duran, 2008; Martiniello, 2008; Ruiz-de-Velasco & Fix, 2000). As a consequence, students' English proficiency can severely limit access to rigorous content courses, which might explain why they are disproportionately underenrolled in college preparatory coursework (Callahan, 2005). Furthermore, even when ELs are reclassified as English language proficient, they may still be denied access to rigorous courses (Valenzuela, 2010).

We end this section with an unfortunate contradiction. As Pandya (2011) found, educators seem to be particularly interested in testing their ELs but often rarely understand the results well enough to make good decisions based on the results. This misunderstanding often leads to additional and superfluous testing. To the extent that testing reduces instructional time, the burden of testing contributes to students' falling further behind in their content area opportunities to learn and their opportunities to develop English skills. If we continue to overtest students, we compromise opportunity to learn, particularly when such tests are not used to inform instruction.

CONTEMPORARY AND FUTURE CONCERNS REGARDING THE PREPARATION OF TEACHERS FOR THE SOUND ASSESSMENT OF EL

One does not learn (or teach) language in a way that resembles instruction in other aspects of human knowledge. Chomsky (1986), for example, was clear enough on this point. We can try to control it, atomize it, objectify it, but it will do what it wants, and always in relation to what its speaker intends. As we consider where assessment of ELs is leading, we find that Bunch's (2013) newly developed theoretical model of EL teacher knowledge has important applications to knowledge of assessment. Building on Shulman's (1987) well-known taxonomy of teacher knowledge, Bunch (2013) argues that

what teachers need is pedagogical *language* knowledge that must be conceived of differently from either the pedagogical content knowledge about language needed by teachers specializing in second language teaching or the pedagogical content knowledge mainstream teachers need in the core subject matters. (p. 304)

Although these comments are directed at mainstream content teachers who teach ELs, we believe that his claim applies to all teachers of ELs, with particular implications for assessment. In his view, language is more than an object, more than a “tool” for learning about content. Bunch (2013) suggests that teachers must consider language learning as linked to students’ ever-enlarging and shifting content knowledge. He maintains that teachers’ knowledge of assessment must include the dimensions of time and context: a recognition that a student’s growth in language will have inflection points and apexes, depending on his or her knowledge of the content. That is, what you know about the world depends on how much English you know, and vice versa. Thus, teachers of ELs must reconsider ways of assessing language and content as though the two were indistinguishable.

This view of assessment is, coincidentally, and perhaps by design, the cornerstone of the Common Core State Standards (CCSS) and the concomitant effort of the Smarter Balanced Assessment Consortium (SBAC; see <http://www.smarterbalanced.org/>).¹⁸ The CCSS emphasizes “language use for communication and learning” in the content areas such as science and mathematics classrooms (Lee, Quinn, & Valdés, 2013, p. 223). Consequently, in the United States, we are entering a new era in the evaluation of K–12 learners. As schools and school systems begin to square their curriculum to the CCSS as well as addressing the daunting technical aspects of the SBAC plan, which will require all students to take the examinations in an online format, researchers and policymakers are considering how SBAC will be modified or otherwise adapted for ELs. A specific concern is whether low-income students from nondominant backgrounds will continue to be underserved in school. The CCSS’s website home page includes a disclaimer that reads as follows:

The Standards set grade-specific standards but do not define the intervention methods or materials necessary to support students who are well below or well above grade-level expectations. It is also beyond the scope of the Standards to define the full range of supports appropriate for English language learners and for students with special needs. At the same time, all students must have the opportunity to learn and meet the same high standards if they are to access the knowledge and skills necessary in their post-school lives. (<http://www.schoolimprovement.com/what-is-not-covered-by-the-standards/>)

This acknowledged lack of attention to instructional and curricular supports for students who traditionally underachieve in schools can lead to a new set of national standards that merely replicates the long-standing disparities in academic achievement, unless strategies are put in place to help improve the educational outcomes of low-income students of color, particularly ELs.

The CCSS (2010) supports states that develop and implement common, high-quality assessments, including formative assessments that help inform classroom

practice. At a minimum, such high-quality assessments are expected to measure the subject areas of reading/language arts and mathematics and provide information for each student annually in Grades 3 through 8.

Changes in assessment requirements under CCSS will offer useful information to middle school teachers and students, particularly ELs. Although policymakers will still use assessment results as summative indicators of learning, practitioners will also incorporate formative assessments of students' learning. Considering that such formative national assessments will include constructed-response items and measure speaking and listening skills, how such assessments affect the teaching and learning of ELs is an important educational policy and practice issue to follow.

With respect to teachers' knowledge of CCSS and ELs, though not specifically addressing the assessment of ELs in a CCSS world, Santos, Darling-Hammond, and Cheuk (2012) outline the shift in EL teacher knowledge by suggesting four primary points that will guide EL instruction under CCSS: (a) language progressions: how students learn language, both in terms of general language acquisition and in terms of the acquisition of discipline-specific academic language, (b) language demands: what kinds of linguistic expectations are embedded within specific texts and tasks with which students are being asked to engage, (c) language scaffolds: how specific representations and instructional strategies can be used to help students gain access to the concepts as well as to the language they need to learn, and (d) language supports: how classrooms and schools can be organized to support students in continually building a deep understanding of language and content.

Although they do not specifically mention a renewed knowledge of assessment strategies, such knowledge is embedded in these changes from previous goals for EL. In particular, if teachers of EL are to know deeply the language demands of academic tasks, they will have to be able to assess with good accuracy the *level* of language demands being placed on ELs (Celedón-Pattichis & Musanti, 2013; Lee et al., 2013). This is a question of sound assessment, of both the content (e.g., readability) and the learners (i.e., Are my students ready for the language demands they'll need to understand the lesson?).

CONCLUDING THOUGHTS

The linguistic turn in contemporary philosophy, which resulted in the discipline we now call poststructuralism, has emphasized that language itself, not our direct experience in the world and certainly not some externally wrought truth discovered independent of experience, is what structures our consciousness (Sarup, 1993). We are our language. It is the distinctiveness of our linguistic histories that makes us unique. And the linguistic histories of our ELs are far richer than most educators imagine.

We have tried to honor this complexity and nuance in our chapter while also recognizing the practical steps we can take to ensure that teachers are prepared and willing to assess their ELs expertly. Pedagogy cannot be made routine, and like the Salvatori (2003) quotation in the beginning of our chapter, we should be proud of the complicated messiness of our discipline. But the complexity often drives us to find

prescriptions instead of compelling questions. Even the most comprehensive so-called solutions for addressing the proper assessment of ELs will fail when applied to individual students, who represent a linguistic and cultural profile that defies categorization.

The best educational assessments are mirrors, reflections of the larger achievement patterns found as a consequence of our priorities. It therefore makes little sense to expect that we can tinker with test content and design and eliminate score differences between ELs and their non-EL counterparts. After all, most ELs attend underfunded schools, cope with abject poverty (in spite of parents and other family members who work several jobs), and sometimes face harsh racism inside and outside the school. We should not expect any test to erase the effects of an inequitable and perhaps unjust society. We establish our priorities, and the measures of academic achievement reflect them. But we can and should be doing much more to help teachers, candidates, and other educators to know more about what makes for reliable and valid assessment for ELs. Assessments quickly turn into evaluations and therefore must be fair to all. This is the ideal we seek.

NOTES

¹We use *teachers* to refer to in-service teachers, *candidates* to refer to preservice teachers earning the teaching license or certificate, and *students* for general elementary- and secondary-age learners.

²The happy advantage of the limited research is that we can devote significant attention to those papers that do address the topic.

³Our task was to complete a review of the available literature. Whereas less structured than a formal meta-analysis of the literature, we nevertheless held certain criteria for inclusion. In general, we tried to limit our review to peer-reviewed papers but we also included dissertations that we deemed to be of high quality. Search terms such as “bilingual” and “assessment” yielded works for the foundation of our review. Other terms used included “assessment literacy,” which led us to the work we review in the section “The State-of-the-Art.” In general, we relied on all the contemporary search techniques available and used indexing tools such as Google Scholar, which allows researchers to find similar articles, citations of key works, and works cited. We had access to all publications available in the University of California system, which holds one of the largest collections in the world

⁴Here we might recall Krashen’s (1985) “i + 1” as a parallel concept to instructional level.

⁵By social capital, we mean the access to support via networks of institutional agents (school personnel) within schools that can potentially be activated by students to advocate on their behalf for more appropriate instructional learning opportunities (Stanton-Salazar, 2001).

⁶Crystal City is the small town in south Texas where, in 1969, Mexican American high school students staged a walkout in protest of harsh discrimination at the hands of an all-White school administration. After decades of discrimination and marginalization, Mexican Americans in Crystal City organized and were elected to the school board and other important political positions in the surrounding county. They soon implemented high school courses on the history of Mexican Americans, the first such course in the United States. They also inaugurated a maintenance bilingual education program, also a first. Scholars such as Angela Valenzuela (2000) have argued that the heroic efforts by the students and community in Crystal City helped to create civil rights era language education programs and ushered in the federal program known as the Office of Bilingual Education and Minority Language Affairs.

Sadly, these reforms have not lasted: During the George W. Bush administration (approximately 2002) the Office of Bilingual Education and Minority Language Affairs had its name changed to the Office of English Language Acquisition and, more or less, stopped funding bilingual education. The Obama administration has not changed this course. And our readers are no doubt aware of state legislation curtailing or outlawing native language instruction in places such as California and Arizona.

⁷Volumes by San Miguel (2004) and others (e.g., Cline, Necochea, & Rios, 2004) have documented the political struggles to retain quality language education in the United States. We recognize the importance of this topic but also know that a treatment of the political debates on language education would take many, many pages. For our purposes, ELs require sound language assessment across languages, regardless of the program in which they participate. Our view on the role of teachers in language policy is reflected in a recent article (Téllez & Varghese, 2013).

⁸Our preferred term would likely be *emerging bilinguals* or *multilinguals*, but we have decided to follow convention in this chapter and use *English Learner*.

⁹But if ELs were excluded from the testing regimes, would we not be concerned that schools would focus their attention on non-ELs, allowing ELs to languish because they were not part of the accountability scheme? If alternative tests were developed for ELs that reduced the language load and, subsequently, the cognitive and content complexity, would we not be concerned that such assessments would encourage lower expectations for ELs. On this point, we should not deceive ourselves: A test that “accommodates” a learner’s developing knowledge of English will almost certainly diminish the complexity of the content to be measured. Our discipline’s recent attention to the development of academic language is a clear admission of this fact.

¹⁰See <http://www.ctc.ca.gov/educator-prep/standards/adopted-TPEs-2013.pdf>, for the text of the California Teaching Expectations.

¹¹Initially, in 2001, the Consortium consisted of 12 professional teacher preparation programs in eight University of California campuses, two California State Universities, Stanford University, and Mills College.

¹²The other domains are (a) Language; (b) Culture; (c) Planning, Implementing, and Managing Instruction; and (d) Professionalism.

¹³Although professional schools and school systems have been slow to develop specialized programs for teachers of EL, we recognize the long history of linguistics (often directed at adult “English as a second language” teachers) and foreign language programs devoted to teacher development.

¹⁴We recognize the overlap with the discussion of standards in the previous section, but what research recommends and what makes it into standards are not always the same.

¹⁵Both authors worked to develop this professional development experience.

¹⁶Jones-Mackenzie (2005) points out that the majority of the population in Jamaica, though often considered to be an English-speaking nation, speaks a creole as the first language. The schools teach English as new language.

¹⁷In advance of the assignment, we are very careful to tell candidates that the assessment is a class requirement only and that students’ names will not be used in the discussion. Furthermore, we insist that they point out that this is an informal assessment designed only to help them learn more about oral language assessment. Even with these qualifications in place, a few of the schools call us concerned about how the results of the “tests” will be used. Such is the heightened scrutiny borne of harsh, punitive accountability schemes (i.e., NCLB).

¹⁸The SBAC is one of two efforts working to create assessments for the CCSS. The other is known as the Partnership for the Assessment of Readiness for College and Careers. We address only the SBAC strategy for several reasons. One is the fact that SBAC will be using a computer adaptive strategy, rather than the effort of the Partnership for the Assessment of Readiness for

College and Careers, which will be computer administered but not adaptive. Understanding computer adaptive testing is a more difficult concept for educators to grasp. Second, our own state of California, which has the largest number of ELs, has selected SBAC as its governing consortia. Texas, the state with the second most ELs, is not yet participating in CCSS.

REFERENCES

- Abedi, J. (2004). The No Child Left Behind Act and English language learners: Assessment and accountability issues. *Educational Researcher, 33*(1), 4–14.
- Abedi, J., & Herman, J. (2010). Assessing English language learners' opportunity to learn mathematics: Issues and limitations. *Teachers College Record, 112*, 723–746.
- Abedi, J., & Lord, C. (2001). The language factor in mathematics tests. *Applied Measurement in Education, 14*, 219–234.
- Alexander, D., Heavside, S., Farris, E., & Burns, S. (1998). *Status of education reform in public elementary and secondary schools: Teachers' perspectives* (NCES No. 1999-045). Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.
- Anderson, L. W. (2000). Time, learning, and school reform: A conceptual framework. In P. C. Gandara. (Ed.), *The dimensions of time and the challenge of school reform* (pp. 13–29). Albany: SUNY Press.
- Angrist, J. D., Dynarski, S. M., Kane, T. J., Pathak, P. A., & Walters, C. R. (2010). Inputs and impacts in charter schools: KIPP Lynn. *American Economic Review, 100*, 239–243.
- Artiles, A. J., Barreto, R. M., Pena, L., & McClafferty, K. (1998). Pathways to teacher learning in multicultural contexts: A longitudinal case study of two novice bilingual teachers in urban schools. *Remedial and Special Education, 19*, 70–90.
- Artiles, A. J., & Klingner, J. K. (2006). Forging a knowledge base on English language learners with special needs: Theoretical, population, and technical issues. *Teachers College Record, 108*, 2187–2194.
- Artiles, A. J., & Ortiz, A. (Eds.). (2002). *English language learners with special needs: Identification, placement, and instruction*. Washington, DC: Center for Applied Linguistics.
- Artiles, A. J., Rueda, R., Salazar, J. J., & Higuera, I. (2005). Within-group diversity in minority disproportionate representation: English language learners in urban school districts. *Exceptional Children, 71*, 283–300.
- Ayala, C. C., Shavelson, R. J., Ruiz-Primo, M. A., Brandon, P. R., Yin, Y., Furtak, E. M., & Young, D. B. (2008). From formal embedded assessments to reflective lessons: The development of formative assessment studies. *Applied Measurement in Education, 21*, 315–334.
- Bachman, L. F. (2000). Modern language testing at the turn of the century: Assuring that what we count counts. *Language Testing, 17*, 1–42.
- Bartolome, L. I. (2004). Critical pedagogy and teacher education. *Teacher Education Quarterly, 31*, 97–122.
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). *Assessment for learning: Putting it into practice*. New York, NY: McGraw-Hill.
- Blank, M., & Frank, S. M. (1971). Story recall in kindergarten children: Effect of method of presentation on psycholinguistic performance. *Child Development, 42*, 299–312.
- Breedlove, L. (2007). Identifying linguistically diverse students as gifted and talented: A qualitative study of adding a new measure (Unpublished doctoral dissertation). Texas A&M University, College Station.

- Brooks, D. (2012, March 22). The relationship school. *New York Times*. Retrieved from http://www.nytimes.com/2012/03/23/opinion/brooks-the-relationship-school.html?_r=0
- Bunch, G. C. (2013). Pedagogical language knowledge: Preparing mainstream teachers for English learners in the new standards era. *Review of Research in Education, 37*, 298–341. doi:10.3102/0091732X12461772
- Bunch, G. C., Abram, P. L., Lotan, R. A., & Valdes, G. (2001). Beyond sheltered instruction: Rethinking conditions for academic language development. *TESOL Journal, 10*(2), 2–3.
- Bunch, G. C., Aguirre, J. M., & Téllez, K. (2009). Beyond the scores: Using candidate responses on a high stakes performance assessment to inform teacher preparation for English learners. *Issues in Teacher Education, 18*(1), 26.
- California Commission on Teacher Credentialing. (2013). *California teaching performance expectations*. Retrieved from <http://www.ctc.ca.gov/educator-prep/TPA-files/TPEs-Full-Version.pdf>
- California State Department of Education. (2012). *English Language Development Standards*. Retrieved from <http://www.cde.ca.gov/sp/el/er/eldstandards.asp>
- Callahan, R. M. (2005). Tracking and high school English learners: Limiting opportunity to learn. *American Educational Research Journal, 42*, 305–328.
- Campbell, C., & Evans, J. A. (2000). Investigation of preservice teachers' classroom assessment practices during student teaching. *Journal of Educational Research, 93*, 350–355.
- Celedón-Pattichis, S. (2004). Rethinking policies and procedures for placing English language learners in mathematics. *NABE Journal of Research and Practice, 2*, 176–192.
- Celedón-Pattichis, S., & Musanti, S. (2013). "Let's suppose that . . .": Developing base-ten thinking. In M. Gottlieb & G. Ernst-Slavit (Eds.), *Academic language demands for language learners: From text to context* (pp. 87–128). Thousand Oaks, CA: Corwin Press.
- Chomsky, N. (1986). *Knowledge of language: Its nature, origin, and use*. New York, NY: Praeger.
- Cizek, G. J. (2000). Pockets of resistance in the assessment revolution. *Educational Measurement: Issues and Practices, 19*(2), 16–23.
- Cizek, G. J. (2010). An introduction to formative assessment: History, characteristics, and challenges. In H. Andrade & G. Cizek (Eds.), *Handbook of formative assessment* (pp. 3–18). New York, NY: Routledge.
- Cizek, G. J., Fitzgerald, S. M., & Rachor, R. A. (1995). Teachers' assessment practices: Preparation, isolation, and the kitchen sink. *Educational Assessment, 3*, 159–179.
- Cline, Z., Necochea, J., & Rios, F. (2004). The tyranny of democracy: Deconstructing the passage of racist propositions. *Journal of Latinos and Education, 3*(2), 67–85.
- Cochran-Smith, M. (2001). The outcomes question in teacher education. *Teaching and Teacher Education, 17*, 527–546.
- Common Core State Standards Initiative. (2010). *Common Core State Standards for English language arts & literacy in history/social studies, science, and technical subjects*. Washington, DC: CCSSO & National Governors Association.
- Cummins, J. (1982). Tests, achievement, and bilingual students. *Focus, 9*, 1–8.
- Daniel, L. G., & King, D. A. (1998). Knowledge and use of testing and measurement literacy of elementary and secondary teachers. *Journal of Educational Research, 91*, 331–344.
- Darling-Hammond, L. (1999). *Reshaping teaching policy, preparation, and practice. Influences of the National Board for Professional Teaching Standards*. Washington, DC: American Association of Colleges for Teacher Education.
- Darling-Hammond, L. (2006). Constructing 21st-century teacher education. *Journal of Teacher Education, 57*, 1–15.
- Darling-Hammond, L., & McLaughlin, M. W. (1995). Policies that support professional development in an era of reform. *Phi Delta Kappan, 76*, 597–604.
- Datta, L.-E. (1981). Damn the experts and full speed ahead: An examination of the study of Federal programs supporting educational change as evidence against directed development and for local problem-solving. *Evaluation Review, 5*, 5–32.

- Díaz-Rico, L. T., & Weed, K. Z. (2009). *The crosscultural, language, and academic development handbook: A complete K-12 reference guide*. Boston, MA: Allyn & Bacon.
- Duran, R. P. (2008). Assessing English-Language Learners' achievement. *Review of Research in Education, 32*(1), 292–327.
- Faltis, C. (1990). Spanish for native speakers: Freirian and Vygotskian perspectives. *Foreign Language Annals, 23*, 117–125.
- Flores, A. (2007). Examining disparities in mathematics education: Achievement gap or opportunity gap? *High School Journal, 91*, 29–42.
- Frederiksen, J. R., & White, B. J. (1997). *Reflective assessment of students' research within an inquiry-based middle school science curriculum*. Paper presented at the annual meeting of the AERA, Chicago, IL.
- Freeman, D. (1991). To make the tacit explicit: Teacher education, emerging discourse, and conceptions of teaching. *Teaching and Teacher Education, 7*, 439–454.
- Freeman, D., & Johnson, K. E. (1998). Reconceptualizing the knowledge base of language teacher education. *TESOL Quarterly, 32*, 397–417.
- Gándara, P. C. (2000). *The dimensions of time and the challenge of school reform*. Albany: SUNY Press.
- Gándara, P. C., & Maxwell-Jolly, J. (2006). Critical issues in developing a teacher corps for English learners. In K. Téllez & H. C. Waxman (Eds.). *Preparing quality educators for English language learners: Research, policies, and practices* (pp. 99–120). Mahwah, NJ: Erlbaum.
- Gándara, P., & Rumberger, R. (2009). Immigration, language, and education: How does language policy structure opportunity? *Teachers College Record, 111*, 750–782.
- Gándara, P., Rumberger, R. W., Maxwell-Jolly, J., & Callahan, R. (2003). English language learners in California schools: Unequal resources, unequal chances. *Education Policy Analysis Archives, 11*(36). Retrieved from http://www.usc.edu/dept/education/CMMR/FullText/ELLs_in_California_Schools.pdf
- García, E. (1990). Educating teachers for language minority students. In W. R. Houston (Ed.), *Handbook of research on teacher education* (pp. 717–729). New York, NY: Macmillan.
- García, O., Kleifgen, J. A., & Falchi, L. (2008). *From English language learners to emergent bilinguals*. New York, NY: Teachers College.
- Garza, C. L. (1990). *Family pictures/Cuadros de familia*. San Francisco, CA: Children's Book Press.
- Gould, S. J. (1996). *The mismeasure of man* (Rev. ed.). New York, NY: W. W. Norton.
- Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain & Language, 36*, 3–15.
- Haag, N., Heppt, B., Stanat, P., Kuhl, P., & Pant, H. A. (2013). Second language learners' performance in mathematics: Disentangling the effects of academic language features. *Learning and Instruction, 28*, 24–34.
- Halliday, M. A. (1969). Relevant models of language. *Educational Review, 22*(1), 26–37.
- Halter, C. P. (2006). *The reflective lens: The effects of video analysis on preservice teacher development* (Unpublished doctoral dissertation). University of California at San Diego.
- Harper, C. A., de Jong, E. J., & Platt, E. J. (2008). Marginalizing English as a second language teacher expertise: The exclusionary consequence of No Child Left Behind. *Language Policy, 7*, 267–284.
- Harris, D. P. (1969). *Testing English as a second language*. New York, NY: McGraw-Hill.
- Heritage, M. (2007). What do teachers need to know and do? *Phi Delta Kappan, 89*, 140–145.
- Heritage, M. (2010). *Formative assessment and next-generation assessment systems: Are we losing an opportunity?* Washington, DC: Council of Chief State School Officers.
- Heritage, M. K., Kim, J., Vendlinski, T. P., & Herman, J. L. (2009). From evidence to action: A seamless process in formative assessment? *Educational Measurement: Issues and Practice, 28*(3), 24–31.

- Heubert, J. P., & Hauser, R. M. (Eds.). (1999). *High stakes: Testing for tracking, promotion, and graduation*. Washington, DC: National Academies Press.
- Hilliard, A. G. (2000). Excellence in education versus high-stakes standardized testing. *Journal of Teacher Education, 51*, 293–304.
- Holmes Group. (1986). *Tomorrow's teachers: A report of the Holmes Group*. New York, NY: Author.
- Hu, G. (2003). English language teaching in China: Regional differences and contributing factors. *Journal of Multilingual and Multicultural Development, 24*, 290–318.
- Johnston, B. (2002). *Values in English language teaching*. Mahwah, NJ: Erlbaum.
- Jones-Mackenzie, C. (2005). Knowledge about language and testing. In N. Bartels (Ed.), *Applied linguistics and language teacher education* (Vol. 4, pp. 313–324). New York, NY: Springer.
- Klingner, J. K., & Solano-Flores, G. (2007). Cultural responsiveness in response-to-intervention models. In C. C. Laitusis & L. L. Cook (Eds.), *Large-scale assessment and accommodations: What works?* (pp. 229–241). Arlington, VA: Council for Exceptional Children.
- Knight, S. L., & Wiseman, D. L. (2006). Lessons learned from a research synthesis on the effects of teachers' professional development on culturally diverse students. In K. Téllez & H. C. Waxman (Eds.), *Preparing quality educators for English language learners: Research, policy, and practice* (pp. 71–98). Mahwah, NJ: Erlbaum.
- Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. London, England: Longman.
- Kumaravivelu, B. (2005). *Understanding language teaching: From method to postmethod*. New York, NY: Routledge.
- Laplante, B. (2000). Learning science is learning to “speak science”: Students in immersion classes talk to us about chemical reactions. *Canadian Modern Language Review-Revue Canadienne Des Langues Vivantes, 57*, 245–271.
- Lee, O., Quinn, H., & Valdés, G. (2013). Science and language for English language learners in relation to next generation Science Standards and with implications for Common Core State Standards for English Language Arts and mathematics. *Educational Researcher, 42*, 223–233.
- Llosa, L. (2011). Standards-based classroom assessments of English proficiency: A review of issues, current developments, and future directions for research. *Language Testing, 28*, 367–382.
- Lucas, T., & Villegas, A. M. (2010). The missing piece in teacher education: The preparation of linguistically responsive teachers. *National Society for the Study of Education, 109*, 297–318.
- Maclellan, E. (2004). Initial knowledge states about assessment: Novice teachers' conceptualizations. *Teaching and Teacher Education, 20*, 523–535.
- Mahoney, K. S., & MacSwan, J. (2005). Reexamining identification and reclassification of English language learners: A critical discussion of select state practices. *Bilingual Research Journal, 29*, 31–42.
- Martiniello, M. (2008). Language and the performance of English-language learners in math word problems. *Harvard Educational Review, 78*, 333–368.
- Martiniello, M. (2009). Linguistic complexity, schematic representations, and differential item functioning for English language learners in math tests. *Educational Assessment, 14*, 160–179.
- McCarty, T. L. (2009). The impact of high-stakes accountability policies on Native American learners: Evidence from research. *Teaching Education, 20*, 7–29.
- McMillan, J. H. (2000). Fundamental assessment principles for teachers and school administrators. *Practical Assessment, Research & Evaluation, 7*(8). Retrieved from <http://pareonline.net/getvn.asp?v=7&n=8>

- McMillan, J. H. (2003). Understanding and improving teachers' classroom assessment decision making: Implications for theory and practice. *Educational Measurement: Issues and Practice, 22*(4), 34–43.
- Menken, K. (2008). *English learners left behind: Standardized testing as language policy* (Vol. 65). Bristol, England: Multilingual matters.
- Mertler, C. A. (2009). Teachers' assessment knowledge and their perceptions of the impact of classroom assessment professional development. *Improving Schools, 12*, 101–113.
- Mertler, C. A., & Campbell, C. (2005, April). *Measuring teachers' knowledge and application of classroom assessment concepts: Development of the Assessment Literacy Inventory*. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Quebec, Canada.
- Minicucci, C. (2000). Effective use of time in the education of English language learners. In P. C. Gandara. (Ed.). *The dimensions of time and the challenge of school reform* (pp. 49–58). Albany: SUNY Press.
- Mosqueda, E. (2010). Compounding inequalities: English proficiency and tracking and their relation to mathematics performance among Latina/o secondary school youth. *Journal of Urban Mathematics Education, 3*, 57–81.
- Mosqueda, E., & Maldonado, S. I. (2013). The effects of English language proficiency and curricular pathways: Latina/os' mathematics achievement in secondary schools. *Equity & Excellence in Education, 46*, 202–219.
- Muñoz-Sandoval, A. F., Cummins, J., Alvarado, C. G., & Ruef, M. L. (1998). *Bilingual verbal ability tests* (Comprehensive manual). Itasca, IL: Riverside.
- National Board for Professional Teaching Standards. (2012a). *Early and middle childhood/English as a new language: Portfolio instructions*. Retrieved from <http://www.nbpts.org/art-emc>
- National Board for Professional Teaching Standards. (2012b). *Early adolescence through young adulthood/English as a new language: Portfolio instructions*. Retrieved from <http://www.nbpts.org/art-eaya>
- Newman, M., & Hanauer, D. (2005). The NCATE/TESOL teacher education standards: A critical review. *TESOL Quarterly, 39*, 753–764.
- Ortiz, A. A., Robertson, P. M., Wilkinson, C. Y., Liu, Y.-J., McGhee, B. D., & Kushner, M. I. (2011). The role of bilingual education teachers in preventing inappropriate referrals of ELLs to special education: Implications for response to intervention. *Bilingual Research Journal, 34*, 316–333.
- Páez, M. M. (2008). English language proficiency and bilingual verbal ability among Chinese, Dominican, and Haitian immigrant students. *Equity & Excellence in Education, 41*, 311–324.
- Pandya, J. Z. (2011). *Overtested: How high-stakes accountability fails English language learners*. New York, NY: Teachers College Press.
- Pennycook, A. (1999). Introduction: Critical approaches to TESOL. *TESOL Quarterly, 33*, 329–348.
- Performance Assessment for California Teachers. (n.d.). *Frequently asked questions*. Retrieved from http://www.pactpa.org/_main/hub.php?pageName=FAQ
- Plake, B. S., Impara, J. C., & Fager, J. J. (1993). Assessment competencies of teachers: A national survey. *Educational Measurement: Issues and Practice, 12*(4), 10–12.
- Popham, W. J. (2003). Seeking redemption for our psychometric sins. *Educational Measurement: Issues and Practice, 22*(1), 45–48.
- Popham, W. J. (2006). Needed: A dose of assessment literacy. *Educational Leadership, 63*(6), 84–85.
- Popham, W. J. (2011). Assessment literacy overlooked: A teacher educator's confession. *The Teacher Educator, 46*, 265–273.

- Porter, J. M. (2010). *Performance Assessment for California Teachers (PACT): An evaluation of inter-rater reliability* (Unpublished doctoral dissertation). University of California, Davis.
- Randall, J., & Engelhard, G. (2010). Examining the grading practices of teachers. *Teaching and Teacher Education*, 26, 1372–1380.
- Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, 73, 417–458.
- Rodriguez, D. (2009). Meeting the needs of English language learners with disabilities in urban settings. *Urban Education*, 44, 452–464.
- Roth, R. A. (1996). Standards for certification, licensure, and accreditation. In J. Sikula (Ed.), *Handbook of research on teacher education* (pp. 242–278). New York, NY: Macmillan.
- Ruiz-de-Velasco, J., & Fix, M. (2000). *Overlooked and underserved: Immigrant students in U.S. secondary schools*. Washington, DC: Urban Institute.
- Ruiz-Primo, M. A., Furtak, E. M., Ayala, C., Yin, Y., & Shavelson, R. J. (2010). Formative assessment, motivation, and science learning. In H. Andrade & G. Cizek (Eds.), *Handbook of formative assessment* (pp. 139–158). New York, NY: Routledge.
- Rumberger, R., & Gándara, P. (2004). Seeking equity in the education of California's English learners. *Teachers College Record*, 106, 2031–2055.
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119–144.
- Salvatori, M. R. (2003). *Pedagogy: Disturbing history, 1819-1929*. Pittsburgh, PA: University of Pittsburgh Press.
- Sandholtz, J. H., & Shea, L. M. (2012). Predicting performance: A comparison of university supervisors' predictions and teacher candidates' scores on a Teaching Performance Assessment. *Journal of Teacher Education*, 63, 39–50.
- San Miguel, G. (2004). *Contested policy: The rise and fall of federal bilingual education in the United States, 1960-2001*. Denton: University of North Texas Press.
- Santos, M., Darling-Hammond, L., & Cheuk, T. (2012). *Teacher development to support English language learners in the context of common core state standards*. Palo Alto, CA: Stanford University.
- Sarup, M. (1993). *An introductory guide to post-structuralism and postmodernism*. Athens: University of Georgia Press.
- Scriven, M. (1991). Beyond formative and summative evaluation. In M. W. McLaughlin & D. C. Phillips (Eds.), *Evaluation and education: At quarter century, 90th yearbook of the National Society for the Study of Education* (pp. 18–64). Chicago, IL: University of Chicago Press.
- Shaw, J. M., Bunch, G. C., & Geaney, E. R. (2010). Analyzing language demands facing English learners on science performance assessments: The SALD framework. *Journal of Research in Science Teaching*, 47, 909–928.
- Shepard, L. A. (2006). Classroom assessment. In R. Brennan (Ed.), *Educational Measurement* (4th ed., pp. 624–646). Westport, CT: Praeger.
- Sherin, M. G., & Han, S. Y. (2004). Teacher learning in the context of a video club. *Teaching and Teacher Education*, 20, 163–183.
- Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57, 1–22.
- Solano-Flores, G. (2006). Language, dialect, and register: Sociolinguistics and the estimation of measurement error in the testing of English language learners. *The Teachers College Record*, 108, 2354–2379.
- Solano-Flores, G. (2011). Language issues in mathematics and the assessment of English language learners. In K. Téllez, J. Moschkovich, & M. Civil (Eds.), *Latinas and mathematics education: Research on learning and teaching in classrooms and communities* (pp. 283–314). Charlotte, NC: Information Age.

- Solano-Flores, G., & Trumbull, E. (2003). Examining language in context: The need for new research and practice paradigms in the testing of English-language learners. *Educational Researcher, 32*(2), 3–13.
- Solano-Flores, G., Trumbull, E., & Nelson-Barber, S. (2002). Concurrent development of dual language assessments: An alternative to translating tests for linguistic minorities. *International Journal of Testing, 2*, 107–129.
- Solórzano, R. W. (2008). High stakes testing: Issues, implications, and remedies for English language learners. *Review of Educational Research, 78*, 260–329.
- Stanton-Salazar, R. D. (2001). Empowering relations of support between students and school personnel. In *Manufacturing hope and despair: The school and kin support networks of U.S.-Mexican youth*. New York, NY: Teachers College Press.
- Stiggins, R. J. (1991). Assessment Literacy. *Phi Delta Kappan, 72*, 534–539.
- Stiggins, R. J. (2001). The unfulfilled promise of classroom assessment. *Educational Measurement: Issues and Practice, 20*(3), 5–15.
- Tang, G. (1992). The effect of graphic representation of knowledge structures on ESL reading comprehension. *Studies in Second Language Acquisition, 14*, 177–195.
- Taubman, P. M. (2009). *Teaching by numbers: Deconstructing the discourse of standards and accountability in education*. New York, NY: Routledge.
- Téllez, K. (1998). Class placement of elementary school emerging bilingual students. *Bilingual Research Journal, 22*, 279–295.
- Téllez, K. (2003). Three themes on standards in teacher education: Legislative expediency, the role of external review, and test bias in the assessment of pedagogical knowledge. *Teacher Education Quarterly, 30*, 9–18.
- Téllez, K. (2010). *Teaching English language learners: Fostering language and the democratic experience*. Boulder, CO: Paradigm.
- Téllez, K. (in press). An analysis of the structure and assessment of standards for teacher candidates and programs. In H. C. Waxman, B. L. Alford, D. Brown, & K. Rollins (Eds.), *Preparing teachers to implement college and career readiness standards: Integrating research, policy, and practice*. Rotterdam, Netherlands: Sense.
- Téllez, K., & Manthey, G. (in press). Teachers' perceptions of collective efficacy of school-wide programs and strategies for English language learners. *Learning Environments Research*.
- Téllez, K., Moschkovich, J., & Civil, M. (Eds.). (2011). *Latinos/as and mathematics education: Research on learning and teaching in classrooms and communities*. Charlotte, NC: Information Age.
- Téllez, K., & Varghese, M. (2013). Teachers as intellectuals and advocates: Professional development for bilingual education teachers. *Theory Into Practice, 52*, 128–135.
- TESOL International Association. (2003). *Standards for the accreditation of initial programs in P-12 ESL teacher education*. Alexandria, VA: Author. Retrieved from [http://www.tesol.org/docs/books/the-revised-tesol-ncate-standards-for-the-recognition-of-initial-tesol-programs-in-p-12-esl-teacher-education-\(2010-pdf\).pdf?sfvrsn=2](http://www.tesol.org/docs/books/the-revised-tesol-ncate-standards-for-the-recognition-of-initial-tesol-programs-in-p-12-esl-teacher-education-(2010-pdf).pdf?sfvrsn=2)
- Theoharis, G., & O'Toole, J. (2011). Leading inclusive ELL social justice leadership for English Language learners. *Educational Administration Quarterly, 47*, 646–688.
- Thompson, A. (2008, December). *Using video technology to provide a professional development forum for reflection on the use of academic language for mathematics in elementary school teachers*. Paper presented at conference of the California Mathematics Council North, Asilomar, CA.
- Trujillo, A. L. (1998). *Chicano empowerment and bilingual education: Movimiento politics in Crystal City, Texas*. New York, NY: Taylor & Francis.
- Trujillo, A. L. (2005). Politics, school philosophy, and language policy: The case of Crystal City schools. *Educational Policy, 19*, 621–654.

- Tuttle, C. C., Teh, B.-R., Nichols-Barrer, I., Gill, B. P., & Gleason, P. (2010). *Student characteristics and achievement in 22 KIPP middle schools*. Washington, DC: Mathematica Policy Research.
- Valdes, G., & Figueroa, R. A. (1994). *Bilingualism and testing: A special issue of bias*. Norwood, NJ: Abex.
- Valencia, R. R., & Suzuki, L. A. (2000). *Intelligence testing and minority students: Foundations, performance factors, and assessment issues* (Vol. 3). Thousand Oaks, CA: Sage.
- Valenzuela, A. (2000). Chicano empowerment and bilingual education: Movimiento politics in Crystal City, Texas by A. Trujillo. *Bilingual Research Journal*, 24, 207–212.
- Valenzuela, A. (2010). *Subtractive schooling: US-Mexican youth and the politics of caring*. Albany: SUNY Press.
- Veenman, S. (1984). Perceived problems of beginning teachers. *Review of Educational Research*, 54, 143–178.
- Volante, L., & Fazio, X. (2007). Exploring teacher candidates' assessment literacy: Implications for teacher education reform and professional development. *Canadian Journal of Education*, 30, 749–770.
- Weinstein, C. S. (1989). Teacher education students' preconceptions of teaching. *Journal of Teacher Education*, 40(2), 53–60.
- Wolf, M. K., & Leon, S. (2009). An investigation of the language demands in content assessments for English language learners. *Educational Assessment*, 14, 139–159.
- Woodcock, R. W., McGrew, K., & Mather, N. (2001). *Woodcock-Johnson tests of achievement*. Itasca, IL: Riverside.
- Xu, Y., & Liu, Y. (2009). Teacher assessment knowledge and practice: A narrative inquiry of a Chinese college EFL teacher's experience. *TESOL Quarterly*, 43, 492–513.
- Zepeda, M., Castro, D. C., & Cronin, S. (2011). Preparing early childhood teachers to work with young dual language learners. *Child Development Perspectives*, 5, 10–14.