UC Riverside

Title
Classification Accuracy of Spanish CBMs for Students in a Dual Immersion Program

Permalink
https://escholarship.org/uc/item/8tx8w1p9

Author
Suchilt, Luisana

Publication Date
2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
RIVERSIDE

Classification Accuracy of Spanish CBMs for Students in a Dual Immersion Program

A Thesis submitted in partial satisfaction
of the requirements for the degree of

Master of Arts

in

Education

by

Luisana Suchilt

June 2019

Thesis Committee:
     Dr. Cathleen A. Geraghty, Chairperson
     Dr. Rollanda O'Connor
     Dr. Asha Jitendra

The Thesis of Luisana Suchilt is approved:

_____

_____

_____
                                                Committee Chairperson

University of California, Riverside

ABSTRACT OF THE THESIS

Classification Accuracy of Spanish CBMs for Students in a Dual Immersion Program

by

Luisana Suchilt

Master of Arts, Graduate Program in Education
University of California, Riverside, June 2019
Dr. Cathleen Geraghty, Chairperson

A vast amount of research has supported the use of English CBMs to identify students at-risk of future reading difficulties. However, less research has focused on using Spanish CBMs with students receiving bilingual instruction. The purpose of this study was to assess the classification accuracy of a Spanish oral reading fluency CBM (FLO) with a group of third grade students participating in a dual immersion (DI) program and receiving 30 percent of instruction in English and 70 percent in Spanish. Overall, findings suggest that Spanish CBMs are effective tools for predicting reading performance for students who are receiving instruction in English and Spanish. Analyses revealed that the FLO measure can also be used with students with varying levels of English proficiency, including students who are English Language Learners (ELLs), English Only (EO), and Reclassified. Despite these findings, additional research is needed in order to further determine best practices when working with this designated population of students. Limitations of the study and future directions are discussed.

Table of Contents

List of Figures

Classification Accuracy of Spanish CBMs for Students in a Dual Immersion Program

The importance of early literacy skills in the development of reading abilities is indisputable. Studies have consistently demonstrated strong associations between early reading skills and later reading outcomes (NRP, 2000). Schatschneider and colleagues (2004) found that phonological awareness, letter naming speed, and letter sound knowledge in kindergarten predicted passage comprehension and word identification skills in first and second grade. Likewise, students' growth in oral reading fluency in first grade was strongly associated with letter-naming and letter sound fluency in kindergarten. Similar findings have been reported for Spanish-speaking English language learners (ELLs). Yesil-Dagli (2011) found that English letter naming fluency, vocabulary skills, and initial sound fluency at the beginning of kindergarten predicted ELL students' English oral reading fluency and growth in the first grade. Findings from Chiappe and colleagues (2002) further indicated that alphabetic knowledge and phonological processing contribute to early reading skills for both native English speakers and ELLs.

Importance of Identifying At-Risk Students

Given the significant implications that early literacy abilities have for future reading achievement, identifying poor readers early on is essential. Students who struggle with reading during their early years in school have a greater disadvantage in catching up with their grade-level peers and becoming proficient readers (Stanovich, 1986). In her seminal study, Juel (1988) demonstrated that students identified as poor readers in first grade had a .88 probability of being poor readers in the fourth grade. Likewise, McNamara and colleagues (2005) found that a group of children identified as requiring reading-based support in kindergarten were more likely to be falling further behind their

grade-level peers in first grade. A study by Ford and colleagues (2013) that examined characteristics that could make students more susceptible to developing reading difficulties indicated that Spanish-speaking ELL students who entered school with lower English reading skills made slower growth in reading from kindergarten to first grade, as opposed to students who entered school with higher reading abilities. Generally, when entering kindergarten, ELL students were also more likely to score lower across all early English literacy measures compared to non-ELL students. This finding further underscores the importance of early identification of struggling readers in order to provide remediation and prevent future reading difficulties.

Aside from being at-risk of continuous reading difficulties, students identified as struggling readers are also susceptible to further disadvantage. For example, Juel (1988) found that poor readers were also more likely to become poor writers. Additionally, students who have reading difficulties are more likely to develop negative attitudes towards reading or display lower levels of motivation (McKenna, Kear & Ellsworth, 1995). Lower expectations of success and finding little value in a reading task and its rewards may further contribute to reading deficits. Lepola and colleagues (2000) found that students who persistently struggled in reading from first to second grade displayed low task orientation and higher ego-defensiveness. That is, they were less willing to engage in a task and more likely to show avoidance behaviors. In contrast, students who start off first grade with strong reading abilities are more likely to engage in more reading activity throughout the elementary years and thereafter (Cunningham & Stanovich, 1997). Overall, students who struggle to read are less likely to engage in reading tasks that will offer the opportunity to practice their reading skills (Anderson, Wilson &

Fielding, 1988), thus further contributing to their lack of growth in reading skills and possibly fueling their negative attitudes toward reading.

ELLs and Early Literacy

Identifying struggling readers early on in order to further prevent reading deficits is critical. Students who are not native English speakers and are learning to speak English constitute a group of students who continue to demonstrate difficulties in literacy. For this reason, identifying ELLs who present early reading underachievement is important. However, in order to best understand how to best address the needs of ELLs, there are various factors that need to be taken into consideration. Some of these factors are subsequently addressed.

Who are ELLs? Approximately 4.8 million students enrolled in US public schools are ELLs (National Center for Education Statistics, 2018), with Spanish speakers representing 77% of this population. As defined by the National Clearinghouse for English Language Acquisition (2006), an English Language Learner is "any K-12 student for whom English is not a first language and who requires language support in the classroom in order to access instructional content." ELL students are simultaneously acquiring their first language and the language of the larger community and thus are also referred to as dual language learners (DLLs). For example, a student who speaks Spanish at home, but is learning English in school, would be considered a DLL. ELL students who have not achieved English language proficiency are said to be Limited English Proficient (LEP) and are thus eligible for Language Instruction Educational Programs supported by the U.S. Department of Education.

Instruction. In order to best address the needs of ELLs, schools typically provide these students with instruction that is either immersive or bilingual. English immersion models of instruction provide instruction in English only, while bilingual programs provide instruction in both English and Spanish. Sheltered English instruction and structured English immersion are both common approaches funded by Title III that provide instruction in English only. Simply put, both of these models focus on making English instruction comprehensible to learners by using simplified forms of the language. The goal is not to teach English, but instead to increase students' content knowledge. In contrast, bilingual programs such as dual language and transitional bilingual provide instruction in both languages. The dual language program was designed to serve both native-Spanish speakers trying to learn English and native-English speakers trying to learn Spanish. The goal is to help students become biliterate. Distinctively, transitional bilingual programs only serve ELL students, as they are primarily taught in their home language and English is taught as a second language. The goal is for students to master educational content in their primary language and then assist students in becoming fluent in their second language. The basis for this approach is that students will more easily transfer skills from their native language to English (Goldenberg, 2008).

Research examining the impact of these programs on ELLs' academic achievement have generally found neutral or positive effects. A longitudinal study found that from kindergarten to second grade, ELLs in a dual language immersion (DI) program had significantly lower ELA test scores than ELLs in an English only immersion program. However, by grade seven, the students in DI and in the transitional program had higher ELA scores than ELLs enrolled in English immersion (Valentino & Reardon,

2015). Findings from a meta-analysis indicated that although there is evidence to support English only and bilingual instructional approaches, more research is needed to "demonstrate the efficacy of any particular curriculum or instructional strategy for use with DLLs" (Buysse et al., 2014).

Academic Achievement. Despite language support provisions provided by schools, ELL students continue to show lower academic achievement compared to native English speakers. When they enter kindergarten, ELLs consistently score lower in reading assessments compared to their English proficient peers (Barrow & Markman-Pithers, 2016). Students who enter kindergarten with limited English proficiency also have large and persistent deficiencies in English reading achievement throughout their elementary years (Kieffer, 2008). Chiappe and colleagues (2002) found that ELLs in kindergarten showed lower performance in phonological tasks, particularly rapid automatic naming and rhyme detection, and were not comparably fluent in manipulating and interpreting English oral language as their native English-speaking peers. More alarming is that gaps in achievement remain even in the 8th grade and can eventually lead to students dropping out of high school. According to the National Center for Education Statistics, in 2007 only 29% of ELLs scored at or above basic on reading achievement tests in the 8th grade compared to 75% of non-ELLs. From 2012-2013 the high school graduation rate of ELLs was 61% compared to 81% of non-ELLs (Barrow & Markman-Pithers, 2016). Such statistics make it evident that having tools to identify struggling ELLs early on is important.

Assessing ELLs

Curriculum-based measurements (CBMs) are brief formative assessments designed to measure academic skills. Originally developed to help special education teachers make modifications to their instruction (Deno, 1985), CBMs are now commonly used for screening and identifying students at-risk of academic difficulties, in order to determine the level of intervention needed by students. A substantial amount of research has supported the use of CBMs for screening and identifying students academically at-risk, as well as for predicting performance on high-stakes assessments (Deno, 2003, Reschly et al., 2009; Wayman et al., 2007). In general, literature reviews have indicated CBMs to be good indicators of how well students will perform on reading achievement tests, including statewide reading tests (Reschly et al., 2009; Wayman et al., 2007).

English CBMs. Various studies have explored whether English CBMs can be effectively used to predict the reading achievement of ELL students. An early study by Baker and Roland (1995) examining the validity of CBMs with English-only and bilingual students found a strong association between CBM reading measures and criterion reading measures for both English-only and bilingual students. Overall, results indicated that English CBMs could be validly and reliably used with both groups of students. In a more recent study, Yesil-Dagli (2011) examined how performance on CBMs of letter naming fluency (LNF) and initial sound fluency (ISF) in kindergarten predicted oral reading fluency in first grade. Findings indicated that LNF and ISF predicted ELLs' first grade oral reading fluency, individually and combined. Yeo (2010) conducted a meta-analysis to assess the relation between CBM and reading performance on a statewide achievement test in reading. CBMs were found to be valid for predicting

future reading performance. However, it was noted that in studies that included a large proportion of ELLs, correlation coefficients were lower than in studies with low proportions of ELLs. Such findings suggest that there may be other factors contributing to the effectiveness of CBMs with ELLs.

For instance, a few studies have also examined a bias for home language within CBMs. Klein and Jimerson (2005) found that in first through third grade, English oral reading fluency probes overpredicted scores on a reading achievement test for students whose home language was Spanish. Likewise, when examining the validity of the MAZE in predicting the performance of Spanish-speaking students on a high-stakes language arts test, Richardson and colleagues (2012) found that the MAZE was more accurate in predicting state test scores for English speakers than Spanish speakers. That is, the MAZE was more likely to under identify Spanish speakers who may have been in need of intervention services.

More recent research has also assessed whether the predictive validity of CBMs varies by ELLs levels of English language proficiency (ELP). In 2013, Gutierrez and Vanderwood assessed the extent to which ELLs' ELP affected reading levels and growth on measures of phonemic awareness (PA), letter-sound correspondence (NSW), and oral reading fluency in English. Findings indicated that throughout second grade, students in the Advanced levels of ELP made significantly more growth on oral reading fluency measures than students with lower ELP levels. While rates of growth for PA for students with Advanced, Early Advanced and Intermediate levels of ELP did not differ, students with Beginning levels of ELP performed significantly lower on PA measures throughout second grade. Students with advanced ELP levels also had steeper growth rates in letter-

sound correspondence. These results not only suggest that ELLs' growth in reading English is dependent upon their English proficiency, but also that CBMs must be sensitive to the heterogeneity that exists among the ELL population.

Kim, Vanderwood, and Lee (2016) investigated the predictive validity and accuracy of CBMs with third-grade Spanish-speaking ELLs of varying English proficiency levels who were receiving instruction in English. Findings indicated that the DIBELS Next Oral Reading Fluency (DORF) measure administered in the Fall significantly predicted performance in the English Language Arts state assessment taken in the Spring. In addition, the authors found that the predictive accuracy of DORF varied by English proficiency level. In the Fall, 69% of students with Beginning and Early Intermediate levels of English proficiency who did not meet state standards in the Spring were identified correctly. For students with Intermediate levels of proficiency, this sensitivity dropped to 54% and further dropped to 26% for students in Early Advanced and Advanced groups. No students in the Beginning and Early Beginning groups scored at or above expectations on the state test so specificity was 0%. The specificity scores went up to 81% for the Intermediate group and 87% for the Early advanced and Advanced groups. These results suggest that although an oral reading fluency measure can be utilized to predict which students are at-risk of not meeting standards on a state test, the reading measure was better at predicting which students would reach proficiency in the Spring. Therefore, it may be necessary to consider utilizing additional measures in order to increase the identification accuracy of at-risk students.

Spanish CBMs. Given the increased awareness that various factors could be contributing to the effectiveness of CBMs in predicting reading outcomes for ELLs,

studies have also taken into account language of instruction when administering CBMs. With Spanish being the home language of 77% of all ELL students in the United States, studies have examined the validity of using reading CBMs (R-CBM) with Spanish-speaking ELL students receiving bilingual instruction.

Ramirez and Shapiro (2006) assessed the oral reading fluency growth rates of Spanish-speaking ELLs in a transitional bilingual program and non-ELL students in general education receiving instruction in English only. As previously described, the transitional bilingual program, which only serves ELLs, consists of content-area instruction in Spanish with a gradual introduction of English. ELL students entering school receive the majority of instruction in Spanish, but this proportion decreases as students move up grades and are exposed to more English instruction and acquire more English literacy skills. Oral reading fluency during the Fall, Winter, and Spring was measured in English and Spanish in a sample of 165 students across first through fifth grade. ELLs were assessed in both languages, while non-ELLs in general education were assessed only in English. Findings indicated that non-ELL students showed greater oral reading fluency growth in English than students in the transitional bilingual group. These students also outperformed Spanish-speaking ELLs in oral reading fluency in English at all time points and across all grades. In addition, the reading gains made by the non-ELL students in English were higher than the reading gains Spanish-speaking ELLs made in Spanish. Therefore, although Spanish-speaking students showed growth in their oral reading fluency, it was not commensurate to the growth seen in the English-speaking general education group. These findings suggest that Spanish-speaking ELLs may be

making progress at a slower rate in English, and thus it may be necessary to monitor their progress in both English and Spanish.

Expanding on previous findings, Ramirez and Shapiro (2007) investigated whether Spanish oral reading fluency in the Fall predicted English oral reading fluency in the Spring. Sixty-eight Spanish-speaking ELLs in a transitional bilingual program were assessed with English and Spanish R-CBMs at three different time points, Fall, Winter, and Spring. The authors note that students new to the United Sates who only spoke Spanish may have been provided with additional language supports, but do not specify any further. Findings indicated that Spanish oral reading fluency and English oral reading fluency were significantly correlated at all times points and across all grades, except fourth grade. It was noted that outlier data points and curriculum changes in the fourth grade may explain why Spanish and English oral reading skills were not correlated in grade four. Moreover, further analyses revealed that Spanish oral reading fluency in the Fall was a significant predictor of English oral reading fluency in the Spring, with Fall performance in Spanish explaining 68.6% of the variation of English performance in Spring. These results provide further support for the use of Spanish R-CBMs with Spanish-speaking ELLs, as assessing Spanish reading skills may be necessary for predicting future English reading skills.

In a more recent study, Keller-Margulis and colleagues (2012) assessed the relation between R-CBMs in Spanish and a statewide measure of achievement, as well as the predictive accuracy of Spanish R-CBMs for identifying students at-risk for reading difficulties in English. Oral reading fluency of third and fourth grade students in a transitional bilingual education program was assessed in Spanish in the Fall, Winter, and

Spring. At the end of the year, students completed the Spanish version of the *Texas Assessment of Knowledge and Skills* (TAKS), a nationally standardized test. Findings indicated significant correlations between the Spanish R-CBM and the Spanish TAKS, with coefficients ranging from .41 to .48 and .37 to .44 in third and fourth grade, respectively. However, diagnostic accuracy analyses revealed that R-CBMs may be better at detecting students who will likely meet standards on the TAKS than those who will not. In other words, the R-CBM in Spanish was not as accurate at identifying students who would not meet standards on the state test.

More recent research by Keller-Margulis and Mercer (2014) assessed the relation between initial benchmark scores and growth on English and Spanish R-CBMs in ELL students in a transitional bilingual program. The sample consisted of 3,559 students in first through fifth grade. Results indicated that across all grades, initial level of performance on the Spanish R-CBM was higher than performance on the English R-CBM, a finding that is consistent with Ramirez and Shapiro (2007). Findings also demonstrated that higher initial scores were related to higher growth rates across the year, except in first grade. In addition, students across all grades made more growth in English than in Spanish during the school year. However, results also revealed a weaker association between initial R-CBMs in Spanish and English for students who perform very poorly or very well. According to the authors, this finding further suggests the need to assess students in both English and Spanish in order to improve data-based decision making.

Present Study

Although a substantial amount of research has demonstrated that CBMs can be effective tools for identifying students at-risk of future reading underachievement, most research has targeted English CBMs and less has focused on Spanish CBMs. In addition, although studies have investigated the relation between English and Spanish CBMs with reading achievement for ELLs receiving transitional bilingual education, less research has focused on ELL students in a dual immersion program (DI). As previously mentioned, transitional bilingual programs only serve ELLs, while DI programs are designed for both native-English speakers and native-Spanish speakers. Consequently, additional research is needed to assess the association between Spanish CBMs and reading achievement for students in a DI program. Given this lack of research, the purpose of this study is to examine whether Spanish CBMs can be used to identify students in a DI program who may be at-risk of future reading difficulties.

Research Questions

1. What is the classification accuracy of a Spanish CBM for identifying students in a DI program who may be at-risk of future reading difficulties?

2. Does the classification accuracy of the Spanish CBM vary based on the student's level of English proficiency?

Method

Participants

Participants included a sample of 55 students in third grade from a school in Southern California. Students were native Spanish-speakers learning to speak English (ELLs; $n = 16$), native English-speakers learning to speak Spanish (EO; $n = 22$) and

students who were previously ELLs but were reclassified to fluent English proficient (Reclassified; $n = 17$). All students included in the study were participating in the school's Dual Language Immersion (DI) program, an instructional method that begins with 90% Spanish and 10% English instruction in kindergarten and gradually decreases to 50% Spanish and 50% English by fifth grade. The goal of the program is for students to be bilingual by the end of fifth grade. That is, be able to read, write, and communicate effectively in English and Spanish. Students from this study were receiving 70% of instruction in Spanish and 30% in English. The ethnic distribution of the school as a whole was 66% Hispanic, 13% White, 10% African American, 2% Asian, and 9% Other. In addition, 66% of the students in the school identified as socioeconomically disadvantaged and 27% as ELLs.

Measures

ELPAC. The English Language Assessments for California (ELPAC) is a test of English language proficiency administered to ELLs annually. The test assesses four, English domains: listening, speaking, reading, and writing. It comprises an initial assessment used to identify ELLs, and a summative assessment designed for both annual progress monitoring and determining a students' level of English proficiency. The initial test yields an overall scaled score and the English performance level of a student. The performance levels are Novice, Intermediate, and Initial Fluent English Proficient (IFEP). The summative test produces a composite score for each English domain and the overall English performance level of a student. Performance levels range from one to four corresponding respectively to the following: minimally developed, somewhat developed, moderately developed, and well-developed English skills. Content on the ELPAC is

aligned with California's 2012 English Language Standards adopted by the State Board of Education.

IDEL. The *Indicadores Dinámicos del Éxito en la Lectura* (IDEL; Baker, Good, Knutson & Watson, 2006) are brief standardized measures designed to assess basic early literacy skills in Spanish. They are derived from the *Dynamic Indicators of Basic Early Literacy Skills* (DIBELS; Good & Kaminski, 2002) a set of standardized measures developed to assess basic early literacy skills in English. IDEL measures are used as screening tools to identify students at-risk for reading difficulties. They have been researched and validated specifically for benchmark testing in Kindergarten through 3rd grade. IDEL measures are not a translation of the DIBELS measures, but instead consider Spanish linguistic and orthographic systems.

FLO. The IDEL measure used in this study was the Fluidez en la Lectura Oral (FLO), a Spanish test of accuracy and fluency with reading connected text, and an indicator of advanced phonics and word attack skills. Students are presented with one unfamiliar grade-level passage and asked to read for one minute. An accuracy score is calculated based on the student's number of words read correctly and the number of errors. This score falls in the Benchmark, Strategic, or Intensive category, with each category corresponding to a distinct probability of a student meeting future benchmark goals and need for support. A score in the Benchmark category indicates that the student has a greater than 80% chance of achieving subsequent benchmark goals and a low probability of needing support. A score in the Strategic category indicates that the student has a 50% chance of achieving subsequent benchmark goals and a moderate probability of needing support. Lastly, a score in the Intensive category indicates that the student has

a less than 20% chance of achieving subsequent benchmark goals and a high probability of needing support. The IDEL technical manual reports alternate-form reliability coefficients for the FLO measure ranging from .87 to .94 for students in first through third grade, and concurrent criterion-related validity with the Woodcock-Muñoz, Batería-R Aprovechamiento en la Lectura ranging from .73 to .80.

i-Ready Diagnostic. The i-Ready Diagnostic online assessment pinpoints students' reading abilities and areas of need in five domain areas: phonological awareness, phonics, high-frequency words, vocabulary, and comprehension. The test yields a score for a student's overall reading performance. This score places students in one of the following three categories: On or Above Grade Level, One Grade Level Below, or Two or More Grade Levels Below. The i-Ready Diagnostic has been previously correlated with various state assessments. These studies have found that the test can be used to accurately predict a student's performance on statewide testing, with correlation coefficients ranging from .72 to .90. The i-Ready Diagnostic is a component of the i-Ready online program, which has previously met the criteria as an evidence-based intervention under the Every Student Succeeds Act (ESSA) Level 3: Promising Evidence.

Data Collection. Data used in this study were collected by trained school staff. The school district routinely collects IDEL data for students in the DI program as part of their adoption of a Response to Intervention (RTI) framework. During the year students are tested at three time points, Fall, Winter, and Spring. Data for this study were collected during the Winter. All students participating in the DI program were administered the FLO measure individually. Additionally, students completed the online i-Ready

Diagnostic assessment as a whole-class. Lastly, data were also obtained for each student to determine their status as an ELL, EO, or Reclassified student.

  Data Analyses. The receiver operating characteristic (ROC) curve is commonly used to assess the classification performance of a measure (Swets, 1988; Fawcett, 2006). The curve provides a visualization of how a well a classifier predicts an outcome by plotting the classifier's true positive rate (sensitivity) against its false positive rate (1-specificity) for every possible classification threshold. The true positive rate refers to how often the classifier predicts a positive outcome when the actual outcome is positive. The false positive rate indicates how often the classifier predicts a positive outcome when the actual outcome is negative. The ROC curve typically falls above, on, or below a diagonal line. This line represents a classifier that does no better than chance at predicting an outcome. A ROC curve above the diagonal line indicates that the classifier does a better job than chance at predicting an outcome.

  In this study, a ROC curve was computed for the entire sample, and for each group of English proficiency—ELLs, EOs, and Reclassifed. The FLO measure was used as the classifier and performance on the i-Ready Diagnostic was the outcome. The IDEL's benchmark goal for the FLO was used as the cut-off score to determine whether a student was considered to be at-risk for reading difficulties. A student scoring below the benchmark goal was considered at-risk of not meeting grade level standards on the i-Ready assessment, a negative outcome. A score above the benchmark indicated a student was likely to meet grade level standards on the i-Ready, a positive outcome. Performance on the i-Ready was set as a dichotomous outcome variable with students either meeting or not meeting grade level standards.

Furthermore, in order to quantify the performance of the classifier, the Area Under the Curve (AUC) was computed for each level of English proficiency. The AUC is an index of accuracy indicating how well the classifier can distinguish between two groups (e.g., at-risk, not at-risk). The closer the AUC is to 1, the better the ROC curve is at predicting an outcome. A classifier with an area of .5 or less would be considered a poor classifier (Swets, 1988). Lastly, the AUC of each group was compared using DeLong, Delong, and Clarke-Pearson's (1988) nonparametric approach. By comparing these areas, it can be assessed whether the classification accuracy of the FLO significantly varied based on student's level of English proficiency.

Results

Research Question 1: Classification Accuracy of the FLO

The AUC was calculated to quantify the classification performance of the FLO. In total, 22 students achieved a passing score on the i-Ready and 33 did not. As shown in Figure 1, the AUC for all the DI students was .914. This means that 91.4% of the time, the FLO correctly discriminated between an at-risk and not at-risk student. Therefore, if the FLO screener classified a student as not at-risk, the likelihood that that student would meet grade level standards on the i-Reading Diagnostic was high. This suggests that the FLO is a valid indicator when used with students in a DI program to predict reading performance on a reading diagnostic assessment.

Research Question 2: Classification Accuracy of the FLO by Level of English Proficiency

Further analyses were conducted to examine the classification accuracy of the FLO for each distinct English proficiency group— ELL, EO, and Reclassified. The AUC

for each group was calculated and comparisons between groups were conducted to test for significance.

AUC Comparisons. In the ELL group, four students achieved a passing score on the i-Ready and 12 did not. The AUC for this group was 1.00, indicating high accuracy. In other words, if an at-risk and not at-risk were randomly chosen from a group, the FLO would correctly classify the two students in this random pair 100% of the time. In the EO group, 11 students achieved a passing score on the i-Ready assessment and 11 did not. Results yielded an AUC of .843 for this group, indicating moderate accuracy. That is, for students within the EO category, the FLO accurately discriminated between an at-risk and not at-risk student 84.3% of the time. Lastly, in the Reclassified group, seven students achieved a passing score of the i-Ready and 10 did not. The AUC for this group was .964, also indicating high accuracy.

ELL and EO. Comparisons of the independent AUC curves for the ELL and EO groups, shown in Figure 2, indicated that the difference between the two areas under the curve was not statistically significant ($p = .104$). That is, the classification accuracy of the FLO was similar for both ELL and EO students.

ELL and Reclassified. Additional comparisons of the independent AUC curves for the ELL and Reclassified groups, shown in Figure 3, indicated that the difference between the two areas under the curve was not statistically significant ($p = .372$). That is, the classification accuracy of the FLO was similar for both ELL and Reclassified students.

EO and Reclassified. Lastly, comparisons of the independent AUC curves for the EO and Reclassified groups, shown in Figure 4, indicated that the difference between the

two areas under the curve was not statistically significant ($p = .246$). That is, the

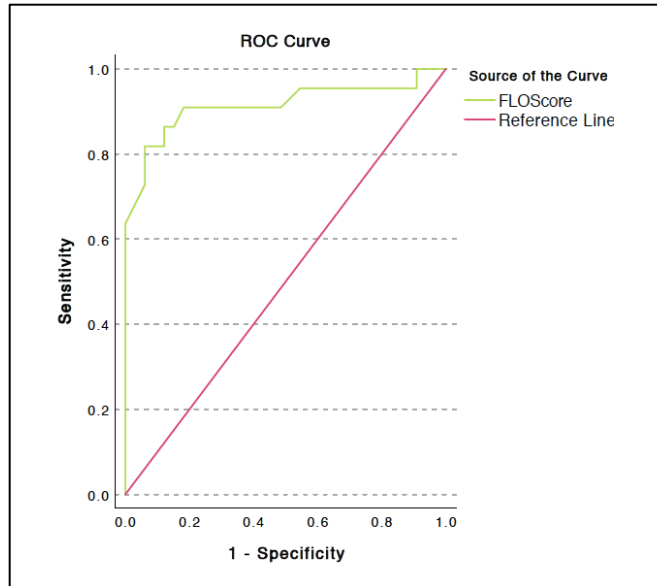classification accuracy of the FLO was similar for both EO and Reclassified students.

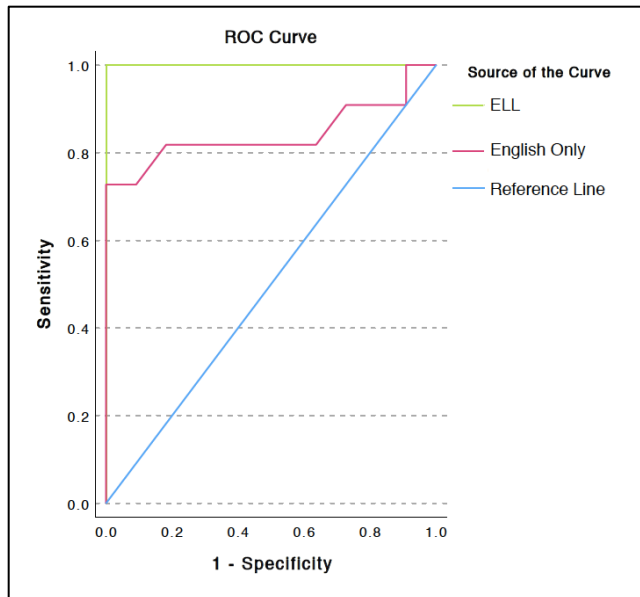*Figure 1.* Non-parametric ROC curve for the total
sample.

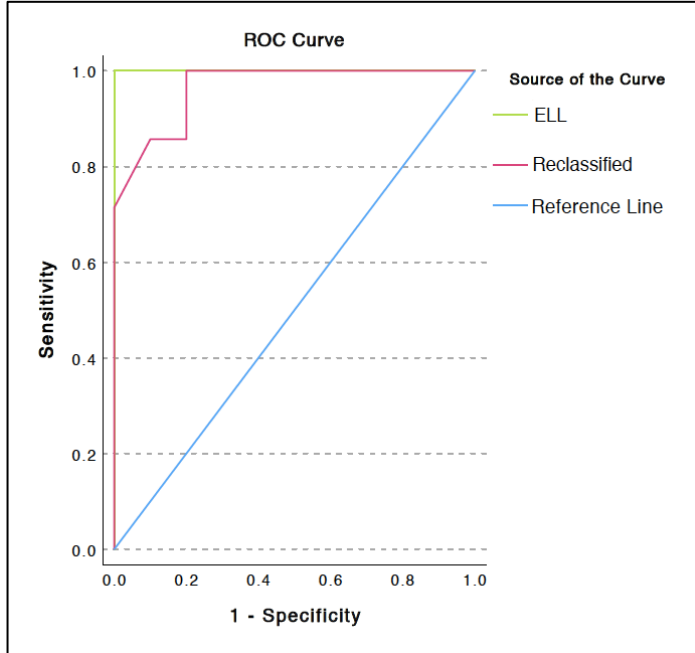*Figure 2.* Non-parametric ROC curves for ELL and
EO groups.

Figure 3. Non-parametric ROC curves for ELL and
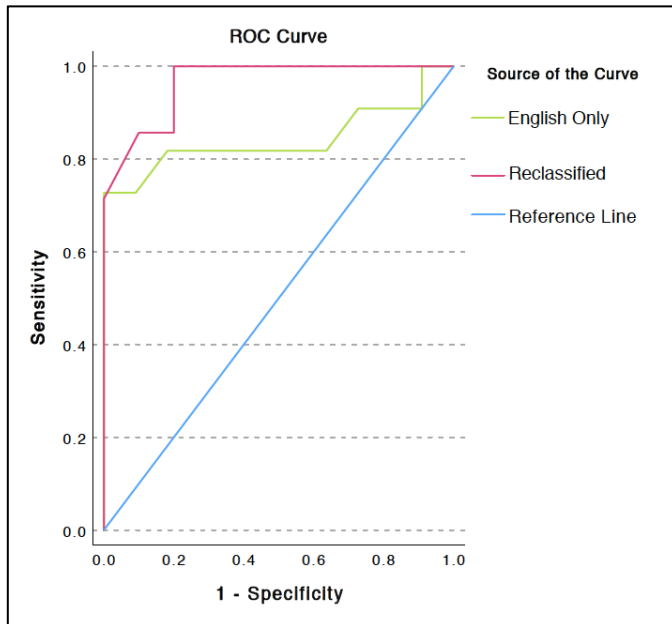Reclassified groups

Figure 4. Non-parametric ROC curves for EO and
Reclassified groups.

Discussion

Given the importance of early literacy skills in predicting later reading achievement, having screening tools to effectively and efficiently identify poor readers early on is essential. A vast amount of research has supported the use of English CBMs to identify students at-risk of future reading difficulties, with a few studies targeting ELL students. However, less research has focused on using Spanish CBMs with students receiving bilingual instruction. The purpose of this study was to assess the classification accuracy of a Spanish CBM with a group of third grade students participating in the DI program and receiving 70 percent of instruction in Spanish and 30 percent in Englsih. In particular, the study aimed to examine the classification accuracy of a Spanish CBM by students' level of English proficiency.

Results indicated that the FLO measure was effective at predicting performance on the i-Ready reading assessment for students participating in the DI program. The FLO measure accurately discriminated between an at-risk and not at-risk student 91.4% of the time, regardless of students' level of English proficiency. This finding suggests that a Spanish CBM could be effectively used when working with students who are learning both English and Spanish at school. Such finding is also consistent with Keller-Margulis, Payan, and Booth (2012) who reported a moderate correlation between reading CBMs in Spanish and a statewide measure of achievement for a group of third and fourth grade students receiving bilingual instruction. However, the sample in Keller-Margulis and colleagues (2012) did not include students who were native English speakers receiving bilingual instruction.

Further analyses revealed that the FLO measurement can also be used with students with varying levels of English proficiency. Within the EO group, the FLO measure accurately classified students 84.3% of the time, thus supporting the use of a Spanish assessment for students who are native English speakers receiving Spanish instruction. This finding is particularly significant as most studies examining the predictive validity of Spanish CBMs have targeted ELL students only. Moreover, for ELLs, the FLO measurement accurately identified students who would pass or not pass the reading assessment 100% of the time. As expected, this finding suggests that Spanish CBMs are valid indicators of how well an ELL student will perform in an English reading assessment. This finding mirrors Ramirez and Shapiro (2007) who reported that Spanish oral reading fluency, as measured by a Spanish CBM, predicted English oral reading fluency for ELL students receiving transitional bilingual instruction. Lastly, for students in the Reclassified group, the FLO had a classification accuracy of 96.4%, indicating that the FLO measurement can also be validly use with this group of students. This finding provides novel insight, as scant research has targeted students who have transitioned from an ELL status to English proficient.

Additional analyses comparing the classification accuracy of the FLO measurement for each level of English proficiency revealed no statistically significant difference among the three groups. Although the FLO had a higher classification accuracy for the ELL and Reclassified students than for the EO students, this difference was not significantly different. That is, the classification accuracy of the FLO was similar for native English-speakers and native Spanish-speakers. Furthermore, this finding

underscores the notion that students should be assessed in their primary language of instruction, as it can also help inform decision-making.

Limitations and Future Directions

The present study was subject to limitations. These limitations reduce the generalizability of the findings and should be addressed in future research. First, the sample was limited to a small group of third grade students. Future research should consider using a larger sample that includes children of different grade levels and reading abilities, who are also receiving bilingual instruction. Students in the DI program are exposed to varying levels of English and Spanish instruction based on their year in school. For instance, a student in kindergarten is exposed to more Spanish instruction and less English instruction than a third grader. As a result of this, the classification accuracy of Spanish CBMs may change for students depending on their school grade. Therefore, additional research should assess the classification accuracy of Spanish CBMs with younger and older students receiving bilingual instruction and have varying levels of Spanish reading skills. In line with this, future studies should take further consideration of EO students receiving bilingual instruction. Native English speakers who are being exposed to Spanish for the first time may perform more poorly on Spanish CBMs compared to English CBMs (Keller-Margulis & Mercer, 2014). Consequently, it may be necessary to test these students in both English and Spanish.

Along the same line, another limitation of the study is that it did not examine the classification accuracy of English CBMs. English proficient students (e.g., EO, Reclassified) also receive bilingual instruction. Because of this, it may be necessary to also test these students in English in order to best assess their skills. For instance, an EO

student's performance on an English CBM, as opposed to a Spanish CBM, may be more predictive of performance on an English reading test. Moreover, assessing students in both languages will allow for comparisons of the classification accuracy of English and Spanish CBMs. As such, future research should examine the classification accuracy of both English and Spanish CBMs with students in a DI program.

Furthermore, the study collected CBM data using only one measure administered at one time-point. The FLO assessment, a measure of oral reading fluency, was the single measure used to predict performance. Future research studies should consider using multiple CBM measures to assess students reading skills. A composite score of multiple CBM measures may be more accurate at predicting future reading outcomes. Lastly, additional research should consider progress monitoring students' performance using English and Spanish CBMs. This process would include tracking students' growth in both languages throughout the school year and examining whether this growth is predictive of performance on a reading assessment. Doing so may be particularly beneficial for students in the DI program as these students may transition from speaking more Spanish to more English, or vice versa, at varying time points throughout the school year or their elementary years.

Implications

Although English CBMs are often used to assess whether a student may be struggling with reading skills, findings from this study suggest that Spanish CBMs may also be effective tools when working with EO, ELL, and Reclassified students who are receiving bilingual instruction. Overall, findings suggest that Spanish CBMs are effective tools for predicting reading performance for students who are receiving instruction in

English and Spanish. This has implications for the school setting as findings indicate that educators can also utilize Spanish CBMs to inform data-based decision making. For instance, data from Spanish CBMs may be particularly helpful for school psychologists who are responsible for evaluating bilingual students and determining their eligibility for special education services. These data can further inform about students reading skills in Spanish and can work in conjunction with English CBMs to help determine the student's language of assessment. However, despite promising findings, additional research is needed in order to further determine best practices when working with this designated population of students, especially given the rising popularity of dual-language education programs within the United States.

References

Anderson, R. C., Wilson, P. T., & Fielding, L. G. (1988). Growth in reading and how children spend their time outside of school. *Reading Research Quarterly*, *23,* 285-303.

Barrow, L., & Markman-Pithers, L. (2016). Supporting young English learners in the United States. *The Future of Children*, 159-183.

Baker, S. K., & Good, R. (1995). Curriculum-Based Measurement Reading with Bilingual Hispanic Students: A Validation Study with Second-Grade Students.

Baker, D., Good, R. H., Knutson, N., & Watson, J. (2006) *Indicadores Dinámicos del Éxito en la Lectura* (6th ed.). Eugene, OR: Institute for the Development of Education Achievement.

Ballantyne, K. G., Sanderman, A. R., & McLaughlin, N. (2008). Dual Language Learners in the Early Years: Getting Ready to Succeed in School. *National Clearinghouse for English Language Acquisition & Language Instruction Educational Programs*.

Burr, E., Haas, E., & Ferriere, K. (2015). Identifying and supporting English learner students with learning disabilities: Key issues in the literature and state practice. *Washington, DC: US Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory West*.

Buysse, V., Peisner-Feinberg, E., Páez, M., Hammer, C. S., & Knowles, M. (2014). Effects of early education programs and practices on the development and learning of dual language learners: A review of the literature. *Early Childhood Research Quarterly*, *29*(4), 765-785.

Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children*, *52*(3), 219-232.

Deno, S. L. (2003). Developments in curriculum-based measurement. *The Journal of Special Education*, *37*(3), 184-192.

Chiappe, P., Siegel, L. S., & Wade-Woolley, L. (2002). Linguistic diversity and the development of reading skills: A longitudinal study. *Scientific Studies of Reading*, *6*(4), 369-400.

Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology*, *33*(6), 934.

DeLong, E. R., DeLong, D. M., & Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*, *44*(3), 837-845.

Farmer, E. (2013). Examining predictive validity and rates of growth in curriculum-based measurement with English language learners in the intermediate grades.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern recognition letters*, *27*(8), 861-874.

Ford, K. L., Cabell, S. Q., Konold, T. R., Invernizzi, M., & Gartland, L. B. (2013). Diversity among Spanish-speaking English language learners: Profiles of early literacy skills in kindergarten. *Reading and Writing*, *26*(6), 889-912.

Goldenberg, C. (2008). Teaching English language learners: What the research does-and does not-say. *American Educator, 32*(2), 8-23, 42-44.

Good, R. H., & Kaminski, R. (2002). *Dynamic Indicators of Basic Early Literacy Skills* (6th ed.). Eugene, OR: Institute for the Development of Education Achievement.

Gutierrez, G., & Vanderwood, M. L. (2013). A growth curve analysis of literacy performance among second-grade, Spanish-speaking, English-language learners. *School Psychology Review*, *42*(1), 3-21.

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology*, *80*(4), 437.

Keller-Margulis, M.A., & Mercer, S.H. (2014). R-CBM in Spanish and in English: differential relations depending on student reading performance. *Psychology in the Schools, 51*(7), 677-692.

Keller-Margulis, M.A., Payan, A., & Booth, C. (2012). Reading curriculum-based measures in Spanish: an examination of validity and diagnostic accuracy. *Assessment for Effective Intervention, 37*(4). 212-223.

Kieffer, M. J. (2008). Catching up or falling behind? Initial English proficiency, concentrated poverty, and the reading growth of language minority learners in the United States. *Journal of Educational Psychology*, *100*(4), 851.

Kim, J. S., Vanderwood, M. L., & Lee, C. Y. (2016). Predictive validity of curriculum-based measures for English learners at varying English proficiency levels. *Educational Assessment*, *21*(1), 1-18.

Klein, J. R., & Jimerson, S. R. (2005). Examining ethnic, gender, language, and socioeconomic bias in oral reading fluency scores among Caucasian and Hispanic students. *School Psychology Quarterly*, *20*(1), 23.

Lepola, J., Salonen, P., & Vauras, M. (2000). The development of motivational orientations as a function of divergent reading careers from pre-school to the second grade. *Learning and Instruction*, *10*(2), 153-177.

McKenna, M. C., Kear, D. J., & Ellsworth, R. A. (1995). Children's attitudes toward reading: A national survey. *Reading Research Quarterly*, 934-956.

McNamara, J. K., Scissons, M., & Dahleu, J. (2005). A longitudinal study of early identification markers for children at-risk for reading disabilities: The Matthew effect and the challenge of over-identification. *Reading Improvement*, *42*(2), 80-98.

National Center for Education Statistics. (2018). English language learners in public schools. Retrieved from https://nces.ed.gov/programs/coe/indicator_cgf.asp.

National Center for Education Statistics. (2007). The Nation's Report Card Reading 2007. Retrieved from https://nces.ed.gov/nationsreportcard/pdf/main2007/2007496.pdf

National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Washington, DC: National Institute of Child Health and Human Development.

Ramírez, R.D., & Shapiro, E.D. (2006). Curriculum-based measurement and the evaluation of reading skills of Spanish-speaking English language learners in bilingual education classrooms. *School Psychology Review, 35*(3), 356-369.

Ramírez, R.D., & Shapiro, E.D. (2007). Cross-language relationship between Spanish and English oral reading fluency among Spanish-speaking English language learners in bilingual education classrooms. *Psychology in the Schools, 44*(8), 795-806.

Reschly, A. L., Busch, T. W., Betts, J., Deno, S. L., & Long, J. D. (2009). Curriculum-based measurement oral reading as an indicator of reading achievement: A meta-analysis of the correlational evidence. *Journal of School Psychology*, *47*(6), 427-469.

Richardson, R. D., Hawken, L. S., & Kircher, J. (2012). Bias using maze to predict high-stakes test performance among Hispanic and Spanish-speaking students. *Assessment for Effective Intervention, 37*(3), 159-170.

Schatschneider, C., Fletcher, J. M., Francis, D. J., Carlson, C. D., & Foorman, B. R. (2004). Kindergarten prediction of reading skills: A longitudinal comparative analysis. *Journal of Educational Psychology*, *96*(2), 265.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, *21*(4), 306-407.

Swets, J. A. (1988). Measuring the accuracy of diagnostic systems. *Science*, *240*(4857), 1285-1293.

Valentino, R. A., & Reardon, S. F. (2015). Effectiveness of four instructional programs designed to serve English learners: Variation by ethnicity and initial English proficiency. *Educational Evaluation and Policy Analysis*, *37*(4), 612-637.

Wayman, M., Wallace, T., Wiley, H. I., Tichá, R., & Espin, C. A. (2007). Literature synthesis on curriculum-based measurement in reading. *The Journal of Special Education*, *41*(2), 85-120.

Yesil-Dagli, U. (2011). Predicting ELL students' beginning first grade English oral reading fluency from initial kindergarten vocabulary, letter naming, and phonological awareness skills. *Early Childhood Research Quarterly*, *26*(1), 15-29.

Yeo, S. (2010). Predicting performance on state achievement tests using curriculum-based measurement in reading: A multilevel meta-analysis. *Remedial and Special Education*, *31*(6), 412-422.