Roundtable Report | July 2023

Building a Research-Based Approach to Bilingual Early Literacy Assessment



Early detection transforms lives. Forever.

In partnership with



EarlyBird and SoapBox Labs' Participation in the Measures for Early Success Initiative

<u>EarlyBird</u> and <u>SoapBox Labs</u>, partners from EarlyBird's beginnings, joined forces for this project. The EarlyBird early literacy assessment uses the SoapBox Labs engine to analyze audio files of children's speech in expressive subtests, and allow for automatic scoring.

About EarlyBird

EarlyBird's easy-to-administer game-based assessment is fun for children and appropriate for pre-readers. Created by leading cognitive neuroscientist Dr. Nadine Gaab, and reading assessment expert Dr. Yaacov Petscher, EarlyBird brings together the tools to easily and accurately identify children at risk of language and literacy challenges in the early grades and help teachers take action in the window when intervention is most effective. Developed and tested at Boston Children's Hospital, EarlyBird is now in use in districts in over 20 states nationwide.

About SoapBox Labs

SoapBox Labs is the leading provider of voice AI for pre-K – 12 education tools in the world. Unlike other players in the market, SoapBox Labs' proprietary voice engine has been built from the ground up to cater to the unique characteristics of children's voices, behaviors, environments, and privacy needs.



Table of Contents

INTRODUCTION	3
The Problem (Why Bilingual? Why Spanish?) The Goal The Roundtable	3 4 4
ROUNDTABLE DISCUSSIONS & FINDINGS	9
1. Content 2. Implementation – Experience and Scalability 3. Data – Psychometrics and Usefulness	11 20 25
CONNECTING THE DOTS	33
Top Five Problems to Address Next Steps	33 34
APPENDICES	35
Additional Descurress	25

Additional Resources	35
References	35
Acknowledgements	38
 Roundtable Participants 	38
– Project Contributors	40
Contact Us	41
-	40 41

Introduction

What if ...

... An early literacy assessment could capture reliable and valid information in an equitable manner about children's needs, competencies, progress, and classroom experiences in the early grades,

... In both English and Spanish,

... While reducing the teacher burden tied to data analysis and reporting,

... And making the data actionable?

The Problem (Why Bilingual? Why Spanish?)

Current assessment practices for Spanish-speaking English learners (ELs*) have limitations with implications for both data collection and data interpretation. And this affects a large segment of our student population; there were 5 million ELs in U.S. public schools in 2020 representing 10.3% of the U.S. student population. Of those, Spanish is the most commonly reported home language of EL public school students, accounting for 76% of all ELs and 8% of the student population in public schools.¹

In states like California, Texas, and New Mexico, nearly half of people are Latine and almost one-third are bilingual. States in the Northeast, Midwest, and South have experienced double-digit growth in their Latine populations since 2010.² Over 75% of Head Start programs serve at least one family that speaks primarily Spanish at home.³

Pre-K programs need tools to support this growing Spanish-speaking student population. By 2050, it is estimated that one in three children in the U.S. will be Latine and at least some Spanish will be spoken in most of their households.⁴

^{*} For purposes of this report, ELs refers specifically to Spanish-speaking English learners.

Introduction | The Goal

The Problem (Why Bilingual? Why Spanish?) (continued)

The lack of current, valid, meaningful, equitable pre-K early literacy assessment tools may also contribute to lagging fourth grade proficiency data: only one-third of U.S. fourth-graders are proficient readers — and English learners' proficiency levels continue to lag behind native English speakers.⁵

The stakes couldn't be higher.

All children have a right to learn to read. It is the foundation that opens doors to other opportunities to learn, choose a career path, make a living, and be successful in life.

The time to improve a child's reading is now.

The Goal

Design and build a groundbreaking, equitable early literacy assessment for Spanish-speaking English learners (ELs) that captures children's distributed skills across English and Spanish; produces reliable, comparable skill estimates across multilingual and monolingual children; and is actionable, guiding teachers to targeted instruction.

The Roundtable

The idea for a bilingual assessment roundtable was inspired by a previous experience with a similar roundtable in 2018, which set the foundation for the current EarlyBird product. Co-founder Dr. Nadine Gaab shared her research and vision for a practical tool that could address unmet needs. During interactive discussions, stakeholders collaboratively envisioned a research-based, game-based early literacy assessment.

From those beginnings, EarlyBird advanced from proof of concept, to beta, to product to market-ready by 2021. The result was a comprehensive early literacy and language assessment that identifies strengths, weaknesses, and potential reading challenges, including dyslexia, early – before they become reading problems.

So, what better way to start the process for another groundbreaking assessment than another roundtable?

Thanks to funding from the Measures for Early Success Initiative led by <u>MDRC</u>, EarlyBird partnered with <u>SoapBox Labs</u>, to convene a group of 24 invited participants on May 24, 2023. The meeting took place at the Executive Conference Center of Babson College, located west of Boston, Mass. The group included a range of stakeholders representing diverse user perspectives, on-the-ground experience, and bilingual learning expertise. (A participant list is provided on pages 38-39). This report was prepared by the EarlyBird team and does not necessarily reflect the perspectives of the Measures Initiative or MDRC.

To develop the right solution, we must first fully define the problem.

During the roundtable, we set out to harvest the expertise of leaders in the field to examine a range of issues, concerns, and sensitivities, and accelerate the development of a better tool for assessing Spanish-speaking ELs. We endeavored to:

 Respond to current research on bilingual assessment, particularly with Spanishspeaking ELs.

- Explore market needs and wants and evaluate the problems and current landscape in pre-K assessment, particularly for ELs.
- Elevate the perspectives of primary user groups and advance the goals outlined in the Measures for Early Success User-Informed Principles.
- Prioritize the most important characteristics of a bilingual (English and Spanish) early literacy assessment.
- Consider how new technologies can support more equitable assessment.
- Accelerate the development of a better tool for assessing Spanish-speaking ELs.

A combination of presentations, whole group discussions, and breakout group discussions culminated in a product development prioritization exercise. A userpersona small group exercise elevated the perspectives of a parent, Head Start director, pre-K teacher, and district administrator. These small groups identified the key assessment-related challenges for each persona and how those challenges are addressed today. With their persona assignments in mind, participants named their assessment must-haves (high priority) and nice-to-haves (lower priority).

To prepare for the roundtable, we developed three content presentations along with pre-readings for participants. These presentations covered current research findings, market demands, and what the Measures for Early Success User-Informed Principles suggest should be our guiding beacons. During the roundtable, the authors provided brief overviews, setting the tone for the day's discussions. You can find several elements from these presentations in the Roundtable Discussions & Findings section, starting on page 9.



1. Research White Paper

Elsa Cárdenas-Hagan, PhD, and Eric Tridas, MD, were tasked with leading the discovery research and crafting a white paper titled "Bilingual Spanish-English Learners: Research Foundation for an Early Identification Assessment" (see Resources page 35).

The primary objective was to identify the most important factors that aid in the early identification of bilingual students at risk for reading challenges.

White Paper authors Eric Tridas and Elsa Cárdenas-Hagan at the roundtable.

Their paper presented a comprehensive summary of research related to culturally and linguistically relevant aspects of development, learning, and competencies that influence reading in Spanish-speaking English Learners, with a particular focus on ages 3 to 5. Additionally, they were asked to consider the five Measures for Early Success User-Informed Principles and explore advancements in technology, including AI and automated speech recognition.

2. Market Needs and Wants

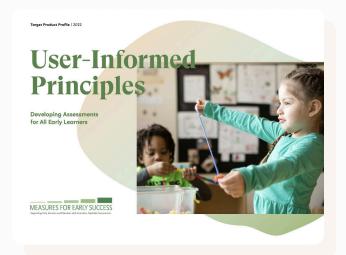
Shatta Mejia from EdSolutions developed and presented an overview of the pre-K Spanish literacy assessment market. Key points include:

- There is significant opportunity in the market, though there are also many unknowns. The variety of users, settings, policy makers, and programs, as well as different approaches to assessment (observational play-based vs. data-based), make it difficult to navigate and break through with new tools.
- Educators need tools to support a growing Spanish-speaking student population in federal Head Start, state, local, and privately-funded pre-K programs.
- Pre-K programs are currently using a range of products, with one product dominant in the market. Current assessments address one or more of the following: identify children who might require more intensive support (screener), inform next steps in teaching (formative), provide benchmark data on how children are progressing (summative), assess kindergarten readiness (KRA).



3. Measures For Early Success <u>User-Informed Principles</u>

<u>The Measures for Early Success Initiative</u> aims to reimagine the landscape of early learning assessments for the millions of 3- to 5-year-olds enrolled in pre-K so that more equitable data can be applied to meaningfully support and strengthen early learning experiences for all young children.



The Measures initiative is guided by a Target Product Profile (2022), developed with input from the target populations, and defined by five User-Informed Principles offering comprehensive and aspirational specifications to support the development of more equitable, useful, and reliable direct early learning assessments: Content, Psychometrics, Experience, Usefulness, Scalability.

These principles are designed to be relevant and usable for pre-K educators, children, and families to support more equitable early learning outcomes for all young children. The User-Informed Principles are organized in support of five Goals.



Goal 1: Content

Instrument(s) comprehensively measure the skills and development of 3- to 5-year-old children in equitable and culturally responsive ways.



Goal 2: Psychometrics

Instrument(s) collect objective information to produce

psychometrically sound and valid data that reflects minimal statistical bias.



Goal 3: Experience

Instrument(s) are enjoyable and engaging for children and easy for educators to collect.

տեւ

Goal 4: Usefulness

Instrument(s) generate timely, easily accessible, readily digestible, and understandable information for several purposes.



Goal 5: Scalability

Instrument(s) can be administered at scale in publicly funded Pre-K systems.

Roundtable Discussions and Findings

The roundtable put a spotlight on the many tensions in the pre-K language and literacy assessment space. While the Measures for Early Success User-Informed Principles offer an ambitious vision of what assessment could be, the roundtable participants grappled with market realities, the impact of public policies at pre-K, areas where research has not fully paved the pathway, and trade-offs that may be required to create an ideal product.



ROUNDTABLE CONCLUSION

At the conclusion of the roundtable, participants ranked the most pressing needs of the problems discussed throughout the day. The group determined that the #1 most pressing need is:

Constructing assessment content that reflects cognitive, linguistic, and cultural features of Spanish and English, and that produces psychometrically reliable and valid data/scores that are actionable for instruction.



DISCUSSIONS AND QUESTIONS

The following pages present the relevant research, market considerations, and key takeaways from the whole group and small group discussions that support this overall conclusion.

These findings are organized into three buckets:

- 1. Content what are we assessing and why?
- 2. **Implementation** how are we assessing those skills? what is the user experience and scalability?
- 3. **Data** What are the psychometrics that give us confidence that the data is valid and reliable? How useful is the data?

Subsections cover the research, the market, and the roundtable discussions, which include questions and quotes* from participants, as well as must-have and nice-to-have assessment suggestions. Each subsection begins with a summary of pertinent Measures for Early Success User-Informed Principles, our Goals, and summary Must-Haves that were generated by a product development sub-group in the wake of the roundtable.

^{*} Some quotes have been condensed or paraphrased for clarity and brevity.

1 Content What are we assessing and why?

User-Informed Principle	Goal	Bilingual Assessment Must-Have
Content – Pertains to children's development, skills, and competencies measured within domains of early learning.	Instrument comprehensively measures the skills and development of 3- to 5-year-old children in equitable and culturally responsive ways.	Content must be reflective of cognitive, linguistic, and cultural features of Spanish and English, and respond to research about what to test in order to ascertain how best to address student needs.

WHAT DOES THE RESEARCH SAY?

Assessing the EL population is complicated.

Finding a one-size-fits-all solution is challenging due to the vast diversity among Spanish-speaking ELs⁶ in terms of:

- Ethnic, geographic, and cultural backgrounds.
- Degrees of Spanish language, English language, and translanguaging proficiency.
- Family socioeconomic status (SES).
- Previous educational experience, including pre-K schooling and type of program.
- Language of instruction in school.

The research shows consensus on what to test.

Current research indicates that there are three major areas that are most useful in assessing early language and literacy in ELs and acting upon the data. These include neurocognitive factors, language proficiency, and environmental factors.⁷

Neurocognitive Factors

Research has shown that the most critical variables that influence the evolution of skills in EL students include the following neurocognitive factors:

- Phonological awareness
- Letter/sound knowledge
- Rapid automatized naming (RAN)
- Oral language skills

Specifically, studies have identified letter/sound knowledge, RAN, verbal short-term memory (VSTM), and phonological awareness as the strongest predictors of reading skills, with pseudoword repetition, receptive/expressive vocabulary, and listening comprehension also important.⁸

When testing skills, assessments must be sensitive to language differences. For children who are acquiring two languages, phonological awareness skills in the first and second language parallel each other. But since Spanish has a shallow phonology, transparent orthography and simple syllabic structure, the level of difficulty in items developed in English may be different for first-language Spanish speakers. Spanish morphology, on the other hand, is significantly more complex when compared to English.

In Spanish, the measures of rhyme and alliteration have not been shown to be clearly predictive of reading success. However, the studies carried out on the development of phonological-syllabic awareness and early reading have shown that the difficulties of pre-readers and beginning readers associated with syllabic analysis lead to difficulties in later learning of the reading.⁹

Language Proficiency

Oral language proficiency in the native language is a predictor of vocabulary and listening comprehension in the second language, and facilitates second language acquisition.

Since oral language proficiency in the native language matters,¹⁰ ideally we should take into account the child's language dominance and previous exposure to English, allowing for an assessment format that evaluates these skills in English, Spanish and/ or bilingually to align the language skills tested with the capabilities of the student.

Roundtable Discussions and Findings | Content

Environmental Factors

Knowing a child's prior learning experiences, language of instruction, exposure to Spanish and English, SES status, and family histories of reading challenges adds value and context to assessment data.

This knowledge can be especially helpful in identifying children who may be at risk and/or require additional support to prevent future reading problems.¹¹

Research also shows that different instructional program models result in varying development and acquisition of both Spanish and English language proficiency.¹² Pre-K bilingual instruction models include:

- Structured English Immersion
- English as a Second Language
- Transitional Bilingual Education
- Dual Language (one-way, two-way, 50/50 model, or other)

Beyond "what to test," lagging research has created contradictory and inconclusive guidance on "how to test."

WHAT DOES THE MARKET NEED AND WANT?

The market needs literacy and language assessments that deliberately account for the needs and strengths of young dual language learners (most of whom are Spanish speakers).

There is limited availability of evidencebased measures that appropriately assess the range of skills that will help teachers target literacy and language instruction for ELs.

"There is so much demand and so little out there that hits the mark."

The market needs an assessment that solves a number of problems in current assessment practice.

Translations of existing English-based assessments are not the answer since they ignore the conceptual, linguistic or semantic content and/or level of difficulty of the translated items across languages. It will be important to include transadaptation in a way that goes deeper than what providers are offering currently.

While there are a few research-based early language and literacy assessments in

Roundtable Discussions and Findings | Content

Spanish that can be used in real-world early learning settings, they do not accurately capture children's abilities across both English and Spanish.

- Administering one test in English and one test in Spanish is burdensome to teachers and children. Also, without good training, teachers may not know what to do with the two scores and how to use them to drive individualized instruction.
- There is often overidentification/overprediction of ELs who are at risk for reading difficulties.

Consideration of significant developmental differences from the beginning of the year to the end of the year in pre-K and kindergarten, and alignment across K, 1, 2 are not adequately addressed in current assessments.

"My pre-K students found the assessments too hard to use at the beginning of the year."

The market needs responsiveness to the specificity of Head Start and state requirements.

To be scalable, the additional subdomains required by Head Start and some state guidelines may need to be added. The most frequently required assessments for literacy and language include:¹³

- Phonological awareness (rhyme, segmenting syllables).
- Print and alphabet knowledge (letter and word recognition, letter naming, letter-sound correspondence, concepts of print, print meaning, print awareness, book handling).
- Comprehension and early reading (fluency, text structure).
- Early writing (emergent writing, writing process, writing application and composition).
- Expressive language (verbal communication, communication through speech, speaking).
- Receptive language (listening to others, understanding, comprehending).
- Social language (social rules of language, social communication).
- Vocabulary (lexical knowledge).

ROUNDTABLE DISCUSSIONS AND QUESTIONS

How do we capture children's abilities across languages? Are we assessing developing literacy skills in English, or in English and Spanish?

Do we need to measure everything in English, Spanish, or both languages? Which subtests are most important for capturing skills in Spanish? What are the differences/

similarities in language and literacy development indicators in each language and for bilingual learners? Is there a way to take into account the child's language dominance or previous exposure to English, allowing for an assessment format that evaluates these skills in English, Spanish and/or bilingually to align the language skills tested with the capabilities of the student?

"I think it would be a big missed opportunity to say our only goal here is to care about English literacy development because dual language teachers would care about Spanish literacy development too."



Can we tap the potential of AI through conceptual scoring? Would this help measure oral language skills objectively and translinguistically?

Al and speech recognition technologies make possible conceptual scoring, which documents responses across languages for each item. In other words, students can respond to each item in their native language, second language, or both. This would allow the students to demonstrate their understanding of different concepts, resulting in scores that can more reliably be compared with those of monolingual Spanish and English speakers. It would also give teachers one score to work with, and account for children's aptitudes in both languages. This concept also aligns well with an assetbased approach to testing and teaching bilingual students.

Can we also tap the potential of AI through speech recognition allowing for a wide range of accents/dialects and also account for dialectal differences in vocabulary, pronunciation, prosody, etc.?

The use of AI shows promise in providing objective and informed scoring that more accurately reflects the student's language and academic skills. This minimizes concerns about dialectal considerations and which type of Spanish will be used to create the assessment. The choice of items should still be sensitive to differences in vocabulary, prosody, etc.

Is dynamic assessment part of the solution?

Dynamic assessment evaluates the underlying abilities/processes that are predictive of a neurodevelopmental disorder and can indicate the capacity for learning. "Dynamic assessment would allow for a test, teach, retest paradigm to assess students' capacity for learning instead of what they know at this one moment in time."

Are there skill areas that do not apply to Spanish speakers due to language/cultural norms? Such as rhyming?

Rhyming and phonemic awareness are cornerstones of early literacy assessment in English, however syllabic structure is a better predictor of reading in 3 and 4 year olds when assessing early literacy skills in Spanish. This prompted the question of whether other phonological tasks, such as sentence repetition or non-word repetition, could be equally valuable in assessing a student's ability to repeat sequenced words/sounds. Roundtable participants also asked how important it might be to include morphology -- a linguistic compenent important in Spanish.

"Assessment would benefit from the perspective of a linguist well versed in EL Spanish phonology and morphology, and the order and milestones in which they are acquired."



Can we build a comprehensive assessment that is also a risk screener? Does it matter?

Assessment focuses on how a child is performing at a point in time, and screening focuses on prediction. Can we do both in a bilingual assessment?

- As the roundtable practitioners pointed out, no matter the cause or origin (inadequate instruction or a neurobiological issue) the intervention is the same. All that changes is the intensity.
- Roundtable participants agreed that there is greater power and opportunities to take action with little to no downside when the assessment is comprehensive and provides predictive screening.

Is it reasonable to apply the same research for flagging a native English speaker at risk for dyslexia to a Spanish speaker, given what research tells us about links to phonological processing and RAN? And what is the timeframe for prediction? Is kindergarten success the appropriate predictive target?

Research suggests that there is enough universality in the underlying cognitive factors of language so that similar variables could be used in the prediction of reading problems regardless of the presentation of language.¹⁴

- Do we need to be careful about plugging scores into the already-built English predictive model?
- Are the skills assessed in English equally applicable to the Spanish language do they become more aligned at 5 years and up when the students become skilled in manipulating phonemes and learning sound-symbol correspondence?

If we include risk assessment, can we mitigate overflagging of ELs?

It is important to find out why ELs are being flagged. Is it because students didn't understand the instructions, or because there is a neurocognitive deficit, or some other reason?

How can we resolve the "environmental factors" dilemma?

Bilingual language experiences and development differ from monolingual language experiences and development. While the neurocognitive functions that impact reading in monolingual students are similar to those in ELs, their language development experiences, including English language proficiency, can significantly impact the development of language and literacy skills.

This is where collecting information on environmental factors can be helpful. In addition to family history of reading challenges (helpful information for all children), and family socioeconomic status, additional factors that are especially helpful with ELs include:

- Previous educational experience, including pre-K schooling and type of program.
- Language of instruction in school.
- Language spoken at home.

Roundtable participants struggled with whether or not collecting information from families is advisable or realistic in real-world educational settings. Obtaining accurate family information within a school setting can be challenging. While some families may have a deep trust in educational systems and institutions, others may be reluctant to share or may not share accurate information.

"Some of our EL families have distrust of 'the system' or have been labeled and stigmatized; that makes it difficult to collect family information and hard to know if it's accurate."



Roundtable participants also struggled with the potential that knowing environmental factors (e.g., SES, parental level of education) may perpetuate bias among educators, prompting many questions about how and when to collect information on students and families:

- Do the benefits outweigh the risks?
- What data should be captured in the assessment?
- What data is most relevant, appropriate, and viable, given the User-Informed Principle(s) and the assessment goal(s)?
- What data will provide the most predictive power (e.g., family history of dyslexia)?

Can we at least capture language/methodology of instruction?

One of the most significant environmental factors that impacts every other aspect of the assessment process is the language/methodology of instruction for the EL:

- Structured English Immersion
- English as a Second Language (ESL)
- Transitional Bilingual Education
- Dual Language (one-way, two-way, 50/50 model, or other)

We know that Spanish-speaking ELs who develop reading proficiency in Spanish first, acquire English reading skills more quickly. We also know that children taught in both Spanish and English perform better on indicators of English reading proficiency than ELs taught in English-only classrooms.

It was noted, however, that a key challenge with these models is that districts do not

deliver instruction the same way. Even within a specific model, there is no standard method of implementation. Instruction can vary widely from district to district, school to school, and classroom to classroom.

ASSESSMENT MUST-HAVES AND NICE-TO-HAVES

Must-Haves

- Assessment items that are culturally and linguistically responsive
- A combination of concepts known in the home language with those concepts being learned in the second language
- Assessment subtests that help teachers drive instruction
- Assessment subtests that meet Head Start and state/local requirements
- Ability to collect as much relevant demographic information as practical (e.g., at a minimum, the home language and prior educational experience, and if possible extend to family history of reading challenges)
- Use of AI (conceptual scoring) to support oral response tests and ascertain the level of language proficiency in Spanish, English, or both

Nice-to-Haves

- Assessment in Spanish and English across subtests
- Al administration in the student's native language (e.g., as students answer a prompt or questions in the assessment, speech recognition technology can detect their native language)
- In-depth follow-up assessments to look at the entire repertoire of literacy and language skills (e.g., having the ability to see the full spectrum of language development/ acquisition in English, Spanish, and/or translanguaging)

Implementation

2 How are we assessing those skills? What is the user experience and scalability?

User-Informed Principles Go	als	Bilingual Assessment Must-Haves
the optimal experience enjoy of important stakeholders for using or benefiting from edu the assessment(s) and the resulting data, including children, educators, and families.	rument(s) are byable and engaging children and easy for cators to administer. rument(s) can be ninistered at scale in licly funded pre-K ems.	 Ease of administration for teachers. Quick, targeted, developmentally appropriate assessments and gaming. Culturally and linguistically responsive assessment items and gaming. Engaging, fun, inviting game-based experience for students. Family involvement.

WHAT DOES THE RESEARCH SAY?

To create an equitable assessment for Spanish-speaking ELs, special attention should be given to the factors that make this population unique.

As mentioned, the research is less definitive about how to assess than what to assess, though there is agreement that the instrument needs to be accessible for students by:

- Accommodating students in various stages of second language acquisition: preproduction, early production, speech emergence, intermediate language proficiency, and advanced language proficiency.
- Focusing on language development in monolingual and bilingual ELs as well as those

Roundtable Discussions and Findings | Implementation

who may speak other dialects or demonstrate "code-switching," which is the use of two languages within a sentence or phrase.

WHAT DOES THE MARKET NEED AND WANT?

Assessments that offer ease of use to teachers – both to administer and to translate into action.

Users are looking for fast, friendly, and accessible assessments. Most current marketleading pre-K Spanish language and literacy assessments are resource- and timeintensive, taking 30-60 minutes at the pre-K level and 1-2 hours at the kindergarten level.

A student-friendly and engaging gaming experience.

Assessments should be brief and developmentally-appropriate, with visuals, instructions, and graphics that support students in completing the assessment with minimal adult support.

ROUNDTABLE DISCUSSIONS AND QUESTIONS

Clearly the time burden for teachers is a big pain point. How do we make assessment quick and easy?

Teachers are key to the administration and utilization of assessment data, and their greatest need is a tool that is quick and simple and does not rely on their knowledge of the student's first language. Designing for teachers first is a good place to start.

- Prioritize the key subtests at various points in the year, and keep the subtests as simple as possible.
- Vary the length of the assessment at various benchmark periods e.g., very short early in the year, then adding subtests, as developmentally appropriate, throughout the year.
- Automate as much as possible.

\sim

"Products that stay in the market are easy to understand and implement. Many times there are a ton of features that teachers say are 'nice' but will never use. It's worth it to make sure teachers can not only understand the 'what,' but also how to make the data actionable. In short, put the teacher experience at the TOP of the list."

How can we support teacher buy-in and understanding of the why and what of the assessment?

Teachers, especially at pre-K, often view assessments as detached from what they see as their real job — the education and care of children. And they may see bilingual assessment as a further step removed. "Too much assessment is compliance-driven. Assessment needs to make sense to teachers and connect to classroom curriculum and expectations."

Can we design in a way that inspires family involvement in the assessment process?

Teachers report that they may not know much about students' families, especially when there is a language barrier. Simply knowing the language spoken at home, the prior educational experiences of the child, including the language of instruction, can make a huge difference in meeting children's needs.

How and where can AI play a role in making administration easier, and at the same time minimize bias?

The current leading pre-K assessments are mostly paper-pencil with observational scoring. New technologies offer opportunities to transform the assessment process. Technologies that both ease the assessment burden on teachers and reduce the bias inherent in observational data, should be explored and adopted where possible.

Different programs tackle assessment administration in different ways. Program leaders need an implementation pathway rubric and flowchart that includes best practices.

More than the K-12 market, pre-K has a myriad of players, processes, requirements, and program structures. To encourage implementation with fidelity, could we design implementation models that meet programs where they are? And take into account the specific language/methodology of instruction?

How can we embed professional development into the platform?

Time for segregated professional development is limited. Professional development is needed that helps educators understand the what and why of the assessment, as well as supports them in interpreting assessment data. Further, educators need to have guidance specific to interpreting the data for ELs, including how to take advantage of

Roundtable Discussions and Findings | Implementation

strengths to accelerate their language development trajectory in both their native and second languages.

It can be difficult for educators to change paradigms and practices. How can we support change management, which can make the difference between a successful implementation and a return to prior practices?

Change management is do-able, however it takes time and resources to plan for and implement a new assessment, garner teacher buy-in, and conduct professional development.

ASSESSMENT MUST-HAVES AND NICE-TO-HAVES

Must-Haves

- Teacher-friendly easy to administer, short and to the point, auto-scored
- Student-friendly gaming and assessment experience
- Developmentally appropriate, including matching the attention span of the child to their age (e.g., maximum of 30 minutes and/or broken up into two sessions for pre-K)
- Scalable and affordable, with usability features that are attractive to a range of buyers such as Head Start and state or regional buyers, pre-K chains, public school districts, or private providers
- Option of subtest instructions in Spanish, even if the subtest is in English.

"Whatever instrument is created, it needs to have the most support possible for teachers and connect to what they are already doing."

Nice-to-Haves

- A video in Spanish that students (and potentially families) can view to help them better understand the assessment instructions
- A choice in the language of the assessment, with the responses scored in the child's native/preferred language

Nice-to-Haves (continued)

- An accompanying app for parents/guardians that detects their spoken language; uses Al to explain the assessment process, demonstrate a game, collect basic information (such as language spoken at home, prior educational experiences and language of instruction); and allows them to submit questions to the teacher
- Guidance around specific phonotactics in Spanish that do not transfer in English and vice versa so that when teachers interpret reports they can clearly identify language variation or reading risk (this will also give insight into instructional steps in between benchmarks)
- Guidance to help schools/districts create the conditions under which the assessment could be utilized in the best way possible (e.g., who to bring in during the assessment, analysis, planning, etc.)
- Best practices for assessment and next steps within each specific language/methodology of instruction (e.g., Structured English Immersion, English as a Second Language, Transitional Bilingual Education, Dual Language)
- Create a new character (like Pip) with a Spanish name that is gender-neutral. Also consider creating:
 - A video (e.g., TikTok or YouTube) in Spanish with the new Spanish character
 - A mascot that can visit schools
 - A puppet that can be used at schools
 - A plush toy of the character

Jata Data What are the psychometrics that give us confidence that the data is valid and reliable? How useful is the data?

User-Informed Principles	Goals	Bilingual Assessment Must-Haves
Psychometrics – The extent to which assessment(s) reflect valid, psychometrically sound, and comparable results for children across races, ethnicities, income levels, early childcare settings, or geographic areas within the United States. Usefulness – The meaningful utilization of the assessment data and findings for important stakeholders, including educators, families, pre-K programs, and policymakers. Data will inform how educators can tailor children's support and instruction appropriately, and also inform program and policy decisions.	 Instrument(s) collect objective information to produce psychometrically sound and valid data that reflects minimal statistical bias. Instrument(s) generate timely, easily accessible, readily digestible, and understandable information for several purposes. 	 Assessment norms for ELs to allow for "true peer" data comparisons Assessments should collect objective information and produce psychometrically sound and valid data. Tools should accurately capture bilingual children's abilities and reflect minimal statistical bias. Tools should ensure reliable and valid assessment for ELs. Assessment data that can easily be used to inform instruction and intervention Gamified assessment and use of automated scoring and AI to minimize bias Data reporting and analysis tools that have different levels of aggregation for various stakeholders, e.g., state and regional oversight agencies, program administrators,

teachers, families

WHAT DOES THE RESEARCH SAY?

The purpose of the assessment must be clearly defined.

Subtests and test items should be designed with the following in mind:

- A test accurately measures a specific function or skill
- The test has the potential to improve student learning based on the test data
- The test items ensure equal access to content for all students
- The test is culturally and linguistically appropriate for the populations being measured

Norm-referenced data needs to reflect the population being tested.

WHAT DOES THE MARKET NEED AND WANT?

The market wants a tool that is:

- Valid and reliable
- Aligned with curriculum
- Includes clear, explicit next steps and resources
- Offers practice with formative assessment elements

A tool, especially a game-based tool, that puts it all together into a single package has market potential in this growing sector. The market is looking for a science-based product that will transform the EL assessment experience, and provide actionable data and next steps instructional resources.

"The bottom line is that whatever assessment instrument is developed, it should lead to instruction."

Scores need to be comparable across English and Spanish, with norms that make sense.

"Currently, we have to give students an English test and a Spanish test, but then we never put the results together. I don't think there's an assessment tool out there that actually looks at this across languages."



*

Assessments should identify student reading difficulties early.

Covid setbacks and the historically low percentage of students who are proficient readers by third grade highlight the need for the early identification of students who may struggle with learning to read, and the need for assessments that have preK-3 alignment. New York City, Rhode Island, and Virginia recently invested in initiatives designed to provide preK-3 alignment. The trend is expected to continue. "I really push back against the 'wait and see' approach that is used with a lot of bilinguals, where we let them skate by for years. Then you have an issue of educational inequity that is compounded for an unnecessarily long time."

Reports need to be online and auto-generated, tailored to the many program stakeholders who use the data (e.g. teachers, program directors), and aligned with regional and state data reporting requirements.

This could include features such as:

- Dropdown menu offering flexible ways to aggregate and disaggregate data for reporting and analysis, and to monitor for equity
- Data available at the student, classroom, grade, school, and district levels
- Side-by-side reporting for beginning, middle, and end of year assessments
- Ability to view by student type (ELs and non-ELs)
- Student reports that clearly show which students need help and what they need

Tools need to help teachers translate data to action, highlighting both strengths and weaknesses.

This includes professional development to help teachers interpret scores in the context of how bilingual students acquire proficiency in a first and second language, using both their strengths and areas for development to guide instruction.

Online, auto-generated reports should be tailored to the many program stakeholders who use the data (e.g., teachers, program directors), and aligned with regional and state data reporting requirements.

Family reporting needs to make sense, be relatable and understandable by families, and encourage school/home partnership in language development in both Spanish and English.

Family reports should highlight each student's strengths and areas for growth, and generate a next steps profile. Ideally the platform facilitates conversations on how student strengths can support the development of other skills and acknowledges the compensatory, resiliency, and translanguaging skills as assets in language and literacy development.

ROUNDTABLE DISCUSSIONS AND QUESTIONS

Is quick and simple at odds with valid and reliable?

Currently available bilingual assessments do not appropriately and comprehensively cover the range of skills needed to create a comprehensive understanding of a child's literacy and language strengths and weaknesses. How do we strike a balance between a tool that is quick and simple and a tool that is giving us the snapshot we need to provide targeted instruction?

How large does the validation sample need to be to generate norms and account for the diversity in the population of Spanish speakers?

Should we consider:

- Using a larger sample size and a normative sample with varying demographics to account for variability among Spanish speaking students. Also consider regional variations; norms may be needed regionally rather than nationally for Spanish speakers.
- Potentially using a greater number of items per subtest to respond to the variability in the performance of young children who are ELs.

Do we need more research to generate models to support predictive reliability and validity for ELs?

Do we know how the current risk models (based on monolingual student data) perform with ELs? Is this something we need to tackle to avoid misidentification (e.g., underestimation or overestimation) of bilingual learners' skills? Further, how do we know the assessment is sensitive enough to separate the issues of bilingualism from literacy impairment? What validation testing do we need to answer this question?

How do we encourage trust in the data such that educators embrace the data and are inspired to take action?

The data dashboard format should help educators make connections between the data and across all students in their classroom, including those students who are tagged as ELs. The assessment should be seen as one measure to be viewed in combination with other measures.



"Educators need to trust the data."

How can we avoid educators viewing the data as punitive?

Centering the assessment experience on equity means that data needs to feel actionable and empowering. Yet, too often, educators view the data as punitive and focused on weaknesses only. It can also be difficult to connect assessment data with what to do next (e.g., at the individual level, group level, etc.) This is another reason why it is important to include strengths and weaknesses together.

"Teachers sometimes feel 'attacked' around data, so they are unwilling to engage in the data collection."

What are the different data needs of each user persona?

Each user needs different aggregations, disaggregations, and levels of data and they bring different expertise to interpreting the data, for example:

- Administrators are looking for ways to easily use data for compliance reporting and spotting trends and overall strengths/weaknesses.
- Teachers don't always know how to interpret the assessment data, and translate data to instruction.
- Families need data that invites trust, celebration, and partnership in their child's learning and in supporting their child's home language.

How can data support students in their pre-K to K transition?

Alignment and data that travels with the student from pre-K to K has significant benefits — both for the continuum of instruction and for an appreciation of the importance of pre-K in early literacy development. "Too often we hear parents say pre-K doesn't matter; students are going to get retaught in kindergarten anyway."

The disconnect between pre-K and K is frustrating for teachers. Teachers are expected to teach all students in an equitable way, but their students may have very

Roundtable Discussions and Findings | Data

different experiences and skill levels (e.g., some students attended pre-K and some did not), and teachers have no data that shows this. This also raises equity of experience concerns for core instruction and intervention.

Could family engagement before, during, and after the assessment support a trusting school/family partner relationship?

To reinforce ways in which parents/caregivers see themselves as partners in their child's education, could we go beyond simply providing a report of results? Could we purposely design family involvement in the assessment process?

How do we help teachers seamlessly link data to instruction and intervention?

Include data that:

- Breaks down skills and gives teachers discrete items to focus on.
- Connects to next step resources and interventions to support ELs in the classroom and at home.
 - With recommendations and supports that are easily accessible, and culturally and linguistically responsive.
 - Next steps that directly and cohesively align with and link to the curriculum.
 - Next steps that make explicit connections for different users.
- Can be used by teachers to group students with similar needs.
- Can be used for progress monitoring.

How can a bilingual assessment address the needs of long-term English learners (LTELs)?

The wait-to-fail approach doesn't work. Early identification of risk can give kids more access to intervention.

- Research shows that learning to read in Spanish is easier than in English. Should we promote having children learn to read in Spanish first and then transfer to English? Results indicate that students do much better in both languages, particularly if their native language is taught systematically.
- Instead of percentiles of skill achievement, what if an assessment provided a percent of certainty of risk (assessed across languages)?

ASSESSMENT MUST-HAVES AND NICE-TO-HAVES

Must-Haves

- Assessments collect objective information and produce psychometrically sound and valid data.
- Assessment norms for ELs to allow for useful comparisons with other ELs and with monolingual students
- Assessment data that can easily be used to guide instruction
- Gamified platform with auto-scoring to objectively account for dialects and accents, avoid subjective observational assessment, and avoid bias inherent in manual scoring
- Flexible dashboard with multiple ways of manipulating and presenting data
- Data and reports that provide easy-tounderstand information for all stakeholders, e.g., state and regional oversight agencies, program administrators, teachers, families
- Reports that highlight each student's strengths and areas for growth, and generate a next steps profile

"There is an incredibly wide range of proficiency levels when we're talking about bilinguals. It may be helpful to organize them into a few manageable profiles and ensure that the assessment covers these profiles."

Nice-to-Haves

- Dashboard has options for school administrators to connect the language and literacy data to the other assessments they deliver, facilitating triangulation of data
- Guidance and training so that teachers see the assessment as being trustworthy, important, worthwhile, and actionable
- Data dashboard that pulls in the perspectives of all of the people involved a student's learning (e.g., ESL teacher, bilingual teacher, classroom teacher, Speech/Language pathologist, special education teacher, reading specialist, principal, district administrator, parents)

Nice-to-Haves (continued)

- Assessment results are aligned with curriculum and assessment data from other tools and/or domains
- Information that is easy to understand for the parent and the child so they can take ownership, advocate for, and have a say in their education. This mght be a "story" that describes what the child's issues are, what that means, what to expect at school, and how parents can help at home and/or a multimedia mode of communication that can be shared with parents who have varying levels of literacy

Ś

"When I can talk with parents about their children's strengths, I'm met with an avalanche of tears. They say, 'You're the first person who has told me that my child has strengths!' It's a very powerful way to establish a link with parents."



Connecting the Dots

During the roundtable, we set out to collect, connect, and clarify research on the markets, buyers, and users we could serve, and the top challenges they face.

We learned that there is significant demand for a bilingual early literacy assessment that addresses the five User-Informed Principles (see page 8). The existing marketleading pre-K Spanish early literacy assessment products address one or more of the principles to varying degrees, but none put it all together.

A groundbreaking, user-informed bilingual assessment for ELs must, at a minimum, address the following areas. The assessment should:

- Address bilingual considerations, including: capturing abilities across languages; evaluating skills in English, Spanish, or both; and using conceptual scoring to account for bilingual aptitudes.
- Gather data on environmental factors that impact the assessment, such as language/ methodology of instruction and relevant, accurate family information.
- Be quick, targeted, developmentally appropriate, and culturally and linguistically responsive.
- Be psychometrically sound, collect objective information, minimize bias, and provide norms specific to EL student data.
- Provide a user experience that is enjoyable for children and easy for educators, and include professional development and training.

At the end of the day, with the benefit of the research and market presentations as well as whole group and small group discussions, participants identified the most important problems to be addressed in a new groundbreaking assessment.

Top Five Problems to Address

As mentioned, roundtable participants concluded that the priority need is:

Constructing assessment content that reflects cognitive, linguistic, and cultural features of Spanish and English, and that produces psychometrically reliable and valid data/scores that are actionable for instruction.

Connecting the Dots | Next Steps

Top Five Problems to Address (continued)

This single priority statement was generated from a whole group discussion about the top five problems to address, listed below in priority order:

- 1. Content that fails to reflect cognitive, linguistic, and cultural features of Spanish that are well-validated
- 2. Lack of appropriate norm, comparative reference for EL student data
- 3. Assessment data that is not connected with what to do next
- 4. Data presentation that lacks support for interpretation; lack of connection with other student information like EL status, background, etc.
- 5. Lack of attention to how different personas need different levels/aggregations of data and how they bring different expertise to interpreting the data

Next Steps

We have a unique opportunity to do for Spanish-speaking and bilingual students what we have done for English-speaking students. We can build a bilingual early literacy assessment that predicts a student's risk for dyslexia and other reading difficulties at the earliest stages, when the window for intervention is most effective.

Utilizing our ongoing roots in scientific research, EarlyBird Education is committed to bringing research to practice, improving screening efficacy, lowering barriers to access, and, as a result, significantly increasing the number of students who are reading at grade level by third grade.



Additional Resources

- MDRC Measures for Early Success Initiative
 - MDRC Measures for Early Success: <u>User-Informed Principles: Developing</u> <u>Assessments for All Early Learners</u>
- "Bilingual Spanish English Learners: Research Foundation for an Early Identification Assessment" white paper by Dr. Elsa Cárdenas-Hagan and Dr. Eric Tridas
- Colorín Colorado ELL Roundtable 2023 Resources
- Rhinehart, L. V., & Gotlieb, R. J. M. (2023). <u>English Learners' Performance on a</u> <u>Measure of Dyslexia Risk</u>. Learning Disabilities Research & Practice. Wiley Online Library. First published: 24 July 2023. https://doi.org/10.1111/ldrp.12316. (Study uses EarlyBird)

References

1 National Center for Education Statistics (NCES, 2023).

- 2 U.S. Census
- 3 National Center for Education Statistics (NCES, 2023).
- 4 Federal Interagency Forum on Child and Family Statistics. (2015). America's children: Key national indicators of well-being, 2015, Table POP3. Washington, D.C.: Government Printing Office. Retrieved from http://www.childstats.gov/americaschildren/tables.asp.
- 5 On the 2022 National Assessment of Educational Progress (NAEP), only 33% of fourth grade students performed at or above the NAEP Proficient level on the reading assessment. This was 2 percentage points lower compared to 2019 and 5 percentage points higher in comparison to 1992, the first reading assessment year. The percentages were even lower for English learners. In 2022, less than 10% of ELs performed at or above the NAEP Proficient level. Given the large number of students who are struggling with reading — including native English speakers and English learners alike — it is important to identify and support them as early as possible.
- 6 Francis, D. J., Santi, K. L., Khalaf, S., Bunta, F., Rojas, R., Gusewki, S., McCardle, P., & Hiebert, L. (2020). Identification of Disabilities in Spanish–Speaking Children in English–Speaking Contexts. In Y. Shtyrov & E. L. Grigorenko (Eds.), All About Language (pp. 155–170). Paul H. Brookes Publishing Co.

References (continued)

- 7 Brice, R. G., & Brice, A. E. (2008). Investigation of phonemic awareness and phonic skills in Spanish-English bilingual and English-speaking kindergarten students. Communication Disorders Quarterly, 30(4), 208–225; Fiestas, C. E., Lugo-Neris, M. J., Pratt, A. S., Peña, E. D., & Bedore, L. M. (2021). Spanish language and literacy intervention for bilingual children at risk for developmental language disorder. Topics in Language Disorders, 41(4), 309–321; Francis, D. J., Santi, K. L., Khalaf, S., Bunta, F., Rojas, R., Gusewki, S., McCardle, P., & Hiebert, L. (2020). Identification of Disabilities in Spanish-Speaking Children in English-Speaking Contexts. In Y. Shtyrov & E. L. Grigorenko (Eds.), All About Language (pp. 155–170). Paul H. Brookes Publishing Co; Gaab, N. & Petscher, Y. (2022). Screening for early literacy milestones and reading disabilities: The why, when, whom, how, and where. Perspectives on Language and Literacy, 48(1).
- 8 August, D., & Shanahan, T. (2006). Developing literacy in second-language learners: Report of the National Literacy Panel on language-minority children and youth. Lawrence Erlbaum;
 Zugarramurdi, C., Fernández, L., Lallier, M., Carreiras, M., & Valle-Lisboa, J. C. (2022). Lexiland: A tablet-based universal screener for reading difficulties in the school context. Journal of Educational Computing Research, 60(7), 1688–1715. https://doi.org/10.1177/07356331221074300
- 9 Carrillo, M. (1994). Development of phonological awareness and reading acquisition: A study in Spanish language. Reading and Writing: An Interdisciplinary Journal, 6(3), 279–298.; Carillo, M., & Marin, J. (1998). Evaluar la conciencia fonológica-silábica, Cómo, cuándo y porqué. "Edetania: Estudios y propuestas de educación"
- 10 Cloud, N., Genesee, F., & Hamayan, E. (2000). Dual language instruction: A handbook for enriched education. Boston, MA: Heinle & Heinle.
- 11 August, D., & Shanahan, T. (2006). Developing literacy in second-language learners: Report of the National Literacy Panel on language-minority children and youth. Lawrence Erlbaum; Ardila, A., Rosselli, M., Matute, E., & Guajardo, S. (2005). The influence of the parents' educational level on the development of executive functions. Developmental Neuropsychology, 28(1), 539–560;
 Willcutt, E. G., Pennington, B. F., Duncan, L., Smith, S. D., Keenan, J. M., Wadsworth, S., DeFries, J. C., & Olson, R. K. (2010). Understanding the complex etiologies of developmental disorders: Behavioral and molecular genetic approaches. Journal of Developmental and Behavioral Pediatrics, 31(7), 533–544.
- 12 Ayre, A., Haynes, C., Hook, P., & Macaruso, P. (2010) Predictors of English Reading Skills in Spanish Speaking English-Language Learners. (SpELLs). In Selected Proceedings of the 12th Hispanic

References (continued)

Linguistics Symposium, ed. Claudia Borgonovo, Manuel Español-Echverría, & Philippe Prévost, 298-311. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #2425.

13 Measures for Early Success Content Blueprint Series (MDRC, 2023)

14 August, D., & Shanahan, T. (Eds.). (2006). Developing literacy in second-language learners: Report of the National Literacy Panel on Language-Minority Children and Youth; Zugarramurdi, C., Fernández, L., Lallier, M., Carreiras, M., & Valle-Lisboa, J. C. (2022). Lexiland: A tablet-based universal screener for reading difficulties in the school context. Journal of Educational Computing Research, 60(7), 1688–1715. https://doi.org/10.1177/07356331221074300; Jiménez, J. E., González, C. J. A., Monzo, A. E., & Hernández-Valle, I. (2000). Onset-rime units in visual word recognition in Spanish normal readers and children with reading disabilities. Learning Disabilities Research & Practice, 15(3), 135–141; Bordoy, S., Luque Vilaseca, J.L., Alegira, J., Carrillo, M., Gimenez, A., & Lopez Perez, J. (2013). Predictors of developmental dyslexia in Spanish-speaking school children: A longitudinal study. San Sebastian, Spain: International Workshop on Reading and Developmental Dyslexia.

Acknowledgements

Roundtable Participants

We wish to express our deepest appreciation to all of the participants who so generously shared their time and expertise during our May 24th roundtable. The wealth of knowledge and experience they contributed was invaluable to this project.

Ashley Adams Sanabria | San Diego State University | Assistant Professor of Speech, Language, and Hearing Sciences

Sandra Barrueco | The Catholic University of America | Professor of Psychology & Director of Clinical Training

Marta Benito-Gomez | MDRC | Research Associate

Susan Bonaiuto | EarlyBird Education | Director of Strategic Initiatives

Lydia Breiseth | WETA/Colorín Colorado | Director, Colorín Colorado

Elsa Cárdenas-Hagan | Valley Speech Language and Learning Center | President

Eileen Catizone | EarlyBird Education | Head Literacy Specialist

Sheena Collier | The Collier Connection | Consultant

Agape Deng | SoapBox Labs | Senior Computational Linguist

Liz Ford | EarlyBird Education | Associate Product Manager / SLP

Grace Gee | EarlyBird Education | Head of Product

Emily Hanno | MDRC | Research Associate

Charles Haynes | MGH Institute of Health Professions | Professor and Clinical Supervisor

MaryKate Heavey | Boston Public Schools | Early Childhood District Literacy Coach

Juan Luis Luque Vilaseca | Malaga University, Spain | Professor

Shatta Mejia | EdSolutions | Consultant

Appendices | Acknowledgements

Roundtable Participants (continued)

Jenny Mosley | Bill & Melinda Gates Foundation | Program Officer

Amy Pratt | University of Cincinnati | Assistant Professor of Communication Sciences and Disorders

Carla Small | EarlyBird Education | CEO

Alexa Sorden | Concourse Village Elementary School | Founding Principal

Cindy Soule | Portland Public Schools | Literacy Teacher Leader

Eric Tridas | Eric Tridas, MD | Developmental Pediatrician

Camila Zugarramurdi | Lexiland and Universidad de la República, Uruguay | Founder and Assistant Professor



Project Contributors

This endeavor would not have been possible without the generous support and participation of our project contributors. We would like to express our deepest gratitude to the following organizations and individuals.

Our team of innovators includes:

- EarlyBird Education
- Collier Connection
- Dr. Elsa Cárdenas-Hagan, Valley Speech Language and Learning Center
- Dr. Eric Tridas, Developmental Pediatrics Consultant
- <u>SoapBox Labs</u>
- <u>MDRC</u>
- EdSolutions

Contact Us



earlybirdeducation.com

A breakthrough for teachers. For students. For reading success.

For Educators: info@earlybirdeducation.com

For Families: <u>families@earlybirdeducation.com</u>

Phone: <u>1 (833) 759 2473</u>

copyright © 2023 EarlyBird Education