

Struggling Reader or Emerging Biliterate Student? Reevaluating the Criteria for Labeling Emerging Bilingual Students as Low Achieving

Journal of Literacy Research

2014, Vol. 46(1) 68–89

© The Author(s) 2013

Reprints and permissions:

sagepub.com/journalsPermissions.nav

DOI: 10.1177/1086296X13504869

jlr.sagepub.com



Susan Hopewell¹ and Kathy Escamilla¹

Abstract

Data systems that use monolingual language frameworks to understand the reading achievement of third-grade students provide inadequate information about emerging bilingual (EB) learners. The authors of this research study apply two competing ideologies (parallel monolingualism and holistic bilingualism) to interpret one set of data. Their findings demonstrate that the same set of scores tells an entirely different story depending on the frames of reference and that these differences are statistically significant. Specifically, they use their analyses to problematize the impact of the Colorado Basic Literacy Act (CBLA) on the categorization of third-grade EB learners. Generalizing from the Colorado data, the authors consider the implications of their findings in a national context of increasing numbers of bilingual learners. Finally, they offer suggestions for site-based school district responses and broader state level policy implications by highlighting one school district's response to their findings.

Keywords

Bilingual, ESL, biliteracy, assessment, diagnosis, at risk, struggling, diversity, minority, SES issues

There are two fundamentally divergent paradigms that drive the discussion about how best to assess bilingual students' competencies. The core difference is whether or not a person who lives life regularly through two languages should be compared with monolingual speakers of either language or should be considered a fundamentally

¹University of Colorado, Boulder, USA

Corresponding Author:

Susan Hopewell, University of Colorado, UCB 249, Boulder, CO 80309, USA.

Email: susan.hopewell@colorado.edu

distinct whole, whose language capacities are distinct yet normal. These distinctions are not insignificant and will be discussed extensively in this article, primarily because incorrect understandings of bilingual students' reading achievement leads to costly and unnecessary remediation initiatives that limit students' opportunities to learn.

The United States boasts an extensive, yet dubious, history of using inexact and often biased assessments to sort people along a continuum of proficiency. IQ tests, for instance, have long been used to provide "scientific" evidence to justify an intelligence hierarchy. In many cases, this hierarchy was intentionally associated with skin color and used to assert that lighter skin tone could be positively associated with higher levels of intelligence. The results have reinforced and sustained the idea that Whites are biologically predisposed to be highly intelligent, while Mexican Americans are less so, and African Americans remain at the bottom (Blanton, 2000; Delpit, 1993). The first decades of the 20th century were replete with examples of such deterministic uses of measurement instruments to provide evidence of the inferiority of those who were not white or monolingual speakers of English (Brigham, 1923; Romaine, 1995; Vera, Feagin, & Gordon, 1995). Evidence abounds of assessment results being used to justify discriminatory practices in such areas as educational opportunity, career promotion, and military service. Challenges to such studies are plentiful, and are primarily rooted in exposing the cultural and linguistic biases inherent in the instruments as well as the methodological flaws in the studies (Hakuta, 1986; Dennis, 1995; Zoref & Williams, 1980). The current study provides evidence that the lens through which we choose to enact policy and examine data contributes to an inadequate system, in particular for students in bilingual education programs. The result is a misapplication of human, curricular, and monetary resources to resolve "problems" that may not actually exist.

Proficiency, competence, expertise, and potential are not finite constructs that can be captured simply. How we interpret and apply test results has a serious and lasting effect, especially in the era of No Child Left Behind (Balfanz, Legters, West, & Weber, 2007; Public Law PL 107-110, 2001). This legislation establishes parameters under which children as young as 8 or 9 years old, and entire schools and school districts can be labeled as failing and targeted for remediation based on assessment results. With such severe consequences at stake, it is imperative that all those involved in making decisions regarding compensatory and remedial education are confident in the validity of the findings they are using to arrive at their judgments. In other words, we must all ask ourselves if the scores we collect are relevant and useful for the interpretations and consequences for which we want to use them (Messick, 1988). Furthermore, given the research regarding student outcomes being directly attributed to teacher expectations, it is particularly critical that we obtain accurate diagnoses regarding student achievement (Brophy, 1983; Jussim, 1989; Rosenthal & Jacobson, 1992).

Reading achievement is a particular area of concern. It is widely accepted that children who do not read on grade level by third grade are not likely to catch up (Juel, 1988; Miller, 2009; Rayner, Foorman, Perfetti, Pesetsky, & Seidenberg, 2002). Witness, for instance, the establishment of *Reading First*, a billion dollar a year U.S. Government initiative, "to ensure that all children learn to read well by the end of third

grade” (U.S. Department of Education, 2010). The amount of money allocated to this effort was directly proportional to the sense of urgency surrounding these dire predictions. These initiatives and beliefs are grounded in the concept of the Matthew Effect, which proposes that initial levels of advantage or disadvantage are likely sustained (Merton, 1968; Stanovich, 1986; Walberg & Tsai, 1983). The Matthew Effect derives from a bible verse in the Gospel of Matthew that reads, “For the one who has will be given more, and he will have more than enough. But the one who does not have, even what he has will be taken from him” (Matthew 25:29–New English Translation). This concept, when applied to reading, predicts that strong readers will continue to reap the rewards of well-developed reading skills and will become stronger over time, while poor readers will be unable to shake a trajectory of disadvantage and will remain relatively poorer readers over time (Morgan, Farkas, & Hibel, 2008).

While the rhetoric about the imperative to read on grade level applies to all students, it is exacerbated for children who are learning English as a second language (August & Shanahan, 2006; Olsen, 2010). The purpose of this article is to use data to demonstrate how the enactment of educational policy results in consequentially disparate inferences regarding the achievement of third-grade Emerging Bilingual (EB) learners, and to suggest an alternative framework for meeting the intent and the letter of the law. Specifically, we use this opportunity to examine the impact of the Colorado Basic Literacy Act (CBLA) on the categorization of third-grade bilingual students and will demonstrate that interpretations of bilingual students’ informal reading scores change dramatically when we apply a lens of holistic¹ bilingualism as opposed to a lens of parallel monolingualism (Dworin, 2003; Escamilla & Hopewell, 2010; Grosjean, 1989; Valdés & Figueroa, 1994). In our analyses, we examine one set of data in three different ways. Each analysis represents an application of one of the two competing ideologies. These separate analyses are then compared and interpreted. Our findings highlight that the same set of scores can tell an entirely different story depending on the frames of reference. These contradictory stories influence students’ schooling opportunities while simultaneously guiding the distribution of district and school resources. Colorado schools offer Spanish–English bilingual education to approximately 11,000 EBs, whose learning opportunities will be severely curtailed should policy implementation label them “at-risk” when in fact they are not. Nationwide, there are 5.1 million EB students, 79% of whom speak Spanish as a home language, who might benefit from bilingual education and bilingual assessment; however, only if assessments are correctly administered and interpreted (Payán & Nettles, 2008).

CBLA and Individual Literacy Plans (ILPs)

The CBLA is a state statute (CRS 301-42, 1997) enacted in 1997 and reauthorized in 2004 that boasts four primary goals (a) to provide students with the literacy skills essential for success in school and life, (b) to promote high literacy standards for all students in K-3rd grade, (c) to help all schools improve the educational opportunities for literacy and performance for all students, and (d) to ensure that all students are adequately prepared to meet Colorado’s fourth-grade reading standards and

benchmarks (Colorado State Board of Education, 2004). These goals are worthy and admirable. Their enactment, however, may categorize students as being “at-risk” for reading failure when in fact they are not.

CBLA requires school districts to screen all students in grades K-3 to identify those who may not be on-track to be reading at grade level at the end of the year. By statute, the only assessments that may be used to make this determination are the *Developmental Reading Assessment*, 2nd edition (DRA2), the *Dynamic Indicators of Basic Early Literacy Skills* (DIBELS), and/or the *Phonological Awareness for Literacy* (PAL). Each student not meeting the grade-level benchmarks is deemed at risk of failure must be placed on an ILP.

The ILP is designed to enable students to meet or exceed third-grade reading proficiency, and is individualized to the extent that it identifies which of the five foundational components of reading (phonological awareness, phonics, fluency, vocabulary, and comprehension) require intervention (National Institute of Child Health and Human Development, 2000). The ILP must include provisions for sufficient in-school reading instruction, an agreement by the student’s family to implement a home reading program that supports (and is coordinated with) the school, and the option for the student to attend a summer reading program should it be deemed necessary at the end of the school year. Administrators, teachers, and parents are required to sign the document as evidence that they are aware of the student’s needs, and as testament to their commitment to accelerate the student’s academic program. This level of mandated identification, targeted individualized resources, and required follow-up is laudable. All students entitled to additional services should receive them.

CBLA and EB Students

Students who speak a home language other than English are not absolved from these regulations; however, they are awarded a 1-year exemption from initial testing provided they are new to the United States and score non-English proficient (NEP) or limited-English proficient (LEP²) on state language exams (Colorado Department of Education, 2007). Furthermore, CBLA states,

As reading comprehension is dependent upon students’ understanding of the language, children with limited English proficiencies, as determined by the individual district’s criteria and documentation, must be assessed in their *language of reading instruction*, leading to their proficiency in reading English” (Colorado State Board of Education, 2004, 5.02, italics added)

Given that there is no Spanish language assessment explicitly named in the three instruments singled out in state statute, English language acquisition (ELA) coordinators have interpreted the mandate to mean that students must demonstrate third-grade reading proficiency in English (Office of Language, Culture, and Equity, Personal Communication, March 2009). This misunderstanding is further reinforced by Colorado Department of Education employees who communicate that the English

assessment is to be used with all students except those who are new to the district (personal communication, June 2009). In other words, while the statute attempts to account for those students who learn to read and write in a language other than English, its interpretation and enactment become muddled by its failure to include and approve assessments in a language other than English. The irony is that each of the mandated assessments has a Spanish language equivalent that could easily be included in the statute to eliminate this confusion.

Theoretical Framework

Current assessments inadequately account for the complexity of assessing academic achievement for bilingual students (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999; Pappamihel & Walser, 2009; Shohamy, 2011). Research to support a bilingual framework for assessing students is virtually nonexistent (August & Shanahan, 2006; Genesee, Lindholm-Leary, Saunders, & Christian, 2006). Despite the dearth of research, it is important to recognize that of the little research that exists, there is a growing body that supports the notion that EB children draw on all of their linguistic resources as they learn to read and write in two languages (Bedore, Peña, García, & Cortez, 2005; García, Kleifgen, & Falchi, 2008; Gort, 2006; Martínez-Roldán & Sayer, 2006). We question the isolated use of monolingual assessment frameworks to evaluate the bilingual learner, and concur with scholars such as Solano-Flores (2008) who suggest that we have not yet developed fair and valid tests that can provide us with more than a fragmentary view of students' language and literacy achievement. It is widely recognized that achievement tests are also inherently tests of language and that standardized assessments privilege particular dialects (American Educational Research Association et al., 1999; Solano-Flores, 2008). While an ideal trajectory would use assessment systems that could measure interactions between and among languages, such tools do not yet exist. Thus, we suggest that a step forward in the quest to provide augmented educational opportunities for bilingual learners is to identify and research new ways of interpreting existing instruments. We offer that trajectories toward biliteracy need to be developed that (a) push us to establish, use, and normalize criteria that recognize that the bilingual is not two monolinguals in one and (b) should not be compared with the trajectories of those who are learning to read and write in only one language (Escamilla & Hopewell, 2010).

Competing Paradigms of Bilingualism

As stated in the introduction, there are two fundamentally divergent paradigms that drive the discussion about how best to assess bilingual students' competencies. These paradigms hinge significantly on the fundamental notion of whether or not monolingual language and literacy development should be the benchmark for students who read, write, speak, and listen in two languages. The crux of the disagreement is whether

or not there is something fundamentally different about normal bilingual language and literacy acquisition that might require us to rethink students' growth and progress.

Proponents of the first orientation subscribe to an ideology of *parallel monolingualism*, and maintain that each of a person's languages can, and should, be developed, assessed, and interpreted separately (Fitts, 2006; Heller, 2001). This is, in fact, the basis for most assessment systems, and derives from the belief that the presence of more than one language in any interaction indicates, or has the potential to cause, cognitive and linguistic confusion. Furthermore, most proponents of parallel monolingualism advocate for the maintenance of English as the hegemonic language of prestige, and justify the use of a language other than English within the educational environment only in as much as it facilitates the acquisition of English. Biliteracy, in other words, is not the desired outcome.

Conversely, proponents of the second orientation subscribe to an ideology of *holistic bilingualism*, in which multiple languages contribute to a syncretic and indivisible whole which cannot be understood by looking at each language in isolation (Escamilla & Hopewell, 2010; Grosjean, 2008). These theories of bilingualism build on Cummins's (1981) notions of common underlying proficiencies which offer that knowledge acquired in one language is always available to help make sense of input provided in second or additional languages. In other words, skills and concepts learned in one language need not be relearned in the other. The melding of languages in the bilingual learner generates a unique entity whose potential is explored through paradigms that consider the totality of what is known and understood across languages (Dworin, 2003; Valdés & Figueroa, 1994). Holistic bilingualism is the goal, and the lens for analyzing proficiency. It attempts to encapsulate a complex, dynamic yet indivisible linguistic system. It requires novel ways of considering language and literacy proficiencies in which trajectories toward biliteracy are established to take into consideration how all languages are vehicles through which learning takes place and is communicated.

Trajectory Toward Biliteracy

During the past 8 years, as part of a research-based biliteracy intervention called Literacy Squared®, we have established and tested a framework for understanding paired language and literacy acquisition for Spanish–English EBs attaining literacy in bilingual settings. The body of research that guides our work concludes that schooling for EB learners is improved when program development and implementation attends thoughtfully to *Language of Instruction*, *Quality of Instruction*, and *Cross-Language Connections* (August & Shanahan, 2006; Genesee et al., 2006; Gersten & Baker, 2000; Koda & Zehler, 2008; Laurent & Martinot, 2009; Slavin & Cheung, 2005). Our framework challenges practitioners to rethink how they design and deliver literacy instruction to best capitalize on students' multiple linguistic resources within bilingual instruction. It begins with the understanding that literacies and languages develop cohesively in reciprocal and mutually supportive ways, and is founded on the idea that Spanish language literacy and English language literacy contribute to a broad

and unified conceptualization of literacy. Lessons are coordinated across language environments, cohesive and complementary, but not duplicative. Students are held accountable to use what they know and can do in one language to further their learning in the other. Our comprehensive biliteracy model is complex, and at its core, it asks teacher to provide equal attention to the following aspects of language: oracy, reading, writing, and metalanguage. Furthermore, it asks that all students be provided paired literacy in which students learn simultaneously to read and write in two languages beginning in kindergarten. We provide teachers with guidance about how to make instructional decisions regarding texts through our work creating a trajectory that defines and guides acceptable relationships among Spanish and English reading levels. Reading and writing progress is measured annually in Spanish and English, and because of this, we are in a unique position to be able to examine and address questions of biliteracy development. Since 2004, the model has been tested in 31 schools in three states with 200 teachers and nearly 4,000 students.

One aspect of our instructional framework developed to help us understand and interpret reading scores across two languages is called the *Trajectory Toward Biliteracy*. We propose that it is a more accurate and appropriate means of assessing and interpreting the progress of EB children than frameworks that look at languages in isolation. Our trajectory is founded on the idea that if students are progressing along a satisfactory path toward biliteracy, their Spanish language literacy will be slightly more advanced than their English language literacy, but that a large discrepancy will not appear between the two. Furthermore, we recognize that the acquisition of literacy in two languages may result in pacing that differs from that expected in monolingual settings; therefore, we acknowledge through our framework that students on a normal trajectory might fall within a range of levels in either language. Currently, the only way to measure and document this trajectory is to assess students in each language and to interpret these assessments side by side. While this does not fully capture language and literacy acquisition holistically, it provides a fuller and more robust picture of a student's achievement than does looking at either assessment result independently. The resulting trajectory, therefore, considers Spanish language literacy achievement plus English language literacy achievement allowing for a broader range of levels to represent adequate progress toward biliteracy proficiency.

Establishing an alternative framework that acknowledges the complexity of learning to read, write, speak, and listen in two languages simultaneously requires rethinking the use of fixed cut scores to determine which students are at risk of failure, especially when such determinations are being made about students acquiring languages and literacies in bilingual programs (Linan-Thompson, Cirino, & Vaughn, 2007; Wagner, Francis, & Morris, 2005). Performance standards are typically determined through the imperfect and subjective science of identification of fixed cut scores used to determine competency and to separate one group of students from another (Horn, Ramos, Blumer, & Madaus, 2000). We have empirical evidence that fixed cut scores can be unstable and unreliable when used to make deterministic high-stakes decisions about students (Francis et al., 2005). By collecting and interpreting data in two languages, we begin to approach Grosjean's (2008) logic that bilingual

Table 1. Trajectory Toward Biliteracy.

EDL2 (Spanish reading level)	DRA2 (English reading level)
A-6	A-2
8-10	3-6
12-16	8-10
18-28	12-16
30-38	18-28
40	30-38
50-60	40+

Note. Literacy Squared® Scaffolded Biliteracy–Targeted Zones.
EDL = *Evaluación del Desarrollo de Lecto-escritura*; DRA = Developmental Reading Assessment.

individuals will be able to express their competencies along a continuum that will differ by language and topic, yet be indicative of a single aptitude. A question we asked ourselves at the start of the process was, “*Is it sensible to use monolingual cut scores to gauge the progress of biliterate students?*” In other words, “Should the minimal standard of competency for a student learning to read and write in two languages be the same as for a student learning to read and write only in one?” On answering negatively, we used expert judgment and biliteracy theory to hypothesize a range of reading levels in Spanish and English that might capture the reading performances we would expect to see across grade levels for students attaining biliteracy (Zieky & Perie, 2006). While this does not move us entirely away from the dilemma of using cut scores to categorize students, it does provide a more nuanced understanding of the complexity of becoming biliterate than the extreme simplicity of applying a fixed cut score developed with one population to make sense of the language and literacy growth for students in bilingual programs. Because establishing cut scores always involves some level of judgment, we cannot be certain that we have identified ideal predictors of any level of proficiency; however, after establishing cut score ranges to define a trajectory toward biliteracy, we assessed students to determine whether or not these ranges were representative of student achievement. We then asked teachers to provide extensive feedback and input as we refined the ranges. We have evidence to support the fact that students performing within these ranges outscore their peers on high-stakes state-mandated standardized tests of reading and writing (Butvilofsky & Escamilla, 2011).

Table 1 illustrates the ranges of reading scores we would expect students to attain when their Spanish language reading and their English language reading are developing in parallel. According to this framework, a student who presents the reading behaviors associated with a Spanish language Level 10 is expected to be reading in English between Levels 3 and 6. The scores in this table correspond to reading levels that students attain when evaluated using the Spanish language *Evaluación del Desarrollo de Lecto-escritura* (EDL2) and the English language Developmental Reading Assessment

(DRA2) (Celebration Press, 2007a, 2007b). These tools measure parallel competencies across languages.

Our research confirms that these ranges are reflective of satisfactory progress for students being instructed in the Literacy Squared paired literacy program (Escamilla, Ruiz-Figueroa, Hopewell, Butvilofsky, & Sparrow, 2010). Though we cannot make definitive statements about how representative they are of satisfactory progress in alternative biliteracy programs, we know that only through bilingual assessment can we approximate an accurate understanding of students' trajectories toward biliteracy. This framework is meant to aid in score interpretation and in the planning of instruction. It provides a foundation for changes in pedagogy, and serves as a tool to guide teachers as they plan appropriate instruction with research-based and research-tested expectations for biliteracy development.

Knowing that the CBLA requires schools to place students on ILPs based on their literacy assessments, we wondered if students' "at-risk" designation would change if their Spanish language reading assessments were considered instead of their English language reading assessment, and further, what would happen if these same literacy scores were interpreted using the Trajectory Toward Biliteracy framework. The following summarizes our investigation and findings.

Research Questions

1. How do ILP designations differ for third-grade EB students in three Colorado schools when assessment data are interpreted using monolingual assessment frameworks versus holistic bilingual assessment frameworks?
 - a. How many third-grade EB students in three Colorado schools require an ILP at the end of third grade when only English language literacy criteria (DRA2) are applied?
 - b. How many of the same students require an ILP at the end of third grade when only Spanish language literacy criteria (EDL2) are applied?
 - c. How many of the same students require an ILP at the end of third grade when a Trajectory Toward Biliteracy is used to interpret Spanish language reading (EDL2) and English language reading (DRA2)?
2. To what extent is there agreement among the three methods (DRA2 only, EDL2 only, or the Trajectory Toward Biliteracy) in determining whether or not a student should be placed on an ILP?

Method

To answer our questions, we examined a subset of data collected within the longitudinal research study our team calls Literacy Squared®. Literacy Squared is a biliteracy intervention that requires connected and coordinated literacy instruction in Spanish and English. All students who participate are assessed in reading and writing in Spanish and English every year from kindergarten to Grade 5.

Because third grade is the year that many consider to be critical in terms of predicting long-term school success, we used the data from this grade to evaluate “at-risk” designations as determined by the need to initiate an ILP. All third-grade students were administered the EDL2 and the DRA2 in spring of 2008 to determine their independent reading levels in Spanish and English. These scores were then analyzed using descriptive statistics to determine the number and percentage of students who would require an ILP based on a variety of benchmark criteria. Each analysis was distinguishable by the employment of either a monolingual assessment lens or a bilingual assessment lens, and which benchmarks were used to determine proficiency. Findings were interpreted with specific attention to the policy issues they raise. Data were then further analyzed to determine the extent to which there was agreement across the three testing conditions and the extent to which there was independence across test conditions.

Participants

The data presented in this article come from two metro area school districts in the state of Colorado. Together these districts educate one in four of Colorado’s EBs. Each uses English language literacy as the exclusive determinant of the need for an ILP (Directors of Language and Literacy, personal communication, 2009).

Our study included 268 Spanish–English bilingual children who were in the third grade in 2007–2008. All children were participants in the larger longitudinal study called Literacy Squared. Thus, all 268 students were learning simultaneously to read and write in Spanish and English in paired literacy instruction. All students qualified for free or reduced price lunch, were Latino (nearly entirely of Mexican heritage), and carried the label of either NEP or LEP, as determined by the Colorado English Language Assessment (CELA) which is administered throughout the state to all bilingual children.

Instruments

A foundational component of the Literacy Squared study was the identification of formal and informal assessments that reliably measure reading growth. Informal assessments included the Spanish EDL2 and the English DRA2 (Celebration Press, 2007a, 2007b). These tools were identified because they were available in Spanish and English and because they are required by the state as tools for monitoring reading progress. That is, these are the instruments the state requires teachers to use to determine whether or not students are making adequate progress in reading achievement. These informal measures of reading are parallel instruments that include miscue analysis, fluency, and reading comprehension. They help educators identify each student’s reading ability, level, and progress. They are used to inform instructional decisions and to analyze targeted and specific areas for literacy intervention. The DRA2 and EDL2 have been studied and determined to be valid and reliable measures of reading in Spanish and English (Weber, 2001). In the EDL2 and DRA2, validity was established

via criterion, construct, and content validity measures, and reliability was established via internal consistency tests, passage equivalency, test-retest and expert rater tests (Weber, 2001). The DRA2 was normed on a student population that was 67% Caucasian, 10% African American, 4% Hispanic, 2% Asian, and 2% Other. Most of these students (66%) attended schools in suburban communities. The EDL2 was “field-tested in bilingual classrooms across the United States” (Pearson Publishing, 2012). In other words, the DRA2 was not normed on a student sample the resembled those in our study, but the ELD2 was.

Analyses

According to Pearson Learning, the publisher of the EDL2 and the DRA2, third-grade students reach grade-level benchmark on these assessments when they are able to read independently at a Level 38. This level is founded on the assumption that students are learning to read and write in a strictly monolingual, or a sequential bilingual, setting. In other words, students are experiencing literacy instruction in either English or Spanish. There are no alternative criteria for those who are learning bilingually.

As mentioned previously, the mandate from the Colorado Department of Education is that students must be reading on grade level (or at a Level 38) to avoid being placed on an ILP. The tests the state lists as acceptable to make grade-level determinations are all assessments of English language literacy; however, as stated above, all assessments are also available in Spanish. Because of this, to answer Research Question 1, we began our analyses by including only the DRA2, (English language reading scores) for all students. Using a monolingual English language lens, we used the state’s extant criteria to determine whether or not a student should be placed on an ILP. All students who fell below a DRA2 Level 38 were labeled “at-risk” and in need of an ILP. Then, in recognition that all of the 268 participants were Spanish–English bilingual students who lived in homes where Spanish was used regularly and were learning simultaneously to read and write in Spanish and English, we reevaluated the data to consider the consequences if the students’ Spanish language EDL2 scores were referenced instead of their English language DRA2 scores. Again, using a monolingual lens, and the grade-level benchmark of a Level 38, we examined only the students’ Spanish language EDL2 scores. All students who fell below an EDL2 Level 38 were labeled “at-risk” and in need of an ILP.

Finally, we applied an analytical lens grounded in a bilingual trajectory requiring us to include the English language DRA2 scores and the Spanish language EDL2 scores to make a determination about students’ achievements in bilingual reading instruction. Knowing that it is rare for a bilingual person to be completely balanced across languages and contexts (Baker, 2001; Valdés & Figueroa, 1994), we rejected the criteria that students would need to be at a Level 38 on both assessments to avoid being placed on an ILP. Instead, we used our Trajectory Toward Biliteracy framework to make these determinations. Our targeted zones ask that third-grade students be reading in the third-grade range of levels (30–38) in Spanish, and in the second-grade range of levels (18–28) in English (see Table 1).

Findings

Operating from the assumption that only DRA2 (English language) scores could be considered when determining whether or not to place a child on an ILP, we found that 224 of the 268 students had independent reading levels in English below Level 38, and would therefore, require an ILP. In other words, 84% of these third graders would be labeled as at risk of failure, and their parents would be asked to sign a contract indicating that they were aware of these determinations and that they would work with the school and the school district to provide interventions designed to accelerate the student's literacy gains.

Continuing with a monolingual assessment lens, we considered what would happen if we looked only at one set of scores, but that instead of the English language DRA2 scores we used the Spanish language EDL2 scores. Analyses of these same children's Spanish language EDL2 scores indicated the 148 out of the 268 students did not meet or exceed a Level 38 on the EDL2. In other words, 55% would require an ILP. Of the 45% of the students ($n = 120$) who did not require an ILP because their Spanish reading scores indicated that they presented the literacy behaviors required of a third-grade student, only 33% ($n = 39$) were also at grade level in English. In other words, 81 students, 67% of those who tested at grade level in Spanish would be placed on an English language ILP even though they had no difficulties reading. These 81 students were not at risk of reading failure, but rather required assistance in ELA.

Finally, we used a bilingual lens and used the targeted zones from the Trajectory Toward Biliteracy to interpret students' Spanish language and English language reading levels as a coordinated whole. A student was considered to be making satisfactory progress if she or he were reading in Spanish in the third-grade range of Levels 30-38, AND reading in English at a Level 18 or higher. Of the 268 children, 194 (72%) were reading in Spanish at a Level 30 or higher, indicating that 74 students (28%) would automatically need to be considered for an ILP. Furthermore, of the 194 reading in the acceptable range in Spanish, only 32 (16%) were not yet reading at a Level 18 in English. Stated differently, 84% of students reading at grade level in Spanish were within the targeted zones for achieving biliteracy as determined by the ranges established on our Trajectory Toward Biliteracy. Knowing that these students read in an acceptable range in Spanish communicates to us that they probably require targeted ELA instruction rather than remediation in reading. Admittedly, 40% is a large number; however, it is many fewer than the 84% originally identified when the English-only lens was applied. Table 2 summarizes the aggregate data.

As can be seen, each analysis resulted in a different interpretation of which students were at risk of reading difficulties. When Spanish language reading assessments replaced the English-only assessments, the number of students deemed proficient more than doubled (nearly tripled). When a trajectory toward biliteracy was used to examine these same students' Spanish and English scores in a holistic manner, the number of students deemed proficient increased by 42 students. In fact, when the holistic interpretation is compared with the English-only results, we see that

Table 2. Students Requiring an ILP (*n* = 268).

	English language reading scores only—DRA2	Spanish language reading scores only—EDL2	Trajectory toward biliteracy—EDL2 + DRA2
% requiring ILP	83.6% (<i>n</i> = 224)	55.3% (<i>n</i> = 148)	39.6% (<i>n</i> = 106)
% not requiring ILP	16.4% (<i>n</i> = 44)	44.7% (<i>n</i> = 120)	60.4% (<i>n</i> = 162)

Note. ILP = Individual Literacy Plan; EDL = *Evaluación del Desarrollo de Lecto-escritura*; DRA = Developmental Reading Assessment.

118 students (44% of the entire sample) move from an “at risk” classification to being deemed making satisfactory progress toward becoming biliterate. This is not inconsequential.

The data presented thus far indicate that for the students in this sample the current means of determining whether or not to place a bilingual student on an ILP are problematic; however, they do not allow us to understand the proportion of agreement among these methods or to determine whether or not these differences can be generalized to other student populations. Cohen’s Kappa (κ) calculates the proportion of agreement that exceeds the concordance that would be expected due to mere chance. While typically used to determine interrater reliability, Cohen’s Kappa has been used to measure the proportion of agreement across instruments, particularly to understand the likelihood of diagnosing a reading disability (Waesche, Schatschneider, Maner, Ahmed, & Wagner, 2011). Therefore, to compare the extent to which there is agreement regarding the need for a student to be placed on an ILP when students’ monolingual scores in either language are used in isolation, the outcomes on the DRA2 were compared with those yielded by the EDL2 resulting in a Kappa of .31, indicating a low level of agreement. To compare the extent to which there is agreement between either of the monolingual assessment systems with the Trajectory Toward Biliteracy, we compared the outcomes on each of the DRA2 and the EDL2 with the Trajectory Toward Biliteracy outcomes using Kappa. When the outcomes of using the EDL2 (Spanish) to determine the need for an ILP was compared with the Trajectory Toward Biliteracy the Kappa was .49, indicating moderate levels of agreement. When the outcomes of using the DRA2 (English) to determine the need for an ILP was compared with the Trajectory Toward Biliteracy the Kappa is .21, indicating poor agreement. Because the outcome of interest (the need for an ILP) can be categorized into yes or no and the data are from a single population, a chi-square test of independence was performed to examine the relationship between the testing framework (English-only, Spanish-only, and bilingual) and the ILP designation. The relationship was significant, $\chi^2(1, N = 267) = 30.83, p < .001$. The choice of a monolingual framework will indicate the need to place a student on an ILP at significantly higher rates than when applying the Trajectory Toward Biliteracy. Monolingual frameworks are most problematic when English is the only language used to make this determination. Furthermore, should students demonstrate proficiency in reading in a language other than English,

in this case Spanish, an English language reading exam that indicates a level less than grade-level reading should be interpreted to mean there is a need for English language development, rather than as an indicator of a reading challenge.

Discussion and Implications

Educational policies, such as CBLA, that result in children being labeled as at risk of reading failure when in fact they are not, may affect the lives of hundreds of thousands of biliterate students throughout the United States. Consider the fact that the number of EBs participating in bilingual education programs in the state of Colorado during the 2007-2008 school year was 11,392 (J. Bruno, Colorado Department of Education, personal communication, February 28, 2011). Assuming that the numbers of EB students deemed proficient in reading in this study are representative of what we might find across the entire population, we see that using a monolingual assessment framework that considers only English language reading scores results in just more than 900 students being recognized as proficient readers; whereas, generalizing from the percentage of students deemed proficient when the trajectory toward biliteracy is applied, we find that more than 6,800 students are likely making satisfactory progress. In other words, it is highly probable that thousands of EB students, across the state of Colorado, were labeled as “at risk,” and placed on ILPs unnecessarily. In fact, when we examine secondary data in the form of the Colorado Student Assessment Program (CSAP), a high-stakes, state-mandated, standardized exam that EBs are able to take in Spanish when they are in third grade, we see that in 2008, 59% scored proficient or advanced. These CSAP scores indicate that only 41% of bilingual learners should have been placed on an ILP. These numbers are quite close to the numbers we identify when using the Trajectory Toward Biliteracy. These data confirm that greater numbers of EBs are making adequate progress in reading than current ILP data might indicate. That is, a great proportion of EB students identified as in need of additional reading instruction are deemed so due to “false positives” (Torgesen, 1998). If we assume that only 10% of the five million bilingual learners in the U.S. are educated in bilingual settings, and we let the percentages of students in this study requiring an ILP serve as a proxy for the numbers nationwide, there are 420,000 students (84%) being targeted for educational remediation when in fact only 200,000 (40%) may need it. In other words, 220,000 bilingual students are relegated to compensatory education that neither builds on what they know and can do nor allows resources to be targeted to those truly in need. The result is harmful to students and society (Center for Mental Health in Schools at University of California, Los Angeles, 2008). Consider as a case in point Florida. Florida is ranked third in the nation in terms of numbers of EB students (at more than 260,000), and since 2003 has required grade retention for third-grade students not able to pass a reading exam (Florida Department of Education). One can only speculate as to the amount of money and resources that Florida is investing in students who do not need it. Currently, a number of states (Iowa, New Mexico, Tennessee, and Colorado) are considering implementing similar retention policies (Smith, 2012). Research is clear that retention is ineffective in accelerating students’

reading achievement and that it increases the likelihood that students will not participate in postsecondary educational opportunities (Hong & Yu, 2007; Ou & Reynolds, 2010). Our data should raise questions for those instituting such policies. These students do not need reading instruction and certainly should not be retained for academic reasons; they need language development.

The consequence of being placed on an ILP in the state of Colorado is relegation to remediation through targeted interventions and to restricted opportunities to learn once labeled “low.” We have a plethora of data that indicate that students who are deemed “low” are often ability grouped to tailor instruction to students’ needs. Often this instruction moves at a slower pace and includes less depth (Ansalone, 2003; Oakes, 2005; Worthly, 2010). Research indicates that these groups are highly skills based, include less demanding material, are extremely repetitive, and result in less overall learning (Heubert & Hauser, 1999; Lleras & Rangel, 2009). Once tracked into low groups, students rarely exit (Gordon & Piana, 1999). In other words, these instructional practices and organizational processes conspire to create conditions to bolster the Matthew Effect, the very theory that sustains the sense of urgency around ensuring that kids reading on grade level before the end of third grade. Many EB students, therefore, are likely denied access to an enriched curriculum because they have been inappropriately labeled as low readers. In essence, it becomes a false, but self-fulfilling, prophecy.

Providing resources for academic interventions is expensive; and targeting these costly resources to students who do not require them denies services to those who are truly in need. This consequently holds students back who might benefit more from richly contextualized classroom instruction. Furthermore, decisions regarding student achievement require that two fundamental assumptions be met: (a) an appropriate instructional environment prior to intervention (e.g., a good “first teach”), and (b) adequate and valid assessments (Reynolds, Wheldall, & Madelaine, 2010). Our study calls into question the validity of current assessment policies and practices that lead to the consideration of only English reading levels when determinations are made regarding the academic achievement of EB children. It raises similar questions should policies allow only consideration of Spanish reading levels. These policies undervalue children’s literacy levels because they do not consider Spanish plus English, a more holistic perspective.

The use of monolingual assessments in either English or Spanish provides inadequate information about the literacy development of EB students in U.S. schools. Placing children who are at normal stages of biliterate development on ILPs negatively labels children “at-risk” when, in fact, they are not underachieving. Inappropriate ILP designations generate costly, unnecessary, inappropriate, and perhaps harmful interventions. In an era of monetary shortfalls in public education, resources that could be better used to assist those with true need are being misdirected. Furthermore, the misuse of test data to label students inappropriately reinforces the idea that two languages are sources of confusion and interference rather than assets to be built on. The message to parents and the community is that learning in two languages may cause cognitive confusion and retard advances in literacy development.

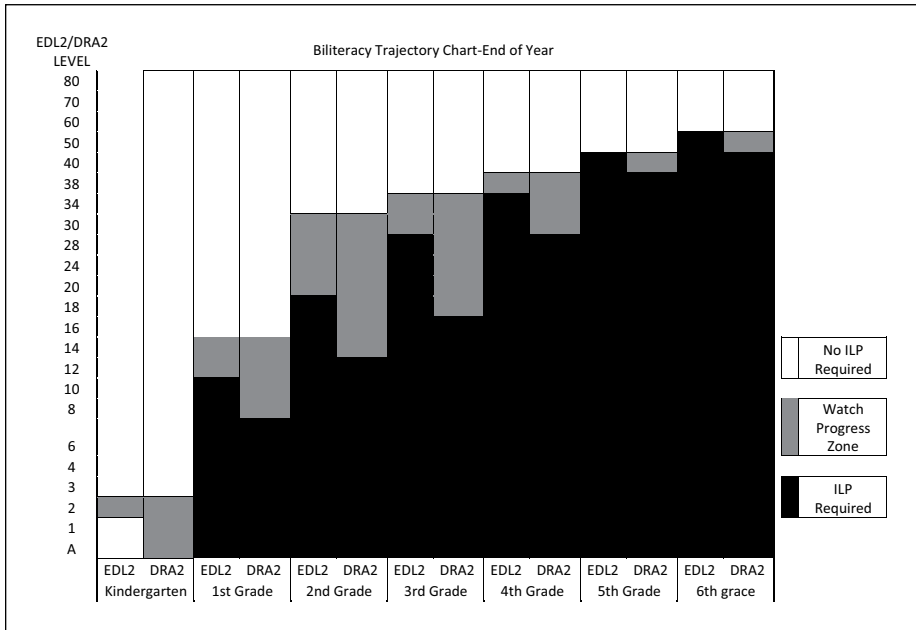


Figure 1. Biliteracy ILP guidelines, 2010.

Note. ILP = Individual Literacy Plan; EDL = Evaluación del Desarrollo de Lecto-escritura; DRA = Developmental Reading Assessment.

Given our history as a nation of using testing to justify discriminatory practices, it behooves us to question the practice of using monolingual frameworks to determine whether or not EB students are “at risk” of long-term reading difficulties. We call on those who determine educational policies to recognize that learning to read in two languages differs significantly from learning to read and write in one, and argue that policy makers should seek, create, and refine trajectories toward biliteracy that use paradigms of holistic bilingualism.

A Change in Policy: One School District’s Response

The findings reported herein inspired one Colorado school district³ to reconsider its criteria for placing EB children who are learning simultaneously to read in Spanish and English on an ILP. Educators in this district are now required to consider a student’s achievement in Spanish language reading and English language reading when making “at-risk” determinations. Teachers use the chart represented in Figure 1 to determine whether or not a child requires an ILP. District administrators originally applied the colors of the stoplight to our trajectory to provide teachers with guidance. Thus, scores falling into a red zone required an ILP, scores falling into the yellow range were designated as needing to be monitored, but were not placed on an official

ILP, and those in the green ranges had were making expected progress and in no need of an ILP. For the purposes of this article, these designations have been recreated using gray scale (see Figure 1).

Students in each grade level are assessed in Spanish language reading and English language reading. Their scores are then charted and compared with the figure above to determine whether or not an ILP designation is warranted. The black bars represent the ranges in each language that require an ILP. The grayed area represents scores that trigger that students be waived from an ILP while being monitored. Finally, the white areas require no ILP and no watch. As can be seen, Spanish language and English language assessment results must be considered holistically. Rather than requiring that students meet a strict monolingual reading criteria, a range of scores are considered to be appropriate for a student developing biliteracy. To read the chart, attend to the color that corresponds to particular reading levels by grade. Students scoring at levels that correspond to the color green/white are meeting monolingual reading criteria and require no special attention. Those scoring at levels that correspond to the color yellow/gray are within the ranges considered appropriate for students learning to read in paired literacy programs as determined by the Literacy Squared Framework for Biliteracy. They are to be monitored carefully, but do not require a formal “at-risk” designation as indicated with the creation of an ILP. Only those students who read at levels corresponding to the color red/black, in either language require an ILP. These guidelines acknowledge that normal biliteracy development may differ in some fundamentally significant ways from monoliteracy development, and that inappropriate labeling may harm students’ opportunities to learn, result in an inappropriate allocation of human and financial resources, and result in unnecessary self-fulfilling prophecies. The results of the implications of this new policy have not yet been reported, but we are hopeful that using a lens of holistic bilingualism will result in students’ assessments being more representative of their actual abilities. This information will be used to tailor instruction to build on their strengths rather than the enactment of interventions aimed at “fixing” deficits that may not actually exist.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. The use of the term *holistic* with regard to bilingualism refers specifically to Grosjean’s conceptualization of the bilingual brain as being a unique and viable whole that cannot be fractionalized into its component languages. This is not to be confused with assessment terminology in which holistic assessment is often referenced in opposition to analytic assessment.

2. The terms *Non-English Proficient (NEP)*, *Limited English Proficient (LEP)*, and *Fully English Proficient (FEP)* are labels that are used by federal and state agencies to indicate language proficiency and eligibility for language support services. We recognize that they are dated and deficit in orientation; however, are choosing to employ them here as they are the terms the Colorado legislation uses.
3. The authors wish to acknowledge and celebrate Angel Stobaugh and Catalina Martis whose thinking and actions have helped us to think through the implications of our research at the school district level.

References

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Ansalone, G. (2003). Poverty, tracking, and the social construction of failure: International perspectives on tracking. *Journal of Children & Poverty*, 9, 3-20.
- August, D., & Shanahan, T. (Eds.). (2006). *Developing literacy in second-language learners: Report of the national literacy panel on language-minority children and youth*. Mahwah, NJ: Lawrence Erlbaum.
- Baker, C. (2001). The development of bilingualism. In C. Baker (Ed.), *Foundations of bilingualism* (pp. 86-107). Clevedon, UK: Multilingual Matters.
- Balfanz, R., Legters, N., West, T. C., & Weber, L. (2007). Are NCLB's measures, incentives, and improvement strategies the right ones for the nation's low-performing high schools? *American Educational Research Journal*, 44, 559-593.
- Bedore, L. M., Peña, E. D., García, M., & Cortez, C. (2005). Conceptual versus monolingual scoring: When does it make a difference? *Language, Speech, and Hearing Services in Schools*, 36, 188-200.
- Blanton, C. K. (2000). "They cannot master abstractions, but they can often be made efficient workers": Race and class in the intelligence testing of Mexican Americans and African Americans in Texas during the 1920s. *Social Science Quarterly*, 81, 1014-1026.
- Brigham, C. (1923). *A study of American intelligence*. Princeton, NJ: Princeton University Press.
- Brophy, J. E. (1983). Research on self-fulfilling prophecy and teacher expectations. *Journal of Educational Psychology*, 75, 631-661.
- Butvilofsky, S., & Escamilla, K. (2011). *Literacy Squared Phase II: Colorado Case Study Technical Report Year One, 2009-2010*. Boulder, CO: Author.
- Celebration Press. (2007a). *Developmental Reading Assessment*. Parsippany, NJ: Author.
- Celebration Press. (2007b). *Evaluación del desarrollo de la lectura*. Parsippany, NJ: Author.
- Center for Mental Health in Schools at University of California, Los Angeles. (2008). *Grade Retention: What's the prevailing policy and what needs to be done?* Los Angeles, CA: Author. Available from <http://smhp.psych.ucla.edu>
- Colorado Department of Education. (2007). *CBLA end of year assessment: Reporting guidelines. Office of learning and results*. Retrieved from http://www.cde.state.co.us/coloradoliteracy/cbla/download/20070907_Updated_CBLA_EOY_Assessment_Guidelines.pdf
- Colorado State Board of Education. (2004). *Rules for the administration of the Colorado Basic Literacy Act*. Retrieved from <http://www.cde.state.co.us/coloradoliteracy/cbla/download/CBLARules2004.pdf>

- Cummins, J. (1981). The role of primary language development in promoting educational success for language minority students. In California State Department of Education (Ed.), *Schooling and language minority students: A theoretical framework* (pp. 2-49). Los Angeles: Evaluation, Dissemination and Assessment Center, California State University.
- Delpit, L. D. (1993). The silenced dialogue: Power and pedagogy in educating other people's children. In L. Weis & M. Fine (Eds.), *Beyond silenced voices: Class, race, and gender in United States schools* (pp. 119-139). New York: State University of New York Press.
- Dennis, R. (1995). Social Darwinism, scientific racism, and the metaphysics of race. *Journal of Negro Education*, 64, 243-252.
- Dworin, J. (2003). Insights into biliteracy development: Toward a bidirectional theory of bilingual pedagogy. *Journal of Hispanic Higher Education*, 2, 171-186.
- Escamilla, K., & Hopewell, S. (2010). Transitions to biliteracy: Creating positive academic trajectories for emerging bilinguals in the United States. In J. Petrovic (Ed.), *International perspectives on bilingual education: Policy, practice, controversy*. Charlotte, NC: Information Age Publishing. (pp. 69-93).
- Escamilla, K., Ruiz-Figueroa, O. A., Hopewell, S., Butvilofsky, S., & Sparrow, W. (2010). *Transitions to biliteracy: Literacy Squared, 2004-2009* (Final technical report). Boulder, CO: Self-published
- Fitts, S. (2006). Reconstructing the status quo: Linguistic interaction in a dual-language school. *Bilingual Research Journal*, 30, 337-366.
- Francis, D. J., Fletcher, J. M., Stuebing, K. K., Lyon, G. R., Shaywitz, B. A., & Shaywitz, S. E. (2005). Psychometric approaches to the identification of LD: IQ and achievement scores are not sufficient. *Journal of Learning Disabilities*, 38, 98-108.
- Garcia, O., Kleifgen, J. A., & Falchi, L. (2008). *From English language learners to emergent bilinguals* (Equity Matters: Research Review Number 1). New York, NY: Teachers College Press. Retrieved from <http://www.schoolfunding.info/news/policy/EmergentBilinguals.pdf>
- Genesee, F., Lindholm-Leary, K., Saunders, W. M., & Christian, D. (2006). *Educating English language learners*. New York, NY: Cambridge University Press.
- Gersten, R., & Baker, S. (2000). What we know about effective instructional practices for English-language learners. *Exceptional Children*, 66, 454-470.
- Gordon, R., & Piana, L. D. (1999). *No exit? Testing, tracking, and students of color in U.S. public schools*. New York, NY: Applied Research Center.
- Gort, M. (2006). Strategic code-switching, interliteracy, and other phenomena of emergent bilingual writing: Lessons from first grade dual language classrooms. *Journal of Early Childhood Literacy*, 6, 323-354.
- Grosjean, F. (1989). Neurolinguists beware! The bilingual is not two monolinguals in one. *Brain and Language*, 36, 3-15.
- Grosjean, F. (2008). *Studying bilinguals*. Oxford, UK: Oxford University Press.
- Hakuta, K. (1986). *Mirror of Language*. New York, NY: Basic Books.
- Heller, M. (2001). Undoing the macro/micro dichotomy: Ideology and categorization in a linguistic minority school. In N. Coupland (Ed.), *Sociolinguistics and social theory* (pp. 261-296). Harlow, UK: Pearson.
- Heubert, J., & Hauser, R. (Eds.). (1999). *High stakes: Testing for tracking, promotion, and graduation*. Washington, DC: National Research Council.
- Hong, G., & Yu, B. (2007). Early-grade retention and children's reading and math learning in elementary years. *Educational Evaluation and Policy Analysis*, 29, 239-261.

- Horn, C., Ramos, I., Blumer, I., & Madaus, G. (2000). *Cut scores: Results may vary* (NBETPP Monographs, Vol. 1, Number 1). Retrieved from <http://www.bc.edu/research/nbetpp/state-ments/M1N1.pdf>
- Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology*, 80, 437-447.
- Jussim, L. (1989). Teacher expectations: Self-fulfilling prophecies, perceptual biases, and accuracy. *Journal of Personality and Social Psychology*, 57, 469-480.
- Koda, K., & Zehler, A. M. (Eds.). (2008). *Learning to read across languages: Cross-linguistic relationships in first- and second-language literacy development*. London, England: Routledge.
- Laurent, A., & Martinot, C. (2010). Bilingualism and phonological awareness: The case of bilingual (French-Occitan) children. *Reading and Writing*, 23, 435-452.
- Linan-Thompson, S., Cirino, P. T., & Vaughn, S. (2007). Determining English language learners' response to intervention: Questions and some answers. *Learning Disability Quarterly*, 30, 185-195.
- Lleras, C., & Rangel, C. (2009). Ability grouping practices in elementary school and African American/Hispanic achievement. *American Journal of Education*, 115, 279-304.
- Martínez-Roldán, C., & Sayer, P. (2002). Reading through linguistic borderlands: Latino students' transactions with narrative texts. *Journal of Early Childhood Literacy*, 6, 293-322.
- Merton, R. (1968). The Matthew effect in Science. *Science*, 159, 56-63.
- Messick, S. (1988). The once and future issues of validity: Assessing the meaning and consequences of measurement. In H. Wainer & H. I. Braun (Eds.), *Test validity* (pp. 33-44). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Miller, S. (2009). Using classroom assistants to tutor struggling third graders: How long will it take to read at grade level? *Literacy Research and Instruction*, 48, 233-246.
- Morgan, P. L., Farkas, G., & Hibel, J. (2008). Matthew effects for whom? *Learning Disability Quarterly*, 31, 187-198.
- National Institute of Child Health and Human Development. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Report of the National Reading Panel (NIH Publication No. 00-4769). Washington, DC: U.S. Government Printing Office.
- Oakes, J. (2005). *Keeping track: How schools structure inequality* (2nd ed.). New Haven, CT: Yale University Press.
- Olsen, L. (2010). *Reparable harm: Fulfilling the unkept promise of educational opportunity of California's long term English learners*. Long Beach: Californians Together. Available from www.californiantogether.org
- Ou, S., & Reynolds, A. J. (2010). Grade retention, postsecondary education, and public assistance. *Educational Evaluation and Policy Analysis*, 32, 118-139.
- Pappamihel, N. E., & Walser, T. M. (2009). English language learners and complexity theory: Why current accountability systems do not measure up. *Kappa Delta Pi*, 73, 133-140.
- Payán, R. M., & Nettles, M. T. (2008). *Current state of English-language learners in the U.S. K-12 student population*. Princeton, NJ: Educational Testing Service. Retrieved from http://www.ets.org/Media/Conferences_and_Events/pdf/ELLSymposium/ELL_factsheet.pdf
- Pearson Publishing. (2012). *Evaluación del desarrollo de lecto-escritura (EDL2)* [Developmental reading assessment (EDL2)]. Retrieved from <http://www.pearsonschool.com/index.cfm?locator=PSZu72&PMDBSUBCATEGORYID=&PMDBSITEID=2781&PMDBSUBSOLU>

- TIONID=&PMDBSOLUTIONID=6724&PMDBSUBJECTAREAID=&PMDBCATEGORYID=1662&PMDBProgramID=30941&elementType=attribute&elementID=142
- Public Law PL 107-110, the No Child Left Behind Act of 2001. Retrieved from <http://www2.ed.gov/policy/elsec/leg/esea02/index.html>
- Rayner, K., Foorman, B. R., Perfetti, C. A., Pesetsky, D., & Seidenberg, M. S. (2002). How should reading be taught? *Scientific American*, 286(3), 85-91.
- Reynolds, M. S., Wheldall, K., & Madelaine, A. (2010). Components of effective early reading interventions for young struggling readers. *Australian Journal of Learning Difficulties*, 15, 171-192.
- Romaine, S. (1995). *Bilingualism* (Language in society) (2nd ed.). Oxford, UK: Blackwell.
- Rosenthal, R., & Jacobson, L. (1992). *Pygmalion in the classroom: Teacher expectation and pupils' intellectual development*. New York, NY: Irvington Publishers.
- Shohamy, E. (2011). Assessing multilingual competencies: Adopting construct valid assessment policies. *Modern Language Journal*, 95, 418-429.
- Slavin, R., & Cheung, A. (2005). A synthesis of research on language of reading instruction for English language learners. *Review of Educational Research*, 75, 247-284.
- Smith, T. (2012). *Schools get tough with third-graders. Read or flunk*. Washington, DC: National Public Radio.
- Solano-Flores, G. (2008). Who is given what tests in what language by whom, when, and where? The need for probabilistic language in the testing of English language learners. *Educational Researcher*, 37, 189-199.
- Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 21, 360-407.
- Torgesen, J. K. (1998). Catch them before they fall: Identification and assessment to prevent reading failure in young children. *American Educator*, 22(1), 32-38.
- U.S. Department of Education. (2010). *Reading First*. Retrieved from <http://www2.ed.gov/programs/readingfirst/index.html>
- Valdés, G., & Figueroa, R. (1994). *Bilingualism and testing: A special case of bias*. Norwood, NJ: Ablex Publishing Corporation.
- Vera, H., Feagin, J. R., & Gordon, A. (1995). Superior intellect? Sincere fictions of the White self. *Journal of Negro Education*, 64, 295-306.
- Waesche, J. S. B., Schatschneider, C., Maner, J. K., Ahmed, Y., & Wagner, R. K. (2011). Examining agreement and longitudinal stability among traditional and RTI-based definitions of reading disability using the affected-status agreement statistic. *Journal of Learning Disabilities*, 44, 296-307.
- Wagner, R. K., Francis, D. J., & Morris, R. D. (2005). Identifying English language learners with learning disabilities: Key challenges and possible approaches. *Learning Disabilities Research & Practice*, 20, 6-15.
- Walberg, H. J., & Tsai, S. (1983). Matthew effects in education. *American Educational Research Journal*, 20, 359-373.
- Weber, W. (2001). *Developmental Reading Assessment and Evaluacion del desarrollo de la lectura: A Validation Study*. Available from www.pearsonlearning.com.
- Worthy, J. (2010). Only the names have been changed: Ability grouping revisited. *The Urban Review*, 42, 271-295.
- Zieky, M., & Perie, M. (2006). *A primer on setting cut scores on tests of educational achievement*. Princeton, NJ: Educational Testing Service. Retrieved from http://www.ets.org/Media/Research/pdf/Cut_Scores_Primer.pdf

Zoref, L., & Williams, P. (1980). A look at content bias in IQ tests. *Journal of Educational Measurement*, 17, 313-322.

Author Biographies

Susan Hopewell is an assistant professor of education in the division of Educational Equity and Cultural Diversity at the University of Colorado. She is interested in issues of language, culture, equity, and identity, especially as they impact—or are affected by—literacy practices.

Kathy Escamilla is a professor of education in the division of Educational Equity and Cultural Diversity. Her research centers on educational issues related to Spanish-speaking language minority students in U.S. schools. She is specifically interested in issues related to the development of bilingualism and biliteracy in early elementary grades for this Spanish-speaking population.