IQ-Achievement Discrepancy for Identification of Disabilities in Spanish-speaking English Learners

Kristi L. Santi, Ph.D.
Department of Educational Leadership and Policy Studies, College of Education, University of Houston

Shiva Khalaf, Ph.D.
Department of Psychology, and Texas Institute for Measurement, Evaluation, & Statistics, University of Houston

Ferenc Bunta, Ph.D.
Department of Communication Sciences and Disorders, University of Houston

Raúl Rojas, Ph.D.
Communication Sciences and Disorders Program, School of Behavioral Sciences, University of Texas at Dallas

David J. Francis, Ph.D.
Department of Psychology, and Texas Institute for Measurement, Evaluation, Statistics, University of Houston

Citation: Santi, K. L., Khalaf, S., Bunta, F., Rojas, R., & Francis, D. J. (2019). IQ-achievement discrepancy for identification of disabilities in Spanish-speaking English learners. In D. J. Francis (Ed.), *Identification, Classification, and Treatment of Reading and Language Disabilities in Spanish-speaking EL Students*. *New Directions for Child and Adolescent Development*, *166*, 111–143.

Acknowledgments

This work was supported by grant R324A160258 funded by the National Center for Special Education Research in the Institute of Education Sciences, and P50 HD052117, Texas Center for Learning Disabilities, from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD). The data reported here were collected under cooperative agreement R305P030031 funded by the US Department of Education and grant P01-HD39521 funded by NICHD. The attitudes and opinions expressed herein are those of the authors and do not reflect the position of the funding agencies or the Federal Government of the United States. We would like to thank the many teachers, students, parents, and colleagues who contributed to the original projects that made this research reported in this volume possible.

Abstract

This article examines the validity of IQ-achievement discrepancy and low-achievement as criteria for the identification of disabilities in Spanish-speaking ELs and the factors that moderate the validity of these approaches as bases for identification. While there has been a long history of examining the validity of different approaches to disability identification in monolinguals, there are no systematic approaches taken for ELs. Data from Grades 1 and 2 of a large longitudinal data set consisting of young Spanish-speaking students attending schools in the U.S. were used to empirically examine criteria for disability identification among language minority children - one of the first large-scale attempts. Findings indicated significant overidentification when the language of assessment was not matched to the language of the instruction, although effects varied predictably over time and by language of instruction. Validation of classifications using measures external to the classification found that low achieving and discrepant children differ from typically developing children, and from one another in predictable ways based on differences in IQ. The study highlights the importance of taking into account the language of instruction and the severity of the cut-off to reduce misidentification of typically developing children.

Keywords: English Learners, Special Education, IQ-achievement discrepancy, Overidentification

Introduction

Scientific advancement is impeded without consistency in definitions and classifications (Blashfield, 1993). The history of research on students with intellectual and developmental disabilities demonstrates that progress accelerated in the 70's and 80's, a time that Hallahan and Mercer (2002) have referred to as the "Solidification Period," following an evolving consensus around theoretical frameworks for identification that could be operationalized, such as discrepancy, and the abandonment of frameworks that could not, such as minimal brain dysfunction. This consensus led researchers during this period and into the early 21st century to focus on the operational definitions they used to include human subjects in research studies, some of which focused directly on the validation of those definitions (Morris, Fletcher, & Francis, 1993; Share, McGee, & Silva, 1989; Shepard, 1980; Siegel, 1992; Stanovich, Siegel, & Gottardo, 1997), on the efficacy of treatments and interventions (Fuchs, Bahr, & Rieth, 1989; Swanson, Hoskyn, & Lee, 1999; Vellutino, Scanlon, & Lyon, 2000), as well as basic research on the development of academic skills (Share, Jorm, MacLean, & Matthews, 1984; Share & Stanovich, 1995) and the failure of these skills to develop adequately or at a normative pace (Francis et al., 1996b).

This research ultimately contributed to the validation of operational definitions of disabilities of all varieties – reading, math, language, attention, and executive functioning. The use of consistent operational definitions across research laboratories and in a wide variety of intervention studies has contributed significantly to an improved understanding of the genetic (Olson, Forsberg, Gayan, & DeFries, 1999) and environmental/instructional factors (Fletcher, Lyon, Fuchs, & Barnes, 2007) that lead to impaired academic development. At the same time, validity research has generally showed a lack of discriminant validity between competing operational definitions of learning disabilities. The reliance on operational definitions in federal legislation eventually led to the President's Commission on Excellence in Special Education (US DOE-OSERS, 2002) and the pursuit of alternative conceptualizations, like response to intervention (RtI; Fletcher, Lyon, Fuchs, & Barnes, 2007; Fuchs & Fuchs, 2006; Vaughn & Fuchs, 2003) and cognitive process models (Johnson, Humphrey, Mellard, Woods, & Swanson, 2010). Interest in these alternatives has resulted in considerable research focused on their operationalization and validity as a means of disability identification, and has ultimately resulted in modifications to federal law in the United States, increasing the number of approaches that schools may use to identify learning disabilities in their students.

Amid this progress, an important area that has largely remained underdeveloped and problematic concerns the identification of disabilities among children who are non-native speakers of the societal language. Presently, federal law allows for the use of any approach to identification that is allowed with native English speakers, but neither the law, nor research has provided guidance on the validity of these operational definitions with language minority students. This paper seeks to fill this gap in the empirical literature by extending research on the validity of the most common disability identification method to Spanish-speaking English-language learners (ELs). Specifically, the purpose of this study is to investigate the validity of IQ-achievement discrepancy and low-achievement as criteria for the identification of disabilities in Spanish-speaking ELs and to examine the factors that moderate the validity of these approaches as bases for identification in this population of students.

Research on ELs in the United States has increased substantially since the turn of the century, including research on language and literacy development, instructional interventions, and educational assessments, as evidenced by the substantial body of research on ELs that has been published since the report of the National Literacy Panel (August & Shanahan, 2006). Conversely, although much has also been written about EL students with disabilities and the challenge of identification, not much data have been brought to the conversation. As such, the historical debate has been that ELs may be bypassed for disability services due to the belief that their achievement difficulties are grounded in language proficiency issues that will resolve with time, or they may be over-represented because of inaccurate measures and/or poor identification procedures (Abedi, 2006; Artiles, Rueda, Salazar, & Higareda, 2005).

A few issues impacting this debate include the lack of a clear and consistent definition across the U.S. and an absence of reliable and valid assessments to determine proper identification of Spanish-speaking ELs (McCardle, Keller-Allen, & Shuy, 2008; McCardle, Mele-McCarthy, & Leos, 2005; August & Hakuta, 1997). Recommendations have been made for stronger theoretical models, better identification tools, better assessment tools, and consistent terminology, definitions, and models (Wagner, Francis, & Morris, 2005). There is also the need to establish a more precise understanding of language and its development in language minority individuals (i.e., those growing up speaking a language other than the language of the society in which they live), literacy development in the majority and minority languages, quality of instruction, and understanding of individual and contextual factors that affect education outcomes (August & Shanahan, 2006; McCardle, Mele-McCarthy, & Leos, 2005). While these authors have addressed important dimensions to the challenge of disability identification in this population, none of these studies empirically investigated the use of approved identification procedures with ELs, as allowed under federal law in the US. Consequently, practitioners faced with the very real issue of having to decide on the disability status of specific children who are also language minorities have very little useful information to guide their decision making.

As a group, ELs are at risk for impaired reading and language development for reasons other than the existence of disabilities, including socioeconomic disadvantage and exposure to poor instruction in one or both languages. Morgan et al. (2015) found that racial, ethnic, and language minority elementary-and middle school students are less likely than similar white, English-monolingual students to be identified as having disabilities and are disproportionately underrepresented in special education. This study also reported that language-minority children were less likely to be identified as having learning disabilities or language impairments. Morgan et al.'s recent findings of underrepresentation stand in contrast to previously reported findings of overrepresentation of language minority students in special education (Artiles & Trent, 1994). Of course, these results reflect the outcome of school practices at the time of the research and serve to highlight the challenges faced by schools in identifying disabilities in language minority students.

To advance understanding of reading and language disability in ELs, a sound theoretical and empirically supported framework for identifying and classifying disabilities in ELs is essential. Researchers (Francis et al., 2005; Wagner, Francis, & Morris, 2005) have often pointed to the disconnect between methods of identification that discretize continuous distributions and conceptualizations of disabilities as qualitatively distinct types of learners. More recently, Swanson, Kudo, and Guzman-Orth (2016) undertook a novel approach to address the prevalence and stability of latent classes at risk for reading disabilities for students in k-3 whose first language is Spanish. While the findings reported aligned with monolingual studies on phonological and cognitive deficits underlying reading disabilities, the authors also point to the need for more work in the area of identification. Specifically, Swanson et al. debate the issues of using the *a priori* cut-off point of the 25th percentile as the basis for identification into a subgroup, an issue continuously highlighted by other researchers in the field (cf., Francis et al., 2005; Wagner, Francis, & Morris, 2005).

Even with the development of a theoretically and empirically supported framework, the problem remains that federal law allows for the use of discrepancy, low achievement, and RtI with ELs, but does not provide guidance on how to implement these approaches with ELs. In part, guidance is lacking because validation research on different operational definitions has not been systematically undertaken with EL students, even though they are the fastest growing subgroup of students in U.S. public schools (Kena et al., 2016: Condition of Education, 2016), comprising 9.3 percent of students nationally in 2013-2014, and as a group, are disproportionately at-risk for poor academic outcomes. In the decade from 2002-2003 to 2013-2014, all but 14 states saw an increase in the percentage of ELs in their student population. Although EL students in the U.S. speak many languages at home, Spanish is the most common language (76.5%) spoken by language minority students in the US. Other languages, such as Arabic, Cantonese, Mandarin Chinese, and Hmong, represent a much smaller fraction of the language minority population, but are nevertheless prominent in select locations around the country (Kena et al., 2016).

Language minority students are not unique to the United States. From a global perspective, the International Commission on Financing Global Education Opportunity (2016) reports ‘that half of all children in low- and middle-income countries are not taught in a language they speak’ (p.7). Given that the European Union (EU) recognizes 23 official languages, it is perhaps not surprising that PISA self-report data show that about 7% of the 15-year-old students in the EU report speaking a language other than the school language at home.

However, these data need to be interpreted with caution as they relate to the percentage of language minority students in a country. First, reporting that one "speaks a language at home" is not synonomous with that language being the student's native language. Second, there was substantial variability across countries in the reported percentage, with the 27 countries participating in 2012 reporting averages ranging from less than two-percent to 89 percent (EACEA, 2012). As one example, the United Kingdom reports that in England alone more than one million children between the ages of 5-18 represent over 360 languages in their school system (NALDIC, 2018). Even though two primary models of support[1] exist for small numbers of students in almost all EU countries, the EACEA also reported that few schools accommodate large numbers of students who do not speak the language of instruction at home. While it is difficult to ascertain the precise numbers of language minority students in school systems worldwide, suffice it to say that such students are not unique to the US, nor to English-speaking countries, nor is the challenge of identifying and classifying language minority students with disabilities.

Most would agree with the assertion that language disabilities are rooted in the neural substrate underlying language and its development. By definition, they are not the result of a poor learning environment, inadequate instruction, nor limited opportunity to learn. Similarly, reading disabilities are rooted in the neural substrate for reading, either in that component of the substrate that affects the acquisition of decoding and fluency, or the higher order cognitive skills that affect the processing of language presented in text. With the exception of disabilities arising

---

[1] EACA (2012) reports that two distinct models for integration are deployed in the EU. These have direct analogues in the US. Specifically, language minority students either receive direct integration with special support, which is analogous to structured English immersion in the US, or separate classes to work towards direct integration, which is more akin to separate ESL instruction prior to mainstreaming in the US.

from focal brain injuries, which might create selective impairments in one of an individual's already acquired languages (Chilosi et al., 2008; Hoeft, F., 2018), one would expect a developmental language disability to present in any language(s) the child speaks, and not be unique to one language. Similarly, reading disabilities are expected to manifest themselves in any language the child learns to read. Although differences across languages and orthographies may affect the specific way(s) in which disabilities in reading and language are evidenced behaviorally, these differences should not be misinterpreted to imply that the disability is present in one language, but not present in others.

Theory on language development (c.f., Genesee, 2008), literacy acquisition (c.f., Zeigler, & Goswami, 2005), and reading disability (c.f., Fletcher et al., 2007) dictate that a child with a language or reading disability will have their learning affected by that disability in any language they speak, read, or write, although the exact behavioral manifestations of their language or reading disability will be expressed differently across languages based on the characteristics of that language and/or its writing/sound system. That is to say, the existence of a reading or language disability will influence all languages read or spoken by the child, albeit in different ways. Precisely how these theoretical expectations can be meaningfully incorporated into operational approaches to the identification of disabilities in language minority children has not been widely researched.

Disability identification is a psychometric process, the validity of which can be investigated empirically for language minority children just as it has been for monolingual speakers of various languages. Whereas the fact that ELs are developing language and literacy in at least two languages simultaneously complicates the process of identification, the theoretical arguments linking language and learning disabilities to problems in the neural substrates that underlie language and/or literacy development of the child suggest that measurements in all of the child's languages should be capable of informing the identification of disabilities in EL children. The challenge is how to leverage this information in ways that lead to valid and reliable decisions about the presence of disability. This article focuses on Spanish-speaking ELs growing up and attending public school in the U.S. to attempt to empirically examine the reliability and validity of different approaches to disability identification in language minority students, and the factors that affect the operationalization of different definitions, such as the language in which the identification is based and the moderating effects of language of instruction, and overlap across approaches and languages will be investigated.

Research Questions

The primary objectives of this study were to examine IQ-achievement discrepancy and low achievement as the bases for identifying students at risk for reading problems. Different approaches to identification and classification of students with problems in the domain of reading were employed, including a comparison and contrast of different models and methods for classification and identification, examining different levels of severity, and examining differences between groups of students identified using these different approaches in the tradition of classical validation work on disability classification systems (Francis et al., 2005).

Unlike disability identification in language majority students, accurate identification in language minority students must consider the possible moderating influence of language of instruction. It stands to reason that a student's performance in any domain of skill in a particular language will be affected by the opportunity to learn that skill in that language. Hence, the first prediction is that language of instruction will significantly moderate the validity of any approach to identification such that failure to condition on the language of instruction will lead to significant misidentification of children using any approach to identification. At the same time, it is expected that severe low performance in a language may be pathognomic regardless of the language of instruction depending on the performance in the other language. For example, severe low performance in both English and Spanish may validly signal problems of learning, regardless of the child's history of instruction in either language. We also predict that students identified under low-achievement and IQ-achievement discrepancy definitions of reading disability will not differ from one another on external measures (i.e., those not used for identification), except to the extent that disabilities identified under discrepancy are more severe conditional on IQ (see Fletcher, Lyon, Fuchs, & Barnes, 2007, for extensive discussion of these issues among language majority children). However, unlike the use of these identification procedures among language majority speakers of English, we further predict that their validity will be significantly impacted by the child's history of instruction in English and Spanish.

Method

The sample for the current study was drawn from two longitudinal large-scale projects: Project BIL was Project 2 from the program project grant titled *Oracy/Literacy Development of Spanish-Speaking Children* in which ELs from kindergarten were recruited and followed through grade 2 and assessed at two time points in each academic school year. The second sample came from a project titled *Success through Academic Interventions in Language and Literacy* (SAILL) in which ELs were recruited in kindergarten and followed into Grade 3 and assessed at two time points in each school year. Because the sampling strategies were largely similar, and the measures, and timing of assessments were comparable across the two projects, we combined the samples for the present study, and focus only on data in Grades 1 and 2.

Across the studies, students were enrolled in various instructional programs, such as structured English immersion (SEI), transitional bilingual education (TBE), dual language (DL), and Spanish maintenance bilingual programs (ME). In the current study, the coding of the language of instruction was simplified to differentiate programs in which instruction was predominantly in English (i.e., SEI) from all other bilingual programs in which instruction was predominantly in Spanish (i.e., TBE, ME, or DL). More specifically, we differentiated English instructed from bilingual instructed programs because (1) the differences in instructional approaches between early and late TBE and ME were small prior to the grade when the transition occurred, typically around Grade 3 for early TBE, and (2) DL and ME were not included in the SAILL project. In the full sample, there are few students in DL or ME programs relative to TBE. For more detailed information about the total number of participants and merging procedure please see Francis et al. (this issue).

Participants

The participants in the current study included 3,440 ELs assessed in Grade 1 and Grade 2 from 506 classrooms and 41 schools, located in urban California, urban Texas, and nonurban Texas. ELs' Fall and Spring reading outcomes were selected for this study. The mean age for the participants was 6.7 years (*SD* = .44) at the beginning of Grade 1, and 7.65 (*SD* = .42) at the beginning of Grade 2. The sample was roughly equally divided by gender (51.6% male). Descriptive data are summarized in Table 1.

Measures

The measures in the current study were selected based on studies conducted by Francis and colleagues since 1992 (e.g., Fletcher et al., 1996; Foorman, Francis, Fletcher, & Lynn 1996; Foorman & Francis, 1994; Foorman et al., 1998; Francis, Fletcher, Shaywitz, Shaywitz, &

Rourke, 1996a; Francis, Shaywitz, Stuebing, Shaywitz, & Fletcher, 1996b; Schatschneider, Francis, Foorman, Fletcher, & Mehta, 1999), which have provided a good knowledge base of what and how to measure the key constructs that represent important reading-related skills. The measures were individually administered in English and in Spanish by trained examiners who spoke both English and Spanish fluently. Students were assessed in their stronger language first, and assessments were typically administered over a two-week period. To examine IQ-achievement discrepancy and low-achievement as the bases for identifying Spanish-speaking ELs at risk for reading problems, the following measures were used:

Classification of IQ-Achievement Discrepancy. To classify ELs as discrepant, their intellectual functioning was assessed using the Raven's Progressive Color Matrices (Raven, Raven, & Court, 2003). This non-verbal assessment of logical reasoning has been widely used in studies involving individuals from cultures speaking different languages. Students were provided instructions in Spanish or in English, based on the student's stronger language and the assessment was administered one-on-one. Internal consistency was estimated at .82 (standardized and unstandardized) using coefficient alpha.

Classification of Low Achieving. The classification measures for low achievement were based on the broad reading subtest of the English (WJ; Woodcock, 1991) and Spanish (WJM; Woodcock & Muñoz-Sandoval, 1995) Woodcock Language Proficiency Battery-Revised. The broad reading cluster is a composite score of Letter-Word Identification and Passage Comprehension, (see Francis et al., this issue, for detailed information about the measures). The composite scores were created separately for the English and Spanish measures according to procedures outlined in the respective test manuals.

Analytical Procedure

The primary method used in the classification analyses was linear regression for operationalizing IQ-achievement discrepancy. More specifically, we regressed the primary reading outcome measure (broad reading) on the measure of IQ (Raven's Progressive Color Matrices) using a standard regression model. The regression model created an expected score and a residual score. Identification was based on the standardized residual. Specifically, students whose standardized residual achievement score in a particular language fell below various pre-specified cut-points were identified as "IQ-Achievement Discrepant" in that particular language at the specified level of severity.

The criterion for defining a significant discrepancy reflected four degrees of severity in the classification process. Historically, research with monolingual students has used a standardized residual of varying degrees of severity. Because the law has changed to specify that the discrepancy not be severe, we specified values for standardized residuals corresponding to *p*-values of $\leq .20$, $\leq .15$, $\leq .10$, and $\leq .05$ as one-sided probabilities for *t* statistics based on 500 degrees of freedom, which correspond to values of -0.842, -1.038, -1.283, and -1.648, respectively. Furthermore, we established a criterion for designating students as low achieving. Traditionally, the literature has used criteria framed as normative percentiles with values ranging from the 25th percentile to the 10th percentile for monolingual students. Because this criterion is not model-based but simply severity-based, and because we are examining ELs, students were classified as low-achieving in reading when the criterion was below the cutoff ranging from the 20th percentile to the 5th percentile to coincide with the discrepancy severity cut-points.

Finally, for all grades and time points (Fall vs. Spring), we classified students using both the regression models and the low-achievement criteria simultaneously. This approach allowed us to determine the degree to which the two classes of criteria identify the same students, or different students. All classification procedures were carried out in SAS 9.4 using PROC GLM for the regression discrepancy criteria.

Results

Classification and the Language of Instruction

Classification results are presented in Tables 2.1a and 2.1b through 2.4a and 2.4b. Each table set (e.g., 2.1a and 2.1b) provides the results for a particular language in a particular grade, with the Table designated *a* providing the results for the Fall and Table designated *b* providing the results for the Spring for that grade and language. Table 2.1 provides results for English in Grade 1, Table 2.2 provides results for Spanish in Grade 1, Table 2.3 provides results for Englsih in Grade 2, and Table 2.4 provides results for Spanish in Grade 2. Each table shows the cross-classification of students according to the discrepancy and low achieving criteria at each of the designated levels of severity for that language, grade, and time point. For example, if a student met the regression discrepancy criterion with p (standardized residual) $< .18$, then that student met the criterion at the $\leq .20$ level, but not at the $\leq .15$ level. Likewise, if a student's achievement score was at the 13th percentile, the student was classified at the $\leq .15$ level, but not at the $\leq .10$ or $\leq .20$ level. In other words, students were classified at the most severe criterion that they met; this classification was carried out independently for the discrepancy criteria and for the low achievement criteria. Note that any child meeting a more severe criterion also met all lesser criteria, but for the purposes of the tables, they are counted in the most severe category whose criteria they met.

Each table provides three sets of joint classifications for low achieving (LA) and regression discrepant (RD). The first joint table provides the overall joint classification of students into discrepant and low achieving levels of severity without regard to language of instruction. The remaining two joint classifications provide the classification of RD and LA conditional on language of instruction. By comparing the three tables, it is possible to demonstrate the effects of ignoring language of instruction in disability identification. Each portion of the table has a similar organization. For example, Table 2.1a presents the joint distribution of students who were classified as discrepant or low achieving in English in the Fall of Grade 1. The table provides both the marignal classifications for RD and LA and their joint distribution. Specifically, the row labeled Total gives the marginal distribution of the RD classification, providing the sum of the unique cases in the column associated with each RD classification, including the column labeled Not RD, which shows the children who did not meet any of the discrepancy criteria. The column labeled Total gives the marginal distribution for LA. We also include a row labeled Subtotal LA that gives the number of students meeting any of the LA criteria at the RD classification associated with a particular column. Thus, the number in the column labeled Total in the row labeled Subtotal LA gives the total number of children with an LA designation, and the rows below that row give the number of LA children meeting each LA criterion. The total number of children meeting any RD classification requires subtracting the number of Not RD in the row labeled Total from the Table Total.

We predicted that language of instruction is a non-ignorable factor when it comes to classification of disabilities among EL students, regardless of the language of assessment. To test this prediction, the first step was to review the simultaneous cross classification of reading discrepant and low achievement within a language of assessment, ignoring the language of instruction and then conditioning on language, differentiating between those instructed in English and those instructed in Spanish. As mentioned, the total number of children meeting any RD classification requires subtracting the number of Not RD in the row labeled Total from the

Table Total. Looking at the Fall of Grade 1 and ingoring the language of instruction, we see in the Table 2.1a that the number of RD students is 2,406 – 1,817 = 589. From the portion of Table 2.1a that ignores Language of Instruction, we see that 54.1% of students were not identified as either discrepant (RD) or low achievig (LA) (1,304 of 2,406 students). Thus, 45.9% of students (1,102 of 2,406 students) were identified as either RD, or LA, or both. Ignoring language of instruction, 513 students (21% of the sample) were identified as LA, but not RD. Importantly, we see from the row Labeled Not LA that no children were identified as discrepant who were not also identified as LA. That is, all children who met a discrepancy criterion were also low achieving. Overall, we see that, ignoring the language of instruction, 46% (1,102 out of 2,406) of students tested in English met one of the two criteria for disabilities with the majority (n=711) scoring below the 5th percentile in achievement, and 589 (53% of children meeting an LA criterion) met both the RD and LA criteria.

The proportion of students meeting one or both criteria in English is absurdly high (over 45%) when ignoring the language of instruction. With well over half of those (n=711 or 1,102 or 64.5%) meeting the LA criterion at the 5th percentile, the problem is not simply having a criterion that is too liberal. Indeed, the problem stems from the failure to match the language of instruction to the language of assessment. Conditioning on the language of instruction shows that 59% of students tested in English, but instructed in Spanish (961 of 1,619), were identified as being LA, RD, or both as compared to just 18% (141 of 787) of students who were tested in English and instructed in English. We see the opposite pattern when we examine Table 2.2a which shows classification in the Fall of Kindergarten using Spanish outcomes. Ignoring the language of instruction, the overall identification rate is (2,441-1,462)/2,441 = 979/2,441 = 40%, whereas taking into account the language of instruction yields an identification rate of 412/1,658

= 24.8% for children instructed in Spanish and tested in Spanish, and 567/783 = 72.4% of students instructed in English and tested in Spanish. While we again see that no children met the discrepancy criteria who did not also meet the low achievement criteria in the Fall of Grade 1, the most compelling finding is the substantial impact of language of instruction on classification rates. Specifically the failure to match the language of instruction to the language of assessment leads to substantial over-identification. Although we cannot rule out that at least some disabled students meet the criteria when language of instruction and language of assessment are not matched, the rates of identification must be too high, because far fewer of these same children are identified when the language of assessment is matched to the language of instruction.

Repeating this process for the Spring of Grade 1 (see Table 2.1b for English and 2.2b for Spanish) indicated that there was an overall decline in the percentage of students identified (672/2,368 = 28.4% in English ignoring the language of instruction; 611/2,333 = 26.2% in Spanish ignoring the language of instruction), but the problem of overidentification when language of instruction was ignored remained. For example, the percentage of identified students in English was only was 92/811, or 11.3% when English was the language of instruction, but 580/1,557, or 37.3%, when Spanish was the language of instruction. In contrast, the precentage of students identified in Spanish was only 147/1,560, or 9.4% when Spanish was the language of instruction, in comparison to 464/773, or 60.0% for students instructed in English. In the Spring of Grade 1, we also saw that students remained unlikely to meet the discrepancy criteria unless they also met the low achieving criteria. However, in this case a handful of cases (n=3 in English and n=4 in Spanish) are identified as discrepant, but not low achieving.

For the Fall (Tables 2.3a and 2.4a) and Spring (Tables 2.3b and 2.4b) of Grade 2, we see identification rates continue to drop overall. Specifically, 555/2,339, or 23.7% of students tested in English met either the RD (n=5), LA (n=112) or both (n=438) criteria in the Fall of Grade 2, and 443/2,221, or 19.9% met either the RD (n=23), LA (n=58), or both (n=362) criteria in the Spring of Grade 2. In Spanish, the rates were similar in the Fall, 523/2276, or 22.98%, and slightly higher in the Spring, 534/2206, or 24.2%, but in both cases all students met either the LA criteria (n=38 in the Fall and n=71 in the Spring), or both the LA and RD criteria (n=485 in the Fall and n=463 in the Spring). That is, in Spanish, no children met the RD criteria only; all students who met the RD criteria also met the criteria for LA. Again, the proportion of students who met the criteria was far greater when students were tested in the language in which they were not being instructed. For English-instructed students tested in Spanish, 50%-54% met the criteria for identification (402/809 = 49.7% in the Fall; 410/763 = 53.7% in the Spring), whereas for Spanish-instructed students tested in English, the percentages were substantially lower (469 1558 = 30.1% in the Fall; and 368/1452 = 25.3% in the Spring), albeit still high compared to identifications based in the language of instruction. For English instructed students tested in English, identification rates were 86/781 = 11% in the Fall of Grade 2 and 75/769 = 9.8% in the Spring. For Spanish instructed students tested in Spanish, identification rates were even lower, 121/1467 = 8.2% in the Fall and 124/1443 = 8.6% in the Spring.

That identification rates of Spanish-instructed students tested in English drop from Fall to Spring and from Grade 1 to Grade 2 is consistent with the fact that students in Spanish-instruction are also learning to read in English, albeit at a slower pace than their English-instructed peers, whereas Engish instructed students are not generally being taught in school to read in Spanish. Thus, we do not see progression to the same degree in Spanish for English instructed students. These developmental relations make identification based exclusively on English outcomes questionable until after second grade for Spanish instructed students.

Nevertheless, the question remains if and how measures in both languages might inform identification for students in either Spanish or English instruction.

To answer this question, we examined the cross-classification of identifications based in English and Spanish. Because so few students met discrepancy criteria in a language and were not simultanesouly low achieving, we simplify the cross-language comparison by focusing on the LA classification, and examine these only while also conditioning on language of instruction. In Tables 3 and 4, we provide the cross-language classifications for students at the beginning of Grade 1 and at the end of Grade 2, respectively. In the top half of each table, we provide the cross-classification for Spanish-instructed students, and in the bottom half of the table is the cross-classification for English-instructed students. In looking at these cross-classifications, it is instructive to focus on those students scoring very low in the language of instruction, e.g., below the 5^{th} or 10^{th} percentile. For Spanish instructed students at the beginning of Grade 1, 277 of 337 students (82%) scoring below the 10^{th} percentile in Spanish (the language of instruction), also scored below the 10^{th} percentile in English, and 263 of those (95% of the 277; 74% of the 337) scored below the 5^{th} percentile in English. In contrast, of the 826 Spanish instructed students scoring below the 10^{th} percentile in English, only 32.0% (277/866) scored below the 10^{th} percentile in the language in which they were instructed. If we focus on English instructed students at the start of Grade 1, we see a similar pattern. In this case, of 89 students scoring below the 10^{th} percentile in English (i.e., the langauge of instruction), 76 (85.4%) scored below the 10^{th} percentile in Spanish, and of those 72 (94.7% of 76; 80.9% of the 89 students) scored below the 5^{th} percentile). It is also instructive to note that only 4.5% of students scoring below the 10^{th} percentile in English, had scores that were above the 20^{th} percentile in Spanish; the other 9 students who did not score below the 10^{th} percentile in Spanish had missing information on the

Spanish assessment, which may have signaled that they were unable to take the test in Spanish. At the same time, only 14.9% of the 510 students scoring below the 10th percentile in Spanish (i.e., the non-instructed language) also scored below the 10th percentile in English (the language of instruction). Thus, although low performance in the uninstructed language is not diagnostic on its own, it may be informative when conditioned on poor performance in the instructed language. In fact, of the 146 English-instructed students scoring below the 20th percentile in English, 110 (75.3%) scored below the 5th percentile in Spanish, and 120 (82.2%) scored below the 10th percentile. However, the strength of this relationship reduces with instruction, although the weakening of the relationship appears to be stronger for students instructed in Spanish. Specifically, by the end of second grade, of the 99 Spanish-instructed children scoring below the 10th percentile in Spanish, only 38 (38.3%) scored below the 10th percentile in English. Indeed, 52 (52.5%) of those scoring below the 10th percentile in Spanish scored above the 20th percentile in English, possibly signaling greater emphasis on English literacy among these students with low performance in Spanish. In contrast, of the 38 English-instructed students scoring below the 10th percentile in English, 30 (78.9%) also scored below the 10th percntile in Spanish, as compared to 7 (18.4%) who scored above the 20th percentile in Spanish.

One might wonder whether those students who are low performing in both the language of instruction and the other language are also those students who are low performing and IQ-Achievement discrepant in the language of instruction. Looking at the data for the Fall of Grade 1, that does not appear to be the case. For students instructed in Spanish who scored below the 10th percentile in Spanish, there were 277 students who also scored below the 10th percentile in English. Of those 277, 11 are missing the IQ score and thus are not scored on IQ-Achievement discrepancy. Of the remaining 266, 83 (31.2%) are not IQ-Achievement discrepant in Spanish, and 183 (68.8%) are discrepant, which is somewhat higher than the percentage of discrepant students (53.2%) among Spanish-instructed students meeting any low achievement criterion (219/412) in Spanish. For English-instructed students, there were 89 students who scored below the 10th percentile in English, of which 76 also scored below the 10th percentile in Spanish. Of those, 4 did not have IQ data. Of the remaining 72 students, exactly half were IQ-Achievement discrepant in English and half were not. This percentage is somewhat higher than the percentage of discrepant students (31.2%) among English-instructed students who met any criterion for low achievement in English (44/141), but we would be hard pressed to say that discrepancy in the language of instruction was strongly associated with being very low achieving in both languages.

Concurrent Validation

As a first step in external validation of the low achievement and discrepancy criteria in the Fall of Grade 1, we examined differences between groups on non-reading measures at the same time point, i.e., concurrent validation. For Spanish instructed students, we looked at both English and Spanish outcomes, and for English instructed students, we looked only at English outcomes. For validation, we focused exclusively on classifications in the language of instruction. That is, Spanish instructed students were classified based on their Spanish Broad Reading in Fall of Grade 1, and English instructed students were classified based on the English Broad Reading. To compare classifications, we constructed three groups based on the Spanish classifications and three groups based on English classifications. Specifically, students were classified as RD if they met any RD criterion, LA if they met any LA criterion and did not meet any of the RD criteria, NRI if they met none of the LA or RD criteria. English and Spanish classifications were examined separately using a mixed model analogue of profile analysis on four language measures: Listening Comprehension, Memory for Sentences, Picture Vocabulary, and Verbal Analogies. The approach used PROC MIXED in SAS 9.4 to obtain tests of the Elevation, Flatness, and Shape hypotheses of profile analysis while allowing for imbalance in the design due to missing data. We controlled for clustering within schools by including random intercepts at the school level and employed an arbitrary (i.e., unstructured) residual covariance matrix that allowed the residuals for the four outcomes to freely covary. We did assume a homogeneous covariance matrix for the residuals across the three groups.

Means and standard deviations and sample sizes for Spanish and English classifications and all outcome measures are presented in Table 5, which shows an fairly typical pattern of means across the three groups in both classifications and in all outcomes. Specifically, the NRI group outperforms on all language measures, followed by the LA group; the RD group has uniformly the poorest performance, which was expected because those meeting the RD classification are the lowest performing at any given level of IQ. Table 5 also shows that, for the most part, variances are quite comparable across groups, with ratios of standard deviations very close to 1.0. The largest difference yields a ratio of just 1.37.

For all three analyses, there was significant variation in school means, indicating that the inclusion of random school intercepts was warranted (school variances of 5.9 ($p = .0119$), 9.1 ($p = .0064$) for Spanish classifications and Spanish and English outcomes, respectively, and 10.9 ($p = .0116$) for English classifications and English outcomes. Given residual variances ranging from 176 to 293 in English outcomes for English classifications, 165 to 400 for Spanish classifications and English outcomes, and 147 to 760 for Spanish classifications and Spanish outcomes, the ICCs ranged from less than .01 to approximately .06, indicating that most of the variability is between students within schools rather than between schools. Residual correlations ranged from

.34 to .59, and from .55 to .73 for Spanish classifications for Spanish and English outcomes, respectively, and from .34 to .54 for English classifications for English outcomes.

Regardless of the classification (English or Spanish) or the language of the outcome, we found significant main effects for Groups and Measures, and a significant interaction of Group by Measure (i.e., we rejected the Shape hypothesis of parallel profiles across groups), all of which were statistically significant at $p < .0001$, with the exception of the interaction of Group by Measure for English classifications on English outcomes, which was statistically significant at $p = .016$. Because the profile shape hypothesis is an interaction hypothesis, rejection of this hypothesis takes precendence over the main effects for Group (i.e., equal Elevation hypothesis) and Measure (i.e., the Flatness hypothesis). In each case, we followed up the significant interaction with tests of simple effects of groups within measures using a Bonferroni adjusted p value of .05/4 = .0125, and followed significant simple effects with pairwise comparisons between the three groups at $p = .0125/3 = .0042$.

For Spanish classifications, all four simple effects were statistically significant a $p < .0001$, which was below the Bonferroni criterion, for both Spanish outcomes and for English outcomes, indicating that means for the three groups differed for each of the four measures in each of the two outcome languages. For English classifications and English outcomes, all four simple effects were also found to be significant at $p < .0001$. Thus, for all outcomes and classifications, pairwise comparisons for each outcome measure were warranted. Examining the Spanish classifications first, we found that the two disability groups (LA and RD) were both significantly different from the NRI group on all measures in Spanish ($p < .0001$), and on all measures in English ($p < .0001$). In contrast, the LA and RD groups differed on some measures in English and in Spanish, but not all. Specifically, we found that LA and RD groups in Spanish differed on measures of Listening Comprehension and Picture Vocabulary ($p < .0001$) in both Spanish and English, and on measures of Verbal Analogies in Spanish ($p = .0004$) and in English ($p = .0023$), they did not differ on Memory for Sentences in Spanish ($p = .3643$) or in English ($p = .8424$). For English classifications and English outcomes we found similar effects when comparing disability groups to typical readers. Specifically, both LA and RD groups differed from NRI ($p < .0001$) on all four outcomes. However, differences between LA and RD were generally not statistically significant at the adjusted *p* – value. Specifically, the difference on Listening Comprehension ($p = .0175$), Memory for Sentences ($p = .0760$), Picture Vocabulary ($p = .1882$), and Verbal Analogies ($p = .0422$), all failed to reach the adjusted criterion of $p < .0042$. This difference between the English classifications and what was found for the Spanish classifications appeared to have been the result of the smaller sample size in the English classifications moreso than differences in effect sizes. Effect size *d* based on the pooled within-groups standard deviation are provided at the bottom of Table 5. For three of four measures, the effect size was largest for the English classification, where effect sizes ranged from .17 to .38, whereas effect sizes for Spanish classifications ranged from .06 to .34 in the language of instruction and from .08 to .35 in English.

Discussion

The primary objective of the study was to examine IQ-achievement discrepancy and low achievement as the bases for identifying students at-risk for reading problems. We examined classifications at four occassions in two languages between the start of Grade 1 and the end of Grade 2. Language of instruction was hypothesized to significantly moderate the validity of any approach to identification such that failure to take into account language of instruction would lead to significant misidentification. The findings support this hypothesis for both grades and for both languages. For example, in the fall of Grade 1, using assessments administered in English, the rate of non-identification was only 54%, which meant that failing to consider the language of instruction would potentially identify 46% of students as disabled. This ‘overidentification’ was not found to be a problem of leniency in the criterion in so far as the preponderance of students were identified at the most severe criterion, namely the .05 criterion which identified 30% of students, or 711 of 2,406 students.

A similar problem was found in the Fall of Grade 1 when basing classification on Spanish language outcomes. Specifically, the rate of non-identification was 60% (1,462 of 2,441), meaning that 40% met either the LA or RD criteria or both. Again, 27% of all students (656 of 2,441) met the most stringent criterion of .05, implying that they scored below the 5th percentile for an LA designation, or had a resdiual that had an associated *p* value below .05, or both. Thus, the problem of overidentification was present regardless of whether identification took place in English or in Spanish. Although identification rates declined in Grade 2, the problem of overidentification was still apparent, with 24% of students identified and 10% of all students (42% of identified students) meeting the .05 criterion in English in the Fall of 2nd Grade, and 20% identified in the Spring of 2nd Grade, with 7% of the total at the .05 criterion. Spanish numbers were quite comparable at 23% and 24% identified (14% and 16% at the .05 criterion) in the in the fall and spring, respectively.

This high rate of identification appeared to stem from a failure to account for the language in which students receive their literacy instruction. Indeed, identification rates decrease substantially when language of instruction is taken into account and matched to the language in which the outcome is assessed. For example, in the Fall of Grade 1, only 18% of English-instructed students were identified on English language outcomes, with 7.9% of students falling below the .05 criterion. For Spanish instructed students assessed in Spanish, 25% of students were identified and 15% fell below the .05 criterion. In stark contrast, Spanish instructed students assessed in English were identified at a rate of 59% with 40% meeting the .05 criterion and English instructed students assessed in Spanish were identified at a rate of 72.4% with 52% identified at the .05 criterion. Although the numbers change somewhat as children age through the system, it seems quite clear that overidentification occurs whenever the language of instruction is not matched to the language of the outcome in which identification is based. Moreover, the problem is more pronounced when English instructed students are assessed in Spanish, and when young Spanish instructed students are assessed in English. As Spanish instructed students age through the system and have more exposure to literacy instruction, and no doubt more exposure to English even though they are being instructed in Spanish, the use of English assessment in identification is less problematic. The percentage of students identified dropped substantially when the language of instruction matched the language of assessment.

Theories of language development would stipulate that a student with a language or reading disability would exhibit this disability in all languages spoken or read by the student, although the precise manifestations of the disability may vary across languages based on characteristics of the languages. Although such theories would seem to make the language of assessment irrelevant because evidence of the disability should be present in all languages that the child speaks, establishing proper cut-offs or other specific criteria for differentiating low performance due to disability from under performance due to limited proficiency and limited opportunity to develop proficiency in that language remains challenging. When we examined performance in Spanish and English jointly within language of instruction, the results showed that low performance in the uninstructed language was informative conditional on poor performance in the language of instruction. Thus, using the uninstructed language to improve identification may be possible, but only if used in conjunction with the instructed language, because limited language proficiency in the language of assessment due to limited exposure and/or opportunity to learn as distinct from disability is a major, potential threat to the valid identification of disabilities based on test performance at a single time point.

The findings from this project have important implications for policy, as well as for teachers, administrators, speech-language pathologists, and school psychology personnel, including diagnosticians, working in school settings. These individuals frequently ask for guidance on identifying disabilities in children whose first language is not English. These individuals understand that factors other than the presence of a disability may affect student performance, but are unsure how to account for these factors when it comes to judging student performance and inferring whether or not a disability is present. The immediate practical implications of the work will be to help these professionals make more informed decisions about normative and non-normative performance in Spanish-speaking ELs and to provide them with clearer criteria for implementing acceptable definitions under 34CFR300.309 under IDEA 2004. This regulation requires demonstration that the child is low-achieving for their age, and continues to allow for IQ-Achievement discrepancy, provided that the state's criteria do not require demonstration of a severe discrepancy. The regulations also permit the use of a process based on the child's response to “scientific, research-based intervention” (i.e., RtI) and may permit the use of “other alternate research-based procedures” for identification prior to Grade 3. However, the state must also demonstrate that the student’s meeting of the criteria is not primarily due to the student’s “limited English proficiency.” This precondition, in and of itself, makes sense at first glance, except when one considers that limited proficiency in English could signal the presence of a disability in language acquisition. What is required is a means to ascertain the extent to which a student's proficiency in the language is consistent with developmental expectations given the specific student's opportunity to learn that particular language.

Certainly, developmental expectations for acquistion of any language will vary depending on whether or not the student is receiving instruction in that language, and the extent to which the student is learning other languages. There is no question that students can acquire proficiency in multiple languages simultaneously (c.f., Paradis, Genesee, & Crago, 2011), but learning rates vary across children and are slower, on average in a given language, for children learning multiple languages compared to children learning only that language. Having better information regarding developmental expectations for students growing up under different instructional contexts would assist researchers and practitioners in developing more effective approaches and criteria for identifying students with disabilities. Although much progress has been made in understanding the development of language and literacy in ELs in recent years, disability identification in ELs has received little empirical attention.

The current work has practical implications for researchers and school personnel interested in developing and testing early interventions for at-risk Spanish-speaking ELs by providing clear-cut selection criteria for including students in studies who are most at-risk for disabilities. Intervention studies based on samples with clearly identified risk characteristics will provide additional evidence for the various definitions as these studies accumulate over time. We expect this work will show, as in the case of research with monolingual students, that differences between low-achieving and IQ-achievement discrepant students are a function of the IQ differences between these groups. That is, for ELs, as for native English-speaking students, disabilities lie on a latent continuum of severity rather than represent latent class categories that differ qualitatively from one another.

Limitations

While the current study used a large, longitudinal dataset to empirically examine IQ-achievement discrepancy and low-achievement as bases for identifying Spanish-speaking ELs at risk for reading problems, it is limited in a number of ways that provide the foundation for several lines of future work. With regard to the generalizability of the findings, the current study is limited in investigating a single language subgroup in three geographical regions in the U.S. While Spanish-speaking children represent the largest proportion of ELs in U.S. public schools, it is critical to extend the current research to other student populations who speak a language other than English at home. For instance, other quickly growing populations include students who speak Arabic and Chinese. Also, this study examined the classification criteria for ELs in Grades 1 and 2. It is important to extend the time frame of the current work beyond grade 2 to the point that students read to learn (minimally grades 4-5). Research that integrates samples from children with different language backgrounds as well as varied instructional and environmental experiences is essential to inform the nature of learning disabilities and how to best assist these students.

This study was further limited in its use of only a single measure for the classification of IQ-achievement discrepancy and low achievement, and relied on a somewhat atypical measure of IQ from the standpoint of reading disability research. Future studies could extend the current findings by examining the classification and severity criteria using other intelligence and reading proficiency measures. Also, in the current study, types of academic instructional programs were simplified to differentiate whether reading instruction was predominantly in English or Spanish.

Thus, it failed to account for the amount of language exposure and instruction in English versus Spanish. Greater precision in characterizing language exposure could prove beneficial for developing more precise expectations for students in each language. Finally, the focus of the current study was identification and classification of reading disability using standard approaches allowed by federal regulations in the US. We used a concurrent validation methodology to examine differences beween identified groups on measures of language functioning at the same time point. Differences were consistent with the kinds of differences between discrepant and low achieving students in lanugage majority students, which comprise the bulk of research on reading disabilities. Further research is needed to validate these classifications across future time points, neurobiological measures, and using treatment outcomes. We expect that such efforts will replicate research with language majority students, but additional empirical research focused on reliability and validity of classifications with language minority students is clearly needed, as well as extension to language disabilities and math disabilities. Valid and reliable identification and classification is fundamental to early risk identification as well as to determining how best to intervene instructionally for ELs with disabilities.

References

Abedi, J. (2006). Psychometric issues in the ELL assessment and special education eligibility. *Teachers College Record, 108*(11), 2282–2303.

Artiles, A.J., & Trent, S.C. (1994). Overrepresentation of minority students in special education: A continuing debate. *The Journal of Special Education*, *27*(4), 410-437.

Artiles, A.J., Rueda, R., Salazar, J.J. & Higareda, I. (2005). Within- group diversity in minority disproportionate representation: English language learners in urban school. *Exceptional Children, 71*(3), 283–300.

August, D., & Hakuta, K. (1997). *Improving schooling for language- minority children: A research agenda*. Washington, DC: National Academy Press.

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

Chilosi, A.M., Cipriani, P., Pecini, C., Brizzolara, D., Biagi, L. Montanaro, D., Tosetti, M., & Cioni, G. (2008). Acquired focal brain lesions in childhood: Effects on development and reorganization of language. *Brain and Langauge, 106*(3), 211-225. https://doi.org/10.1016/j.bandl.2007.12.010

Blashfield, R.K. (1993). Models of classification as related to taxonomy of learning disabilities. In G. R. Lyon, D. B. Gray, J. F. Kavanagh, & N. A. Krasnegor (Eds.), *Better understanding learning disabilities* (pp. 17-26). Baltimore: Brookes.

Education, Audiovisual, and Culture Executive Agency (2012). *Key data on teaching languages at school in Europe: 2012*. Retrieved from: http://eacea.ec.europa.eu/education/eurydice. DOI:10.2797/83967

Fletcher, J.M., Francis, D.J., Stuebing, K.K., Shaywitz, B.A., Shaywitz, S.E., Shankweiler, D.P., Katz, L., & Morris, R. (1996). Conceptual and methodological issues in construct definition. In G.R. Lyon (Ed.), *Validating the constructs of attention, memory, and executive functions* (pp. 17-42). Baltimore: Paul H. Brookes.

Fletcher, J.M., Lyon, G.R., Fuchs, L.S., & Barnes, M.A. (2007). *Learning disabilities: From identification to intervention.* New York: Guilford.

Foorman, B.R., Fletcher, J.M., Francis, D.J., Carlson, C.D., Chen, D.T., Mouzaki, A., … Taylor, R.T. (1998). *Texas Primary Reading Inventory.* Houston, TX: Center for Academic and Reading Skills.

Foorman, B.R., & Francis, D.J. (1994). Exploring connections among reading, spelling, and phonemic segmentation during first grade. *Reading and Writing: An Interdisciplinary Journal*, 6, 65-91.

Foorman, B.R., Francis, D.J., Fletcher, J.M., & Lynn, A. (1996). Relation of phonological and orthographic processing to early reading: Comparing two approaches to regression-based, reading level-match design. *Journal of Educational Psychology, 88*(4), 639-652, doi: 10.1037/0022-0663.88.4.639.

Francis, D.J., Fletcher, J.M., Shaywitz, B.A., Shaywitz, S.E., Rourke, B.P. (1996a). Defining learning and language disabilities conceptual and psychometric issues with the use of IQ tests. *Language, Speech, and Hearing Services in Schools, 27*(2), 132-143, doi: 10.1044/0161-1461.2702.132

Francis, D.J., Fletcher, J.M., Stuebing, K.K., Lyon, G.R., Shaywitz, BA., & Shaywitz, S.E. (2005). Psychometric approaches to the identification of LD: IQ and achievement scores are not sufficient. *Journal of Learning Disabilities, 38*(2), 98-108, doi: 10.1177/00222194050380020101.

Francis, D.J., Shaywitz, S.E., Stuebing, K.K., Shaywitz, B.A., & Fletcher, J.M. (1996b). Developmental lag versus deficit models of reading disability: A longitudinal, individual growth curves analysis. *Journal of Educational Psychology, 88*(1), 3-17, doi: 10.1037/0022-0663.88.1.3.

Fuchs, L.S., Bahr, C.M., & Rieth, H.J. (1989). Effects of goal structures and performance contingencies on the math performance of adolescents with learning disabilities. *Journal of Learning Disabilities, 22*(9), 554-560, doi: 10.1177/002221948902200907.

Fuchs, D., & Fuchs, L.S. (2006). Introduction to responsiveness-to-intervention: What, why, and how valid is it? *Reading Research Quarterly, 4*(1), 93–99, doi: 10.1598/RRQ.41.1.4.

Genesee, F. (2010). Dual language development in preschool children. *Young English language learners: Current research and emerging directions for practice and policy*, 59-79.

Hallahan, D.P., & Mercer, C.D. (2002). Learning disabilities: Historical perspectives. In R. Bradley, L. Danielson, & D.P. Hallahan (Eds.), *Identification of learning disabilities: Research to practice* (pp. 1-67). Mahwah, NJ: Lawrence Erlbaum.

Hoeft, F. (2018). Is dyslexia the same in different countries and cultures? Understood.org.

International Commission on Financing Global Education Opportunity. (2016). *The learning generation: Investing in education for a changing world*. Free distribution online at http://report.educationcommission.org/report/.

Johnson, E.S., Humphrey, M., Mellard, D.F., Woods, K., & Swanson, H.L. (2010). Cognitive processing deficits and students with specific learning disabilities: A selective meta-analysis of the literature. *Learning Disability Quarterly, 33*(1), 3-18, doi: 10.1177/073194871003300101.

Kena, G., Hussar W., McFarland J., de Brey C., Musu-Gillette, L., Wang, X., Zhang, J., Rathbun, A., Wilkinson-Flicker, S., Diliberti M., Barmer, A., Bullock Mann, F., & Dunlop Velez, E. (2016). *The Condition of Education 2016*. (NCES 2016-144). U.S. Department of Education, National Center for Education Statistics. Washington, DC. Retrieved [date] from http://nces.ed.gov/pubsearch.

McCardle, P., Keller-Aleen, C., & Shuy, T. (2008). Learning disability identification. In E. Grigorenko (Ed.), *Educating Individuals with Disabilities* (pp. 137–164). New York, NY: Springer.

McCardle, P., Mele-McCarthy, J., & Leos, K. (2005). English language learners and learning disabilities: Research agenda and implications for practice. *Learning Disabilities Research and Practice, 20*(1), 68-78.

Morgan, P.L., Farkas, G. Hillemeier, M. M., Mattison, R, Maczuga, S. Li, H., & Cook, M. (2015). Minorities are disproportionately underrepresented in special education. *Educational Researcher, 44*(5), 278-292. doi: 10.3102/0013189X15591157

Morris, R.D., Fletcher, J.M., & Francis, D.J. (1993). Conceptual and psychometric issues in the neuropsychological assessment of children: Measurement of ability discrepancy and change. In I. Rapin & S. Segalovitz (Eds.), *Handbook of neuropsychology* (pp. 341–352). Amsterdam: Elsevier.

NALDIC. (2018). National Assoication for Language Development in the Curriculum. Retrieved from: https://naldic.org.uk/the-eal-learner/research-and-statistics/

Olson, R.K., Forsberg, H., Gayan, J. & DeFries, J.C. (1999). A behavioral-genetic analysis of reading disabilities and component processes. In R. M. Klein & P.A. McMullen (Eds.), *Converging methods for understanding reading and dyslexia* (pp. 133-153). Cambridge, MA: MIT Press.

Paradis, J., Genesee, F., & Crago, M. (2011). Dual language development and disorders: A handbook on bilingualism & second language learning. Baltimore, MD: Paul H. Brookes Publishing.

Raven, J., Raven, J.C., & Court, J.H. (2003). *Manual for Raven's Progressive Matrices and Vocabulary Scales*. San Antonio, TX: Harcourt Assessment.

Schatschneider, C., Francis, D.J., Foorman, B.F., Fletcher, J.M., & Mehta, P. (1999). The dimensionality of phonological awareness: An application of item response theory. *Journal of Educational Psychology, 91*(3)*,* 467–478. doi: 10.1037/0022-0663.91.3.439

Share, D.L., Jorm, A.F., MacLean, R. & Matthews, R. (1984). Sources of individual differences in reading acquisition, *Journal of Educational Psychology 76*(6), 1309-1324, doi: 10.1037/0022-0663.76.6.1309.

Share, D.L., McGee, R., & Silva, P.A. (1989). IQ and reading progress: A test of the capacity notion of IQ. *Journal of the American Academy of Child & Adolescent Psychiatry, 28*(1), 97-100, doi: 10.1097/00004583-198901000-00018.

Shepard, L. (1980). An evaluation of the regression discrepancy method for identifying children with learning disabilities. *Journal of Special Education, 14*(1), 79-91. doi: 10.1177/002246698001400108.

Siegel, L.S. (1992). An evaluation of the discrepancy definition of dyslexia. *Journal of Learning Disabilities, 25*(10), 618-629. doi: 10.1177/002221949202501001.

Stanovich, K.E., Siegel, L.S., & Gottardo, A. (1997). Converging evidence for phonological and surface subtypes of reading disability. *Journal of Educational Psychology, 89*(1), 114-127, doi: 10.1037/0022-0663.89.1.114

Swanson, H.L., Hoskyn, M., & Lee, C. (1999). *Interventions for students with learning disabilities: A meta-analysis of treatment outcome*. New York: Guilford.

Swanson, H.L., Kudo, M., & Guzman-Orth, D. (2016). Cognition and literacy in English language learners at risk for reading disabilities: A latent transition analysis. *Journal of Educational Psychology, 108*(6), 830-856. http://dx.doi.org/10.1037/edu0000102

U.S. Department of Education Office of Special Education and Rehabilitative Services (2002. *A New Era: Revitalizing Special Education for Children and Their Families.* Washington, DC, 2002.

Vaughn, S., & Fuchs, L.S. (2003). Redefining learning disabilities as inadequate response to instruction: The promise and potential problems. *Learning Disabilities Research and Practice, 18*(3), 137–146.

Vellutino, F.R., Scanlon, D.M., & Lyon, G.R. (2000). Differentiating between difficult-to-remediate and readily remediated poor readers: More evidence against the IQ-achievement discrepancy definition for reading disability. *Journal of Learning Disabilities, 33*(3), 223-238, doi: 10.1177/002221940003300302.

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 and Practice, 20*(1), 6-15.

Woodcock, R.W., & Johnson, M.B. (1991). *Woodcock–Johnson psycho-educational battery-Revised*. Allen, TX: DLM Teaching Resources.

Woodcock, R.W., & Muñoz-Sandoval, A.F. (1995). *Woodcock language proficiency battery-revised, Spanish form.* Chicago, IL: Riverside.

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across langauges: A psycholinguistic grain size theory. *Psychological Bulletin, 131*(1), 3-29.

Table 1
Demographic Characteristic of the Sample

		N	Percent
Total Participants		3440	100
Gender	Male	1775	51.6
	Female	1665	48.4
Ethnicity	Hispanic	3440	100
Geographical	Urban California:		
Region	Los Angeles	588	17.1
	Urban Texas:		
	Austin	235	6.8
	Houston	319	9.3
	Nonurban Texas:		
	Brownsville	2298	66.8
Instructional	Spanish Instructed	2260	65.7
Program	English Instructed	1180	34.3

Table 2.1a
Cross Classification: Regression Discrepancy (RD) vs. Low Achieving (LA) in Broad Reading in English for Fall Semester of Grade 1

	RD Criterion[a]					
LA Criterion[b]	**Not RD**	.20	.15	.10	.05	**Total**
Ignoring the Language of Instruction						
Not LA	**1304**	**0**	**0**	**0**	**0**	**1304**
Subtotal LA	**513**	**148**	**155**	**183**	**103**	**1102**
20	*99*	*0*	*0*	*0*	*0*	*99*
15	*116*	*0*	*0*	*0*	*0*	*116*
10	*162*	*12*	*2*	*0*	*0*	*176*
5	*136*	*136*	*153*	*183*	*103*	*711*
Total	**1817**	**148**	**155**	**183**	**103**	**2406**
Within English Instructed Programs						
Not LA	**646**	**0**	**0**	**0**	**0**	**646**
Subtotal LA	**97**	**17**	**18**	**6**	**3**	**141**
20	28	0	0	0	0	*28*
15	28	0	0	0	0	*28*
10	22	3	1	0	0	*26*
5	19	14	17	6	3	*59*
Total	**743**	**17**	**18**	**6**	**3**	**787**
Within Spanish Instructed Programs						
Not LA	**658**	**0**	**0**	**0**	**0**	**658**
Subtotal LA	**416**	**131**	**137**	**177**	**100**	**961**
20	71	0	0	0	0	*71*
15	88	0	0	0	0	*88*
10	140	9	1	0	0	*150*
5	117	122	136	177	100	*652*
Total	**1074**	**131**	**137**	**177**	**100**	**1619**

Note. Numerical criteria are mutually exclusive (i.e., children are counted in the table at the lowest criterion that they meet). [a]Student meets regression discrepancy criterion if the probability of the standardized residual is less than or equal to the criterion probability, and greater than the next lower criterion; [b]Achievement score as a percentile is less than or equal to the low achievement criterion, and greater than the next lower criterion.

Table 2.1b

Cross Classification: Regression Discrepancy (RD) vs. Low Achieving (LA) in Broad Reading in English for Spring Semester of Grade 1

	RD Criterion[a]					
LA Criterion[b]	**Not RD**	.20	.15	.10	.05	**Total**
Ignoring the Language of Instruction						
Not LA	**1696**	**3**	**0**	**0**	**0**	**1699**
Subtotal LA	**184**	**77**	**97**	**146**	**165**	**669**
20	*84*	*8*	*1*	*0*	*0*	*93*
15	*58*	*26*	*2*	*1*	*0*	*87*
10	*37*	*30*	*34*	*11*	*0*	*112*
5	*5*	*13*	*60*	*134*	*165*	*377*
Total	**1880**	**80**	**97**	**146**	**165**	**2368**
Within English Instructed Programs						
Not LA	**719**	**1**	**0**	**0**	**0**	**720**
Subtotal LA	**35**	**13**	**14**	**18**	**11**	**91**
20	*11*	*3*	*0*	*0*	*0*	*14*
15	*8*	*4*	*0*	*0*	*0*	*12*
10	*13*	*2*	*6*	*4*	*0*	*25*
5	*3*	*4*	*8*	*14*	*11*	*40*
Total	**754**	**14**	**14**	**18**	**11**	**811**
Within Spanish Instructed Programs						
Not LA	**977**	**2**	**0**	**0**	**0**	**979**
Subtotal LA	**149**	**64**	**83**	**128**	**154**	**578**
20	*73*	*5*	*1*	*0*	*0*	*79*
15	*50*	*22*	*2*	*1*	*0*	*75*
10	*24*	*28*	*28*	*7*	*0*	*87*
5	*2*	*9*	*52*	*120*	*154*	*337*
Total	**1126**	**66**	**83**	**128**	**154**	**1557**

Note. Numerical criteria are mutually exclusive (i.e., children are counted in the table at the lowest criterion that they meet). [a]Student meets regression discrepancy criterion if the probability of the standardized residual is less than or equal to the criterion probability, and greater than the next lower criterion; [b]Achievement score as a percentile is less than or equal to the low achievement criterion, and greater than the next lower criterion.

Table 2.2a

Cross Classification: Regression Discrepancy (RD) vs. Low Achieving (LA) in Broad Reading in Spanish for Fall Semester of Grade 1

	RD Criterion[a]					
LA Criterion[b]	**Not RD**	.20	.15	.10	.05	**Total**
Ignoring the Language of Instruction						
Not LA	**1462**	**0**	**0**	**0**	**0**	**1462**
Subtotal LA	**388**	**155**	**163**	**184**	**89**	**979**
20	*77*	*0*	*0*	*0*	*0*	*77*
15	*88*	*0*	*0*	*0*	*0*	*88*
10	*135*	*20*	*3*	*0*	*0*	*158*
5	*88*	*135*	*160*	*184*	*89*	*656*
Total	**1850**	**155**	**163**	**184**	**89**	**2441**
Within English Instructed Programs						
Not LA	**216**	**0**	**0**	**0**	**0**	**216**
Subtotal LA	**195**	**82**	**101**	**120**	**69**	**567**
20	*27*	*0*	*0*	*0*	*0*	*27*
15	*46*	*0*	*0*	*0*	*0*	*46*
10	*69*	*13*	*2*	*0*	*0*	*84*
5	*53*	*69*	*99*	*120*	*69*	*410*
Total	**411**	**82**	**101**	**120**	**69**	**783**
Within Spanish Instructed Programs						
Not LA	**1246**	**0**	**0**	**0**	**0**	**1246**
Subtotal LA	**193**	**73**	**62**	**64**	**20**	**412**
20	*50*	*0*	*0*	*0*	*0*	*50*
15	*42*	*0*	*0*	*0*	*0*	*42*
10	*66*	*7*	*1*	*0*	*0*	*74*
5	*35*	*66*	*61*	*64*	*20*	*246*
Total	**1439**	**73**	**62**	**64**	**20**	**1658**

Note. Numerical criteria are mutually exclusive (i.e., children are counted in the table at the lowest criterion that they meet). [a]Student meets regression discrepancy criterion if the probability of the standardized residual is less than or equal to the criterion probability, and greater than the next lower criterion; [b]Achievement score as a percentile is less than or equal to the low achievement criterion, and greater than the next lower criterion.

Table 2.2b
Cross Classification: Regression Discrepancy (RD) vs. Low Achieving (LA) in Broad Reading in Spanish for Spring Semester of Grade 1

	RD Criterion[a]					
LA Criterion[b]	**Not RD**	.20	.15	.10	.05	**Total**
Ignoring the Language of Instruction						
Not LA	**1722**	**4**	**0**	**0**	**0**	**1726**
Subtotal LA	**74**	**101**	**112**	**130**	**190**	**607**
20	*46*	*17*	*0*	*0*	*0*	*63*
15	*22*	*38*	*3*	*0*	*0*	*63*
10	*6*	*37*	*35*	*2*	*0*	*80*
5	*0*	*9*	*74*	*128*	*190*	*401*
Total	**1796**	**105**	**112**	**130**	**190**	**2333**
Within English Instructed Programs						
Not LA	**309**	**2**	**0**	**0**	**0**	**311**
Subtotal LA	**45**	**78**	**76**	**102**	**161**	**462**
20	*30*	*11*	*0*	*0*	*0*	*41*
15	*12*	*28*	*3*	*0*	*0*	*43*
10	*3*	*31*	*25*	*2*	*0*	*61*
5	*0*	*8*	*48*	*100*	*161*	*317*
Total	**354**	**80**	**76**	**102**	**161**	**773**
Within Spanish Instructed Programs						
Not LA	**1413**	**2**	**0**	**0**	**0**	**1415**
Subtotal LA	**29**	**23**	**36**	**28**	**29**	**145**
20	*16*	*6*	*0*	*0*	*0*	*22*
15	*10*	*10*	*0*	*0*	*0*	*20*
10	*3*	*6*	*10*	*0*	*0*	*19*
5	*0*	*1*	*26*	*28*	*29*	*84*
Total	**1442**	**25**	**36**	**28**	**29**	**1560**

Note. Numerical criteria are mutually exclusive (i.e., children are counted in the table at the lowest criterion that they meet). [a]Student meets regression discrepancy criterion if the probability of the standardized residual is less than or equal to the criterion probability, and greater than the next lower criterion; [b]Achievement score as a percentile is less than or equal to the low achievement criterion, and greater than the next lower criterion.

Table 2.3a

Cross Classification: Regression Discrepancy (RD) vs. Low Achieving (LA) in Broad Reading in English for Fall Semester of Grade 2

	RD Criterion[a]					
LA Criterion[b]	**Not RD**	.20	.15	.10	.05	**Total**
Ignoring the Language of Instruction						
Not LA	**1784**	**5**	**0**	**0**	**0**	**1789**
Subtotal LA	**112**	**119**	**91**	**94**	**134**	**550**
20	*67*	*28*	*3*	*0*	*0*	*98*
15	*41*	*43*	*14*	*1*	*0*	*99*
10	*4*	*47*	*53*	*18*	*0*	*122*
5	*0*	*1*	*21*	*75*	*134*	*231*
Total	**1896**	**124**	**91**	**94**	**134**	**2339**
Within English Instructed Programs						
Not LA	**695**	**0**	**0**	**0**	**0**	**695**
Subtotal LA	**22**	**20**	**16**	**12**	**16**	**86**
20	*12*	*6*	*0*	*0*	*0*	*18*
15	*10*	*6*	*1*	*0*	*0*	*17*
10	*0*	*8*	*13*	*3*	*0*	*24*
5	*0*	*0*	*2*	*9*	*16*	*27*
Total	**717**	**20**	**16**	**12**	**16**	**781**
Within Spanish Instructed Programs						
Not LA	**1089**	**5**	**0**	**0**	**0**	**1094**
Subtotal LA	**90**	**99**	**75**	**82**	**118**	**464**
20	*55*	*22*	*3*	*0*	*0*	*80*
15	*31*	*37*	*13*	*1*	*0*	*82*
10	*4*	*39*	*40*	*15*	*0*	*98*
5	*0*	*1*	*19*	*66*	*118*	*204*
Total	**1179**	**104**	**75**	**82**	**118**	**1558**

Note. Numerical criteria are mutually exclusive (i.e., children are counted in the table at the lowest criterion that they meet). [a]Student meets regression discrepancy criterion if the probability of the standardized residual is less than or equal to the criterion probability, and greater than the next lower criterion; [b]Achievement score as a percentile is less than or equal to the low achievement criterion, and greater than the next lower criterion.

Table 2.3b

Cross Classification: Regression Discrepancy (RD) vs. Low Achieving (LA) in Broad Reading in English for Spring Semester of Grade 2

	RD Criterion[a]					
LA Criterion[b]	**Not RD**	.20	.15	.10	.05	**Total**
Ignoring the Language of Instruction						
Not LA	**1778**	**22**	**1**	**0**	**0**	**1801**
Subtotal LA	**58**	**74**	**97**	**76**	**115**	**420**
20	*42*	*24*	*10*	*2*	*0*	*78*
15	*16*	*31*	*43*	*7*	*0*	*97*
10	*0*	*19*	*42*	*27*	*1*	*89*
5	*0*	*0*	*2*	*40*	*114*	*156*
Total	**1836**	**96**	**98**	**76**	**115**	**2221**
Within English Instructed Programs						
Not LA	**694**	**4**	**0**	**0**	**0**	**698**
Subtotal LA	**13**	**14**	**17**	**10**	**17**	**71**
20	*10*	*6*	*2*	*0*	*0*	*18*
15	*3*	*5*	*7*	*1*	*0*	*16*
10	*0*	*3*	*7*	*5*	*0*	*15*
5	*0*	*0*	*1*	*4*	*17*	*22*
Total	**707**	**18**	**17**	**10**	**17**	**769**
Within Spanish Instructed Programs						
Not LA	**1084**	**18**	**1**	**0**	**0**	**1103**
Subtotal LA	**45**	**60**	**80**	**66**	**98**	**349**
20	*32*	*18*	*8*	*2*	*0*	*60*
15	*13*	*26*	*36*	*6*	*0*	*81*
10	*0*	*16*	*35*	*22*	*1*	*74*
5	*0*	*0*	*1*	*36*	*97*	*134*
Total	**1129**	**78**	**81**	**66**	**98**	**1452**

Note. Numerical criteria are mutually exclusive (i.e., children are counted in the table at the lowest criterion that they meet). [a]Student meets regression discrepancy criterion if the probability of the standardized residual is less than or equal to the criterion probability, and greater than the next lower criterion; [b]Achievement score as a percentile is less than or equal to the low achievement criterion, and greater than the next lower criterion.

Table 2.4a

Cross Classification: Regression Discrepancy (RD) vs. Low Achieving (LA) in Broad Reading in Spanish for Fall Semester of Grade 2

	RD Criterion[a]					
LA Criterion[b]	**Not RD**	.20	.15	.10	.05	**Total**
Ignoring the Language of Instruction						
Not LA	**1753**	**0**	**0**	**0**	**0**	**1753**
Subtotal LA	**38**	**94**	**102**	**130**	**159**	**523**
20	*37*	*19*	*0*	*0*	*0*	*56*
15	*1*	*56*	*0*	*0*	*0*	*57*
10	*0*	*19*	*70*	*0*	*0*	*89*
5	*0*	*0*	*32*	*130*	*159*	*321*
Total	**1791**	**94**	**102**	**130**	**159**	**2276**
Within English Instructed Programs						
Not LA	**407**	**0**	**0**	**0**	**0**	**407**
Subtotal LA	**29**	**66**	**83**	**104**	**120**	**402**
20	*28*	*8*	*0*	*0*	*0*	*36*
15	*1*	*41*	*0*	*0*	*0*	*42*
10	*0*	*17*	*57*	*0*	*0*	*74*
5	*0*	*0*	*26*	*104*	*120*	*250*
Total	**436**	**66**	**83**	**104**	**120**	**809**
Within Spanish Instructed Programs						
Not LA	**1346**	**0**	**0**	**0**	**0**	**1346**
Subtotal LA	**9**	**28**	**19**	**26**	**39**	**121**
20	*9*	*11*	*0*	*0*	*0*	*20*
15	*0*	*15*	*0*	*0*	*0*	*15*
10	*0*	*2*	*13*	*0*	*0*	*15*
5	*0*	*0*	*6*	*26*	*39*	*71*
Total	**1355**	**28**	**19**	**26**	**39**	**1467**

Note. Numerical criteria are mutually exclusive (i.e., children are counted in the table at the lowest criterion that they meet). [a]Student meets regression discrepancy criterion if the probability of the standardized residual is less than or equal to the criterion probability, and greater than the next lower criterion; [b]Achievement score as a percentile is less than or equal to the low achievement criterion, and greater than the next lower criterion.

Table 2.4b
Cross Classification: Regression Discrepancy (RD) vs. Low Achieving (LA) in Broad Reading in Spanish for Spring Semester of Grade 2

	RD Criterion[a]					
LA Criterion[b]	**Not RD**	.20	.15	.10	.05	**Total**
	Ignoring the Language of Instruction					
Not LA	**1672**	**0**	**0**	**0**	**0**	**1672**
Subtotal LA	**71**	**101**	**90**	**113**	**159**	**534**
20	*53*	*0*	*0*	*0*	*0*	*53*
15	*18*	*46*	*0*	*0*	*0*	*64*
10	*0*	*55*	*15*	*0*	*0*	*70*
5	*0*	*0*	*75*	*113*	*159*	*347*
Total	**1743**	**101**	**90**	**113**	**159**	**2206**
	Within English Instructed Programs					
Not LA	**353**	**0**	**0**	**0**	**0**	**353**
Subtotal LA	**52**	**76**	**71**	**91**	**120**	**410**
20	*39*	*0*	*0*	*0*	*0*	*39*
15	*13*	*35*	*0*	*0*	*0*	*48*
10	*0*	*41*	*13*	*0*	*0*	*54*
5	*0*	*0*	*58*	*91*	*120*	*269*
Total	**405**	**76**	**71**	**91**	**120**	**763**
	Within Spanish Instructed Programs					
Not LA	**1319**	**0**	**0**	**0**	**0**	**1319**
Subtotal LA	**19**	**25**	**19**	**22**	**39**	**124**
20	*14*	*0*	*0*	*0*	*0*	*14*
15	*5*	*11*	*0*	*0*	*0*	*16*
10	*0*	*14*	*2*	*0*	*0*	*16*
5	*0*	*0*	*17*	*22*	*39*	*78*
Total	**1338**	**25**	**19**	**22**	**39**	**1443**

Note. Numerical criteria are mutually exclusive (i.e., children are counted in the table at the lowest criterion that they meet). [a]Student meets regression discrepancy criterion if the probability of the standardized residual is less than or equal to the criterion probability, and greater than the next lower criterion; [b]Achievement score as a percentile is less than or equal to the low achievement criterion, and greater than the next lower criterion.

Table 3

Cross-Classification of LA in Spanish and English as a Function of Language of Instruction for Fall of Grade 1

Spanish Instructed

LA Criteria in English	LA Criteria in Spanish						
	Missing	**> 20**	**≤ 20**	**≤ 15**	**≤ 10**	**≤ 5**	**Total**
Missing	48	41	4	4	8	27	132
> 20	16	642	5	3	7	13	686
≤ 20	0	74	0	1	0	0	75
≤ 15	0	86	1	0	0	5	92
≤10	2	125	7	4	5	9	152
≤ 5	18	326	34	33	58	205	674
Total	84	1294	51	45	78	259	1811

English Instructed

LA Criteria in English	LA Criteria in Spanish						
	Missing	**> 20**	**≤ 20**	**≤ 15**	**≤ 10**	**≤ 5**	**Total**
Missing	51	1	0	2	2	39	95
> 20	35	216	27	44	73	276	671
≤ 20	2	3	1	0	3	20	29
≤ 15	0	6	0	1	3	18	28
≤10	2	4	0	0	2	18	26
≤ 5	7	0	0	0	2	54	63
Total	97	230	28	47	85	425	912

Table 4

Cross-Classification of LA in Spanish and English as a Function of Language of Instruction for Spring of Grade 2

Spanish Instructed

LA Criteria in English	LA Criteria in Spanish						
	Missing	**> 20**	**≤ 20**	**≤ 15**	**≤ 10**	**≤ 5**	**Total**
Missing	1	13	0	0	0	2	16
> 20	17	1058	10	9	11	41	1146
≤ 20	0	58	1	1	1	1	62
≤ 15	1	79	1	0	2	3	86
≤10	2	64	0	2	1	6	75
≤ 5	7	92	2	7	2	29	139
Total	28	1364	14	19	17	82	1524

English Instructed

LA Criteria in English	LA Criteria in Spanish						
	Missing	**> 20**	**≤ 20**	**≤ 15**	**≤ 10**	**≤ 5**	**Total**
Missing	0	0	0	0	0	2	2
> 20	8	350	37	51	57	221	724
≤ 20	0	5	1	0	1	12	19
≤ 15	0	2	0	0	1	13	16
≤10	0	2	1	0	0	12	15
≤ 5	0	5	0	0	1	17	23
Total	8	364	39	51	60	277	799

Table 5

Means and Standard Deviations for Concurrent Validation of Fall Grade 1 Reading Classifications

Spanish Classification – English Outcomes

GROUP	**Listening Comprehension**			**Memory for Sentences**			**Picture Vocabulary**			**Verbal Analogies**		
	N	**Mean**	**s.d.**	**N**	**Mean**	**s.d.**	**N**	**Mean**	**s.d.**	**N**	**Mean**	**s.d.**
LA	192	40.9	18.4	115	57.4	13.4	187	42.8	23.2	191	73.3	13.8
RD	216	36.7	19.8	112	56.6	13.2	192	38.6	23.7	215	68.8	12.5
NRI	1,238	48.8	20.2	494	63.1	12.4	1,231	50.5	22.0	1,232	80.9	14.5

Spanish Classification – Spanish Outcomes

	Listening Comprehension			**Memory for Sentences**			**Picture Vocabulary**			**Verbal Analogies**		
	N	**Mean**	**s.d.**	**N**	**Mean**	**s.d.**	**N**	**Mean**	**s.d.**	**N**	**Mean**	**s.d.**
LA	193	79.1	15.0	117	74.8	11.3	192	76.5	26.2	192	83.5	14.8
RD	218	73.0	19.2	114	73.8	12.7	214	67.9	27.3	218	78.8	17.1
NRI	1243	89.4	15.6	498	83.1	12.3	1237	89.3	27.5	1244	94.8	14.1

Table 5 (cont.)

English Classification – English Outcomes

GROUP	Listening Comprehension			Memory for Sentences			Picture Vocabulary			Verbal Analogies		
	N	**Mean**	**s.d.**	**N**	**Mean**	**s.d.**	**N**	**Mean**	**s.d.**	**N**	**Mean**	**s.d.**
LA	95	58.2	17.0	56	75.3	13.5	97	68.6	18.1	95	80.0	12.7
RD	44	51.3	22.7	29	71.7	13.2	43	65.3	21.7	44	75.1	12.9
NRI	644	72.2	16.6	293	84.6	13.7	645	82.6	17.5	644	90.7	13.6

Effect Sizes Comparing LA and RD

Classification –Outcome Langauge	***d***	***d***	***d***	***d***
Spanish-Spanish	.22	.06	.18	.34
Spanish-English	.35	.08	.32	.29
English-English	.36	.27	.17	.38

NOTE: Outcomes are shown in English and in Spanish for Spanish Classifications, and in English only for English Classifications. Classifications and outcomes are from the Fall of Grade 1. LA = Low Achievement; RD = IQ-Achievement Discrepant; NRI = Typical Reader not meeting either LA or RD criterion. Students are classified RD if they meet any RD criterion, LA if they meet any

LA criterion and do not meet any RD criteria, NRI if they meet none of the LA or RD criteria. Effect sizes are computed using the pooled within-groups standard deviation.