

Independent Evaluation of Colorado READ Act Materials



Approved Assessment List and Advisory List for
Instructional Programming and Professional Development

Linda Friedrich, Bejanae Kareem, Matthew Brunetti, Priya Kannan, Andrea Rolla,
and Rachel Tripathy, WestEd
Jill Pierce and Robin Wisniewski, RTI International

May 20, 2025

Contract Number: 148306



© 2025 WestEd. All rights reserved.

Suggested citation: Friedrich, L., Kareem, B., Brunetti, M., Kannan, P., Rolla, A., Tripathy, R., Wisniewski, R., & Pierce, J. (2025). *Independent Evaluation of the Colorado READ Act Materials*. WestEd.

WestEd is a nonpartisan, nonprofit agency that conducts and applies research, develops evidence-based solutions, and provides services and resources in the realms of education, human development, and related fields, with the end goal of improving outcomes and ensuring equity for individuals from infancy through adulthood. For more information, visit [WestEd.org](https://www.wested.org). For regular updates on research, free resources, solutions, and job postings from WestEd, subscribe to the E-Bulletin, our semimonthly e-newsletter, at [WestEd.org/subscribe](https://www.wested.org/subscribe).



Table of Contents

Executive Summary	ES-1
Overall Findings	ES-3
Assessments	ES-5
Instructional Programs	ES-6
Professional Development Programs	ES-7
Review of Spanish-Language Instructional Program Rubric	ES-8
Summary of Recommendations for Each Type of Material	ES-8
Introduction	1
The Colorado READ Act.....	2
Evaluation of the READ Act.....	3
Purpose and Organization of This Report	4
Scientific Foundations of Reading Proficiency in Early Elementary Grades	6
Assessments	12
Criteria Used for Review	13
Description of Review Categories for READ Act Compliance	13
Description of Criteria Drawn from Additional Professional Standards	15
Data Collection and Methods	17
Ratings and Reconciliation Process	19
Results and Discussion	20
Discussion of Main Themes and Issues That Emerged from the Assessment Reviews.....	21
Instructional Programs on the Advisory List	28
Independent Evaluation Review Process	29
Data Collection and Methods	33
Results and Discussion	34
CDE Processes for Identifying Items for Approved and Advisory Lists	41
Professional Development Programs on the Advisory List	44
Description of Review Categories	45
Data Collection and Methods	48
Results and Discussion	49
Review of Spanish-Language Rubric.....	54
Introduction.....	55
Development of the Revised Rubrics.....	55
Key Findings and Analysis	59
Conclusion.....	59
Conclusions and Recommendations	60
Conclusions.....	61
Specific Recommendations.....	63
References	R-1
List of Appendices	A-1



Executive Summary

The Colorado State Legislature passed the READ Act in 2012 and updated the Act in 2019. The revised Act requires an independent evaluation to identify and assess strategies that the state and local districts and schools have taken to support Colorado students in achieving proficiency in reading.

This report focuses on the findings related to assessments on the Approved list, Advisory Lists of Instructional Programming and Professional Development, and CDE's processes for selecting materials for these lists.

Key Findings

- The materials the Colorado Department of Education (CDE) approved for use with Reading to Ensure Academic Development (READ) Act funds meet the minimum requirements of Senate Bill (SB) 19-199.
- All but one of the approved assessments meet the minimum summary threshold for compliance with the elements required by SB19-199: One “fully meets,” 23 “largely meet,” and six “partially meet” the requirements of the READ Act.
- For our overall summary rating, 30 instructional programs received “fully meets,” 29 received “largely meets,” and three programs received “partially meets.”

The importance of achieving early-grade reading proficiency for later student academic success is well documented. The Colorado State Legislature responded to this challenge by passing the Colorado Reading to Ensure Academic Development (READ) Act (2012) to provide school districts with funding and support to aid literacy development for students in kindergarten through 3rd grade (K–3), especially those identified with “significant reading deficiencies” (SRDs) at risk of not reading at grade level by the end of 3rd grade. In 2019, the Legislature updated the READ Act with revisions that included requirements for an independent evaluation of READ Act implementation. This report focuses on the findings related to approved assessments, Advisory Lists of Instructional Programming and Professional Development, and the Colorado Department of Education’s (CDE’s) processes for selecting materials for these lists.

The overall conclusion from the review of assessments, instructional materials, and professional development programs is **that the materials CDE approved for use with READ Act funds meet the minimum requirements in Senate Bill (SB) 19-199** (see Exhibits ES.1A and 1B).

Exhibit ES.1A. Summary of Ratings				
Type	Fully Meet	Largely Meet	Partially Meet	Do Not Meet/ Not Rated
Assessments	1	23	6	1
Instructional Programs	30	29	3	0
Professional Development Programs	7	0	0	0

Exhibit ES.1B. Explanation of Summary Ratings ¹	
Type	Explanation
Fully meets	Received a rating of “Fully meets” on all indicators
Largely meets ²	Received a rating of at least “Partially meets” on all indicators.
Partially meets	Received a rating of “Does not meet” on at least one, but not all indicators
Does not meet	Received a rating of “Does not meet” on all indicators.

The remainder of this executive summary summarizes overall findings, describes findings for each type of material, and presents some broad recommendations. The concluding chapter of this report contains more detailed findings and recommendations.

Overall Findings

- Review processes and identified resources align with READ Act requirements.** The review processes for assessments, instructional materials, and professional learning have resulted in resources that align with the guidance of the READ Act. The State Board, Legislature, and the public can be confident that CDE is working to faithfully implement the READ Act in both letter and spirit. All but one assessment, all instructional programs, and all professional development programs that received a rating of “fully meets,” “largely meets,” or “partially meets” meet the minimum threshold for alignment with the READ Act.
- Resources provide local education providers (LEPs) with choices.** The identified assessments, instructional materials, and professional development programs provide LEPS with a great deal of choice in how to use READ Act funds to support

¹ Detailed explanations of the ratings for each material type can be found in Appendices A and B.

² The rating “largely meets” is only used as a summary rating. Individual criteria can receive a rating of “Fully meets,” “Partially meets,” or “Does not meet.” The ratings are further explained in Appendices A and B.

the implementation of scientifically based reading instruction. The breadth of resources aligns with Colorado's local control focus. As noted in the Year 4 Per Pupil report, the choices also create a need for additional guidance about how to use materials and professional learning in ways that enhance each other.

- **Most instructional programs and professional learning programs are supported with the lowest tier of empirical evidence, a logic model.** The overwhelming majority of instructional materials and professional learning opportunities that appear on CDE's Advisory List are supported by Every Student Succeeds Act (ESSA) Evidence Tier 4 ("Demonstrates a rationale"). This is reflective of the national materials and professional learning market, which largely does not currently have systematic evidence of impact on student outcomes.
- **Approved resources provide limited guidance about supporting English learners (ELs) and students with disabilities.** Assessments, for example, provide limited guidance about appropriate accommodations. Further, Spanish-language instructional materials focus on building reading skills in Spanish but do not offer guidance about how these programs might be integrated into different models of bilingual education adopted by LEPs. Finally, approved professional learning programs provide limited guidance in how to support ELs in learning to read in English.
- **CDE's Spanish-language materials rubric and selection processes meet READ Act requirements.** CDE's updated rubric for Spanish-language instructional materials, like its English-language counterpart, aligns with the READ Act and with empirical literature about Spanish-language reading

instruction. We find that the rubric and review processes address the challenges that we identified in the Year 3 Instructional Materials report.

Assessments

All but one of the approved assessments meet the minimum summary threshold for compliance with the SB19-199 required elements: one “fully meets,” 23 “largely meet,” and six “partially meet” the requirements of the READ Act³ (see Exhibit ES.2 for details). Some assessment vendors provided more robust, organized, and comprehensive evidence than others; differences in the extent and quality of evidence explain, in part, differences in the evaluation’s ratings of the assessments. The most challenging criteria for vendors were those related to validity and reliability. Many assessments that received a “partially meets” summary rating did so because of a lack of comprehensive evidence to support the breadth and depth of the validity and reliability criteria.

Exhibit ES.2. Overall Summary Ratings for Assessments’ Compliance with Required Criteria					
Assessment Type	Fully Meets	Largely Meets	Partially Meets	Does Not Meet	Total
Interim	1	6	1	0	8
Diagnostic	0	13	3	1	17
Summative	0	4	2	0	6
Total	1	23	6	1	31

All approved assessments also either “largely meet” (21) or “partially meet” (9) additional professional and technical criteria examined in this evaluation. No assessments fully meet the additional professional and technical criteria. Some assessment vendors provided

³ According to the evaluation rubric, a Spanish-language diagnostic assessment received a “does not meet” rating for compliance with READ Act requirements because the assessment publisher provided minimal or no evidence to support the requirements.

more robust, organized, and comprehensive evidence than others; differences in the extent and quality of evidence explain, in part, differences in the evaluation’s ratings of the assessments. The most challenging criteria for vendors were those related to the removal of bias, appropriate accommodations, and score reports. Most assessments that received a “partially meets” summary rating did so due to lack of comprehensive evidence for these three criteria.

A recent update of the READ Act requires that approved interim reading assessments include indicators that screen for characteristics of dyslexia. **Of the interim assessments reviewed, all but one included a Dyslexia Indicator Worksheet indicating the assessment can screen for characteristics of dyslexia such as phonemic awareness and rapid automatic naming.** Two interim assessment vendors provided the same worksheet for both English- and Spanish-language versions, indicating similar functionality in screening for dyslexia traits. However, these vendors did not furnish additional evidence regarding the Spanish-language versions’ effectiveness in screening for dyslexia characteristics.

Instructional Programs

CDE’s instructional program review process is rooted in empirical evidence and reflects both historical and current understandings of how the science of reading can be applied effectively in classroom practice. **CDE fully or partially approved 62 instructional programs during its most recent review cycle. The evaluation of these instructional programs concluded that all 62 programs meet the core requirements outlined in the READ Act** (see Exhibit ES.3).

For the evaluation’s overall summary rating, 30 instructional programs received a rating of “fully meets,” 29 received a rating of “largely meets,” three programs received a rating of “partially meets,” and none received a rating of “does not meet.” All instructional programs meet the minimum threshold for evidence—a clear logic model rooted in the science

of reading—of having the potential to make a positive impact on students’ reading outcomes. All but one program demonstrated the presence of skill development in the areas of phonemic awareness, phonics, reading fluency, and reading comprehension (as applicable), with 30 fully meeting criteria for explicit and systematic skill development. All core programs meet the minimum requirements for including texts on core academic content to assist students in maintaining or meeting grade-appropriate proficiency in academic subjects in addition to reading. These summary statistics include five core programs, 28 supplemental programs, 23 intervention programs in English, and six Spanish-language programs.

Exhibit ES.3. Instructional Materials Summary Rating			
Program Type	Summary Rating: Compliance with SB19-199 Requirements		
	Fully Meets	Largely Meets	Partially Meets
Core programs in English	3	1	1
Supplemental programs in English	10	17	1
Intervention programs in English	15	7	1
Programs in Spanish (all)	2	4	0
OVERALL	30	29	3

Professional Development Programs

All seven professional development programs on the Advisory List are in compliance with all SB19-199 required elements. All professional development programs meet the minimum threshold for evidence by providing a logic model, theory of action, or synthesis of research that outlines how and why the program expects to have impact. One vendor submitted formal research documenting moderate evidence of positive impact on student outcomes. Each program is rooted in the science of reading and has the potential to make a positive impact on students’ reading outcomes. Each of the programs has rigorous evaluation throughout the course, testing teacher knowledge of reading instruction and pedagogy.

Exhibit ES.4. Professional Development Program Summary SB19-199 Requirement				
Summary Rating	Fully meets	Largely meets	Partially meets	Does not meet
Compliance with SB19-199 requirements	7	0	0	0

Review of Spanish-Language Instructional Program Rubric

The revised rubric for evaluating Spanish-language instructional programs effectively incorporates expert and community feedback. It integrates Spanish-specific literacy components, such as phonological awareness and metalinguistic skills, and aligns with research-supported practices, including those in the Colorado READ Act's *Minimum Reading Competency Skills Matrix* (2017).

Summary of Recommendations for Each Type of Material⁴

Although all but one assessment meet READ Act requirements, we found that the quality of evidence submitted by vendors varied and that this variation explains, in part, differences in the evaluation's ratings of programs. Our recommendations related to **assessments** are to:

- Adjust the vendor submission process to focus on the different purposes for assessment (i.e., interim, diagnostic, summative.)
- Ask vendors to improve their provided evidence and documentation.
- Request that vendors enhance their documentation of standards alignment.
- Encourage vendors to strengthen approaches to bias removal and cultural representation.
- Invite vendors to expand accommodations and score report usability.

⁴ Chapter 6 presents more detailed recommendations.

Now that over 100 instructional programs that meet READ Act requirements appear on CDE’s Advisory List, LEPs will benefit from additional guidance about selecting the materials programs that best meet their students’ needs. Our recommendations related to **instructional programs** are to:

- Provide additional guidance to LEPs focused on selecting instructional programs.
- Identify additional supplemental and intervention programs that focus on vocabulary and reading comprehension.
- Offer additional comprehensive bilingual guidance focused on the effective use of both English- and Spanish-language instructional programs with ELs.

Although the new Spanish-language instructional materials rubric aligns with READ Act requirements, both Spanish- and English-language programs offer limited guidance about supporting ELs. Our recommendations related to **Spanish-English bilingual guidance** are to:

- Highlight materials that provide explicit supports for integration across bilingual models.
- Incorporate a focus on approved instructional materials into professional learning.
- Provide guidance to support coordinated literacy instruction.

CDE’s review process has resulted in professional development programs that fully meet READ Act requirements, but there are opportunities to further strengthen the professional learning opportunities available to Colorado’s educators. Our recommendations related to **professional development programs** are to:

- Encourage vendors to provide more rigorous evidence of program impact that meets ESSA Tiers of Evidence 1, 2, or 3, and make logic models explicit.

- Ask vendors to submit copies of all teacher evaluation tasks embedded in the program to improve review.
- Ask vendors to provide additional information on how their programs prepare educators to address the specific needs of ELs and address culturally responsive instructional approaches.



1

Introduction

The 2019 revision of the READ Act (Senate Bill [SB] 19-199) includes a provision mandating that an independent, external evaluation of the READ Act program be conducted over a 5-year period.

The multi-year evaluation is now underway and is being conducted by an independent research team led by WestEd that includes APA Consulting and RTI International.

The key legislative goals for this evaluation are to:

- **Help state policymakers and district leaders understand the impacts of Colorado Reading to Ensure Academic Development (READ) Act funding and support on students, families, schools, and districts.**
- **Determine the extent to which the Colorado Department of Education's (CDE's) processes resulted in approved assessments and Advisory Lists of Instructional Programming and Professional Development that are consistent with READ Act requirements.**
- **Provide feedback on how CDE's processes for selecting assessments, instructional programming, and professional learning might be improved.**

The Colorado READ Act

The importance of achieving early-grade reading proficiency for later student academic success is well documented. In fact, researchers and education leaders consider achievement of reading proficiency by the end of the 3rd grade to be crucial to a child's future academic success and financial independence (Hernandez, 2012; Fiester, 2013). To help schools and districts support all children in achieving this goal, the Colorado State Legislature passed the Colorado Reading to Ensure Academic Development (READ) Act in 2012, replacing the Colorado Basic Literacy Act.⁵ The READ Act provides local education providers (LEPs), including school districts, with funding and support to aid literacy development for kindergarteners through 3rd-grade students, especially those identified with “significant reading deficiencies” (SRDs) who are at risk of not reading at grade level by the end of 3rd grade.

Under provisions of the READ Act, schools test students using reading assessments approved by the Colorado State Board of Education (CDE, 2025b). Schools are then required to develop individual READ Act plans that identify a pathway for reaching grade-level proficiency for those designated as having an SRD. The READ Act specifies certain components as required in all students’ READ Act plans; however, each plan must be tailored to meet individual student needs.

In addition to specifying that the Colorado State Board of Education approves a set of reading assessments, the READ Act also charged CDE with creating Advisory Lists of Instructional Programming (CDE, 2024b) and Professional Development (CDE, 2025a) that are scientifically based and evidence-based. LEPs may use READ Act funds to purchase instructional programming from the Advisory List of Instructional Programming. With the

⁵ The READ Act includes many of the same elements as the Colorado Basic Literacy Act, including a focus on K–3 literacy, assessment, and individual plans for students reading below grade level with the addition of (1) funding to support these efforts, (2) requirements for parent communication, and (3) an explicit focus on students designated as having significant reading deficiencies.

2019 revision of the READ Act, the legislature required all K–3 teachers to complete evidence-based training in teaching reading by January 31, 2022. Teachers who have successfully completed the professional development (PD) programs on CDE’s Advisory List meet this requirement.

Evaluation of the READ Act

The 2019 revision of the READ Act (SB19-199) includes a provision mandating that an independent external evaluation of the READ Act program be conducted over a 5-year period (see *2020 Annual Report on the Colorado READ Act* for an overview of updates in SB19-199) (CDE, 2024a). The multi-year evaluation is now underway and is being conducted by an independent research team led by WestEd that includes APA Consulting and RTI International.

The key legislative goals for this evaluation are as follows:

- Help state policymakers and district leaders understand the impacts of READ Act funding and support on students, families, schools, and districts.
- Determine the extent to which CDE’s processes resulted in approved assessments and Advisory Lists of Instructional Programming and Professional Development that are consistent with READ Act requirements.
- Provide feedback on how CDE’s processes for selecting assessments, instructional programming, and professional learning might be improved.

This report summarizes findings and data gathered during the first year of the legislatively mandated evaluation for Goals 2 and 3. The report relies on multiple sources of information, including

- Materials collected by the independent evaluation from vendors.
- Materials submitted by vendors to CDE as part of the review process, as available.

- Publicly available documentation of CDE’s review processes and timelines.
- Interviews with CDE staff who led the review processes, CDE staff who participated in review processes, and individuals not employed by CDE who participated in review processes.

Future reports will examine the implementation and impact of instructional programming and professional learning on student outcomes. Because this report focuses on newly approved programs, it is not feasible to examine impact in this report.

Purpose and Organization of This Report

In this report, the evaluation team describes the evaluation of 31 approved assessments, 62 instructional programs on the Advisory List,⁶ and seven PD programs on the Advisory List. Key data and information presented in this summary report for assessments, instructional programs, and PD programs include (a) the evidence base for the assessments or programs; (b) a focus on scientifically based reading skills; and (c) assessment- and program-specific requirements. The summary report describes the processes used, results with lessons learned, and recommendations. This report focuses on providing initial answers to research questions for each type of material. It starts with a general literature review, then describes the processes used in the evaluation and findings with discussion of lessons learned, and ends with conclusions and recommendations.

It is also important to note that this report does not address the question of whether individual instructional programs result in growth to standard. All analyses related to the impact of READ Act policies and

⁶ CDE approved 62 unique instructional programs, but some were approved for multiple categories and therefore counted more than once in our total.

investments are addressed in the *Independent Evaluation of the Colorado READ Act: Per Pupil Funding Year 5 Summary Report*.

Research Questions

This 5th-year report addresses the following questions pertaining to assessments, instructional programs, and PD programs.

Approved Assessments

- Do all items on the Approved assessment list meet the requirements of the READ Act?
- Do all items on the Approved assessment list meet additional professional standards of quality?

Advisory List of Instructional Programming

- Do all items on the Advisory List of Instructional Programming meet the requirements of the READ Act?
- Do all items on the Advisory List of Instructional Programming meet additional professional standards of quality?

Advisory List for Professional Development

- Do all items on the Advisory List of Professional Development meet the requirements of the READ Act?
- Do all items on the Advisory List of Professional Development meet additional professional standards of quality?

Spanish-Language Instructional Materials Review Rubric

1. Does the Spanish-language Instructional Materials review rubric address required instructional materials selection criteria in the READ Act?
2. Does the Spanish-language Instructional Materials review rubric reflect current research-based approaches to the development of reading skills in Spanish?

Analytic Frameworks Used for Review

The criteria used in this evaluation for reviewing assessments, instructional programs, and professional development programs derive from READ Act statutory language, updated regulatory and nonregulatory guidance, the Colorado Academic Standards for Reading, Writing, and Communicating, and other related policies and guidance. Additional criteria used in the evaluation derive from information provided by an expert advisory panel convened for this project in 2020 as well as professional standards for evaluating assessments, instructional programs, and professional development quality (see Appendix sections A.1–A.3 for detailed rubrics). Anchored in related statutes, guidance, and regulations, these evaluation protocols provide a transparent and consistent framework to determine READ Act compliance. The rubrics (see Appendix sections A.1–A.3) list the required elements, criteria for the elements, ratings for the evidence, what evidence is needed, and the workflow for the reviewers.

Scientific Foundations of Reading Proficiency in Early Elementary Grades

Decades of research have demonstrated the importance of reading proficiency in the early elementary grades. Around 3rd grade, students transition from developing foundational reading skills (“learning to read”) to using reading as a tool for acquiring information (“reading to learn”) (Adams, 1990). These early years are a critical time for intervening to support struggling readers, since students who do not have the ability to read

independently by 3rd grade are at risk of falling behind academically in subsequent grades. Longitudinal studies have shown that students with low reading test scores in 3rd grade are less likely to complete high school (Lloyd, 1978), failing to graduate on time at a rate four times higher than their proficient peers (Hernandez, 2012).

Recognizing the importance of reading in the early grades, the United States Congress asked the National Institute of Child Health and Human Development to establish a National Reading Panel (NRP) to perform a comprehensive and informed synthesis of the research around effective methods for teaching children to read. In 2000, the 14-member panel released its report, identifying five instructional components that are essential for early-grade reading development: phonemic awareness, phonics, reading fluency, vocabulary, and reading comprehension (NRP, 2000). In a minority view included with the report, panel member Joanne Yatvin cautioned Congress about interpreting the NRP findings as definitive, claiming that the scope of topics that NRP examined was biased and narrow and that the panel had neither the time nor resources to conduct analyses with the rigor required to answer their research questions with certainty. Still, the NRP findings have had substantial influence on both policy and practice, as the five essential components of reading have become widely accepted as best practices in reading instruction.

Following the 2001 passage of the No Child Left Behind Act and its emphasis on increased instructional time for reading, numerous funding and policy initiatives emerged aimed at raising early-grade reading proficiency rates. At the federal level, Reading First provided roughly one billion dollars in grants annually from 2002 through 2008 to support the NRP-recommended instructional practices (Gamse et al., 2015). At the state level, at least 26 states have passed reading laws since 2000 that are aimed at providing financial support, accountability measures, procedural requirements, and interventions that will improve 3rd-grade reading proficiency rates (Council of Chief State School Officers, 2019). Most of

these laws reference or require “scientifically based” reading instruction, interventions, and curricula; however, by the time many of these laws were passed, major publishers and teacher training programs had already adopted the five essential components of reading in response to the NRP report (Kemple et al., 2009).

With the proliferation of curricula, interventions, teacher professional development programs, and assessments centered around these five essential components has come a large body of empirical research aimed at determining the efficacy of targeting them. The studies on early reading instruction and intervention are so numerous that researchers have been able to conduct meta-analyses attempting to identify all high-quality studies on a given topic and use statistical modeling to produce a more accurate impact estimate than any single study could provide. What follows is a short summary of recent meta-analytic findings on each of the five essential components of reading for students in pre-K through 3rd grade; all five components are included in the READ Act.

- *Phonemic awareness* is the ability to notice, distinguish, and manipulate the individual sounds in spoken words (Liberman et al., 1974) (e.g., the word “juice” has three phonemes: “j-,” “ooo,” and “sss”) and is a strong predictor of students’ later reading abilities (e.g., Share et al., 1984; Snider, 1997). Research indicates that explicit instruction is highly effective in promoting the development of phonemic awareness skills and leads to moderate improvements in reading overall (Bus & Van Ijzendoorn, 1999; Ehri, Nunes, Willows, et al., 2001). Longitudinal studies have shown that interventions focused specifically on supporting phonemic awareness were found to have lasting impacts on student reading proficiency, showing a greater effect one year after the end of an intervention than interventions that were more focused on phonics (Suggate, 2016).

Phonics is an instructional approach in which students learn to sound out and blend letters in order to decode a word (which is a different skill than understanding what that word means). Explicit and systematic teaching of phonics has been shown to improve student decoding, spelling, and comprehension to a statistically greater degree than instruction without a focus on phonics (Ehri, Nunes, Stahl, et al., 2001; Jeynes, 2008). Research on phonics instruction specifically for low-performing readers similarly has found systematic phonics instruction to improve reading outcomes (McArthur et al., 2018). Explicit phonics instruction was found to have a smaller effect over time than instruction focusing on phonemic awareness and comprehension (Suggate, 2016).

Fluency refers to the relative degree of ease and automaticity with which letters are understood as words, words are understood for their meaning, and comprehension of a subject is derived from that meaning (Wolf & Katzir-Cohen, 2009). At higher levels of reading fluency, mental attention can be devoted to comprehension rather than to the mechanics of reading—fluency is therefore considered a critical link between word analysis and text comprehension. The developmental definition of fluency makes it difficult to study empirically, and evidence around the effectiveness of interventions and approaches to support fluency is mixed. There is some evidence that repeated reading and the modeling of reading (either in person or via audiobook) can improve fluency and comprehension (Chard et al., 2002; Stevens et al., 2017), but more rigorous empirical research is needed to understand how to best improve reading fluency in the early grades.

Vocabulary instruction represents an important component of reading comprehension because understanding text requires the construction of meaning from known words (Kamil, 2004). There

is strong consensus that the size of a student's vocabulary is predictive of how well they will understand what they read (e.g., Scarborough, 2001). Recent research indicates that interventions supporting vocabulary development are effective at improving expressive and receptive vocabulary (Marulis & Neuman, 2010). There is evidence that such interventions are also effective at improving comprehension of texts aligned with the intervention, but there are fewer studies that have found that these interventions improve generalized reading comprehension (Elleman et al., 2009; Wright & Cervetti, 2016). Multidimensional approaches to learning words (e.g., providing contextual information around a set of words) tend to have a stronger impact on student reading comprehension than instruction focused on definitions (Stahl & Fairbanks, 1986; Wright & Cervetti, 2016). *Reading comprehension* is the overall goal of reading instruction. It occurs when students can process the text they read, derive meaning from it, and integrate that meaning with what they already know. Gough & Tunmer's (1986) influential model describes successful reading comprehension as dependent upon two foundational components: decoding and linguistic comprehension. Others have argued that fluency is a third critical component for supporting text comprehension (Joshi & Aaron, 2000; Solari et al., 2018). Although meta-analytic evidence supports decoding (García & Cain, 2014) and linguistic comprehension as key predictors of reading comprehension, findings on the impact of phonics instruction—aimed at improving decoding—on comprehension outcomes are less consistent, with some reviews reporting only modest or inconclusive effects (McArthur et al., 2018). Part of the challenge in studying the effect of foundational components on reading comprehension is that the most important components for reading change with students'

ages. In elementary school, for example, reading ability is largely based on print knowledge and phonological awareness, whereas in middle school reading accuracy and linguistic comprehension play a larger role in overall comprehension (Ehri, Nunes, Stahl, et al., 2001; Storch & Whitehurst, 2002). It is not surprising, then, that studies show interventions focused on phonemic awareness to be most appropriate for students entering elementary school; interventions focused on phonics and fluency to have greatest effect in 1st and 2nd grade; and interventions targeting overall comprehension to be most effective for 3rd grade and beyond (Suggate, 2016).



2

Assessments

The evaluation team used the criteria established by the READ Act to build its evaluation and review rubric.

The team added criteria to the existing evaluation and review rubric for the purpose of gathering supplemental information on the assessments. These additional criteria draw from established research on principles and characteristics for identifying high-quality assessments.

Key Findings

- **Overall, the approved assessments largely meet or partially meet the requirements outlined in the READ Act.**
- **Twenty-four of the 31 assessments reviewed fully meet or largely meet all READ Act requirements.**
- **However, for each required element, vendor-provided evidence was of varying levels of quality.**

Criteria Used for Review

This section defines the criteria used for consistently reviewing all assessments. The assessments rubric (see Appendix section A.1) lists the criteria (required and additional technical quality), the required elements for the criteria, the evidence needed to support the criteria and required elements, and the possible ratings for each element.

Description of Review Categories for READ Act Compliance

As the foundation for the review and evaluation of the assessments, the evaluation team used the criteria established by the READ Act to build the evaluation and review rubric. Exhibit 2.1 summarizes the rubric and sample evidence, which is based on READ Act requirements, while the text following the exhibit provides definitions for each criterion as well as a more extensive list of evidence considered.

Exhibit 2.1. READ Act Requirements for Assessments and Example Evidence

Criterion Number	Short Name	SB19-199 Requirement	Example Evidence Reviewed
1	Evidence-based	Is evidence-based or scientifically based (22-7-1209 (2)(b)(I)(A))	<ul style="list-style-type: none"> • Theory of action about assessment • Assessment purpose and documentation of theoretical basis • Assessment development process and evaluation
2	Standards-aligned	Is aligned with the preschool through elementary and secondary education standards for reading adopted by the State Board (22-7-1209 (2)(a)(II)(A))	<ul style="list-style-type: none"> • Assessment alignment to Colorado learning standards • Minimum Reading Competency Skill levels • Component areas measured at each grade level and time of year with evidence of content standard alignment
3	Validity	Each of the recommended reading assessments is valid proven to effectively measure students' reading skills in the areas of phonemic awareness; phonics; vocabulary development; reading fluency, including oral skills; and reading comprehension (22-7-1209 (2)(a)(II)(B))	<ul style="list-style-type: none"> • Evidence that assessment addresses targeted areas of reading • Evidence that experts, including teachers, developed and reviewed the content of the assessment • Evidence of the appropriateness of tasks • Evidence that the internal structure supports score interpretations • Evidence that the total score and sub-scores are related to external variables • Evidence of an appropriate standard-setting method and process • Evidence for classification accuracy to identify students with an SRD (for interim assessments) • Evidence that information is provided to identify strengths and areas of support (for diagnostic assessments)
4	Reliability	Each of the recommended reading assessments is reliable, proven to accurately measure students' reading skills in the areas of phonemic awareness; phonics; vocabulary development; reading fluency, including oral skills; and reading comprehension (22-7-1209 (2)(a)(II)(B))	<ul style="list-style-type: none"> • Evidence of appropriate reliability estimates to support score interpretation • Reliability estimates met or exceeded 0.70 • Evidence that classification decisions are reliable • Evidence for reliability estimates is drawn from a representative sample of students • Evidence of alternate forms, including evidence for comparability of scores from alternate forms, if applicable • Standard errors of measurement reported for score ranges and cut scores

Exhibit 2.1. READ Act Requirements for Assessments and Example Evidence			
Criterion Number	Short Name	SB19-199 Requirement	Example Evidence Reviewed
5	Spanish-language assessment ^a	“At least one of the recommended reading assessments for kindergarten and 1st, 2nd, and 3rd grades is normed for the performance of students who speak Spanish as their native language, which assessment is available in both English and Spanish” (22-7-1209 (2)(a)(II)(D))	<ul style="list-style-type: none"> • Evidence that the sample utilized for norming is representative of students who speak Spanish as their native language • Evidence that experts in Spanish language and literacy were included in the development • Evidence that demonstrates culturally representative content, including accounting for dialectical differences among Spanish speakers
6	Paper-and-pencil assessment ^b	“The list of recommended reading assessments and reading diagnostics includes at least one assessment and one diagnostic that a student can complete using paper and pencil rather than using a computer” (22-7-1209 (2)(a)(II)(E))	<ul style="list-style-type: none"> • Evidence of comparability of all forms for each grade level, using a representative sample of students

^a Assessments that included a Spanish-language version or were submitted as a Spanish-language version of the English-language assessment were evaluated against this criterion; however, this criterion did not contribute to the overall rating for assessments that were not submitted as a Spanish-language version or Spanish-language assessment.

^b The paper-and-pencil criterion did not contribute to the overall rating for each assessment, as the READ Act (22-7-1209 (2)(a)(II)(E)) requires that at least one assessment and one diagnostic can be complete using pencil and paper rather than using a computer.

Description of Criteria Drawn from Additional Professional Standards

In addition to the required criteria, the evaluation team included additional criteria to the evaluation and review rubric to provide supplemental information on the assessments. These additional criteria draw from established research on principles and characteristics for identifying high-quality assessments (AERA et al., 2014; National Center on Intensive Intervention, n.d.). Exhibit 2.2 lists the criteria and includes sample evidence. The criteria are further defined and additional evidence is listed in the Data Collection and Methods section.

Exhibit 2.2. Additional Criteria for Assessments and Example Evidence

Criterion Number	Short Name	Criterion	Example Evidence Reviewed
7	Removal of bias	The assessment development and review processes are designed and implemented to remove bias against all students.	<ul style="list-style-type: none"> • Statistical analyses were conducted to detect possible bias • Bias reviews were conducted with a representative panel • Evaluation for culturally representative content • Application of Universal Design for Learning principles
8	Assessment administration guidance	The administration of the assessment is supported by appropriate guidance and resources.	<ul style="list-style-type: none"> • An administration guide (or comparable set of resources) with a scripted administration protocol or guidelines for administration • Average administration time is reasonable and balanced • Training for administration is readily available to teachers • Training materials provide clear instructions for determining which students should potentially be identified with an SRD
9	Appropriate accommodations	The assessment offers appropriate accommodations so all students can be fairly and accurately assessed.	<ul style="list-style-type: none"> • Evidence that the assessment items and accessibility features permit all students to demonstrate their knowledge and abilities • Accommodations for students with disabilities, including information to support the evidence base for accommodations • Evidence that accommodations do not compromise the interpretation or purpose of the assessment • Evidence that training materials include specific guidelines on accommodations

Exhibit 2.2. Additional Criteria for Assessments and Example Evidence			
Criterion Number	Short Name	Criterion	Example Evidence Reviewed
10	Score report usability	The assessment produces assessment data and information, such as student scores and score reports, that are usable for the intended audiences.	<ul style="list-style-type: none"> • Evidence of user-focused evaluations to demonstrate the utility of the reports • Score reports indicate whether students should potentially be identified with an SRD • Scoring guidelines are clear and easily interpreted • Cut scores, score ranges, and/or confidence intervals are clearly specified • Reports describe specific areas for additional student support (for diagnostic assessments) • Reports are designed for specific audiences • Reports are available in languages other than English • Description of the process and technology used to issue reports in a timely manner

Data Collection and Methods

Information Used to Review Programs

The evaluation team reviewed all materials submitted by the assessment vendor applications in response to CDE’s Solicitation for Assessments READ Act Advisory List: Review Period 2022–2026. This solicitation requested that assessment vendors provide the following materials in their applications:

- Section C: Assessment Overview
 - Vendor Information
 - Stated Purpose of Assessment
 - Mapping Assessed Content to READ Act Literacy Areas
 - Assessment Administration
 - SRD Cut Score Determination and Interpretation
 - Assessment Type and Content Areas for Review
- Section D: Worksheets
 - Vendor Assessment Worksheet
 - Minimum Reading Competency Skill Levels Worksheet

- Dyslexia Indicators Worksheet (optional)⁷

The documents and worksheets provided by the vendors in Sections C and D of their applications were used during our evaluation.

Training for Independent Evaluators

The lead assessment evaluator conducted the training for the assessment reviewers. As part of the training, the evaluation team reviewed the requirements of the Colorado READ Act; the assessment evaluation goals; the materials provided by the vendors in response to the solicitation; the assessment types, as defined by the READ Act (i.e., interim, diagnostic, and summative); the assessment evaluation process; and the assessment evaluation rubric. Shortly after the initial training, members of the evaluation team independently reviewed a common set of applications, then met to calibrate on the independent ratings provided to a common set of applications, clarify definitions and types of evidence as described in the rubric, and document decision rules.

Review of Rubric and Calibration Process

During the initial training session, the lead assessment evaluator walked the assessment review team through the entire rubric and allowed them to ask clarifying questions. After the initial training session, the assessment review team independently reviewed a common set of vendor applications, applying the rubric as they reviewed the evidence. Assessment review team staff (hereafter referred to as evaluators) were tasked with not only documenting their decisions based on the rubric, but also noting instances where they had questions or concerns, where they were not able to clearly apply the rubric criteria, or where evidence that did not align with the rubric criteria provided.

⁷ The State Board of Education rule that included characteristics of dyslexia was adopted after the CDE review of the assessments. The list of assessments and characteristics of dyslexia were adopted at the same time; therefore, the vendor worksheet was optional at the time of application.

Once all evaluators on the assessment review team finished their individual reviews, they reconvened for a collaborative session led by the lead assessment evaluator. During this meeting, they calibrated their findings by discussing and comparing their ratings and notes from the initial reviews. Together they clarified the definitions and rubric elements to ensure consistency across the evaluations. After the calibration meeting, each evaluator was assigned a set of applications, including a subset of applications that were assigned to all evaluators (for continued calibration and reconciliation across the evaluators). The lead assessment evaluator supported the evaluators throughout the review process, answering any questions or concerns about rubric criteria that arose throughout the review process. After all evaluators had completed their independent reviews, the assessment review team reconvened to compare and share ratings and notes from their reviews.

Ratings and Reconciliation Process

After the evaluators completed their review of all assessments, the lead evaluator reviewed the ratings, compared ratings for assessments that were common across evaluators, and documented any differences in ratings. The evaluators then convened for a final meeting to discuss and reconcile any disagreements. In instances where the evaluators had different ratings for the same evidence, the lead evaluator discussed the differences and attempted to bring the reviewers to a consensus rating. The assessment reviewers were able to come to an agreement for all ratings.

Summary Ratings

After raters had determined ratings for each criterion, the criterion-level ratings were aggregated into two summary ratings, one for compliance with the required criteria (i.e., SB19-199 requirements) and a second for the additional criteria. The evaluation team applied the following decision rules.

*Summary Rating: Compliance with the Required Criteria*⁸

- Fully meets: Received a rating of “fully meets” on all criteria.
- Largely meets: Received a rating of at least “partially meets” on all criteria.
- Partially meets: Received a rating of “does not meet” on at least one, but not all, criteria.
- Does not meet: Received a rating of “does not meet” on all criteria.

Summary Rating: Compliance with the Additional Criteria

- Fully meets: Received a rating of “fully meets” on all criteria.
- Largely meets: Received a rating of at least “partially meets” on all criteria.
- Partially meets: Received a rating of “does not meet” on at least one additional criteria.
- Does not meet: Received a rating of “does not meet” on all criteria.

A summary rating of “partially meets” is considered to meet the minimum threshold for compliance with the READ Act evaluation.

“Partially meets” ratings were often given when a vendor’s documentation was incomplete.

Results and Discussion

Results of the Assessment Evaluation

Exhibit 2.3 shows the overall summary ratings for compliance with the required criteria (additional detail of the ratings are shown in Appendix Table C.2.1).

⁸ The full rubric used for rating compliance is included in Appendix B. Reports for each assessment explain the rationale for each rating.

Exhibit 2.3. Overall Summary Ratings for Compliance with the Required Criteria					
Assessment Type	Fully Meets	Largely Meets	Partially Meets	Does Not Meet	Total
Interim	1	6	1	0	8
Diagnostic	0	13	3	1	17
Summative	0	4	2	0	6
Total	1	23	6	1	31

Exhibit 2.4 shows the overall summary ratings for compliance with the additional criteria (additional detail is shown in Appendix Table C.2.2).

Exhibit 2.4. Overall Summary Ratings for Compliance with the Additional Criteria					
Assessment Type	Fully Meets	Largely Meets	Partially Meets	Does Not Meet	Total
Interim	0	7	1	0	8
Diagnostic	0	10	6	1	17
Summative	0	4	2	0	6
Total	0	21	9	1	31

Discussion of Main Themes and Issues That Emerged from the Assessment Reviews

Overall, the assessments largely or partially meet the requirements outlined in the READ Act. Although only one of the 31 assessments fully meets all READ Act requirements, 23 largely meet the READ Act requirements. Only one assessment does not meet the requirements of the READ Act. Of the eight interim assessments submitted for review, seven assessments fully or largely meet READ Act requirements; similarly, of the 17 assessments submitted for review as a diagnostic, 13 largely meet the READ Act requirements. Finally, of the six assessments submitted for

review considerations as a summative assessment, four largely meet READ Act requirements.

A similar pattern emerged through the evaluation of the additional criteria. All but one assessment largely or partially meet the additional criteria. No assessment fully meets the additional criteria; however, 21 largely meet and nine partially meet the additional criteria. Of the eight interim assessments submitted for review, seven assessments fully or largely meet the additional criteria; similarly, of the 17 assessments submitted for review as a diagnostic, 10 assessments largely meet the additional criteria. Finally, of the six assessments submitted for review considerations as a summative assessment, four largely meet the additional criteria.

Notably, for each required and additional criterion, vendors provided evidence of varying levels of quality. Most assessments fully meet the criteria related to being evidence-based and aligning to learning standards, as well as those related to including appropriate guidance and resources for administration. The criteria for which most assessments partially meet the requirements are related to validity, reliability, processes to mitigate bias, accommodations, and usability of assessment data and reports.

According to the READ Act (1 CCR 301-92, 9.0), approved interim reading assessments should include embedded indicators that screen for characteristics of dyslexia. Among the interim assessments reviewed, all but one included a Dyslexia Indicator Worksheet, which was optional per Section D of the CDE's Solicitation for Assessments READ Act Advisory List: Review Period 2022–2026. In these worksheets and associated documentation, the interim assessment vendors claimed that their instruments were validated to screen for characteristics of dyslexia, such as phonemic awareness and rapid automatic naming (RAN). Two interim assessments provided one worksheet for both the English- and Spanish-language versions of the assessment. This suggests that the

Spanish-language version functions similarly to the English version when screening for characteristics of dyslexia; however, the vendors did not provide explicit evidence to support the adequacy of the Spanish versions in screening for characteristics of dyslexia.

For each evaluated criterion, we provide a summary of the lessons learned from our evaluation. These summaries are intended to provide general information about the assessments that were evaluated; however, the information may not apply to all assessments.

- *Evidence-based*
 - a. Most of the assessments received ratings of “fully meets” for this criterion. Some assessments exhibited gaps in information and evidence supporting their development based on a well-articulated theory of action. For some assessments, minimal evidence was provided to support the theoretical basis, with only brief descriptions of assessment use. Some assessments showed a lack of iterative development and evaluation processes, such as cognitive laboratories and pilot studies. Overall, there is a need for some assessment vendors to provide more robust evidence and detailed articulation of the theoretical frameworks guiding their assessments.
- *Standards-aligned*
 - a. The assessment evaluation identified gaps in standards alignment for some assessments, and especially for the Spanish-language versions of the assessments. Although alignment efforts were evident for the English-language version of the assessments, their Spanish-language counterparts lacked clear documentation and evidence of similar alignment to standards. Attempts to map tasks to grade levels, times of year, and component areas measured for some assessments were incomplete or unclear, often missing necessary references to the Colorado Academic Standards.

- *Validity*
 - a. Most assessments received ratings of “partially meets” for this criterion. Many showed insufficient evidence of using representative samples in pilot or field testing, particularly in the areas of including clear data on the inclusion of English learners (ELs) and students with disabilities. Although some assessments involved expert review in the development phases, there was often a lack of transparency about the composition of these expert panels and whether they included practicing teachers. Additionally, for some assessments there was minimal evidence of robust standard-setting methods or detailed explanations of how the assessments meet criteria for convergent, discriminant, and internal validity. Some assessments do not comprehensively cover all targeted areas of scientifically based reading instruction, such as reading fluency and vocabulary. In addition, the Spanish-language versions of assessments particularly lacked evidence supporting their validity and ability to diagnose strengths and weaknesses in reading skills.
- *Reliability*
 - a. Like the validity criterion, most assessments received “partially meets” ratings for this criterion. Several assessments reported a number of reliability estimates that fell below the threshold of 0.70. Further, there was often evidence that reliability estimates were derived from samples reported as being representative of gender, race, and socioeconomic status, yet it was unclear whether these samples truly encompassed all student groups, such as students with disabilities or ELs. Additionally, although some assessments confirmed the inclusion of diverse groups in their samples, detailed demographic breakdowns or sample sizes for specific student groups were frequently missing. In addition, standard errors of

measurement were frequently reported at the grade level and by domain but were often lacking for cut scores. Evidence of alternate forms and interrater reliability were often not applicable or not well documented due to the nature of computer-adaptive testing. Overall, more robust documentation and details of reliability estimates and sampling practices are needed to ensure that some assessments adequately support diverse student groups and performance levels.

- *Spanish-language assessment*
 - a. There is a need for vendors to submit more comprehensive and culturally representative evidence, including additional information that norming samples are representative of Spanish-speaking populations and that the content reflects the cultural and linguistic diversity of Spanish speakers, including dialectical variations. Across the Spanish-language assessments, there is also a need for a more thorough analysis to ensure that both linguistic and cultural aspects of the Spanish-speaking student population are appropriately considered and integrated into the assessment processes.
- *Paper-and-pencil assessment*
 - a. Of the 31 assessments that were evaluated, nine (three interim assessments and six diagnostic assessments) offered paper-and-pencil options for administration. Although paper-and-pencil versions are available, there is a persistent gap in demonstrating their comparability to the computer-based versions. Specifically, there was insufficient evidence showing whether the paper-and-pencil versions have been evaluated using a representative sample, particularly among students identified with SRDs.

- *Removal of bias*
 - a. In general, across the assessments, there is a need for more comprehensive and transparent processes related to the removal of bias. One recurring issue noted across assessments was that bias reviews were conducted without clear information on the panel composition and detailed results were often not provided for these bias reviews. Further, the application of Universal Design for Learning (UDL) principles was often mentioned but not explicitly demonstrated in assessment development. There was limited evidence of assessments being evaluated for content that truly represents diverse backgrounds, particularly in terms of dialectal diversity and cultural context. The lack of detailed results from bias analyses, combined with the absence of comprehensive studies such as differential item functioning and measurement invariance makes it challenging to ensure fairness across diverse student groups. Overall, most assessments would benefit from more explicit, documented evidence of practices aimed at ensuring culturally representative content and mitigating bias.
- *Assessment administration guidance*
 - a. Most of the assessments received ratings of “fully meets” in this criterion. For those that did not, there was insufficient evidence that training materials provide clear instructions for identifying students with SRDs.
- *Appropriate accommodations*
 - a. All the assessments received “partially meets” or “does not meet” ratings for this criterion, showing several significant gaps. Across all assessment types, there was a lack of evidence supporting the research basis for allowable accommodations. This missing evidence raises concerns about whether these accommodations

might compromise the intended interpretations for the intended purposes of the assessments. The training materials also lacked sufficient clarity regarding specific guidelines for selecting and implementing accommodations. Additionally, although some documentation mentions UDL principles, explicit evidence of the integration of UDL principles in the development process was not provided for most assessments—particularly for the Spanish-language versions. Accommodations for aspects like timing were noted, yet clarity on their applicability for students with disabilities was often lacking. Including more comprehensive and transparent practices regarding accommodations would ensure that all students can adequately demonstrate their knowledge without compromising assessment validity.

- *Score report usability*
 - a. Almost all assessments received a “partially meets” rating for this criterion. A primary concern was the lack of user-focused evaluations to assess the utility of the score reports for different audiences, such as parents, teachers, and other users at the school and district levels. There was limited to no evidence of focus groups being conducted to confirm the interpretation or use of score reports by various intended user groups. The processes for issuing reports in a timely manner was unclear for many assessments. Further, for many assessments, there was no evidence provided of the availability of reports in multiple languages (as appropriate for the test-taker population the assessment is intended to serve).



3

Instructional Programs on the Advisory List

The review of instructional programming considered three types of materials: core programs, supplemental programs, and intervention programs. Core programs are those used in general instruction. Supplemental programs are used in classrooms where support beyond the core program is needed to enhance reading instruction. Intervention programs are used to support individual students who are identified as needing additional intensive support for their reading development.

Key Findings

- **Nearly all the approved instructional materials (core, supplemental, and intervention programs in English and Spanish) either fully meet or largely meet SB19-199 requirements.**
- **Just over 16% of the programs fully meet the criterion of being “evidence-based” or “scientifically based.”**
- **Most core academic programs include academic content as required; 48% of programs fully meet this criterion.**
- **All six Spanish-language programs either fully or largely meet SB19-199 requirements; however, they do not provide guidance related to use in bilingual settings.**

Independent Evaluation Review Process

This section summarizes findings related to whether the materials that appear on CDE's Advisory List for Instructional Programming meet READ Act requirements. The review of instructional programming considered three types of materials: core programs, supplemental programs, and intervention programs. *Core programs* are those that are used in general instruction and must target all five areas of scientifically based reading: phonemic awareness, phonics, reading fluency, vocabulary development, and reading comprehension. *Supplemental programs* are used in classrooms where additional instructional materials beyond the core program are needed to supplement reading instruction. For example, although the core program includes systematic and sequential instruction focused on reading fluency, assessments indicate that students would benefit from additional instruction beyond that included in the core program. *Intervention programs* are used to support individual students who need are identified as needing additional intensive support for their reading development. Supplemental and intervention programs were subject to fewer review criteria, under the assumption that all students have access to a core instructional program.

The instructional materials review followed a rubric (see Appendix A.2) consisting of three main categories: (a) whether the program is evidence-based; (b) whether the program provides explicit and systematic skill development in the areas of phonemic awareness, phonics, vocabulary development, reading fluency, and reading comprehension, and is aligned with preschool through elementary and secondary state standards for reading adopted by the State Board; and (c) whether the program includes texts on core academic content to assist the student in maintaining or meeting grade-appropriate proficiency levels in academic subjects in addition to reading.

The following describes each of the three areas with their criteria:

- *Is evidence-based (22-7-1209 (2) (b) (I)).* Vendors were invited to submit up to three research studies or a logic model or theory of action as evidence. The independent evaluators then evaluated the evidence provided using the Every Student Succeeds Act (ESSA) evidence levels. As independent evaluators read through each study, they documented methodology, key findings, and effect sizes.

ESSA Evidence Levels

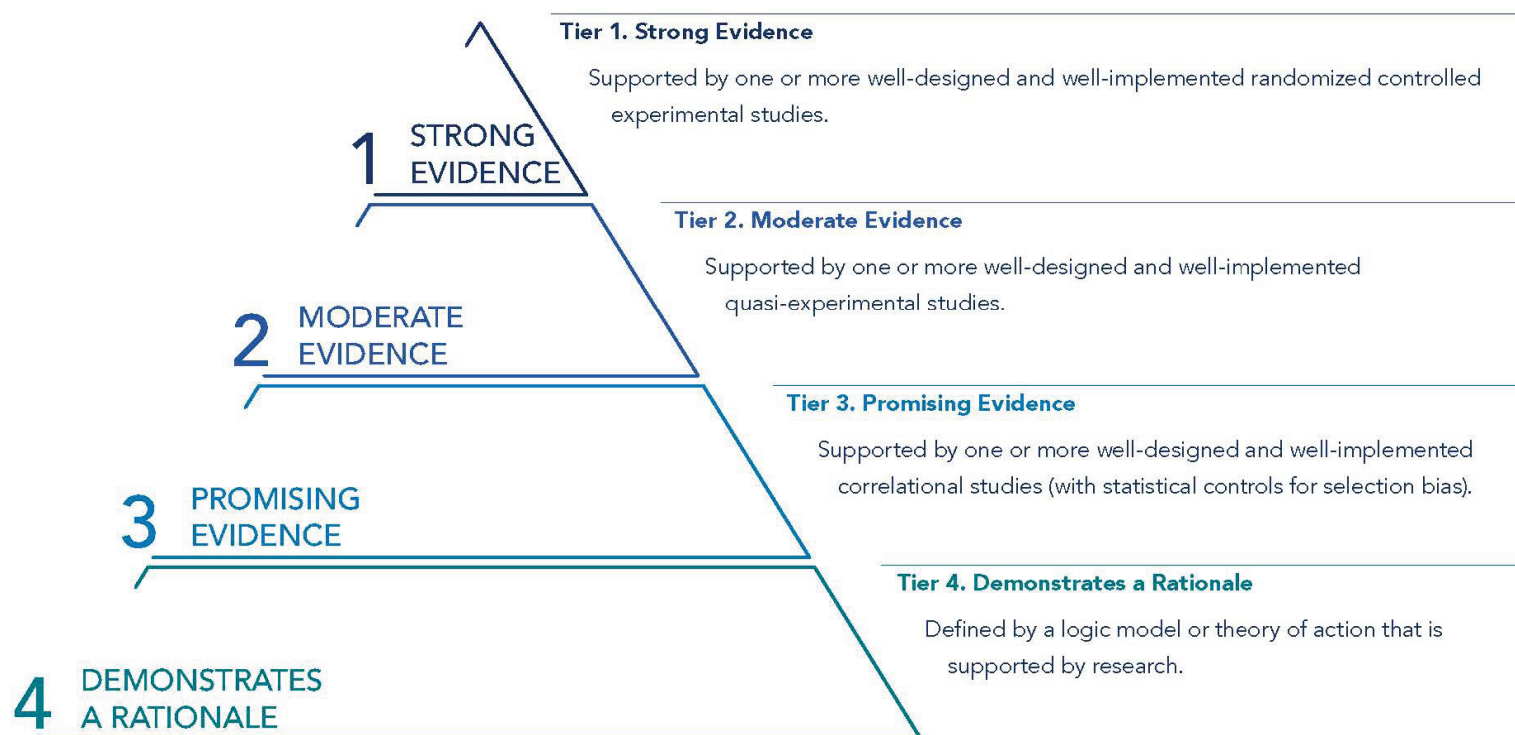
ESSA (2015) established a four-tiered method of evaluating evidence. This framework is designed to ensure that states, districts, and schools can identify programs that work. Stronger research methods provide stronger evidence for a program, resulting in higher tiers of ESSA evidence levels. When a program has a higher-tier rating, we can be more confident that it works. See Exhibit 3.1.

Independent evaluators also reviewed study designs for sample size, attrition, bias reduction, and baseline equivalence; the What Works Clearinghouse (WWC) reviews include these as important study design characteristics. The reviews presented here approximate but are not as in-depth as WWC reviews.

After reviewing the available evidence, researchers assigned each program an evidence rating ranging from 1 to 4. Evidence ratings were guided by the ESSA Tiers of Evidence (see Exhibit 3.1). An ESSA level 1 or 2 earned a rating of “fully meets,” while an ESSA level of 3 or 4 earned a rating of “partially meets.” Programs that could not demonstrate an ESSA evidence level of at least 4 would “fail to meet.” For the summary rating, a program could earn “fully meets” if it partially meets on this indicator.

Exhibit 3.1. ESSA Four Tiers of Evidence

The Every Student Succeeds Act (ESSA; 2015) establishes a four-tiered method of evaluating evidence. This framework is designed to ensure that states, districts, and schools can identify programs that work. Stronger research methods provide stronger evidence for a program, resulting in higher tiers of ESSA evidence levels. When a program has a higher tier rating, we can be more confident that it works.



- *Provides explicit and systematic skill development in the areas of phonemic awareness; phonics; vocabulary development; reading fluency, including oral skills; and comprehension (22-7-1209 (2) (b) (II)), and is aligned with the preschool through elementary and secondary state standards for reading adopted by the State Board (22-7-1209 (2) (b) (II.5)).* We evaluated whether skill development across reading areas was present, explicit, and systematic, using vendor-supplied information and EdReports, when available. Core programs were evaluated for all five areas of reading (phonemic awareness, phonics, reading fluency, vocabulary development, and reading comprehension), whereas supplemental and intervention programs were evaluated only for the areas that vendors claimed to specifically target.

A core program received a rating of “fully meets” if it was clear that all elements were explicitly and systematically taught. A program received a rating of “partially meets” if all elements were present but it was not clear that they were presented both explicitly and systematically. Finally, a core program received a rating of “does not meet” if at least one of the elements of reading failed to meet the criteria outlined in the rubric.⁹ These decision rules were the same for supplemental and intervention programs, except that these programs were not required to address all the elements of reading. These decision rules applied only to those elements that the vendor claimed to address.

- *Includes texts on core academic content to assist the student in maintaining or meeting grade-appropriate proficiency levels in academic subjects in addition to reading (22-7-1209 (2) (b) (V)).* The evaluators operationalized this requirement into two criteria: (a) grade-appropriate text complexity and (b) a range of content

⁹ For reading comprehension, both elements would have to fail to meet.

areas (e.g., history, science) and genres (e.g., fiction, nonfiction).

The evaluators used the documentation that vendors had supplied to CDE in their initial applications.

To fully meet the criteria of this indicator, both (a) and (b) had to fully meet. If both (a) and (b) partially met or if one fully met and the other did not meet, then this indicator would be rated as “partially meets.” If neither (a) nor (b) fully met and at least (a) or (b) did not meet, then this indicator would receive a rating of “does not meet.”

Finally, the following decision rules were used to create an overall instructional program rating:

- Fully meets: Received a rating of at least “partially meets” on the evidence-based indicator and received a rating of “fully meets” on all other indicators.
- Largely meets: Received a rating of at least “partially meets” on all indicators.
- Partially meets: Received a rating of “does not meet” on at least one but not all indicators.
- Does not meet: Received a rating of “does not meet” on all indicators.

Data Collection and Methods

Information Used to Review Instructional Programs

The evaluation team read three sets of materials in reviewing instructional programs: (a) the vendor application to CDE, (b) any associated documents the vendor submitted with the application, and (c) vendor-supplied access to online program platforms. For a program’s evidence base, the evaluation team reviewed up to three empirical research studies submitted by vendors in their application to CDE. When vendors did not submit empirical research, the evaluation team reviewed either a logic model or a theoretical rationale.

Training for Independent Evaluators

In partnership with our external expert advisory panel, the evaluation team created a rubric that specified criteria and rating options for each required program element. CDE approved the rubric (see Appendix A.2). WestEd staff with expertise in research, curriculum, and/or Spanish-language reading instruction were the evaluators.

Ratings and Resolution of Discrepancies in Ratings

Teams of evaluators reviewed 62 CDE-approved programs. Reviewers with expertise in research were responsible for rating the evidence base and reviewers with broader expertise in curriculum were responsible for rating the remainder of the criteria. Two sets of reviewers reviewed 10% of the programs; the reviewers came together to address any discrepancies and submit final ratings. The first reviewer evaluated the vendor-provided evidence and documented their findings based on the provided rubric. The second reviewer independently assessed the same material without prior knowledge of the first reviewer's conclusions. The results from both reviewers were compared to measure consistency (inter-rater reliability). The reviewers discussed and resolved any discrepancies to improve accuracy and objectivity.

Processes Used for Reviews

The evaluation was conducted on 62 CDE-approved core, supplemental, and intervention programs, targeting vendor-supplied curriculum, materials, and documentation.

Results and Discussion

Nearly all the approved instructional materials (core, supplemental, and intervention programs in English and Spanish) were found to either fully meet or largely meet SB19-199 requirements. Exhibit 3.2 contains results for the four review categories.

Exhibit 3.2. Summary Rating			
Program Type	Summary Rating: Compliance with SB19-199 Requirements		
	Fully Meets	Largely Meets	Partially Meets
Core programs in English	3	1	1
Supplemental programs in English	10	17	1
Intervention programs in English	15	7	1
Programs in Spanish (all)	2	4	0
OVERALL	30	29	3

Is Evidence-Based or Scientifically Based (22-7-1209 (2) (a) (II) (A))

Although all instructional programs meet the minimum standard for being evidence-based or scientifically based, just over 16% of the programs fully meet this criterion (see Exhibit 3.3). The programs that fully meet the evidence-based standard demonstrated impact on students' reading outcomes using rigorous research designs that qualified as ESSA Tier 1 or Tier 2 evidence.

Is Evidence-Based or Scientifically Based (22-7-1209 (2) (b) (I))

All 62 programs meet at least the minimum standard for being evidence-based or scientifically based. Ten programs fully meet this criterion through their rigorous research designs that demonstrate impact on students' reading outcomes (Exhibit 3.3).

Exhibit 3.3. Summary Rating for Evidence-Based or Scientifically Based		
Program Type	Rating	
	Fully Meets	Partially Meets
Core programs in English	3	2
Supplemental programs in English	3	25
Intervention programs in English	4	19
Programs in Spanish (all)	0	6
OVERALL	10	52

Provides Explicit and Systematic Skill Development in the Elements of Scientifically Based Reading Instruction (22-7-1209 (2) (b) (II)) and Is Aligned with the Preschool Through Elementary and Secondary State Standards for Reading Adopted by the State Board (22-7-1209 (2) (b) (II.5)).

Overall, the materials on CDE’s Advisory Lists for core, supplemental, and intervention programs offer explicit and systematic instruction in the elements of scientifically based reading instruction (Exhibit 3.4). Notably, three English-language core programs and one Spanish-language core program fully meet the independent evaluation’s criteria on all elements of scientifically based reading instruction. This is important because core programs are used to provide instruction to all students, including those who struggle with reading. Vendors who received a rating of “partially meets” did not provide clear, sufficient evidence to demonstrate that the program had systematic instruction throughout the school year.

Exhibit 3.4. Skill Development			
Program Type	Rating		
	Fully Meets	Partially Meets	Does Not Meet
Core programs in English	3	2	0
Supplemental programs in English	18	9	1
Intervention programs in English	17	6	0
Programs in Spanish (all)	2	4	0
OVERALL	40	21	1

Across all of 62 instructional programs, the independent evaluation found only one program that had one or more elements of scientifically based reading instruction that did not meet the independent evaluation’s criteria for explicit and systematic instruction (Exhibit 3.5). *Working with English Language Learners* (WELLS) 3rd Edition, an English-language program approved for supplemental use, did not meet the criteria for explicit and systemic instruction of phonics and fluency.

The evidence WELLS provided for phonics instruction did not provide a clear, research-based phonics scope and sequence, and outlined phonics integrated within broader language activities rather than through explicit, structured lessons. For fluency, the evidence provided by WELLS does not emphasize explicit strategies for improving reading speed and expression; thus, it obtained “does not meet” ratings in both categories for phonics and fluency.

Exhibit 3.5. Programs That Fully, Partially, or Do Not Meet for Each Component of Scientifically Based Reading Instruction, by Program Type					
Component of Scientifically Based Reading	Program Type	Fully Meets	Partially Meets	Does Not Meet	Not Applicable*
Phonemic Awareness	Core – English	4	1	0	0
	Supplemental – English	14	1	0	13
	Intervention – English	11	0	0	12
	All – Spanish	4	2	0	0
Phonics	Core – English	4	1	0	0
	Supplemental – English	23	1	1	3
	Intervention – English	18	1	0	4
	All – Spanish	4	1	0	1
Fluency	Core – English	3	2	0	0
	Supplemental – English	5	5	1	17
	Intervention – English	7	5	0	11
	All – Spanish	2	3	0	1
Vocabulary	Core – English	3	2	0	0
	Supplemental – English	2	4	0	22
	Intervention – English	2	1	0	20
	All – Spanish	2	3	0	1
Reading Comprehension	Core – English	3	2	0	0
	Supplemental – English	1	3	0	24
	Intervention – English	1	0	0	22
	All – Spanish	2	3	0	1

* Some programs received a summary rating of “Not applicable” because the vendor application indicated the program did not target that particular component of reading.

Includes Texts on Core Academic Content to Assist the Student in Maintaining or Meeting Grade-Appropriate Proficiency Levels in Academic Subjects in Addition to Reading (22-7-1209 (2) (b) (V)).

Most core academic programs include academic content as required (Exhibit 3.6); 88% of programs fully meet this criterion. One core English-language program was rated as “does not meet” due to lack of documentation to demonstrate a range of subject areas and text types of academic content.

Exhibit 3.6. Core Academic Content			
Program Type	Rating		
	Fully Meets	Partially Meets	Does Not Meet
Core programs in English	4	0	1
Core programs in Spanish	4	0	0
OVERALL	8	0	1

Text complexity and quality. This element was only evaluated for core programs, as our aim was to assess grade appropriateness of texts and the evaluation team felt that an appropriate supplemental or intervention text would likely not be at grade level. Sixty-six percent of English core instructional programs and 50% of Spanish core instructional programs, respectively, included texts written at grade level, as evidenced by vendor-supplied quantitative and qualitative measures of text complexity (Exhibit 3.7). One core English-language and two core Spanish-language programs were rated as “partially meets” due to a lack of qualitative evidence of grade-appropriate text complexity.¹⁰ Having opportunities to read appropriately complex text is a pre-requisite for maintaining or meeting grade-level proficiency standards. If students only access texts at easier proficiency levels, it would be impossible for them to meet grade-level standards.

¹⁰ The programs were Amplify CKLA, 2nd Edition (Amplify Education, Inc.), Amplify Caminos, 2022 (Amplify Education, Inc.), and HMH ¡Arriba La Lectura!, 2020 (Houghton Mifflin Harcourt).

Exhibit 3.7. Programs That Fully, Partially, or Do Not Meet for Text Complexity and Quality, by Program Type			
Program Type	Rating		
	Fully Meets	Partially Meets	Does Not Meet
Core – English	4	1	0
Core – Spanish	2	2	0
OVERALL	6	3	0

Discussion of Spanish-Language Instructional Programs Reviewed

WestEd completed an external evaluation of the six Spanish-language instructional programs that partially or fully meet the CDE's rubric criteria.¹¹ Two WestEd evaluators participated; they double-coded a Spanish-language program to calibrate. One person evaluated the other programs. The evaluators held a calibration meeting to review any questions and ensure continued agreement on ratings. All six approved Spanish-language instructional materials (core, supplemental, and intervention) were rated as either “fully meets” or “largely meets” SB19-199 requirements.

The external evaluation of Spanish-language instructional programs highlighted program strengths as well as opportunities for enhancement across the six programs reviewed. Programs demonstrated varying degrees of alignment with foundational literacy components, such as phonemic awareness, phonics, vocabulary, and reading comprehension. Several programs showcased strong cross-linguistic connections and evidence-based practices, while others presented opportunities to further strengthen vocabulary instruction and oral language development for Spanish-English bilinguals. This variation underscores the potential

¹¹ Both WestEd reviewers for the Spanish-language programs are completely bilingual, bicultural, and biliterate in English and Spanish. They received training from the broader WestEd team on implementing the rubrics with instructional programs.

for refining instructional materials to better support Colorado-based Spanish-English bilingual students.¹²

CDE Processes for Identifying Items for Approved and Advisory Lists

The process CDE used to review instructional programs reflected the READ Act's components and intent. At the beginning of the process, CDE hired Dr. Stephanie Stollar, an educational consultant in the early literacy sphere and former Vice President for Professional Learning at Acadience Learning Inc., to assist in developing an evidence-based rubric for evaluating instructional programs. Dr. Stollar has expertise in the mechanics of early literacy development and is deeply familiar with research on early reading intervention and success.

Nearly 100 empirical studies, reports, and scientific articles were referenced in the rubric design process. The resulting rubric comprises elements that research has shown are central to learning to read. Additionally, best practices were derived from rubrics used by other states in successful material vetting processes (see Chapter 5 for an explanation of the processes that CDE used to create its Spanish-language review rubric).

The review process consisted of two phases. The first phase used a rubric primarily focused on how programs align with the science of reading; whether instruction is explicit, sequential, systematic, and cumulative; and whether the programs were supported by research. Programs that were reviewed favorably in the first phase of the rubric were invited to submit additional information for the second phase, which included individual academic components by grade level and close examination of all five components of scientifically based reading instruction (phonemic awareness, phonics, reading fluency, vocabulary, and reading

¹² The findings related to uneven guidance for use of Spanish-language materials in bilingual settings parallel findings from the evaluation's Year 1 review. There, we found few English-language core instructional programs with adequate guidance for supporting ELs.

comprehension). The rubrics were made publicly available, and CDE hosted an accompanying webinar and solicited public feedback. The contractor adjusted the rubrics based on the feedback received, and the CDE team provided final approval.

Reviewers were selected via a competitive application process. Selected reviewers were teachers representing districts that ranged in size and urbanicity. Many of the reviewers were instructional coaches, university professors, or special education teachers. Some were bilingual in English and Spanish. Reviewers were not paid, and they were required to sign a confidentiality and conflict-of-interest statement.

CDE staff cast a wide net to reach instructional programming vendors. They contacted all vendors on the existing Approved list, posted the information on their website, discussed the process during several monthly READ Act webinars, made listserv announcements, and encouraged districts to reach out to any vendors they wanted to be considered. CDE also provided a technical assistance webinar for vendors to explain the process and answer questions.

Of 98 instructional programs reviewed by CDE, 67 passed Phase I of the rubric. Of those, 58 passed Phase 2 of the rubric and were therefore approved. Exhibit 3.9 contains a breakdown of the programs that passed each phase of the rubric.

Exhibit 3.9. Number of Programs That Passed Each Phase of the Rubric, by Program Type						
Program Type	Phase 1			Phase 2		
	Pass	No Pass	Pass Rate	Pass	No Pass	Pass Rate
Core	4	4	50%	3	1	75%
Supplemental	30	14	68%	27	3	90%
Intervention	27	10	73%	22	5	81%
Spanish	6	3	67%	6	0	100%

The instructional program review process reflects the goals of READ Act legislation. Both phases reflect an emphasis on the components of

scientifically based reading instruction. The Phase 2 rubric criteria for each grade level are rooted in evidence, are clearly specified, and reflect both historical and current understandings of how the science of reading can be applied effectively in classroom practice. The review was executed in a thoughtful, systematic way that produced consistent ratings and allowed program vendors to appeal and clarify program content and approaches as needed.

The independent evaluation and CDE's evaluation had very similar approaches. Both processes developed criteria using foundational requirements from policy and included steps in the process for review and feedback from individuals outside of the vendor's organization. The independent evaluation solicited feedback on its rubric from an expert advisory panel of independent literacy experts and scholars.

Some operational criteria differed between CDE's and the independent evaluation. Specifically, the independent evaluation set a very high bar for research evidence to fully meet, as a means of supporting stakeholders in understanding the range of evidence across programs.



4

Professional Development Programs on the Advisory List

The PD program review focused on four categories: whether the PD program

- (1) is evidence-based;
- (2) provides for explicit and systematic skill development in the five reading elements;
- (3) includes rigorous evaluations of teacher learning throughout and at the end of the course; and
- (4) has support for ELs.

Key Findings

- Overall, all PD programs fully meet SB19-199 requirements.
- One PD program fully meets the evidence base requirement.
- All seven PD programs reviewed meet requirements around explicit and systematic skill development, including evaluations of learning.

Description of Review Categories

The PD program review focused on four categories: whether the PD program (1) is evidence-based; (2) provides for explicit and systematic skill development in the five reading elements; (3) includes rigorous evaluations of teacher learning throughout and at the end of the course; and (4) has support for ELs. The evaluation team created a rubric for each of these four areas (see Appendix A.3 for the initial PD evaluation rubric). Although the fourth area is not a requirement of SB19-199, it is of interest to CDE and vendor applications requested information about support for ELs. Themes from vendor-provided information about support for ELs are summarized in this chapter.

The following are the four areas with their criteria.

- *Is evidence-based (22-7-1209 (2) (c)).* The team evaluated whether program evidence reflected one of the four ESSA evidence levels through (a) formal research studies that demonstrate impact on teacher practice and student outcomes or (b) a logic model, theory of action, or synthesis of research that outlines how and why the program expects to have impact or (c) alignment to PD program elements associated with impacts on teacher practice and student outcomes based on four research-based PD criteria (content focus, models of effective practice, feedback and direction, and ongoing support with sufficient duration of at least 45 hours).

If vendors had formal research studies, the evaluators reviewed the reports to determine whether the evidence was in alignment with ESSA evidence levels 1 or 2 (“fully meets” the evaluation criteria for being evidence-based) or level 3 (“partially meets”). Vendors were considered to “partially meet” the evidence-based criteria through promising evidence from research studies (ESSA evidence level 3) or through research-based logic models, theories

of action, syntheses of research, or alignment to research-based criteria as described in 1c. Vendors that had none of the evidence described to provide a research-based rationale for their program received a rating of “does not meet” for the evidence-based area.

- *Provides explicit and systematic skill development in the areas of phonemic awareness; phonics; vocabulary development; reading fluency, including oral skills; and comprehension (22-7-1209 (2) (c) (I)).* The team evaluated whether skill development across reading areas was present, explicit, and systematic using vendor-supplied information. Reading areas were phonemic awareness, phonics, vocabulary development, reading fluency, and reading comprehension. For reading comprehension, evaluators focused on “close reading,” which was emphasized by READ Act minimum skill competencies and required in the evaluation of PD programs.
- *Includes rigorous evaluations of learning throughout and at the end of the course that a person taking the course must pass to successfully complete the course (22-7-1209 (2) (c) (II)).* The team evaluated the presence of (a) evaluation of teacher knowledge of program content, (b) evaluations both during and at the end of the course, (c) indication of rigor, and (d) specific criteria and indicators of successful course completion (e.g., a certificate). The team conceptualized rigor in alignment with the upper levels of Bloom’s Taxonomy and with Webb’s Depth of Knowledge levels 3 and 4 (because of the use of Webb’s model in creating and assessing rigor in state standards such as the Common Core State Standards, 2010).
- *Has support for ELs.* Support for ELs was not a requirement for vendors and therefore was not included in the final rating of the PD programs. As in previous years, the evaluation team added an indicator of evidence of differentiation and support

for ELs to the review criterion because this continues to be an area of interest to CDE.

Rationale for Inclusion of Additional Professional Standards

CDE's vendor solicitation referenced research-based PD elements described by Joyce & Showers (2002). These elements included the presentation or theory and strategy as a rationale for active engagement, demonstration of new learning, practice with feedback, and ongoing support. The evaluation team aligned these research review elements with two other more recent reviews by Desimone (2009) and Darling-Hammond and colleagues (2017) to determine a final set of evidence-based criteria. Specifically, Desimone's five research-based areas were a focus on content with modeling, active learning with feedback, coherence, duration, and collective participation. Darling-Hammond and colleagues' research review yielded six areas of PD: is content-focused and incorporates active learning, uses models of effective practice, offers feedback and direction, provides coaching and expert support, is of sustained duration, and supports collaboration. Both the Desimone (2009) and Darling-Hammond and colleagues' (2017) PD research syntheses based their findings on research that shows impact on changes in teacher practice and positive impact on student learning outcomes. Drawing from this research, the evaluation team constructed a set of criteria to determine alignment with elements of impactful PD: (a) has a content focus (focuses on five components of reading and incorporates active professional learning); (b) uses models of effective practice (e.g., demonstration); (c) offers feedback and direction; and (d) features ongoing support (e.g., coaching) with sufficient duration (minimum of 45 hours).

Data Collection and Methods

Description of Information Used to Review Professional Development Programs

The evaluation team reviewed two documents or sets of documents to assess the programs: (a) the vendor application to CDE and (b) any associated documents the vendor submitted with the application. The sequence of review started with the response to the evaluation's vendor request form because it included key categories for the evaluation. The evaluators then used the vendor application to CDE for supplemental information and to corroborate findings.

Training for Independent Evaluators and Rating Process

Evaluation team members engaged in trainings, reviewed the rubric, and calibrated on how a sample vendor's materials should be rated across rubric dimensions. The process of rating a sample provider's materials led the team to clarify several definitions within the rubric. When rating the remaining vendors' programs, members of the evaluation team took responsibility for particular sections of the rubric across vendors to ensure consistency in rating within sections and develop understandings of trends across vendors (e.g., the level of rigor within program/course evaluations). Team members met to discuss evidence for ratings as well as emergent themes.

The evaluators rated the three program criteria, as well as evidence to support ELs, which was not included in the final rating. If programs had ratings of "partially meets" or "fully meets" for being evidence-based, "fully meets" for skill development, and "fully meets" for assessment, the summary rating was calculated as "fully meets." The evidence-based criteria allowed for "partially meets" because this meant the programs meet at least ESSA evidence level 3 or 4.

Results and Discussion

Overall, all PD programs fully meet SB19-199 requirements (Exhibit 4.1). One fully meets the evidence-based requirement and all seven meet requirements around explicit and systematic skill development, including evaluations of learning. Findings are explained in detail below.

Is evidence-based (22-7-1209 (2) (c)). Six PD programs partially met the standard of evidence-based by meeting ESSA evidence levels 3 or 4. One PD program was rated as “fully” meeting the standard of evidence-based because it met ESSA evidence level 2. This program provided findings from a quasi-experimental study involving 1,914 students in kindergarten through 4th grade. The vendor’s study found statistically significant gains in reading scores for students in a number of areas—including phonological awareness, phonemic segmentation fluency, and beginning and advanced decoding—and medium-to-large effect sizes on all measures for all student populations in the study. The vendor did not include information about its study’s funding source.

The six PD programs that partially meet the standard provided a logic model, theory of action, or synthesis of research outlining how and why the program expects to have impact and/or demonstrate alignment to PD program elements associated with impacts on teacher practice and student outcomes based on the four research-based PD criteria described in Review Category 1.

Exhibit 4.1. Professional Development Program Summary SB19-199 Requirement				
Program Summary	Rating			
	Fully meets	Partially meets	Does not meet	
Is evidence-based (22-7-1209 (2) (c)).	1	6	0	
Provides explicit and systematic skill development in the areas of phonemic awareness; phonics; vocabulary development; reading fluency, including oral skills; and comprehension (22-7-1209 (2) (c) (I)).	7	0	0	
Includes rigorous evaluations of learning throughout and at the end of the course that a person taking the course must pass to successfully complete the course (22-7-1209 (2) (c) (II)).	7	0	0	
Summary Rating	Fully meets	Largely meets	Partially meets	Does not meet
Compliance with SB 19-199 requirements	7	0	0	0

Provides explicit and systematic skill development in the areas of phonemic awareness; phonics; vocabulary development; reading fluency, including oral skills; and comprehension (22-7-1209 (2) (c) (I)).

All seven programs fully meet the criteria for explicit and systematic skill development in all five reading areas. The programs all emphasize building teachers' knowledge of the five core components of scientifically based reading instruction. For comprehension, the programs focus on close reading within a single text.

Includes rigorous evaluations of learning throughout and at the end of the course that a person taking the course must pass to successfully complete the course (22-7-1209 (2) (c) (II)). All seven programs fully meet the criteria for inclusion of rigorous evaluations of learning throughout and at the end of the course with criteria and an indication of completion.

All programs include evidence of rigor within the evaluation items. Although some items involve straightforward recall of information and recognition of concepts, sample evaluation questions and exercises from every vendor reflect Webb's Depth of Knowledge levels 3 and/or 4. Examples include assessments that ask teachers to reflect on which elements of reading and which potential interventions apply in sample classroom scenarios; reflect on case studies; interpret screening assessment data; create a lesson plan reflecting course learnings; demonstrate instructional practices learned in the program within their classrooms; and self-evaluate on a lesson using a program-provided framework and engage in other self-critiques.

Differentiation for English Learners in Professional Development

Learners whose native languages are not English and who are learning to read in English have needs that are both similar and different to native English speakers learning to read. Researchers have pointed out that ELs need a focus on Tier 1, or general English vocabulary, due to limited exposure in early childhood (August et al., 2005) and Spanish speakers would benefit from the integration of Spanish knowledge in learning and understanding English words and sentences (Pearson et al., 2007). Recently, Goldenberg (2020) reviewed the science of reading knowledge base and research on effective instruction in reading for ELs and proposed an emerging science of reading specifically for them that prioritizes English literacy and oral language proficiency.

All vendors provided some information related to differentiation and support for ELs. The pedagogical approaches (e.g., explicit and systematic phonics instruction, oral language strategies, practices for vocabulary development) to teaching foundational literacy skills were identified by all vendors as being similar for both native and non-native speakers. The evaluation team synthesized other themes related to differentiation and support for ELs from the evidence vendors supplied in their applications:

- Several focused on understanding phonology of native languages to bridge the gap to English sounds. Strategies named for doing so included identifying similarities in phoneme production and phoneme/grapheme matching and extending additional explicit and systematic practice to sounds in English that do not exist in the learner's native language or interfere with English phoneme production. However, there was limited evidence of strategies embedded within professional learning for doing so. Two vendors mentioned aligning to or working with teachers of EL learners. One vendor mentioned Sheltered Instruction Observation Protocol strategies and several referenced resources from WIDA at the Wisconsin Center for Education Research, University of Wisconsin-Madison.
- There was limited evidence of the actual content of professional learning for staff on second language acquisition or reading for non-native speakers. For all seven vendors, the information was listed as contained within one or two modules or chapters depending on the mode of delivery. When content was further explained, the focus was on the differences in how EL students learn to read when compared with students whose first language is English with little to no evidence that teachers would receive professional learning that explicitly addresses the needs of these students.
- Four of the seven vendors included a focus on culturally responsive instruction and instructional materials. This asset-based approach recognizes the need to build on the valuable knowledge that students bring to the table. Vendor applications included information about leveraging both content knowledge and the lived experiences of students and their families as a way to connect to curricular materials and vocabulary within literacy instruction.

- The evidence of EL differentiation for all seven vendors was presented together with meeting the needs of multiple groups of diverse learners, such as students with disabilities. It was not clear if the professional learning described addressed these groups as a monolith or it was presented together simply due to this area not being the focus of the request for applications.



5

Review of Spanish- Language Rubric

Strong literacy instruction builds on the bilingual strengths of our students, ensuring every learner has the tools to thrive in a multilingual world. This review highlights the successes and opportunities within the Spanish Instructional Program Rubric, providing recommendations to further support literacy development for Spanish-English bilingual learners.

Key Findings

- **Evidence-based practices:** Revised rubrics align with Spanish language-specific literacy components, such as phonological awareness, oral language development, and metalinguistic skills, and Colorado READ Act standards.
- **Gaps in bilingual literacy:** Rubrics do not currently provide guidance for instructional materials used in different instructional programs (dual-language and transitional bilingual programs).

Introduction

The Spanish Instructional Program Rubric was developed to evaluate and support high-quality literacy instruction for Spanish-English bilingual students under the Colorado READ Act. This section provides an in-depth review of the revised rubric, focusing on its alignment with evidence-based, inclusive practices and its relevance to the multilingual realities of Colorado classrooms.

Spanish-English bilingual students in the United States often face unique challenges in literacy development due to the challenging bilingual environments in which they are educated (Mora & Lopez, 2023). To address these challenges, the Spanish language-specific rubric emphasizes key literacy components such as phonological awareness, phonics, vocabulary development, oral language, metalinguistic awareness, and reading comprehension. These foundational skills, paired with high-quality instructional materials, are crucial for fostering academic success and positive student outcomes.

Additionally, this review highlights the critical importance of bilingual alignment in instructional programming. Effective literacy instruction must not merely translate English-language materials into Spanish but in addition integrate instructional practices that reflect the bilingual and biliterate realities of students, schools, and programs. By examining the development process, stakeholder feedback, and the final rubric structure, this analysis identifies key strengths and areas for improvement to enhance literacy outcomes for Spanish-speaking learners.

Development of the Revised Rubrics

The development of the revised Spanish Instructional Program Rubric followed a structured, iterative process to ensure alignment with evidence-based practices and responsiveness to the needs of Spanish-

speaking students in Colorado. Below is a detailed summary of the timeline and key deliverables:

Timeline and Process

- **Planning Phase (July–August 2023):** The groundwork for the revised rubrics was established through extensive collaboration and research. Key activities included outreach to literacy networks, collaboration with the New Mexico Public Education Department to review its Spanish-language rubrics, and research on foundational Spanish-language literacy skills aligned with Spanish Language Arts standards (e.g., the California Common Core State Standards in Spanish [CA CCSS en Español]).
- **Draft Development Phase (August–September 2023):** Initial drafts of the rubrics for core, supplemental, and intervention programs were created, incorporating input from WestEd and bilingual literacy experts. Special attention was given to phonological awareness, metalinguistic components, and cross-linguistic connections to support biliteracy development. Emphasis was placed on aligning phonological and phonics skill in Spanish, with a particular focus on syllabication and oral language development, ensuring appropriate pacing for Spanish language learners.
- **Stakeholder Feedback Phase (October 2023):** Stakeholders emphasized the need for instructional materials to be evidence-based and relevant to U.S.-based Spanish-speaking students. Cultural and linguistic relevance was a critical focus.
- **Validation Phase (October–November 2023):** Feedback was integrated to ensure cultural and linguistic appropriateness; validation by Dr. Linda Cavazos confirmed alignment with evidence-based best practices.

Incorporation of Recommendations from the First Rubric

The revised Spanish Instructional Program Rubric reflects critical updates based on WestEd’s recommendations in its Year 3 Instructional Materials report, enhancing its relevance and effectiveness in evaluating programs for Spanish-speaking students. WestEd’s recommendation was to convene a panel of experts in early literacy learning and instruction in Spanish for Spanish speakers to review CDE’s rubrics and then revise these rubrics to reflect best practices of evidence-based early literacy instruction in Spanish.

One significant update to the revised rubric, based on the incorporation of expert feedback, is the stronger emphasis on Spanish language–specific literacy components, particularly syllable-level phonological awareness. This approach aligns with linguistic research, that highlights the importance of syllable-based instruction in transparent orthographies like Spanish. Studies by Bravo-Valdivieso (2004) and Anthony and Francis (2005) emphasize that teaching syllable awareness lays the groundwork for decoding skills in Spanish, providing a foundation for literacy development before progressing to phoneme-level tasks. Furthermore, cross-linguistic transfer research, including the work of López and Greenfield (2004), supports the integration of these skills to facilitate bilingual literacy development in English and Spanish.

Listening comprehension was another key addition, incorporated as an essential component across all grades (kindergarten to 3rd grade). This ensures the rubric supports oral language development tailored to Spanish-English bilingual students, in addition to addressing a critical area for literacy in ELs.

Materials were revised to align with evidence specific to the Colorado context, ensuring relevance to transitional and dual-language instruction. For example, the rubric acknowledges that simply translating English-language programs into Spanish is insufficient. Effective instructional

programs must align literacy instruction across languages while reflecting the goals of the transitional and dual-language models commonly used in Colorado. By drawing on biliteracy frameworks, such as New Mexico Public Education Department's Biliteracy Guidance (NMPED & Cavazos, 2022), the rubric emphasizes structured literacy principles, integrating oral language, phonological awareness, vocabulary development, fluency, and reading comprehension.

Additionally, the rubric expands on the existing framework of the Colorado READ Act's Minimum Reading Competency Skills Matrix to better address the needs of Spanish-English bilingual learners. It underscores the importance of alignment between Spanish and English literacy instruction to ensure that programs are linguistically appropriate, as highlighted by Durgunoğlu et al. (1993) and Carlo et al. (2005).

Structure and Components of the Revised Rubrics

The revised rubrics address literacy needs through tailored frameworks for core, supplemental, and intervention programs. Core rubrics prioritize a comprehensive focus on the Colorado READ Act's Minimum Reading Competency Skills as adapted to the Spanish language, with a particular focus on cross-linguistic connections between Spanish and English. Supplemental rubrics focus on instructional materials that enhance specific literacy skills. Intervention rubrics focus on instructional materials that provide differentiated strategies and informal assessments for struggling Spanish-speaking readers.

Key Deliverables

The finalized rubrics for core, supplemental, and intervention reading programs were completed by November 8, 2023. These rubrics provide a comprehensive tool for evaluating Spanish-language instructional materials, ensuring alignment with state standards and the unique needs of Colorado's Spanish-speaking learners.

Key Findings and Analysis

The revised Spanish Language Instructional Program Rubrics demonstrate critical strengths and highlight areas for further improvement in supporting Spanish-speaking students in Colorado.

Strengths of the Revised Rubrics

The rubrics align with Spanish-specific literacy practices, emphasizing key components such as syllabic awareness in the developmental progression of phonological awareness. In addition, stakeholder feedback was integrated to ensure cultural and linguistic relevance, with materials tailored to the sociocultural realities of U.S.-based Spanish-speaking students. Additionally, the rubrics align with the Colorado Academic Standards for Reading, Writing, and Communicating (CDE, 2020), ensuring alignment with state standards.

Conclusion

The revised Spanish Instructional Program Rubrics represent significant progress in addressing the literacy needs of Spanish-speaking students in Colorado. By incorporating evidence-based practices and stakeholder input, the rubrics provide a robust framework for evaluating instructional materials.

However, instructional materials necessitate continuous updates to reflect the realities of transitional and dual-language classrooms. Moving forward, further enhancing alignment with bilingual instructional models and between English and Spanish will be important. These efforts will ensure Spanish-English bilingual learners have access to high-quality instructional materials tailored to their unique sociolinguistic and educational contexts.



6

Conclusions and Recommendations

The Colorado State Legislature passed the READ Act in 2012 and updated the Act in 2019. The revised Act requires an independent evaluation to identify and assess strategies that the state and local districts and schools have taken to support Colorado students in achieving proficiency in reading. This report focuses on the findings related to assessments on the Approved list, Advisory Lists for Instructional Programming and Professional Development, and CDE's processes for the Spanish rubric.

Key Findings

- The materials CDE approved for use with READ Act funds meet the minimum requirements in SB19-199.
- The revised Spanish-language rubric and approved materials meet READ Act requirements.
- The resources identified provide Colorado's LEPs with a range of choices for meeting READ Act requirements.
- Few instructional and professional learning programs provided evidence of impact on student outcomes.
- Most resources offer limited guidance on accommodations and supports for ELs and students with disabilities.

Conclusions

- **Review processes and identified resources align with READ Act requirements.** The review processes for assessments, instructional materials, and professional learning result in resources that align with the guidance of the READ Act. The State Board, Legislature, and the public can be confident that CDE is working to faithfully implement the READ Act in both letter and spirit. All but one assessment, all instructional programs, and all professional development programs that received a rating of “fully meets,” “largely meets,” or “partially meets” meet the minimum threshold for alignment with the READ Act.
- **CDE’s Spanish-language materials rubric and selection processes meet READ Act requirements.** CDE’s updated rubric for Spanish-language instructional materials, like its English-language counterpart, aligns with the READ Act and with empirical literature about Spanish-language reading instruction. We find that the rubric and review processes address the challenges that we identified in the Year 3 Instructional Materials report. Moreover, the rubric reflects a comprehensive effort to align instructional materials with the sociolinguistic realities and academic needs of Spanish-speaking students in Colorado. Informed by extensive feedback from stakeholders and expert analyses, the revised rubric addresses the unique needs of ELs while emphasizing cultural and linguistic appropriateness.
- **Resources provide LEPs with choices.** The identified assessments, instructional materials, and professional development programs provide LEPS with a great deal of choice in how to use READ Act funds to support the

implementation of scientifically based reading instruction. The breadth of resources aligns with Colorado's local control focus. As noted in the Year 4 Per Pupil report, the choices also create a need for additional guidance about how to use materials and professional learning in ways that enhance each other.

- **Most instructional programs and professional learning programs are supported with the lowest tier of empirical evidence, a logic model.** The overwhelming majority of instructional materials and professional learning opportunities that appear on CDE's Advisory List are supported with ESSA Tier 4 evidence ("Demonstrates a rationale"). This is reflective of the national materials and professional learning market, which largely does not currently have systematic evidence of impact on student outcomes. CDE may wish to provide additional guidance to LEPs, educators, and the public about how to understand and interpret this evidence.
- **Approved resources provide limited guidance about supporting ELs and students with disabilities.** Assessments, for example, provide limited guidance about appropriate accommodations. Further, Spanish-language instructional materials focus on building reading skills in Spanish, but do not offer guidance about how these programs might be integrated into different models of bilingual education adopted by LEPs. Finally, approved professional learning programs provide limited guidance in how to support ELs in learning to read in English.

Specific Recommendations

Recommendations for Approved Assessments

Based on our findings, we recommend the following additions to the assessment application and approval process:

- **Adjust the vendor submission process to focus on the different purposes for assessment (i.e., interim, diagnostic, summative.)** A single assessment should not be used for multiple purposes unless explicit evidence is provided to support those multiple purposes. For example, if the vendor submits the same assessment for both interim and diagnostic, they need to provide evidence supporting each assessment purpose. In addition, if a vendor submits English- and Spanish-language versions of an assessment, they should provide evidence for each. As CDE increases its focus on screening for indicators of dyslexia characteristics, ensure that interim assessments provide explicit evidence that they validly and reliably provide both universal and dyslexia screening.
- **Ask vendors to improve the evidence and documentation they provide.** Ask vendors to provide robust and detailed evidence that supports the intended use of each assessment, based on CDE's definitions of the required assessment types (interim, diagnostic, summative). The specific types of additional evidence that CDE may ask for are outlined in the rubric WestEd used to conduct its review of the assessments. For example, ask vendors to provide more evidence that representative samples include ELs and students with disabilities. Further, ask vendors to provide information about convening bias review committees and conducting qualitative bias reviews.

- **Request that vendors enhance their documentation of standards alignment.** Ask vendors to strengthen alignment with Colorado standards, especially for Spanish-language versions of the assessments.
- **Encourage vendors to strengthen approaches to bias removal and cultural representation.** Ask vendors to enhance bias analysis and ensure cultural and dialectical representation through diverse sampling.
- **Invite vendors to expand accommodations and score report usability.** Encourage vendors to develop clear guidelines for accommodations and improve usability of score reports, ensuring that they are developed for the intended score user and making them available in multiple languages.

Recommendations for Instructional Programs

Based on our findings, we recommend that CDE enact the following recommendations:

- **Provide additional guidance to LEPs focused on selecting instructional programs.** With over 100 CDE-approved instructional programs, we recommend CDE provide clear guidance to LEPs on selecting programs that best support student outcomes within local contexts. Educators need evidence-based criteria to identify programs aligned with student literacy needs. By offering structured guidance, CDE can help districts and schools make informed decisions, maximize instructional impact, and ensure access to high-quality learning resources that best fit the needs of the LEP. Additional guidance might focus on topics such as:
 - recommending processes for program selection;
 - highlighting information about ESSA evidence tiers and explanations for what evidence means in terms of impact on

students (e.g., Tier 4 evidence indicates alignment with evidence-based reading instruction but does not indicate that a program has had measurable impact on student outcomes); and

- offering guidance to LEPs on how to choose an optimal combination of core, supplemental, and intervention programming.

- **Identify additional supplemental and intervention programs that focus on vocabulary and reading comprehension.** Of the 28 supplemental programs approved in the most recent review cycle, six specifically target vocabulary development and four target reading comprehension. Of the 23 intervention programs approved in the most recent review cycle, three target vocabulary and one target reading comprehension—both essential skills for reading success. We recommend CDE encourage vendors to submit supplemental and intervention programs that explicitly support vocabulary acquisition and reading comprehension. Expanding these offerings will help LEPs address gaps in instruction and ensure all students receive the targeted support they need to become proficient readers.
- **Address gaps in bilingual literacy.** The use of both the English- and Spanish-language lists of instructional programs can be further strengthened by offering additional comprehensive bilingual guidance for LEPs. The purpose of this is to ensure better alignment across dual-language and transitional instructional models that reflect the diverse realities of districts, schools, and classrooms.

Recommendations for Additional Spanish-English Bilingual Guidance

This section outlines targeted recommendations to enhance CDE's guidance to LEPs and ensure LEPs' efficacy in supporting literacy outcomes for Spanish-speaking students. Key recommendations include:

- **Highlight materials that provide explicit supports for integration across bilingual models.** When reviewing instructional materials, CDE may wish to highlight materials that provide alignment for the variety of bilingual models common in Colorado, including Dual Language Immersion, Transitional Bilingual Education, Developmental Bilingual Education, and Two-Way Immersion (these program models are defined in Appendix D.) Regardless of the program model, guidance should be tied to goals for ELs. For example, districts implementing a transitional Spanish-English program should start by selecting a program that promotes strong Spanish literacy foundational skills and offers help in the transition to English literacy. Providing this information in an easily accessible way would support LEPs in selecting the materials that are most appropriate to their settings.
- **Incorporate a focus on approved instructional materials into professional learning.** As Colorado adopts professional learning on promoting foundational literacy skills among ELs, we recommend that one component of that professional learning focus on adopting and using both Spanish- and English-language approved instructional materials to provide rigorous, high-quality instruction so that students can meet or exceed standards.
- **Provide guidance to support coordinated literacy instruction.** We recommend that CDE provide a guidance document to help bilingual programs integrate oral and written literacy development in both English and Spanish through a

coordinated and intentional approach. This includes aligning instruction across languages with a clear language allocations plan that supports the development of biliteracy, rather than treating each language in isolation. This approach aligns with frameworks, such as New Mexico Public Education Department's Biliteracy Guidance (NMPED & Cavazos, 2022), which emphasizes the integration of phonological awareness, oral language, and vocabulary development across languages. For example, the Biliteracy Guidance illustrates how to implement various program models. Some examples include:

- "The purpose of this guide is to support the literacy and biliteracy instruction of ELs in the four Bilingual Multicultural Education Program (BMEP) models in which they receive their instruction." (p. 3)
- "The goal is to help New Mexico teachers provide improved instruction in the areas of oral language development and the components of reading (phonological awareness, phonics, fluency, vocabulary, and comprehension) and their specific application in English and the home language. Emphasis will be given to oral language development, cross-linguistic connections, and metalinguistic awareness to improve students' language development in L1 and L2. Please note that the oral language development strategies can be used for any language, but for the purpose of providing concrete examples, Spanish is used." (p.3)

By providing additional guidance, CDE can further support equitable reading instruction for Spanish-speaking learners while reflecting the sociocultural and linguistic realities of Colorado's classrooms.

Recommendations for Professional Development Programs

- **Encourage vendors to provide more rigorous evidence and make logic models explicit.** Only one professional development program presented empirical research that demonstrated its potential to have an impact on student outcomes (ESSA evidence level 1 or 2). We recommend encouraging vendors to collect and provide empirical research that documents evidence of program impact that meets ESSA evidence levels 1, 2, or 3.

All programs provided a logic model, theory of action, or synthesis of research that outlines how and why the program expects to have impact or alignment to PD program elements associated with impacts on teacher practice and student outcomes based on four research-based PD criteria (content focus, models of effective practice, feedback and direction, and ongoing support with sufficient duration of at least 45 hours) (ESSA evidence level 4). We recommend, at minimum, that CDE encourage vendors to make the logic model and/or theory of action guiding their programs explicit. In addition to encouraging vendors to be explicit about the evidence base that supports the approach to reading, ask them to be explicit about how their programs use models of effective practice; offer educators feedback and direction; and provide ongoing support such as coaching.

- **Ask vendors to submit copies of all teacher evaluation tasks embedded in the program to improve review.** Having access to these materials would help the independent evaluation team to better determine the extent to which vendors include assessment items related to pedagogical knowledge and classroom performance in teacher evaluations.

- **Ask vendors to provide additional information on how their programs prepare educators to address the specific needs of ELs and address culturally responsive instructional approaches.** This would include having vendors provide information about topics such as distinct needs related to phonemic awareness and offering guidance on how educators might learn about sounds in English that do not exist in their students' native languages or interfere with English phoneme production. In addition, CDE should ask all vendors to indicate how their programs reflect culturally responsive, asset-based instructional approaches.




References

- Adams, M. J. (1990). *Beginning to read: Thinking and learning about print*. MIT Press.
- American Educational Research Association (AERA), American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association.
https://www.testingstandards.net/uploads/7/6/6/4/76643089/standards_2014edition.pdf
- Anthony, J. L., & Francis, D. J. (2005). Development of phonological awareness. *Current Directions in Psychological Science*, 14(5), 255–259.
<https://doi.org/10.1111/j.0963-7214.2005.00376.x>
- August, D., Carlo, M., Dressler, C., & Snow, C. (2005). The critical role of vocabulary development for English language learners. *Learning Disabilities Research and Practice*, 20(1), 50–57. <https://doi.org/10.1111/j.1540-5826.2005.00120.x>
- Bravo-Valdivieso, L. (2004). El desarrollo de habilidades de segmentación en español: Factores influyentes en niños preescolares. *Revista Latinoamericana de Psicología*, 36(1), 33–50. <https://www.redalyc.org/pdf/805/80536103.pdf>
- Bus, A. G., & Van Ijzendoorn, M. H. (1999). Phonological awareness and early reading: A meta-analysis of experimental training studies. *Journal of Educational Psychology*, 91(3), 403–414. <https://doi.org/10.1037/0022-0663.91.3.403>
- Carlo, M. S., August, D., & Snow, C. E. (2005). Sustained vocabulary-learning strategy instruction for English-language learners. *Teaching and learning vocabulary: Bringing Research to Practice*, 137–153.
- Chard, D. J., Vaughn, S., & Tyler, B. J. (2002). A synthesis of research on effective interventions for building reading fluency with elementary students with learning disabilities. *Journal of Learning Disabilities*, 35(5), 386–406.
<https://doi.org/10.1177/00222194020350050101>
- Colorado Department of Education (CDE). (2017). *Minimum Reading Competency Skills Matrix*. Colorado Department of Education. Retrieved April 18, 2025 from <https://www.cde.state.co.us/coloradoliteracy/minimumcompetencylinkedmatrix>
- Colorado Department of Education (CDE). (2020). *Colorado academic standards: Reading, writing, and communicating*. Retrieved April 18, 2025 from <https://www.cde.state.co.us/coreadingwriting>
- Colorado Department of Education (CDE). (2024a). *Colorado READ Act: 2024 annual report*. <https://www.cde.state.co.us/cdedepcom/readactreport>
- Colorado Department of Education (CDE). (2024b). *Advisory List of Instructional Programming*.
<https://www.cde.state.co.us/coloradoliteracy/advisorylistofinstructionalprogramming2020>

- Colorado Department of Education (CDE). (2025a). *Approved professional development*. <https://www.cde.state.co.us/coloradoliteracy/readactprofessionaldevelopmentevideceteachertraining>
- Colorado Department of Education (CDE). (2025b). *READ Act approved assessments*. <https://www.cde.state.co.us/coloradoliteracy/readactassessments>
- Council of Chief State School Officers. (2019, July). *Third grade reading laws: Implementation and impact*. <https://files.eric.ed.gov/fulltext/ED603144.pdf>
- Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017, June). *Effective teacher professional development*. Learning Policy Institute. Retrieved from [https://learningpolicyinstitute.org/sites/default/files/product-files/Effective Teacher Professional Development REPORT.pdf](https://learningpolicyinstitute.org/sites/default/files/product-files/Effective_Teacher_Professional_Development_REPORT.pdf)
- Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher*, 38(3), 181–199. <https://doi.org/10.3102/0013189X08331140>
- Durgunoğlu, A. Y., Nagy, W. E., & Hancin-Bhatt, B. J. (1993). Cross-language transfer of phonological awareness. *Journal of Educational Psychology*, 85(3), 453–465. <https://doi.org/10.1037/0022-0663.85.3.453>
- Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, B. V., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's meta-analysis. *Reading Research Quarterly*, 36(3), 250–287. <https://doi.org/10.1598/rrq.36.3.2>
- Ehri, L. C., Nunes, S. R., Stahl, S. A., & Willows, D. M. (2001). Systematic phonics instruction helps students learn to read: Evidence from the National Reading Panel's meta-analysis. *Review of Educational Research*, 71(3), 393–447. <https://doi.org/10.3102/00346543071003393>
- Elleman, A. M., Lindo, E. J., Morphy, P., & Compton, D. L. (2009). The impact of vocabulary instruction on passage-level comprehension of school-age children: A meta-analysis. *Journal of Research on Educational Effectiveness*, 2(1), 1–44. <https://doi.org/10.1080/19345740802539200>
- Fiester, L. (2013). *Early warning confirmed: A research update on third-grade reading*. The Annie E. Casey Foundation.
- Gamse, B., Boulay, B., Fountain, A. R., Unlu, F., Maree, K., McCall, T., & McCormick, R. (2011). *Reading First implementation study 2008–09*. Office of Planning, Evaluation and Policy Development, U.S. Department of Education.
- García, J. R., & Cain, K. (2014). Decoding and reading comprehension: A meta-analysis to identify which reader and assessment characteristics influence the strength of the relationship in English. *Review of Educational Research*, 84(1), 74–111. <https://doi.org/10.3102/0034654313499616>
- Goldenberg, C. (2020). Reading wars, reading science, and English learners. *Reading Research Quarterly*, 55(S1), S131–S144. <https://doi.org/10.1002/rrq.340>

- Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, 7(1), 6–10.
<https://doi.org/10.1177/074193258600700104>
- Hernandez, D. J. (2012). *Double jeopardy: How third-grade reading skills and poverty influence high school graduation*. The Annie E. Casey Foundation. Retrieved from <https://files.eric.ed.gov/fulltext/ED518818.pdf>
- Jeynes, W. H. (2008). A meta-analysis of the relationship between phonics instruction and minority elementary school student academic achievement. *Education and Urban Society*, 40(2), 151–166. <https://doi.org/10.1177/0013124507304128>
- Joshi, R. M., & Aaron, P. G. (2000). The component model of reading: Simple view of reading made a little more complex. *Reading Psychology*, 21(2), 85–97.
<https://doi.org/10.1080/02702710050084428>
- Joyce, B. R., & Showers, B. (2002). *Student achievement through staff development* (3rd ed.). Association for Supervision & Curriculum Development.
- Kamil, M. (2004). Vocabulary and comprehension instruction: Summary and implications from the National Reading Panel findings. In P. McCardle & V. Chhabra (Eds.), *The voice of evidence in reading research* (pp. 213–235). Brookes.
- Kemple, J., Herlihy, C., Bloom, H., Zhu, P., & Berlin, G. (2009). *Understanding Reading First: What we know, what we don't, and what's next*. MDRC.
- Lieberman, I., Shankweiler, D., Fischer, F., & Carter, B. (1974). Explicit syllable and phoneme segmentation in the young child. *Journal of Experimental Child Psychology*, 18(2), 201–212. [https://doi.org/10.1016/0022-0965\(74\)90101-5](https://doi.org/10.1016/0022-0965(74)90101-5)
- Lloyd, D. N. (1978). Prediction of school failure from third-grade data. *Educational and Psychological Measurement*, 38(4), 1193–1200.
<https://doi.org/10.1177/001316447803800442>
- López, L. M., & Greenfield, D. B. (2004). Cross-language transfer of phonological skills of Hispanic Head Start children. *Bilingual Research Journal*, 28(1), 1–18.
<http://dx.doi.org/10.1080/15235882.2004.10162609>
- Marulis, L. M., & Neuman, S. B. (2010). The effects of vocabulary intervention on young children's word learning. *Review of Educational Research*, 80(3), 300–335.
<https://doi.org/10.3102/0034654310377087>
- McArthur, G., Sheehan, Y., Badcock, N. A., Francis, D. A., Wang, H. C., Kohnen, S., Banales, E., Anandakumar, T., Marinus, E., & Castles, A. (2018). Phonics training for English-speaking poor readers. *Cochrane Database of Systematic Reviews*, 11. John Wiley and Sons Ltd.
<https://doi.org/10.1002/14651858.CD009115.pub3>
- Mora, L., & Lopez, M. H. (2023). *Latinos' views of and experiences with the Spanish language*. Pew Research Center. Retrieved from <https://www.pewresearch.org/race-and-ethnicity/2023/09/20/latinos-views-of-and-experiences-with-the-spanish-language/>

- National Center on Intensive Intervention. (n.d.). *Tools chart review process*. Retrieved from <https://intensiveintervention.org/tools-charts/review-process>
- National Reading Panel (NRP). (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. National Institute of Child Health and Human Development, National Institutes of Health, Public Health Service, U.S. Department of Health and Human Services.
- New Mexico Public Education Department (NMPED), & Cavazos, L. (2022). *New Mexico Public Education Department Biliteracy Guidance: The science of reading for English learners*. Retrieved April 18, 2025 from https://web.ped.nm.gov/wp-content/uploads/2025/01/NMPED-Biliteracy-Guidance_12.4.22.pdf
- Pearson, P. D., Hiebert, E. H., & Kamil, M. L. (2007). Vocabulary assessment: What we know and what we need to learn. *Reading Research Quarterly*, 42(2), 282–296. <https://doi.org/10.1598/rrq.42.2.4>
- Scarborough, H. S. (2001). Connecting early language and literacy to later reading (dis)abilities: Evidence, theory, and practice. In S. B. Neuman & D. Dickinson (Eds.), *Handbook of Early Literacy Research* (pp. 97–110). Guilford Press.
- Share, D., Jorm, A., Maclean, R., & Matthews, R. (1984). Sources of individual differences in reading acquisition. *Journal of Educational Psychology*, 76, 1309–1324. <https://doi.org/10.1037/0022-0663.76.6.1309>
- Snider, V. E. (1997). The relationship between phonemic awareness and later reading achievement. *The Journal of Educational Research*, 90(4), 203–211. <https://doi.org/10.1080/00220671.1997.10544574>
- Solari, E. J., Grimm, R. P., McIntyre, N. S., & Denton, C. A. (2018). Reading comprehension development in at-risk vs. not at-risk first grade readers: The differential roles of listening comprehension, decoding, and fluency. *Learning and Individual Differences*, 65, 195–206. <https://doi.org/10.1016/j.lindif.2018.06.005>
- Stahl, S. A., & Fairbanks, M. M. (1986). The effects of vocabulary instruction: A model-based meta-analysis. *Review of Educational Research*, 56(1), 72–110. <https://doi.org/10.3102/00346543056001072>
- Stevens, E. A., Walker, M. A., & Vaughn, S. (2017). The effects of reading fluency interventions on the reading fluency and reading comprehension performance of elementary students with learning disabilities: A synthesis of the research from 2001 to 2014. *Journal of Learning Disabilities*, 50(5), 576–590. <https://doi.org/10.1177/0022219416638028>
- Storch, S. A., & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology*, 38(6), 934–947. <https://doi.org/10.1037/0012-1649.38.6.934>

- Suggate, S. P. (2016). A meta-analysis of the long-term effects of phonemic awareness, phonics, fluency, and reading comprehension interventions. *Journal of Learning Disabilities*, 49(1), 77–96. <https://doi.org/10.1177/0022219414528540> 
- Wolf, M., & Katzir-Cohen, T. (2009). Reading fluency and its intervention. *Scientific Studies of Reading*, 5(3), 211–239. https://doi.org/10.1207/S1532799XSSR0503_2 
- Wright, T. S., & Cervetti, G. N. (2016). A systematic review of the research on vocabulary instruction that impacts text comprehension. *Reading Research Quarterly*, 52(2), 203–226. <https://doi.org/10.1002/rrq.163> 



List of Appendices

A. Independent Evaluation Rubrics

A.1. Assessments Review Rubric

A.2. Instructional Programming Rubric

A.3. READ Act Evaluation of Professional Instructional Programming and Technology, Rubric for Winter 2025

B. Summary of Individual Program Ratings

B.1. Assessments

B.2. Professional Development Programs on CDE's Advisory List

B.3. Instructional Programs

C. Assessments

C.1. Description of Criteria Drawn from Additional Professional Standards

C.2. Summary of Ratings Tables

D. Bilingual Model Definitions