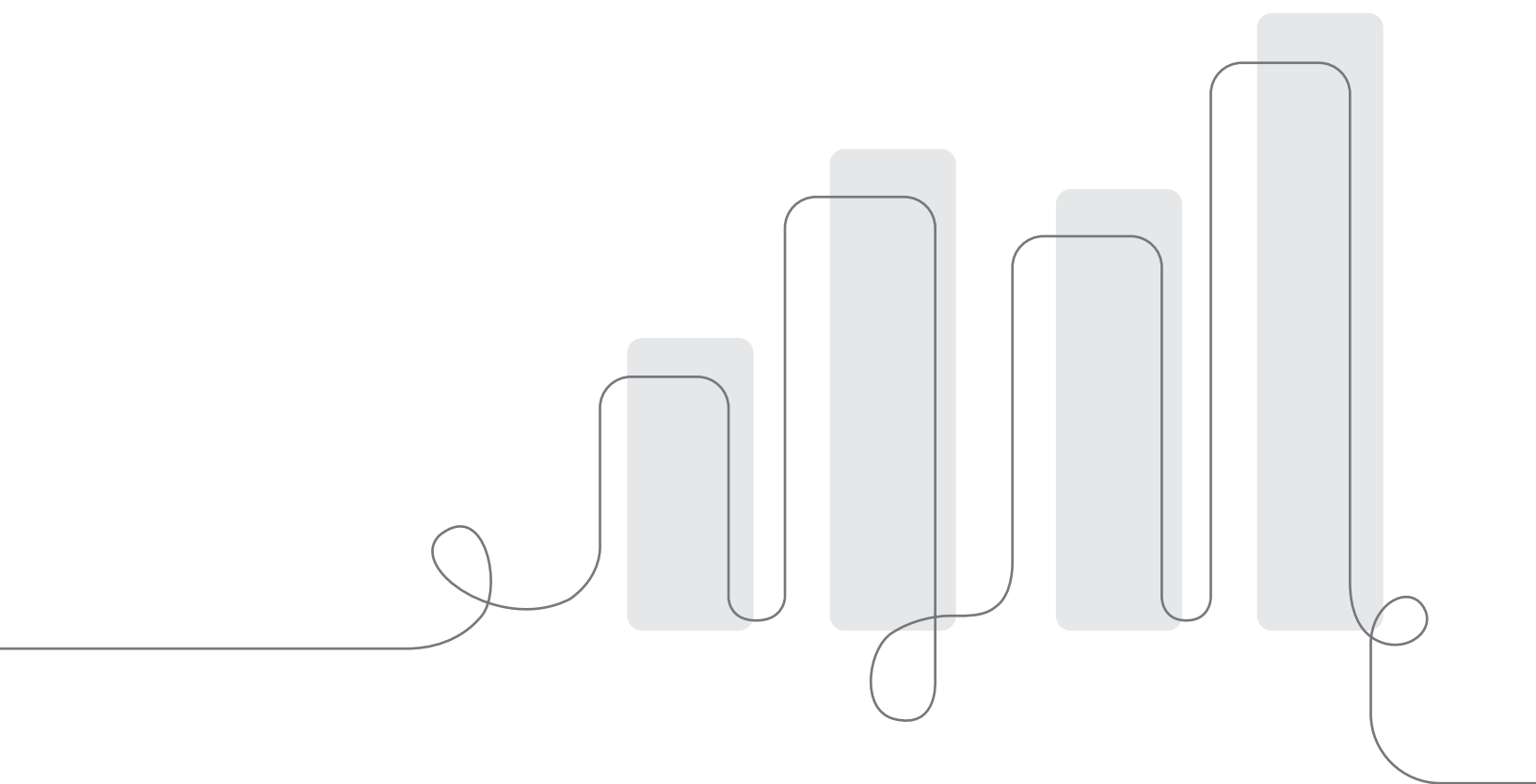# Different outcomes for Spanish-speaking students assessed in English and in Spanish demonstrate the importance of bilingual assessment.

June 2024

Amplify.

Executive summary

Changing demographics in the United States public school population necessitate the availability of reliable and valid universal screening assessments for literacy in languages other than English. In 2020, for example, approximately 5 million public school students—10 percent of all students—identified as multilingual/English learners (ML/ELs) [1]. For the majority of these students (75 percent), Spanish is one of the top five languages spoken in their state[2]. Unfortunately, an English-only approach to screening bilingual students may underestimate their knowledge and skills, given that limited English proficiency can adversely affect a student's performance on English measures[3]. Moreover, an English-only approach to assessment fails to capture the knowledge and skills MLs/ELs have in their home language that can be leveraged to support their English literacy development through processes of cross-linguistic transfer[4].

A new Amplify study of K–3 Spanish-speaking multilingual learners (MLs) shows that more of these students were identified as at risk for not learning to read when their early literacy skills were measured only in English.

Across all grades and times of year, more Spanish-speaking MLs were identified as needing intervention when they were assessed in English compared to Spanish. Differences vary by grade level, from 6 percent more students identified as at risk at the end of grade 1 to a staggering 51 percent more students identified as at risk at the beginning of kindergarten; however, the data show consistently fewer Spanish-speaking MLs identified as needing intervention when assessed in Spanish. These results demonstrate the importance of measuring Spanish-speaking students' skills in Spanish and English.

For the study, Amplify analyzed data from an intact group of Spanish-speaking MLs in grades K–3 who have Spanish literacy data from mCLASS® Lectura assessments and English literacy data from mCLASS DIBELS® 8th Edition assessments at the beginning, middle, and end of the 2023–2024 school year.

> When MLs are assessed in English only, the skills students have from another language may go unrecognized. Educators may see these students as generally struggling readers simply because they are not able to demonstrate fully their literacy skills in English. However, these students likely have knowledge and skills in Spanish and may need support transferring those skills from Spanish to English[5]. Assessing students in their home languages enables educators to identify areas of strength and areas that require further support, ensuring a more comprehensive assessment of their overall language and literacy abilities.

1   Irwin, V., Zhang, J., Wang, X., Hein, S., Wang, K., Roberts, A., ... & Parker, S. (2021). Report on the Condition of Education 2021 (NCES 2021-144). U.S. Department of Education: Washington, DC: National Center for Education Statistics. Retrieved 10/1/2022 from https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2021144.

2   Office of English Language Acquisition. (2023). The Top Languages Spoken by English Learners in the United States. Washington, DC: Author.

3   Pitoniak, M., Young, J. W., Martiniello, M., King, T. C., Buteux, A., & Ginsburg, M. (2009). Guidelines for the assessment of English language learners. Education Testing Service.

4   Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. Review of Educational Research, 49(2), 222–251. https://doi. org/10.3102/00346543049002222.

5   Below, K., Duran, L. (2023). Embracing Bilingual Assessment. Language Magazine. https://www.languagemagazine.com/2023/10/06/embracing-bilingual-assessment/.

Insights

To what extent does the identification of the need for reading intervention vary for Spanish-speaking MLs when they are assessed in Spanish and English?

Figure A shows the percentage of Spanish-speaking MLs in grades K–3 at each time of year who were identified by their composite score as needing intervention support when assessed in Spanish (purple dot) and when assessed in English (blue dot). The gray line between the dots represents the difference in the percentage of students identified as needing intervention across languages.

Figure A. Percentages of Spanish-speaking MLs in grades K–3 identified as needing intervention based on overall reading risk in Spanish and English
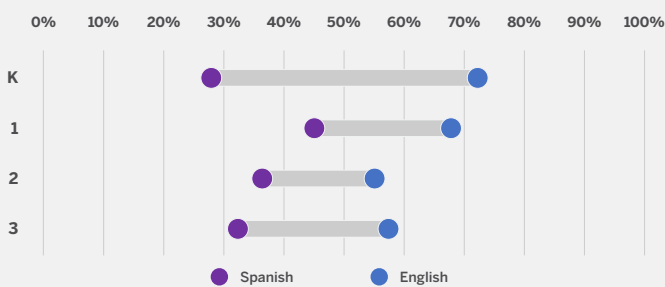
Percentage of Students Needing Intervention Support at BOY
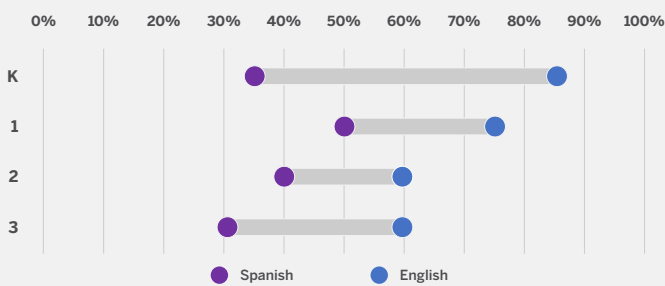
Beginning of Year

Grade	English	Spanish	Difference
K	86%	35%	-51 pts
1	75%	50%	-25 pts
2	59%	39%	-20 pts
3	59%	32%	-27 pts

Percentage of Students Needing Intervention Support at MOY

Middle of Year

Grade	English	Spanish	Difference
K	73%	27%	-46 pts
1	67%	45%	-22 pts
2	55%	36%	-19 pts
3	58%	33%	-25 pts

Percentage of Students Needing Intervention Support at EOY
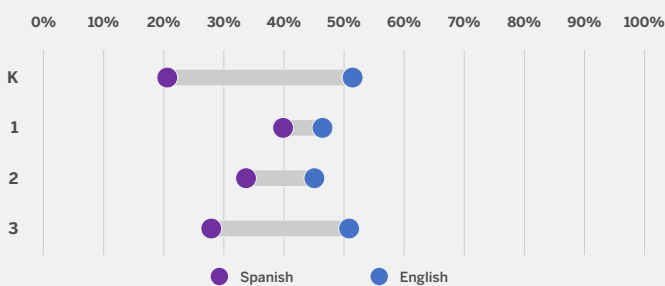
End of Year

Grade	English	Spanish	Difference
K	52%	21%	-31 pts
1	46%	40%	-6 pts
2	45%	34%	-11 pts
3	51%	27%	-24 pts

First, across all grades and times of year, more Spanish-speaking MLs were identified as needing intervention when they were assessed in English compared to Spanish. Differences vary by grade level, from 6 percent more students identified as at risk at the end of grade 1 to a staggering 51 percent more students identified as at risk at the beginning of kindergarten; however, the data show consistently fewer Spanish-speaking MLs identified as needing intervention support when assessed in Spanish. These results highlight the importance of gathering contextual information about instruction, such as the amount of Spanish and English literacy instruction Spanish-speaking MLs are receiving, because it is likely that instruction is influencing student universal screening performance in both languages.

Second, these data illustrate that across grade levels and times of year, high percentages of Spanish-speaking MLs are identified as needing intervention in English (45 percent–86 percent) and in Spanish (21 percent–50 percent). These results necessitate examining the instructional supports that are offered, with respect to the quality of the instructional materials, the instructional delivery, and the coordination and alignment of the instruction students receive across languages.

Third, scores across grades show that differences in the percentages of Spanish-speaking MLs identified as needing intervention are the greatest in kindergarten and significantly smaller in grades 1–3.

- At the beginning of year kindergarten, for example, we see that approximately 51 percent more Spanish-speaking MLs were identified as needing intervention when assessed in English (86 percent) compared to Spanish (35 percent). This difference decreased by almost half (20 percent) at the end of the year, at which point 52 percent of Spanish-speaking MLs were identified as needing intervention when assessed in English and 21 percent were identified when assessed in Spanish.

- In grades 1 and 2, the difference in percentages of students identified as needing intervention at the beginning of the year is smaller than in kindergarten (25 percent in grade 1 and 20 percent in grade 2) and only narrows slightly at middle of year before dropping significantly at end of year (6 percent in grade 1 and 11 percent in grade 2).

Fourth, within each grade, the difference in the percentage of Spanish-speaking MLs identified as needing intervention when assessed in Spanish and English decreased across time; that is, within each language, the difference is the greatest at the beginning of the year and smallest at the end of the year.

Implications

A dual-language approach for assessing the early literacy skills of multilingual learners learning to read in two languages has powerful implications for educators.

Assessing Spanish-speaking MLs in Spanish significantly reduces the percentage of students identified as struggling readers. Across grade levels and times of year, fewer students whose home language is Spanish were identified as needing intervention support when assessed in Spanish. Conversely, significantly more students were identified as needing additional instructional support when assessed only in English. These findings align with previous research suggesting that an English-only approach to the assessment of MLs may underestimate their knowledge and skills because limited English proficiency can adversely affect a student's performance on English measures[1].

Assessing Spanish-speaking MLs in Spanish and English provides educators with a more comprehensive understanding of multilingual learners' literacy skills that can yield meaningful instruction. Data obtained from assessments in Spanish and English could indicate how much core or intervention instruction is necessary across one or both languages. For example, if a student whose home language is Spanish demonstrates proficiency with Spanish letter sound correspondences but is identified as needing intervention support with English letter sound correspondences, instruction does not need to begin as if the student has no letter-sound correspondence knowledge at all; rather, instruction can focus on the letter-sound correspondences that are different between Spanish and English, such as those for vowels and some consonants (e.g., h, j, r, v, x, z).

1   Pitoniak, M., Young, J. W., Martiniello, M., King, T. C., Buteux, A., & Ginsburg, M. (2009). Guidelines for the assessment of English language learners. Education Testing Service.

A dual-language assessment approach yields more precise instructional profiles based on students' needs. As illustrated by Figure B, there are critical differences in the percentage of Spanish-speaking MLs needing core support versus intervention support across English and Spanish. From this, we can discern four potential performance profiles (the quadrants of Figure B), each of which necessitate different approaches to instruction and potential intervention. Rather than needing to determine the Spanish and English instructional needs for each student individually, profiles of student performance can be determined by aggregating patterns of intervention support needed at the classroom or grade level, which can then be used to support conversations about resource allocation needs (i.e., instructional time, group size, staffing needs, etc.).

Figure B. Percentage of kindergarten students at BOY who would benefit from different levels of support (core and intervention) in English and Spanish

	English Core Support	English Intervention Support
Spanish Core Support	13%	**52%**
Spanish Intervention Support	1%	**34%**

- Spanish-speaking MLs in the upper-left quadrant who are identified as needing core support in English and Spanish could benefit from receiving core literacy instruction in both languages that focuses on foundational skills, vocabulary, and reading comprehension, including transfer from one language to another. Students in the upper-right quadrant (green box), who are identified as needing core support in Spanish and intervention support in English, likely have sufficient Spanish literacy skills they can leverage to support the acquisition of English literacy skills. Students in the lower-right quadrant (blue box), who are identified as needing intervention support in English and Spanish, need more explicit, systematic instructional support building their foundational literacy skills in both languages. Skill-level data from both assessments provides additional detail about where educators should focus their support.

Side-by-side assessment data in both English and Spanish provides educators with valuable information about the skills students already have. Educators need to understand how students are performing in each language; without these data, differences in skills in each language may not be readily apparent[2].

---

2   Hopewell, S., & Escamilla, K. (2014). Struggling reader or emerging biliterate student? Reevaluating the criteria for labeling emerging bilingual students as low achieving. Journal of Literacy Research, 46(1), 68-89. https://doi.org/10.1177/1086296X13504869.

Recommendations

Supporting student literacy development in both languages

A dual-language approach to literacy assessment for Spanish-speaking multilingual learners provides educators with a more comprehensive understanding of students' literacy development in both languages, highlighting student strengths and bringing to light opportunities for improvement. As the data included in this brief indicate, significantly greater percentages of Spanish-speaking MLs were identified as needing intervention when assessed only in English, providing further evidence that a monolingual approach to assessment fails to capture students' true range of literacy skills.

To help educators support MLs who are learning to read, it is important that schools and districts invest in a reliable universal screener, high-quality core curriculum, evidence-based interventions, and professional development. Although the availability of universal screening assessments for Spanish literacy has increased in recent years, information about the reliability, validity, and diagnostic accuracy of these Spanish screening assessments is limited. All educators serving Spanish-speaking students deserve access to assessments that yield consistent information about student performance (are reliable) and support trustworthy inferences about student Spanish literacy skill development[1] (are valid). See Figure C.

Results from universal screening assessments in English and Spanish should provide educators with consistent and trustworthy information about students' biliteracy development. In particular, information about student skill development in phonological awareness, alphabetic understanding, fluency, vocabulary, and comprehension in Spanish and English are important for monitoring student literacy development in both languages.

---

1   Durán, L., Basaraba, D. L., Baker, D. L., & O'Brien, G. (2023). The importance of dual language assessment in early literacy, 16-18. amplify.com/pdf/uploads/2023/02/mCLASS_Paper_Biliteracy_021023_web.pdf.

Figure C. Key features of instructionally relevant Spanish literacy assessments

Features	Details	Examples/Notes
Focus on critical early literacy skills	• Phonological awareness • Alphabetic understanding/decoding • Fluency with connected text • Vocabulary • Comprehension	Subtests of the screening assessment are developed to measure each of these domains that are based on Spanish development and are not translations of English.
Attend to specific linguistic features of Spanish	Measures should be written in Spanish to authentically assess Spanish literacy skills.	Letter name/letter sound measures must reflect the letters/sounds of Spanish, including letters/sounds that are unique to Spanish. Measures of phonological awareness and word reading explicitly account for syllabic and morphological structures of Spanish (e.g., consonant digraphs such as ll and rr that are unique to Spanish, as well as other common digraphs, such as *tr*, *dr*, and *gl*). Passages of connected text for measuring reading fluency should be written using syntactical, lexical, and grammatical rules of Spanish (and passage difficulty should be measured accordingly).

Figure C. (continued) Key features of instructionally relevant Spanish literacy assessments

Features	Details	Examples/Notes
Strong technical adequacy	Should include information about: • Reliability: Consistency of scores across alternate test forms, measurement occasions, raters, and/or between items and overall score. • Criterion validity: Extent to which performance on measure of interest is related to performance on other similar measures. • Classification accuracy: Consistency with which scores on screening assessment classify students and predict a decision regarding performance on an outcome assessment. • Generalizability: Extent to which reliability, validity, and classification accuracy generalize to other demographic groups of students.	Reliability (alternate-form, test-rest, interrater, and/or internal consistency) Criterion validity (concurrent, predictive) Classification accuracy (sensitivity, specificity, accuracy, Area Under the Curve)
Appropriateness of tasks	Tasks used to measure Spanish literacy skills should be appropriate for: • The construct being assessed. • Age/skill level of students being assessed.	Phonological awareness in Spanish will focus on syllable awareness and phoneme awareness. Tasks should be sufficiently sensitive to measure skills and detect change, even among the youngest learners.
Brief, easy, and efficient to administer	Tasks should be: • Brief in duration, because all students should be assessed. • Efficient with respect to the time, personnel, and resources required to administer the assessments and interpret results. • Easy to administer, because minimal training is required to administer the assessments using standardized administration procedures (Deno, 2003).	Items are scripted for the examiner and scoring is well specified. Each subtest is no more than 3 minutes long.

About the study

This research brief describes the results from one study. The goals for the study were to better understand the extent to which the identification of reading intervention support varies for Spanish-speaking MLs when they are assessed in Spanish and English. Amplify analyzed data from an intact group of Spanish-speaking MLs in grades K–3 who have Spanish literacy data from mCLASS Lectura assessments and English literacy data from mCLASS DIBELS 8th Edition assessments at the beginning, middle, and end of the 2023–2024 school year.

During the 2023–2024 school year, more than 212,000 students in grades K–6 were assessed at least once with mCLASS Lectura. These students were enrolled in schools located across 30 states and the District of Columbia (representing four census regions and all nine census divisions) and one country outside of the United States[1]. Of these, the data for this study were from students who had assessment scores from beginning, middle, and end-of-year in both Spanish and English. Our sample includes more than 21,000 Spanish-speaking MLs in grades K–3 who have data on both assessments at all three times of year.

We were specifically interested in exploring the results of dual-language assessment on students whose home language was Spanish and who were learning to read in both Spanish and English. Note: We do not have access to key contextual variables that could contribute to observed variability in student performance, such as language proficiency and exposure to both languages or instructional program information (e.g., bilingual program model or amount of Spanish literacy instruction).

For the figures cited by the Office of English Language Acquisition, we use the term multilingual/English learners (ML/EL) to align with English learner (EL), the term used in the Condition of Education reports. For technical accuracy throughout the study and research brief, we use the term Spanish-speaking multilingual learner (ML).

About mCLASS and mCLASS Lectura

The data for this study was collected with mCLASS, Amplify's teacher-administered literacy assessment and intervention suite for grades K–6. mCLASS, powered by DIBELS 8th Edition, automates the data collection of Dynamic Indicators of Basic Early Literacy Skills (DIBELS), a widely-used series of short tests that assess K–8 literacy. Developed by the University of Oregon, DIBELS is an observational assessment collected by teachers interacting with students one-on-one, either live or over video. DIBELS is typically administered three times a year (beginning, middle, and end of year), and is used to identify reading difficulty, monitor progress, and inform instruction, especially for struggling readers.

---

1   U.S. Department of Commerce. (2020). Census regions and divisions in the United States. United States Census Bureau, Geography Division.

mCLASS Lectura uses the same assessment performance levels as DIBELS 8th Edition. However, for the purposes of this study, students whose performance was Below Benchmark and Well Below Benchmark were identified as needing intervention support. These features support implementation of mCLASS Lectura in combination with DIBELS 8th Edition across dual-language and other bilingual settings to support educators' understanding of students' Spanish and English biliteracy development.

Explanation of assessment performance levels

mCLASS Lectura performance levels	Status	Instructional implications
🟦 Above benchmark	On track	Ready for core instruction, likely to meet grade-level reading standards at end of year
🟩 At benchmark		
🟨 Below benchmark	At risk	In need of strategic/intervention support, unsure of the likelihood of meeting grade-level reading standards at end of year
🟥 Well below benchmark		

About Amplify

A pioneer in K–12 education since 2000, Amplify is leading the way in next-generation curriculum and assessment. Our captivating core and supplemental programs in literacy, math, and science engage all students in rigorous learning and inspire them to think deeply, creatively, and for themselves. Our formative assessment products turn data into practical instructional support to help all students build a strong foundation in early reading and math. All of our programs provide teachers with powerful tools that help them understand and respond to the needs of every student. Today, Amplify reaches more than 15 million students in all 50 states. To learn more, visit amplify.com.

For more information on mCLASS,
visit **amplify.com/mclass**.

Amplify.