

Draft Independent Evaluation of Colorado READ Act Materials



Advisory List of Instructional Programming

Linda Friedrich, Bejanae Kareem, Rachel Lagunoff, and Rachel Tripathy, WestEd; Robin Wisniewski and Kershini Naidu, RTI International

May 2023

Contract Number: 148306



Colorado READ Act EVALUATION



© 2023 WestEd. All rights reserved.

Suggested citation: Friedrich, L., Kareem, B., Lagunoff, R., Tripathy, R., Wisniewski, R., & Naidu, K. (2023). *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	1
Instructional Programs	2
Processes for Reviewing Instructional Programs.....	3
Processes for Reviewing Spanish-Language Instructional Programs	7
Recommendations	7
1. Introduction	9
Colorado READ Act	10
Evaluation of READ Act	11
Purpose and Organization of Report.....	12
Scientific Foundations of Reading Proficiency in Early Elementary Grades	15
2. Advisory List of Instructional Programming	23
Introduction.....	24
Data Collection and Methods	29
Results and Discussion	30
Comparison of CDE and Independent Evaluation Results	34
3. CDEs for Reviewing Instructional Materials	36
Introduction.....	37
Results and Discussion	39
Overview of CDE’s Application and Review Process.....	39
CDE Program Review Rubrics	43
Reviewer Recruitment and Selection.....	55
Reviewer Training Process and Content.....	59
The Program Review Structure and Processes for Assigning Points.....	62
Summary of Findings.....	68
4. Spanish-Language Instructional Programs	69
Introduction.....	70
Summary of General Issues in Applicability of Rubrics to Spanish Literacy Materials	70
Overview of Applicability Issues in Rubrics	71
5. Conclusion	74
Summary of Findings.....	75
Recommendations.....	78
References	83
List of Appendices	89



Executive Summary

The Colorado State Legislature passed the Reading to Ensure Academic Development (READ) Act in 2012 and updated it 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 Approved Assessments, Advisory Lists of Professional Development and Instructional Programming, and Colorado Department of Education (CDE) processes for selecting materials for these lists.

Key findings

- **Materials CDE approved for use with READ Act funds meet minimum requirements of SB 19-199.**
- **In an overall summary rating, 22 instructional programs received a rating of “fully met,” 13 received a rating of “largely met,” and 1 received a rating of “partially met.”**
- **CDE’s rubrics and processes for vetting instructional programs are aligned with provisions of the READ Act and are grounded in evidence-based practices for teaching reading.**
 - Phase 1 of CDE’s selection process can reasonably be applied to both Spanish- and English-language programs.
 - Phase 2 of CDE’s selection process includes criteria that are not appropriate for teaching foundational reading skills in Spanish.



Executive Summary

The importance of students achieving early-grade reading proficiency for later academic success is well documented. The Colorado State Legislature responded to this challenge by passing the Colorado (Reading to Ensure Academic Development) READ Act in 2012. Most recently, the Legislature updated READ Act requirements with its 2019 revision. These revisions included requirements for an independent evaluation of the READ Act program. This report focuses on the findings related to the Colorado Department of Education’s (CDE) Advisory List of Instructional Programming and CDE’s processes for reviewing English- and Spanish-language instructional programs.

The overall conclusion from the review of instructional materials is **that materials CDE approved for use with READ Act funds meet minimum requirements of SB 19-199** (Exhibit ES.1).

Exhibit ES.1. Summary of Ratings

	Fully Met	Largely Met	Partially Met	Did Not Meet / Not Rated
Instructional Programs	22	13	1	0

Note: Includes ratings for instructional programs in both English and Spanish.

The remainder of this executive summary describes findings related to instructional programming and presents recommendations for refining review processes. The concluding chapter to this report contains more detailed findings and recommendations.

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. Of 72 instructional programs that CDE reviewed in 2022, 36 were fully or partially approved. The evaluation of these instructional programs concluded that, by and large, these programs met the core requirements outlined in the READ Act.



Executive Summary

For the evaluation’s overall summary rating, 22 instructional programs received a rating of “fully met,” 13 received a rating of “largely met,” and 1 received a rating of “partially met.” Of the 36 instructional programs reviewed, 34 met the minimum threshold for evidence of the potential to make a positive impact on students’ reading outcomes—having a clear logic model rooted in the science of reading. All programs demonstrated the presence of skill development in phonemic awareness, phonics, reading fluency, and reading comprehension (as applicable), with 22 fully meeting criteria for explicit and systematic skill development. All core programs met 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.

Processes for Reviewing Instructional Programs

Consistent with provisions of the READ Act, CDE reviews vendor applications for an Advisory List of Instructional Programming every 2 years. The Advisory List includes core, supplemental, and intervention programs. CDE employs a two-part process. In Part I, vendors submit a letter of interest that provides initial evidence of a program’s alignment with the components of scientifically based reading instruction named in the READ Act (i.e., phonemic awareness, phonics, fluency, vocabulary, and comprehension). This evidence is screened by CDE staff. In Part II, vendors submit a body of evidence to demonstrate the alignment of their programs with the READ Act and that their programs use research-based practices for teaching reading. CDE convenes a panel of educators who have a demonstrated background in scientifically based reading instruction and can certify that they have no conflicts of interest. CDE then trains reviewers to rate programs using a two-phase, consensus-based review process. Both phases employ rubrics and scoring guides that clearly articulate specific criteria and evidence for rating programs. Vendors receive a copy of their scores and have an opportunity following each part and phase of

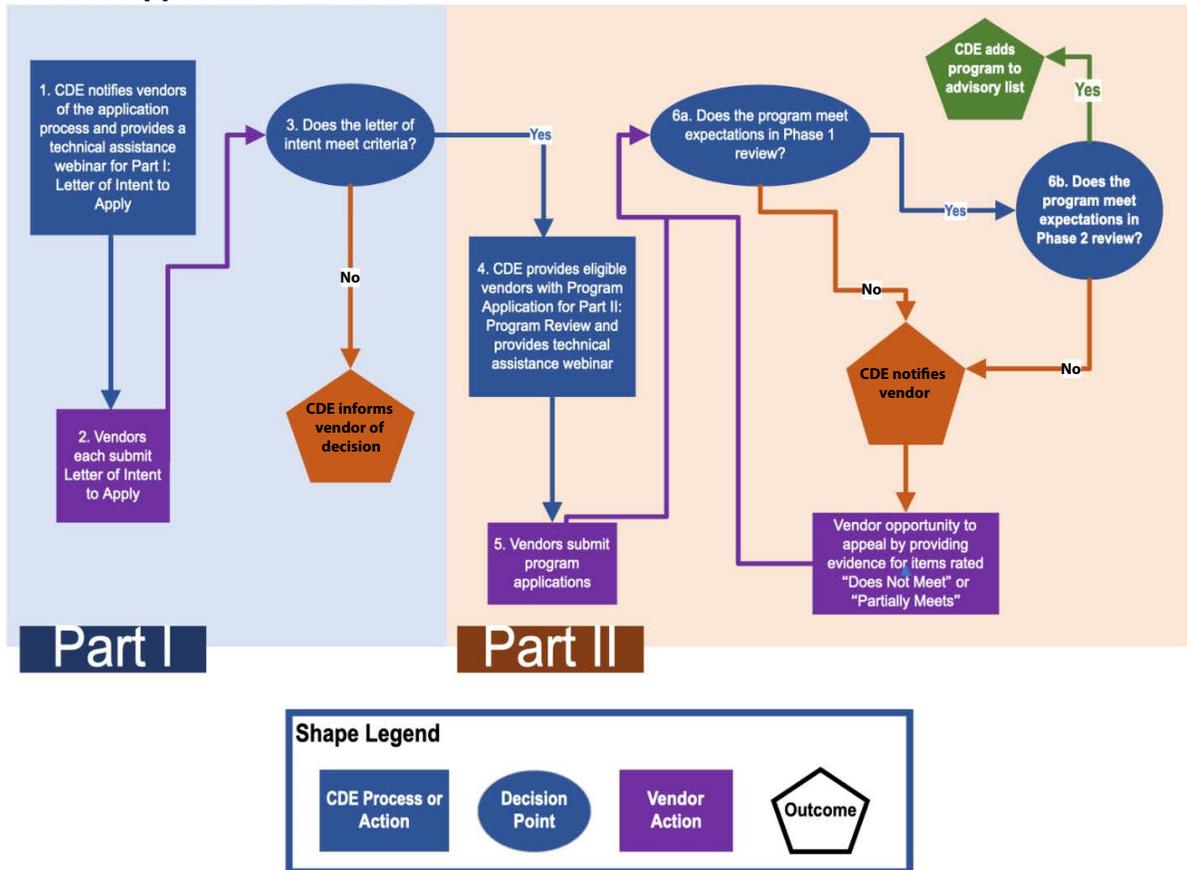


Executive Summary

the process to submit an appeal; all appeals are reviewed by CDE staff. This process is summarized in Exhibit ES.2.

Exhibit ES.2. Colorado Department of Education Vendor Application Process for Inclusion on Advisory List

Vendor Application Process



This year’s report includes an in-depth qualitative review of CDE’s processes for recruiting and training instructional materials reviewers, processes and criteria for reviewing instructional materials, and considerations for the selection of Spanish-language programs. In its review of the Advisory List of Instructional Programming, the evaluation team identified several cross-cutting findings.

- **The Advisory List offered clear guidance about explicit and systematic instruction in the elements of scientifically based**



Executive Summary

reading instruction. This finding is consistent with the 2021 evaluation.

- **Vendors submitted empirical evidence or promise of effectiveness to CDE that varied significantly in type and quality.** Only a handful of instructional program vendors submitted high-quality randomized controlled trials (Every Student Succeeds Act [ESSA] Tier 1 evidence level) and quasi-experimental studies (ESSA Tier 2). In contrast, some vendors that submitted evidence to claim ESSA Tier 3 included research summaries or data tables that lacked contextual information or interpretation, case studies from a small number of observations, or research that relied on poorly designed empirical studies. The evaluation team found that 5% of vendors submitted Tier 1 or Tier 2 evidence in 2022, which is a smaller proportion of vendors than the 10% that submitted such evidence in 2020. This finding speaks to the difficulty and expense of conducting rigorous evaluation research, which was compounded by the COVID-19 pandemic.
- **The instructional program review process clearly reflected requirements of READ Act legislation.** Each step of CDE's vetting process emphasized the components of scientifically based reading instruction in English. 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.
 - Vendors at both the Letter of Intent (Part I) and Application (Part II) stages were required to submit evidence that their approaches to reading instruction were aligned with the five components of reading instruction named in the READ Act.
 - The Phase 1 and Phase 2 rubric criteria for each grade level were rooted in evidence, were clearly specified, and



Executive Summary

reflected both historical and current understandings of how the science of reading could be applied effectively in classroom practice.

- Reviewers in both the 2020 and 2022 review cycles were well qualified, demonstrating an understanding of and educational experience implementing scientifically based reading instruction, and well trained to apply CDE’s explicit criteria focused on scientifically based reading instruction.
- **While all aspects of the instructional program review were clearly consistent with READ Act requirements, reviewers noted some limitations of the review process for programs in English.**
 - Most reviewers said that the training was “just right,” but a few wished it had been longer. CDE leaders discussed different supports needed for reviewers, and experienced reviewers suggested that those not as experienced may need more support.
 - Some reviewers found it challenging to find the right evidence for each area, with one suggesting that vendors may not have content experts supporting application submission.
 - Some reviewers believed that their colleagues in districts did not understand processes that CDE uses to vet and recommend instructional materials.
 - Some reviewers noted that the Part II, Phase 1 criterion, “Differentiation and support are provided for supporting English learners, students who are struggling, and those who need acceleration,” was difficult to score because it combines supports for students with disparate needs.



Executive Summary

- In 2020, the evaluation revealed that CDE was less successful in identifying instructional materials in English that offer differentiated support for English learners; this was especially true of intervention and supplemental programs. While 2022 programs were not reviewed for this criterion, several reviewers from outside CDE mentioned that one of the most difficult Phase 1 criteria to score was the one on differentiation. These reviewers noted that English learners, students with learning differences, and students who require advanced learning opportunities have different needs.

Processes for Reviewing Spanish-Language Instructional Programs

CDE was successful in identifying Spanish-language materials that included the required components of scientifically based reading that met the evaluation's thresholds. However, the evaluation identified a significant challenge for reviewing Spanish-language programs: reviewers were required to use rubrics and scoring guides for English-language programs. Several Phase 2 criteria were defined in ways specific to English-language reading instruction and were not applicable to Spanish-language reading instruction.

Recommendations

Recommendations, which are elaborated upon in Chapter 5, include the following:

Recommendation 1: Keep rubrics and the two-phase process for reviewing English-language programs the same.

Recommendation 2: Develop and list separate rubric indicators in the Phase 1 rubric for meeting the needs of English learners and consider adding indicators for cultural representativeness.

Recommendation 3: Convene a panel of experts in early literacy learning and instruction in Spanish for Spanish speakers to review CDE's rubrics and then



Executive Summary

revise these rubrics to reflect best practices of evidence-based early literacy instruction in Spanish.

Recommendation 4: Refine the process for reviewing English-language programs as follows:

- a) Add differentiated supports for less experienced reviewers
- b) Codify the consensus-based process that is currently used for program selection
- c) Provide vendors with additional guidance on submitting materials that align with rubric criteria

Recommendation 5: Clearly identify the degree to which vendors' evidence meets ESSA evidence levels, make studies available, and provide guidance for how districts and schools might consider this research when selecting instructional programs.

Recommendation 6: Provide districts with streamlined and accessible information about how programs are vetted.



1

Introduction

The 2019 revision of the READ Act (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 underway and is being conducted by an independent research team led by WestEd and 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 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 Professional Development and Instructional Programming 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.**



1. Introduction

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.¹ To help schools and districts support all children in achieving this goal, the Colorado State Legislature passed the Colorado Reading to Ensure Academic Development Act (READ Act) in 2012; this replaced the Colorado Basic Literacy Act.² 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” 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.³ Schools are then required to develop individual READ Act plans that identify a pathway for reaching grade-level proficiency for those designated as having a significant reading deficiency. The READ Act specifies certain components required in all student 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 approve a set of reading assessments, the READ Act charges the Colorado

¹ Hernandez, D. J. (2011). *Double jeopardy: How third-grade reading skills and poverty influence high school graduation*. The Annie E. Casey Foundation; Fiester, L. (2013). *Early warning confirmed: A research update on third-grade reading*. The Annie E. Casey Foundation. <https://www.aecf.org/resources/double-jeopardy>

² 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 (a) funding to support these efforts, (b) requirements for parent communication, and (c) an explicit focus on students identified as having a significant reading deficiency.

³ See <https://www.cde.state.co.us/coloradoliteracy/readact/resourcebank>.



1. Introduction

Department of Education (CDE) with creating Advisory Lists of Professional Development⁴ and Instructional Programming⁵ that are scientifically based and evidence based. LEPs may use READ Act funds to purchase instructional programming from the advisory list; they may purchase instructional programs that are not on CDE's advisory list if they do not use READ Act funds. With the 2019 revision of the READ Act, the legislature requires all K–3 teachers to complete evidence-based training in teaching reading by the beginning of the 2022-2023 school year and each school year thereafter. The professional development programs on CDE's advisory list allow teachers who successfully complete the professional development to meet this requirement.

Evaluation of READ Act

The 2019 revision of the READ Act (SB 19-199) 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 SB 19-199).⁶ The multiyear evaluation is now underway and is being conducted by an independent research team led by WestEd and that includes APA Consulting and RTI International. The key legislative goals for this evaluation are as follows:

1. Help state policymakers and district leaders understand the impacts of READ Act funding and support on students, families, schools, and districts
2. Determine the extent to which CDE's processes resulted in approved assessments and Advisory Lists of Professional Development and Instructional Programming that are consistent with READ Act requirements

⁴ See <https://www.cde.state.co.us/coloradoliteracy/readactprofessionaldevelopmentevidenceteachertraining>.

⁵ See <https://www.cde.state.co.us/coloradoliteracy/advisorylistofinstructionalprogramming2020>.

⁶ See <https://www.cde.state.co.us/cdedepcom/readactreport>.



1. Introduction

3. 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 third year of the legislatively mandated evaluation for Goals 2 and 3, focusing on instructional programming since new assessments and professional development programs were not added in 2022. The report relies on multiple sources of information, including

1. materials submitted by vendors to CDE as part of the review process;
2. documentation of CDE’s review processes and timelines provided by CDE, including publicly available resources; and
3. interviews with CDE staff who led 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 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 Report

In this report, the evaluation team describes the evaluation of 36 instructional programs on the advisory list.⁷ Key data and information presented in this summary report for instructional programs include (a) the evidence base for the programs; (b) scientifically based reading skills; and (c) 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

⁷ CDE approved 36 unique instructional programs, but some were approved for multiple categories and therefore counted more than once in the total (e.g., a program was approved as a core program, a supplemental program, and as an intervention program, so it was counted three times).



1. Introduction

processes used in the evaluation as well as findings with discussion of lessons learned, and ends with conclusions and recommendations. Of special interest this year is an in-depth qualitative review of CDE's processes for recruiting and training instructional materials reviewers, processes and criteria for reviewing instructional materials, and considerations for the selection of Spanish-language programs (Chapters 3 and 4).

It is also important to note limitations in this report. First, the question of whether programs resulted in a growth to standard is not addressed. In Year 2 of the READ Act evaluation, the evaluation convened a Technical Advisory Group (TAG). The TAG concluded that a single growth to standard model using equated READ act assessments scores is not advisable. Second, this report does not include an analysis of which instructional materials districts report using. These analyses are included in the Independent Evaluation of the READ Act: Per Pupil Funding Summary Reports for years 1, 2, and 3. Districts' adoption and use of programs will lag behind their inclusion on CDE's Advisory List.

Research Questions

This 3rd-year report addresses the following questions pertaining to assessments, instructional programs, and professional development programs.

Advisory List of Instructional Programming

1. Do all items on the Advisory List of Instructional Programming meet the requirements of the READ Act?
2. Do all items on the Advisory List of Instructional Programming meet additional professional standards of quality?
3. To what extent do items that pass Phase 1 but not Phase 2 of CDE's selection process meet the requirements of the READ Act?



1. Introduction

CDE's Processes for Identifying Items for Advisory List of Instructional Programming

The evaluation of CDE's process for reviewing English-language instructional materials was guided by two overall questions:

1. To what extent were CDE's processes for conducting reviews of instructional programming aligned with the READ Act and with evidence-based practices for teaching reading?
2. What was CDE's process for conducting the last two rounds of instructional programming reviews?

Subquestions were as follows:

1. What was the effectiveness of the rubrics used to review all core, supplemental, and intervention programming?
2. What was CDE's process for recruiting, training, and supporting reviewers?
3. What processes did CDE use to select instructional programs for the advisory list?
4. What were the steps for vendor material submission, approval, and appeal?
 - o What technical guidance did vendors receive?

Analytic Frameworks Used for Review

The criteria used in this evaluation for reviewing instructional 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 as well as professional standards for evaluating instructional program quality (see Appendix A.1 for detailed rubrics). By anchoring the evaluation protocols in related statute, guidance, and regulations, these evaluation protocols provide a transparent and consistent framework to



1. Introduction

determine READ Act compliance. The rubric (see Appendix A.1) lists the required elements, criteria for the elements, ratings for the evidence, and evidence needed.

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).

History of Early Reading Research and Policy

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) that would perform a comprehensive and informed synthesis of the research around effective methods for teaching children to read. In 2000, the 14-member NRP released its report, identifying five instructional components that are essential for early-grade reading development: phonemic awareness, phonics, reading fluency, vocabulary development, and reading comprehension (Langenberg et al., 2000). In a minority view included with the report, NRP member Joanne Yatvin cautioned Congress about interpreting the findings as definitive, claiming that the scope of topics that NRP examined was biased and narrow and that NRP 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



1. Introduction

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 passage of the No Child Left Behind Act in 2001, 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 instructional practices recommended by NRP (U.S. Department of Education, 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 (Center on Enhancing Early Learning Outcomes & Council of Chief State School Officers, 2019). Most of these laws reference or require “scientifically based” reading instruction, interventions, and curricula, although by the time many of these laws were passed the five essential components of reading had already been adopted by major publishers and teacher training programs in response to the NRP report (Herlihy et al., 2009).

Evidence-Based Components of Reading Instruction

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 these components. In fact, there have been so many studies on early reading instruction and intervention that researchers have been able to conduct meta-analyses whereby the authors attempt to identify all high-quality studies on a given topic and use statistical modeling to produce a more accurate impact estimate than any one study alone could provide. What follows is a short summary of recent meta-analytic findings on each of the five essential components of reading for pre-K through 3rd-grade students; all five components are programmed in the READ Act.



1. Introduction

1. *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”). It 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 1 year after the end of the interventions than interventions focused on phonics (Suggate, 2016).
2. *Phonics* is an instructional approach where 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 finds 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).
3. *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 the mechanics of reading, and fluency is therefore considered a critical link between word analysis and text comprehension. The developmental definition of fluency makes it



1. Introduction

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 best to improve reading fluency in the early grades.

4. *Vocabulary development* represents an important component of reading comprehension because understanding text requires the construction of meaning from known words (Kamil, 2004). There is a strong consensus that the size of students' 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 in improving expressive and receptive vocabulary (Marulis & Neuman, 2010). There is evidence that such interventions are also effective in improving comprehension of texts aligned with the intervention, but there are fewer studies finding that these interventions improve generalized reading comprehension (Elleman et al., 2009; Wright & Cervetti, 2017). 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, 2017).
5. *Reading comprehension* is the overall goal of reading instruction and 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). While some meta-analytic reviews show



1. Introduction

that decoding (García & Cain, 2014) and linguistic comprehension are each important predictors of reading comprehension, others find the effects to be small or inconclusive (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' age. 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 (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 effects in 1st and 2nd grade; and interventions targeting comprehension overall to be most effective for 3rd grade and beyond (Suggate, 2016).

Evidence-Based Practice Supporting English Learners and Students with Disabilities

Effective reading comprehension is dependent upon a complex and not entirely understood network of foundational skills that shift in their importance with a student's age and individual learning needs. In other words, when it comes to reading instruction one size does not fit all—and certain groups that have historically struggled with reading in the early grades require support and intervention beyond the typical reading curriculum. Effective reading instruction for English learners and students with disabilities, for example, shares many elements of reading instruction for proficient readers but also includes additional practices and supports for these groups. Research shows that English learners benefit from frequent and intentional instruction focused on oral language development—in other words, including modifications and support to ensure that students understand the words and concepts they read (Goldenberg, 2020). Additionally, multiple systematic reviews of research have found that models focused on simultaneously strengthening students' home language and their



1. Introduction

English skills have been more effective than models that focus on English alone (Greene, 1998; Slavin & Cheung, 2005). Consequently, effective reading instruction for young English learners would include modifications that help them understand a language that is new to them, likely by using native language supports or bilingual resources.

While students with disabilities comprise a heterogenous group with different challenges and needs, research has shown specific instructional strategies benefit reading outcomes for many students in this group, including sustained multiyear interventions, one-on-one or small group instruction, systematic instruction on foundational reading components, and abundant opportunities for practice and feedback (Berkeley et al., 2010; Vaughn & Wanzek, 2014). Effective reading instruction for young students with disabilities would incorporate personalized, targeted reading interventions that allow for supported practice of foundational skills.

Despite efforts to tailor instruction and improve reading outcomes for groups like English learners and students with disabilities, national reading outcomes for these groups have not improved in the last decade; on average, English learners and students with disabilities in 4th grade score far below even the “Basic” reading benchmark as measured by the National Assessment of Educational Progress (NAEP) (National Center for Education Statistics, 2019). Unfortunately, race and socioeconomic background are also predictors of student reading ability. While Asian and White students’ 4th-grade reading scores have hovered at or around the NAEP “Proficient” benchmark, Black and Hispanic students’ scores fall around or below the NAEP “Basic” benchmark. Students who are not classified as economically disadvantaged tend to score near the NAEP “Proficient” benchmark, while students from economically disadvantaged backgrounds score, on average, around the NAEP “Basic” benchmark. These disparities in early elementary reading scores are alarming and the achievement gaps are not narrowing, underscoring the need for effective instruction and resources that work specifically to support at-risk groups.



1. Introduction

Even with an ever-expanding body of research on reading mechanics and instructional best practices, most large-scale early literacy interventions have not produced the desired positive impacts on student reading achievement. Only a handful of rigorous impact evaluations have been conducted for large federal- and state-level reading initiatives, and they present mostly similar findings—some impact on instructional practices, but no impact on student reading performance. Following the Reading First funding initiative, for example, the U.S. Department of Education commissioned a study to examine the impact of Reading First on student reading proficiency. While the study found that teachers in Reading First schools received more professional development for reading instruction and spent more instructional minutes on the five essential components of reading, no impact on student reading performance was detected (Gamse et al., 2008). More recently, North Carolina State University evaluated the impact of the state’s Read to Achieve program, aimed at grade-level reading mastery for all 3rd-grade students. The study found no significant impacts on student reading achievement for students altogether or for demographic subgroups (e.g., low-income students or students with disabilities) (Weiss et al., 2018).

One exception to these interventions that seemingly failed to impact student reading performance was Oregon’s Reading First program, implemented from 2003 through 2009, which was shown by a rigorous multiyear evaluation to have improved student reading scores for students in kindergarten through 3rd grade (Baker et al., 2007). This comprehensive evaluation analyzed data from three different cohorts of students over 3 years. A staggered implementation rollout (i.e., the first cohort began their Reading First activities in Year 1, the second cohort began in Year 2, etc.) allowed researchers to examine not only year-to-year impact but also to analyze the magnitude of impact as schools became more experienced with the intervention. The Oregon Reading First evaluation found that schools receiving Reading First funding were more effective in improving student reading outcomes each year they implemented the



1. Introduction

intervention—in other words, they got better with experience. This finding is consistent with literature on effective educational interventions that has found consistent, sustained interventions to produce impacts of greater magnitude than short interventions (Borman & D’Agostino, 1996). These findings suggest evaluations of state reading policies and programs may need to focus on longer-term outcomes in order to identify impacts on student reading performance.



2

Advisory List of Instructional Programming

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

Key findings

- All but one of the approved instructional materials (core, supplemental, and intervention programs in English and Spanish) either fully met or largely met SB 19-199 requirements.
- Just over 5% of the programs reviewed by the evaluation team fully met the criterion of being “evidence-based” or “scientifically based.”
- All core academic programs approved in 2022 fully met the criterion of including required academic content.
- For programs on the Advisory list, CDE’s and the evaluation’s ratings for five components of scientifically based reading instruction were the same or one category apart 98% of the time.



2. Advisory List of Instructional Programming

Introduction

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 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 more support beyond the core program is needed to supplement reading instruction. *Intervention programs* are used to support individual students who need intervention support for their reading development. Supplemental and intervention programs were subject to fewer review criteria, under the assumption that all students had access to a core instructional program.

The instructional materials review followed a rubric (see Appendix A.1) 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 students in maintaining or meeting grade-appropriate proficiency levels in academic subjects in addition to reading. The following describes the three areas and their criteria:



2. Advisory List of Instructional Programming

1. *Is evidence-based (22-7-1209 (2) (b) (I)).*

The independent evaluators reviewed the logic models and empirical research studies that vendors submitted as part of their application to CDE. The evidence for each program was evaluated using Every Student Succeeds Act (ESSA) evidence levels criteria defined by What Works Clearinghouse (WWC)⁸ (Exhibit 2.1). As independent evaluators read through each study, they documented methodology, key findings, and effect sizes. Considerations in designating an ESSA evidence level included sample size, attrition and related statistical adjustments, bias reduction, and baseline equivalence; WWC reviews include these as important study design characteristics. The reviews presented here, then, approximate but are not as in-depth as WWC reviews.

After reviewing the available evidence, evaluators assigned each program an evidence rating ranging from 1 to 4. An ESSA Tier 1 or 2 earned a rating of “fully meets,” and an ESSA Tier 3 or 4 earned a rating of “partially meets.” If a program could not demonstrate an ESSA Tier 4, or did not submit appropriate evidence documentation, then it received a rating of “does not meet.” For the summary rating, a program could earn a rating of “fully meets” if it partially met criteria on this indicator.

ESSA Evidence Levels

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 identify programs that work. Stronger research methods provide stronger evidence for a program, resulting in higher tiers of ESSA evidence levels. Confidence in a program’s effectiveness rises with its tier rating. See Exhibit 2.1.

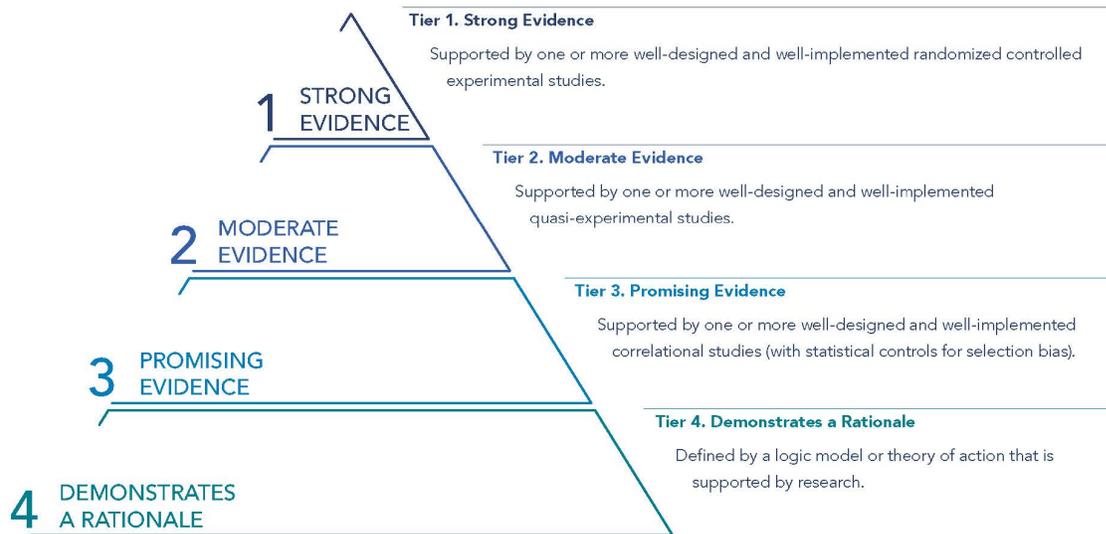
⁸ See <https://ies.ed.gov/ncee/wwc/essa>.



2. Advisory List of Instructional Programming

Exhibit 2.1. Every Student Succeeds Act 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.



2. *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)).* The independent evaluators examined whether skill development across reading components was present, explicit, and systematic, using vendor-supplied information. Core programs were evaluated for all five components 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 all components fully met criteria (i.e., demonstrated that components were explicitly and systematically taught). A program received a rating of “partially meets” if all components at least partially met criteria (meaning the component was



2. Advisory List of Instructional Programming

present but did not suggest that it was presented both explicitly or systematically). Finally, a core program received a rating of “does not meet” if at least one component of reading did not meet criteria. These decision rules were the same for supplemental and intervention programs, except that these programs were not required to address all components 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 areas (e.g., history, science) and genres (e.g., fiction, nonfiction). The evaluators used vendor-supplied documentation. A core program receives a rating of “fully meets” if the vendor-supplied analysis of quantitative and qualitative text complexity have the appropriate level of complexity for the grade for textual/linguistic demands and content demands. To receive a rating of “fully meets” both the content and types of texts in the program need to draw on a range of subject areas. A core program receives a rating of “partially meets” if quantitative and qualitative text complexity have appropriate level of complexity for either textual/linguistic or content demands. To receive a rating of “partially meets” either the content or the types of texts in the program need to draw on a range of subject areas. A core program “does not meet” if the vendor-supplied analyses of text complexity demonstrate limited opportunities for students to access grade-appropriate text. In addition, a core program receives a rating “does not meet” if both the content and types of texts do not draw on a range of subject areas.



2. Advisory List of Instructional Programming

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.



2. Advisory List of Instructional Programming

Data Collection and Methods

Information Used to Review Instructional Programs

The evaluation team read three sets of documents 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 or a logic model or theoretical rationale for the program.

Training for Independent Evaluators

The evaluation team created a rubric that specified criteria and rating options for each of the four instructional program ratings described above. The rubric was developed in partnership with an external expert advisory panel and was approved by CDE (see Appendix A.1). Evaluators were WestEd staff with expertise in research, curriculum, and/or Spanish-language reading instruction. All evaluators were trained as part of the Year 1 process.

Process Used for Reviews

The evaluation was conducted on 36 CDE-approved core, supplemental, and intervention programs, targeting the 1st-grade curriculum, materials, and documentation. The following three programs were reviewed at the 2nd-grade level because they were not designed for 1st grade per the vendor's application: (a) Curriculum Associates: Magnetic Reading 2021, (b) McGraw Hill: Corrective Reading, and (c) Really Great Reading: HD Word, First Edition.

In addition, evaluators examined 10 additional programs that met CDE's Phase 1 requirements but did not meet requirements for Phase 2 (see additional details in Chapter 3). Evaluators used the same process for reviewing these additional programs and were unaware of whether programs had passed one or both phases of CDE's review.



2. Advisory List of Instructional Programming

Results and Discussion

Nearly all approved instructional materials (core, supplemental, and intervention programs in English and Spanish) either fully met or largely met SB 19-199 requirements (Exhibit 2.2). (See Appendix B.1 for a list of all programs added to the advisory list with their ratings for each element.)

Exhibit 2.2. Summary Rating for Compliance with SB 19-199 Requirements, by Program Type

Program Type	Rating		
	Fully meets	Largely meets	Partially meets
Core programs in English	4	1	0
Supplemental programs in English	10	6	0
Intervention programs in English	6	6	1
Programs in Spanish (all)	2	0	0
OVERALL	22	13	1

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

Thirty-four of the 36 programs met the minimum standard for being evidence-based or scientifically based. Two programs fully met this criterion through rigorous research designs that demonstrated impact on students' reading outcomes (Exhibit 2.3).

Exhibit 2.3. Summary Rating for Evidence-Based or Scientifically Based Standard, by Program Type

Program Type	Rating		
	Fully meets	Partially meets	Does not meet
Core programs in English	0	5	0
Supplemental programs in English	1	14	1
Intervention programs in English	1	11	1
Programs in Spanish (all)	0	2	0
OVERALL	2	32	2



2. Advisory List of Instructional Programming

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, materials on CDE’s Advisory List for core, supplemental, and intervention programs offered explicit and systematic instruction in components of scientifically based reading instruction (Exhibit 2.4). Notably, four of the five English-language core programs and the one Spanish-language core program fully met 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 continued over the year.

Exhibit 2.4. Summary Rating for Skill Development in Components of Reading, by Program Type

Program Type	Rating		
	Fully meets	Partially meets	Does not meet
Core programs in English	4	1	0
Supplemental programs in English	10	6	0
Intervention programs in English	9	4	0
Programs in Spanish (all)	2	0	0
OVERALL	25	11	0

All 36 instructional programs approved by CDE had one or more components of scientifically based reading instruction that met the independent evaluation’s criteria for explicit and systematic instruction (Exhibit 2.5).



2. Advisory List of Instructional Programming

Exhibit 2.5. Summary Rating for Component of Reading, by Program Type

Component of Reading	Program Type	Rating			
		Fully meets	Partially meets	Does not meet	Not applicable*
Phonemic awareness	Core – English	4	1	0	0
	Supplemental – English	14	1	0	1
	Intervention – English	9	1	0	3
	All – Spanish	2	0	0	0
Phonics	Core – English	4	1	0	0
	Supplemental – English	13	1	0	2
	Intervention – English	11	1	0	1
	All – Spanish	2	0	0	0
Vocabulary development	Core – English	5	0	0	0
	Supplemental – English	3	4	0	9
	Intervention – English	5	3	0	5
	All – Spanish	2	0	0	0
Reading comprehension	Core – English	5	0	0	0
	Supplemental – English	3	4	0	9
	Intervention – English	5	2	0	6
	All – Spanish	2	0	0	0
Reading fluency	Core – English	5	0	0	0
	Supplemental – English	7	2	0	7
	Intervention – English	9	1	0	3
	All – Spanish	2	0	0	0

* 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)).

One hundred percent of programs met the criterion of core academic programs including academic content as required (Exhibit 2.6).



2. Advisory List of Instructional Programming

Exhibit 2.6. Summary Rating for Core Academic Content, by Program Type

Program Type	Rating		
	Fully meets	Partially meets	Does not meet
Core programs in English	5	0	0
Core programs in Spanish	1	0	0
OVERALL	6	0	0

Text complexity. This element was only evaluated for core programs, as the aim was to assess grade appropriateness of texts, and the evaluation team believed that an appropriate supplemental or intervention text would likely not be at grade level. All core instructional programs, in both English and Spanish, included texts written at grade level as evidenced by vendor-supplied quantitative and qualitative measures of text complexity (Exhibit 2.7). Having opportunities to read appropriately complex text is a prerequisite for maintaining or meeting grade-level proficiency standards. If students only access texts at lower complexity levels, it would be impossible for them to meet grade-level standards.

Exhibit 2.7. Summary Rating for Text Complexity, by Program Type

Program Type	Rating		
	Fully meets	Partially meets	Does not meet
Core – English	5	0	0
Core – Spanish	1	0	0

Range of texts. The range of texts element was also only evaluated for core programs. Supplemental and intervention programs are intended to complement the use of a core program, and appropriate supplemental and intervention programs are likely to have a limited range of genres and content. To determine whether core programs included a sufficient range of texts, the range of content was considered as well as genres included in programs. All five English-language, and the one Spanish-language, core programs fully met expectations for the range of genres and content. Engaging students with such a range is important in building a foundation for comprehending, interpreting, and using a range of texts as students move into upper elementary grades where the focus of reading shifts to reading to learn.



2. Advisory List of Instructional Programming

Comparison of CDE and Independent Evaluation Results

Of 72 instructional programs reviewed by CDE, 41 passed Phase 1 of the review. Of those, 36 passed Phase 2 of the review. In total, 50% of reviewed programs were ultimately approved or partially approved. Exhibit 2.8 shows the number of each type of program that passed each phase of the CDE review. (CDE’s review phases are described in Chapter 3.)

Exhibit 2.8. Program Performance on Rubric, by Program Type

Program Type	Phase 1			Phase 2		
	Pass	No Pass	Pass Rate	Pass	No Pass	Pass Rate
Core	7	7	50%	6	1	86%
Supplemental	19	15	56%	17	2	89%
Intervention	15	9	63%	13	2	87%
Total	41	31	57%	36	5	88%

Across the approved instructional programs independently reviewed by the evaluation team a total of 132 program components (i.e., Phonemic Awareness, Phonics, Vocabulary, Fluency, and Comprehension) were reviewed by both the evaluation team and CDE. For 78% of the program components reviewed, the evaluation team and CDE agreed on the extent to which the components met the criteria for evidence-based instruction. Of the program components for which the evaluation team and CDE assigned different ratings, about three-quarters of the time the evaluation team rated a component higher than CDE, and about one-quarter of the time the evaluation team rated a component lower than CDE. Agreement across the five components ranged from 65% to 88% (Exhibit 2.9), and agreement across different types of programs (core, supplemental, intervention) ranged from 65% to 93% (Exhibit 2.10). Where the evaluation team and CDE assigned program components different ratings, the assigned ratings differed by only one degree (e.g., “fully meets” and “partially meets”) in all but two instances. In most of cases where different scores were



2. Advisory List of Instructional Programming

assigned, the evidence supporting the rating landed the program component close to the threshold between the two different ratings it was given by each organization.

Exhibit 2.9. Comparison of CDE and Evaluation Program Component Ratings for 36 Approved Programs by Component Type

	Phonemic Awareness	Phonics	Vocabulary	Fluency	Compre-hension	TOTAL
Total number of ratings	32	33	21	26	20	132
Rated equally by CDE and Evaluation	28	26	14	22	13	103
Rated lower by Evaluation	2	3	1	1	0	7
Rated higher by Evaluation	2	4	6	3	7	22
Percent agreement	88%	79%	67%	85%	65%	78%

Exhibit 2.10. Comparison of CDE and Evaluation Program Component Ratings for 36 Approved Programs by Program Type

	Core	Supplemental	Intervention	TOTAL
Total number of ratings	30	55	47	132
Rated equally by CDE and Evaluation	28	36	37	103
Rated lower by Evaluation	2	1	4	7
Rated higher by Evaluation	0	18	4	22
Percent agreement	93%	65%	83%	78%



3

CDE's Processes for Reviewing Instructional Materials

The review of CDE's processes for reviewing instructional materials focused on the alignment of CDE's processes with the READ Act and scientifically based reading instruction. It included an analysis of all materials created by CDE for soliciting and reviewing vendor applications, interviews with CDE and external reviewers, interviews with CDE leadership for READ Act implementation, and an analysis of the research used to develop review criteria.

Key findings

- **CDE's rubrics and processes for vetting instructional programs are aligned with the provisions of the READ Act and are grounded in evidence-based practices for teaching reading.**
- **Reviewers have a background in scientifically based reading instruction. CDE works to recruit external reviewers that reflect the geographic and demographic diversity of the state. While most reviewers find CDE's training on its review processes and rubrics to be sufficient, some requested additional training.**
- **The depth and quality of materials submitted by vendors varies.**



3. CDE's Processes for Reviewing Instructional Materials

Introduction

This section summarizes findings and data gathered for the evaluation of CDE's program review process. The purpose of this part of the evaluation was to evaluate CDE's instructional program selection process over the past two cycles (2020 and 2022) for English-language instructional materials. In both the 2020 and 2022 cycles, CDE conducted reviews of core, supplemental, and intervention programs submitted by vendors to approve programs to include on the Advisory List of Instructional Programming.⁹

The findings in this chapter are derived from multiple sources of data, including

- materials created by CDE for training and guidance to reviewers and vendors;
- materials submitted by vendors to CDE as part of the review process of instructional programming;
- publicly available literature on the science of reading, including literature reviewed by CDE to develop the instructional programming rubrics;
- interviews with CDE staff who led the review process; and
- interviews with CDE staff and external reviewers who participated in the review process.

Following the 2022 review cycle, evaluators completed interviews or focus groups with a total of 16 reviewers. Of these reviewers, three were members of CDE's leadership team; six were CDE staff, two of whom reviewed Spanish-language programs only; and seven were external reviewers, one of whom reviewed Spanish-language programs only (Exhibit 3.1).

⁹ <https://www.cde.state.co.us/coloradoliteracy/advisorylistofinstructionalprogramming2020>.



3. CDE's Processes for Reviewing Instructional Materials

Exhibit 3.1. Reviewers who participated in interviews

	CDE Leadership Team	CDE staff	External Reviewers	Total Reviewers
English Language Programs	1	6	6	13
Spanish language programs	0	2	1	3

Of the sixteen interviewees, nine interviewees reported participating in both the 2020 and 2022 review cycles. Seven respondents participated in the 2022 review cycle only (Exhibit 3.2).

Exhibit 3.2. Involvement in Review Cycles

	CDE Leadership Team	CDE staff	External Reviewers	Total Reviewers
2020 and 2022 Review Cycle	3	3	3	9
2022 Review Cycle Only	0	3	4	7

The following sections present key data and information about CDE's application and selection processes, which included (a) rubrics for reviewing core, supplemental, and intervention instructional programs; (b) recruitment and selection of reviewers who demonstrated expertise in the science of reading; (c) training of reviewers for how to apply the rubric criteria to vendors' applications; and (d) the consensus-based review process and documentation of evidence that supported reviewers' conclusions, including vendors' applications containing evidence of alignment with the READ Act.



3. CDE’s Processes for Reviewing Instructional Materials

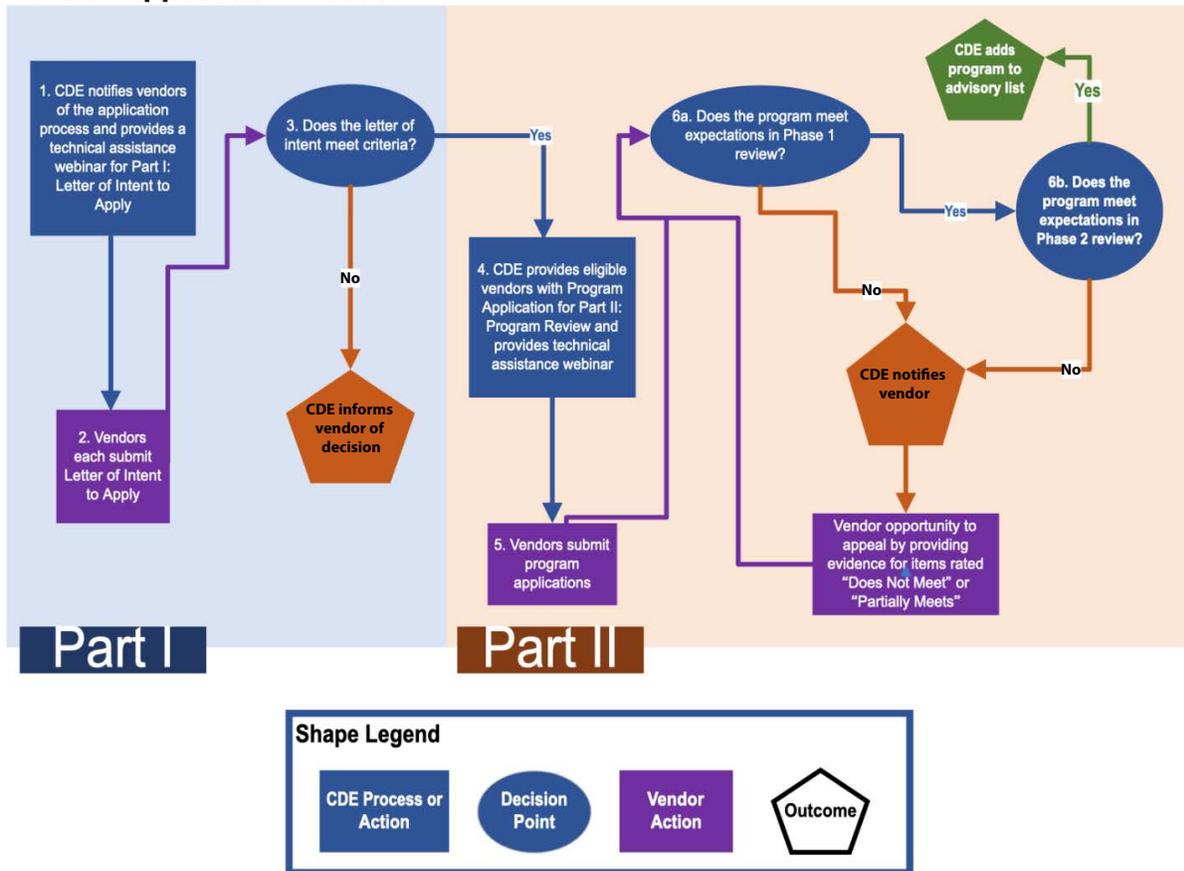
Results and Discussion

Overview of CDE’s Application and Review Process

CDE’s Advisory List Application Process, visualized in Exhibit 3.3, was divided into two parts. In Part I, vendors submitted a letter of intent (LOI). In Part II, vendors that passed Part I were eligible to submit their programs for review.

Exhibit 3.3. Colorado Department of Education Vendor Application Process for Inclusion on Advisory List

Vendor Application Process



Part I: LOI to Apply

- CDE notifies vendors and conducts technical assistance webinar.**
CDE informed vendors of the Instructional Programming Review Cycle and application process for vendors interested in submitting an instructional program for inclusion on the advisory list. CDE invited



3. CDE's Processes for Reviewing Instructional Materials

vendors to a technical assistance webinar which explained the background and purpose of the READ Act and advisory list, the process and timeline for the Part I submission, and what to include and how to submit the LOI. CDE also explained the two-part process (LOI then program review) and the appeal window.

2. **Vendors submit an LOI.** The LOI included the background of the program and a program summary. The LOI also included guidance about material submission that “hard copies of materials cannot be accepted and will not, under any circumstances, be reviewed.” Further program-related guidance stressed READ Act provisions and that if certain program elements were present then they may not advance to Part II:

“Programs included on the Advisory List of Instructional Programming must be scientifically and evidence-based as defined by statute and rule to meet the requirements of the READ Act.... Programs that use the Three Cueing Systems Model of Reading, also known as Meaning, Syntax, Visual (MSV) as their primary model for instruction may not advance to Part II Program review. If a program is aligned to or promotes Balanced Literacy or Whole Language instructional practices in the program materials or on the program website, it will not be approved.”

3. **CDE evaluates LOIs.** CDE evaluated the LOI in two main categories: form completion and program summary. For form completion, vendors needed to complete the required fields, meet READ Act requirements, and submit all materials in an online format. For the program summary, vendors needed to upload a program summary and describe the program's instructional focus, target audience, alignment with READ Act requirements, change from a previous submission if applicable, and whether the approach to teaching reading changed since previous editions and in what way. Vendors had to include all requested information in Part



3. CDE's Processes for Reviewing Instructional Materials

I: LOI to Apply to advance on to Part II: Program Review of the Advisory List Submission. CDE notified vendors whether they met the criteria.

Part II: Program Review

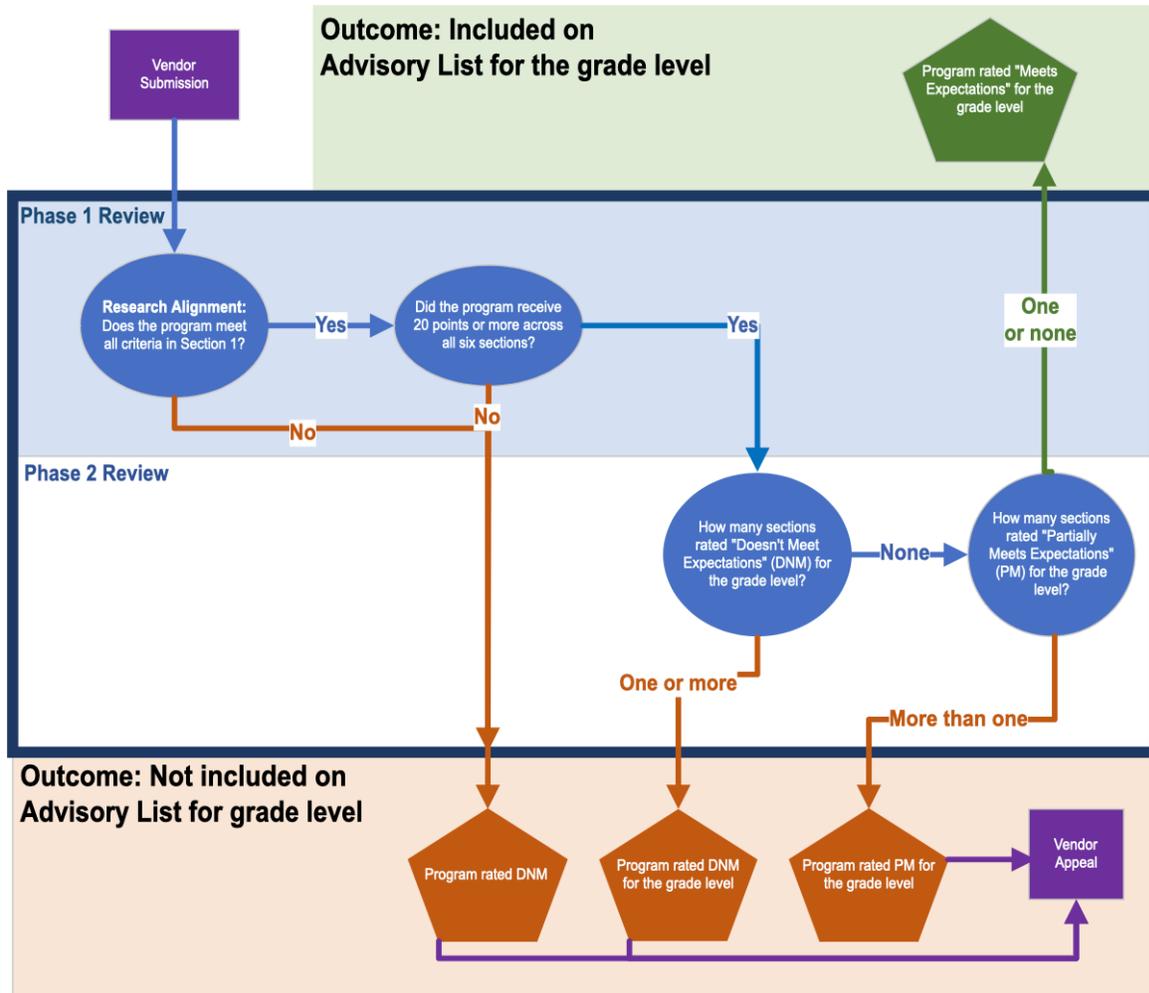
4. **CDE provided Part II applications to eligible vendors and conducted a technical assistance webinar.** CDE provided the Part II application to vendors who fulfilled Part I criteria and invited them to a technical assistance webinar. The webinar focused on required components of the application, describing what to include for scope and sequence, definitions of ESSA evidence levels 1–4 and the required evidence, essential program components, the theoretical model guiding the program, and alignment with state standards. CDE also presented the required format and submission details.
5. **Vendors submit programs for review.** Interested vendors submitted documentation, including sample program materials, about how their programs aligned with READ Act criteria.
6. **CDE conducts program review.** Submitted applications were reviewed by teams of CDE staff and qualified external reviewers including district staff, professors, and advocates. In evaluating vendors' Part II applications, CDE employed a two-phase process (Exhibit 3.4). In Phase 1, teams of CDE and external reviewers employed a consensus process to determine whether programs submitted provided sufficient evidence of alignment with READ Act criteria to warrant full review. The same teams then used the Phase 2 rubric to determine whether instructional programs met criteria for inclusion on the advisory list. When reviewers determined that programs met the Phase 2 review criteria, CDE notified the vendor and added the program to the advisory list.



3. CDE's Processes for Reviewing Instructional Materials

Exhibit 3.4. Program Review Process Using Two-Phase Rubrics

Instructional Program Review



7. **CDE provides vendors with appeal opportunities.** If programs did not meet or partially met established criteria, CDE informed vendors of the decision and provided them with guidance to appeal the decision. Vendors that submitted an appeal completed CDE's appeal documents which included information about where to locate additional evidence in the original submission when reviewers reported not finding sufficient evidence. Upon submitting an appeal, CDE's review team reconsidered the instructional program and rendered a decision following the postappeal review.



3. CDE's Processes for Reviewing Instructional Materials

Appeals Process

The program review concluded with vendor opportunities to appeal. CDE's internal review team reviewed the appeals. One CDE leader presented the rationale for CDE itself conducting the review: "We had no idea of knowing how many [appeals] we would receive back, ...the volume of materials that needed to be reviewed." Additionally, CDE staff members could conduct appeal reviews as part of their regular work assignments as opposed to external reviewers for whom conducting instructional programming reviews involves committing time over and above their regular work assignments. CDE explained that vendors could appeal for any elements in Phase 1 or Phase 2, and they could appeal again if, for example, they met criteria in Phase 1 after appeal but not Phase 2. Vendors could appeal the decision for one or both Phases, with the appeal submission process concluded after Phase 2. A CDE leader added that appeals were set up as "different ways that [vendors] could continue to send us evidence if they felt that we missed something in their application." CDE had 14 working days to notify the vendor of receipt of the appeal and 30 working days to complete the appeals process and inform the vendor of the decision.

The appeals process was the same in both the 2020 and 2022 review cycles except that in the latter CDE required vendors to locate evidence in the original submission specific to sections rated "partially meets" or "doesn't meet" rather than allowing vendors to submit additional evidence with their appeal. One CDE leader commented on this change: "In the past, vendors just sent a flood of information" so the change helped to match the evidence provided with the area of appeal.

CDE Program Review Rubrics

CDE's process for reviewing vendors' Part II applications was grounded in two-phase rubrics that operationalized READ Act requirements and provided operational definitions of the five components of the science of reading. CDE developed separate rubrics for core, supplemental, and intervention programs.



3. CDE’s Processes for Reviewing Instructional Materials

This section first describes the rubrics that CDE used to review programs in the 2020 and 2022 review cycles. The section highlights how points were allocated for each rubric section, criteria used to determine whether programs were eligible to move from Phase 1 to Phase 2 reviews, and revisions made to rubrics between the 2020 and 2022 review cycles. The section then analyzes how the rubrics aligned with the research on the science of reading and the weighting of Part II, Phase 2 rubric components by grade level. The section concludes by summarizing reviewers’ feedback on the rubrics.

Description of Rubrics

Reviewers used CDE’s Phase 1 rubric to evaluate programs on key elements and features of scientifically based reading instruction as well as other criteria defined by the READ Act, such as whether programs met ESSA evidence levels. If a program met the criteria in Phase 1, reviewers then used the Phase 2 rubric to evaluate the extent to which programs implemented effective instructional practices for teaching essential early literacy skills relevant to each grade level (K–3). See Exhibit 3.5 for the list of rubrics used.

Exhibit 3.5. List of CDE’s Rubrics for Part II - Instructional Materials Review

Rubric	Description
Phase 1 Rubric	The Phase 1 rubric contains criteria for all programs. Programs need to meet criteria in the Phase 1 rubric to move to one of the three Phase 2 rubrics.
Phase 2 Rubric for Core and Supplemental Programs	The Phase 2 rubric for core and supplemental programs contains criteria for programs by grade level then by reading skills within each grade. Programs need to meet criteria in the Phase 2 rubric to qualify as an approved vendor.
Phase 2 Rubric for Intervention Programs	The Phase 2 rubric for intervention programs contain criteria for programs by reading skill, then by grades within each skill. Programs need to meet criteria in the Phase 2 rubric to qualify as an approved vendor.

Part II, Phase 1 Rubric

Part II, Phase 1 rubric criteria were the same for core, supplemental and intervention programs. The Phase 1 rubric consisted of six sections with between



3. CDE's Processes for Reviewing Instructional Materials

3 and 7 points in each section for a possible score of 25 points. See Exhibit 3.6 for the sections and points possible for each of the sections.

Exhibit 3.6. Sections of the Phase 1 Rubric

Six Sections of the Phase 1 Rubric	Points Possible
Research Alignment	5
Explicit Instruction	3
Sequential Instruction	3
Systematic and Cumulative Instruction	7
Coordinated Components	4
Related Elements	3
Total Points Possible	25

The minimum passing score for Phase 1 was 20, with a requirement of receiving all possible points in the Research Alignment section. The following lists the five indicators in the Research Alignment section, each with a requirement of 1 point. See Exhibit 3.7 for the indicators in the Research Alignment section and the points possible for each indicator.

Exhibit 3.7. Phase 1 Required Research Alignment Indicators

Research Alignment Indicator	Points Possible and Required
For grades for which the program is submitted, the program must include evidence of alignment with ESSA Tiers 1, 2, 3, or 4; if Tier 4, then a logic model must be submitted.	1
The program provides evidence of grounding in conceptual research and theoretical models with reference to research articles and websites. If the program is constructed for learning to read in a language other than English, a conceptual model and research foundation, as well as evidence that it is not merely a translation of an English program, should be provided.	1
There is an obvious emphasis on teaching and learning the five essential early literacy skills. This should be scored only for literacy components the vendor selected within the application.	1
The program reflects the understanding that reading is a language-based skill and learning to read depends on mapping sounds to print.	1
Word recognition is explicitly taught through relating sounds to letters, and not visual memory, guessing, the shape of the word, or the use of context clues to decode words.	1
Word recognition is explicitly taught through relating sounds to letters, and not visual memory, guessing, the shape of the word, or the use of context clues to decode words	1
Total Points Possible	5



3. CDE's Processes for Reviewing Instructional Materials

All vendors were to submit the grade-level focus of their program. The Phase 1 rubric referenced grade levels in each of the six sections, starting with Section 1 in which vendors provided evidence from ESSA Tiers 1, 2, 3, or 4 for the grade levels submitted. Core programs, therefore, needed evidence of all five reading components at each grade level. Supplemental or intervention programs needed evidence for their stated grade level or instructional emphasis. Examples in other Phase 1 sections relating to grade levels were instructional routines or scripts that relate back to grade-level outcomes and standards for explicit instruction, progression within and sequence across grade levels for sequential instruction, grade alignment for systematic and cumulative instruction, and procedures and language across grades for coordinated elements. Finally, specific criteria within the related elements section requested explicit links to state standards and grade-level expectations, with a crosswalk showing the links.

If the program met criteria for Phase 1, then reviewers moved to Phase 2. If the program did not pass Phase 1, then the review was complete. While reviewing criteria for Phase 1, reviewers used evidence that vendors provided for both Phase 1 and Phase 2. CDE also trained reviewers that if they found information while reviewing Phase 2 that did not support a passing rating in Phase 1, then the reviewers provided the specific evidence and changed the rating.

Part II, Phase 2 Rubrics: Core and Supplemental Programs

Phase 2 rubrics for **core and supplemental** programs had the same sections with slight differences across grades (Exhibit 3.8). Sections for kindergarten were phonological and phonemic awareness, phonics and word study, vocabulary, and listening comprehension. Sections for 1st grade were the same with the addition of text reading and fluency, the only grade level with five sections instead of four. The 1st grade also listed reading comprehension with listening. Sections for 2nd and 3rd grades were phonics and word study, vocabulary, text reading and fluency, and reading comprehension. The Phase 2



3. CDE’s Processes for Reviewing Instructional Materials

rubric also included noninstructional practice areas of usability and professional development topics across grade levels for the program as a whole.

Exhibit 3.8. Part II, Phase 2 Rubric Sections for Core and Supplemental Programs by Grade Level

Section	Section by Grade Level			
	Kindergarten	1st Grade	2 nd Grade	3rd Grade
Phonological and phonemic awareness	Yes	Yes	No	No
Phonics and word study	Yes	Yes	Yes	Yes
Vocabulary	Yes	Yes	Yes	Yes
Text reading and fluency	No	Yes	Yes	Yes
Listening and/or reading comprehension	Yes, listening	Yes, both*	Yes, reading	Yes, reading
Usability and professional development (noninstruction sections)	Yes	Yes	Yes	Yes

* This 1st-grade section was labeled “listening and reading comprehension” for core programs and “reading comprehension” in the rubric for supplemental programs.

To help with scoring in all sections, CDE provided reviewers with a description of ESSA evidence levels and examples of evidence for each rubric section that did and did not meet expectations.¹⁰ For example, in the research section, CDE described ESSA evidence levels, listed sample theoretical models, like the Simple View of Reading (Hoover & Gough, 1990) and Ehri’s Phases of Word Reading (Ehri & McCormick, 1998), as examples of evidence that would meet expectations, and shared examples of evidence that would not meet expectations, like the Three Cueing System (Adams, 1998).

For a grade level to be rated as “meets expectations,” all but one section needed to receive such a rating. That last section needed to be rated as “partially meets.” If more than one section was rated as “partially meets,” the Phase 2 rating for that grade level was “partially meets.” If any one section was rated as “doesn’t meet,” the grade level was also rated as such.

¹⁰ CDE refers to vendor-supplied evidence that would not meet expectations as nonexamples.



3. CDE’s Processes for Reviewing Instructional Materials

Core and supplemental programs that were rated as “meets expectations” could have a full recommendation or a partial recommendation for inclusion on the advisory list. A full recommendation meant that programs received a rating of “meets expectations” in all grade levels areas submitted for review. A partial recommendation meant that a grade level (core and supplemental) in all areas submitted and/or a component area (supplemental) received a rating of “meets expectations.” Supplemental programs could be recommended for inclusion on the advisory list by grade level and/or reading component.

Part II, Phase 2 Rubric: Intervention Programs

The Phase 2 rubric for **intervention** programs had five sections focused on reading components: phonological and phonemic awareness, phonics and word study, vocabulary, text reading and fluency, and listening and reading comprehension, and sections for usability and professional development sections. Different from core and supplemental program rubrics, reviewers provided ratings for grade levels within each topic area rather than for reading skills within each grade level. Exhibit 3.9 shows each reading skill section and the number of indicators in each section for grade levels.

Exhibit 3.9. Part II, Phase 2 Grade Level Indicators for Intervention Programs by Skill Sections

Topic Area	Number of indicators per grade level
Phonological and phonemic awareness	12 indicators: 6 for K-3, 1 for K only, 5 for K-1
Phonics and word study	26 indicators: 12 for K-3, 4 for 1-3, 4 for K-1, 2 for 2-3, 5 for 3 only
Vocabulary	13 indicators: 10 for K-3, 3 for 2-3
Text reading and fluency	6 indicators: 1 for K-3; 5 for 1-3
Comprehension	Listening comprehension: 9 indicators, 7 for K, 1 for K-3, 1 for K-1 Reading Comprehension: 13 indicators, 1 for K-1, 1 for 1 only, 7 for 1-3, 4 for 2-3
Usability and professional development (noninstructional sections)	1 indicator for usability and 1 indicator for professional development



3. CDE's Processes for Reviewing Instructional Materials

Just like for core and supplemental programs, CDE provided reviewers with a description of ESSA evidence levels. For example, an explanation for reviewers in Section 1, phonological and phonemic awareness, was that vendors include a description of activities for a detailed scope and sequence. CDE also provided examples of materials that met expectations (e.g., blending is taught before segmenting) and those that did not (e.g., blending syllables, manipulating phonemes, and segmenting compound words all appear in the same lesson).

Specific to Phase 2 rubric for intervention programs, reviewers considered the (a) individual items within the foundational skill section by grade, (b) completeness of the foundational skill section for continuity across grades, when applicable, and (c) program in its entirety when it covered more than one grade and/or more than one foundational skill area. An additional area for reviewers to consider while reviewing the intervention programs was “intensifying intervention” for each reading component. Specifically, when students move from a core reading program to an intervention reading program, there is more intensity in the instruction, such as in the frequent monitoring of student progress and providing explicit instruction in small steps.

Each topic area was scored as “meets expectations,” when there was evidence for the topic area, or “doesn’t meet” when evidence was missing from the vendor submission for the topic area or the evidence submitted was contrary to the criteria for topic area. Intervention programs could be recommended for inclusion on the Advisory List of Instructional Programming by full program, grade level, and/or reading component area. To be “fully recommended,” the program had to receive a rating of “meets expectations” for all grade levels and all reading components submitted for review. To be recommended by **grade level** for all areas submitted for review, the program had to receive a rating of “meets expectations” (80% of points possible) for the grade level submitted. To be recommended for inclusion on the advisory list by **reading component area** submitted for review, the program had to receive a rating of “meets expectations” (80% of points possible) for the submitted reading component area.



3. CDE's Processes for Reviewing Instructional Materials

Changes to Part II Review Rubrics Between 2020 and 2022

CDE modified the Part II, Phase 1 rubric between the 2020 to 2022 review cycles. For the 2022 review cycle, CDE added a requirement that programs meet all 5 possible points in the Research Alignment section to demonstrate they are aligned with evidence-based and scientifically based instructional practices. To reinforce this requirement, CDE added an instruction for reviewers about the need for earning all points in the research section: "To move forward, a program must be marked as MET in all criteria in section 1 as well as receive a score of 20 points or higher." In interviews about the addition of the requirement, one CDE leader reported instituting the change to make sure the "research base was solid right away" in the first section of Phase 1. A second CDE leader added that it was previously "more loose" and now, with all 5 points required, the research alignment was "nonnegotiable."

CDE did not change Part II, Phase 2 rubrics between the 2020 and 2022 review cycles.

How Rubrics Align with Research, Including Grade-Level Weighting

CDE based the development of the instructional program rubrics in empirical studies, reports, and scientific articles. Prior to the first review cycle in 2020, CDE hired Dr. Stephanie Stollar, an education consultant in reading and former vice president for Professional Learning at Acadience Learning Inc., to assist in developing the rubric.

The initial literature used to support rubric development was organized into the following categories: (a) the five scientifically based reading components, (b) explicit instruction, (c) sequential and sequenced instruction, (d) systematic instruction, (e) ineffective reading instruction, (f) cognitive science and reading, (g) special populations of those with learning disabilities in reading and English learners, (h) multitiered systems of support and response to intervention, (i) writing connected to reading, and (j) assessment for reading instruction. Each



3. CDE's Processes for Reviewing Instructional Materials

category and its respective research studies aligned with rubric topics and criteria that CDE created.

Analysis of Rubric Alignment with Additional Literature

The evaluation team reviewed the initial list and identified additional literature that addressed rubric components. This literature supported the approach taken by CDE in each category. In the category of special populations, CDE may wish to consider refining rubrics to address English learners specifically to differentiate and support readers who are struggling with reading and those who need acceleration. Section 5 of the Phase 1 rubric reads, “differentiation and support are provided for supporting English learners, students who are struggling, and those who need acceleration.”

Research on students who are English speakers struggling with reading as well as those whose first language is not English identifies similar difficulties in phonological awareness, word reading, and spelling among the two groups (August & Shanahan, 2006). At the same time, there is strong evidence for two approaches to assisting English learners: explicit teaching of vocabulary words using a variety of instructional activities intensively over several days and the integration of oral and written language in content areas (Baker et al., 2014). Programs that address considerations for English learners may help verify that the program has support for this population. A separate indicator for English learners would also allow for eliciting evidence needed for differentiating them from native English speakers who need extra support or enrichment.

Current topics such as advanced phonemic awareness and cultural representativeness were not addressed in the rubrics. Advanced phonemic awareness is described as skills that some consider more challenging than basic phonemic awareness skills such as phoneme segmentation and phoneme blending to include phoneme deletion or replacing medial sounds or sounds in blends (Kilpatrick & O'Brien, 2019). Current research has yet to support the need for advanced phonemic awareness training as a need for core, supplemental, or



3. CDE's Processes for Reviewing Instructional Materials

intervention reading programs to be deemed scientifically based (Brown et al., 2021; Clemens et al., 2021).

Cultural representativeness is defined as texts in an instructional program that are authored by individuals and include characters who reflect the ethnicities and cultures of students. In the 2021 review of cultural representativeness in vendor programs on the advisory list, 80% of vendors that provided a response to the question about cultural representation shared their commitment to diversity, but only 43% provided evidence that characters in their programs represent diverse populations, and 3% of vendors provided quantitative evidence of diverse representation within their programs. Furthermore, 56% of vendors did not answer the question of diverse authorship of texts in their programs.

Research supports having culturally representative texts. Representation positively contributes to child development (Hughes-Hassell & Cox, 2010). Children interact with text features while reading, and when texts are diverse, children engage in acceptance, exposure to different cultures, and perspectives that challenge their worldviews (Okoye-Johnson, 2011; Shachar, 2012). Publication of diverse texts have increased over the past two decades, as shown by the University of Wisconsin—Madison's Cooperative Children's Book Center, which has been tracking cultural representativeness since 1985. Recent statistics show gains in books about and by diverse authors (Dahlen, 2020). However, the numbers are still significantly below the population of children in school, whose demographics are majority Black, Indigenous, or other people of color (National Center for Education Statistics, 2022). A separate indicator for cultural representation can help CDE support the needs of children across the state in learning to read with texts that represent diverse characters and experiences.

In the Phase 2 rubric, the research supports having all five components of scientifically based reading instruction evaluated across grades, with phonological and phonemic awareness phasing out after 1st grade and text reading and fluency beginning in 1st grade. However, the research does not provide guidance for how much weight to assign to each component per grade in



3. CDE’s Processes for Reviewing Instructional Materials

a selection process or in terms of instructional time (Foorman et al., 2016; Foorman et al., 2017; Hall et al., 2022; Shanahan et al., 2010). Currently, literature from the special education field suggests an emphasis on specific reading components by grade level (e.g., decoding in second grade; listening comprehension in third grade) may help skill deficits (Berkeley et al, 2010; Connor et al., 2014). Future investigation can consider to what extent the relative weights align with the needs of those with reading deficiencies, by grade level.

Exhibit 3.10 shows weights of the five components for core programs by grade level in Phase 2. There was variability in how much weight was given to each component. However, the percentage of emphasis on the rubric does not indicate the amount of time an instructional program spent on that reading component.

Exhibit 3.10. Weights of Components for Core Programs, by Grade Level

Component of Reading Instruction	Kindergarten – Points out of 55	1st Grade – Points out of 58	2nd Grade – Points out of 49	3rd Grade – Points out of 52	Total Program Points
Phonological and phonemic awareness	12 (22%)	11 (19%)	0 (0%)	0 (0%)	
Phonics and word study	23 (42%)	18 (32%)	18 (37%)	18 (35%)	
Text reading and fluency	0 (0%)	6 (10%)	6 (12%)	6 (11%)	
Vocabulary	11 (20%)	10 (17%)	13 (27%)	14 (27%)	
Listening and reading comprehension	9 (16%)	13 (22%)	12 (24%)	14 (27%)	
Total points per grade level	55 (100%)	58 (100%)	49 (100%)	52 (100%)	
Usability					5 (100%)
Professional development					2 (100%)

Note: Supplemental and intervention programs were not included because they may only focus on one or two reading components.



3. CDE's Processes for Reviewing Instructional Materials

Reviewer Feedback on Rubrics

In addition to examining the research alignment of CDE's rubrics, reviewers' feedback on rubrics was solicited. Overall, reviewers reported that the rubrics were thorough, clear, and comprehensible. Most reviewers said that the detail that CDE provided was necessary and helpful.

All reviewers had to demonstrate expertise in scientifically based reading instruction. Reviewers acknowledged that bringing background knowledge in the science of reading was essential to being able to apply the rubric.

Reviewers made few suggestions for updating rubrics, and there was no broad consensus for making changes. One reviewer recommended that all rubric sections include examples. Another reviewer said that some sections' criteria were very discrete, which made these criteria difficult to score. This reviewer recommended that such items could be consolidated.

In general, reviewers found the ease of use of the rubric to be "just right." One reviewer saw usability as difficult to rate, but another found it too easy.

Individual reviewers reported criteria for unique sections or areas that were too easy or difficult for vendors to meet. For example, reviewers stated that

- the review would be more rigorous if a program was automatically disqualified if it included "ineffective practices," such as guessing strategies;
- criteria for differentiation for English learners, struggling readers, and advanced readers were too easy to meet in Phase 1 because the three groups with very different needs were grouped together; and

“So if you're asking about the quality of the rubric, I would say it was thorough... We knew specifically from grade to grade and from element to element what we were looking for and whether it was easily identified within the [program].”
— Reviewer from 2022



3. CDE's Processes for Reviewing Instructional Materials

- criteria for spelling and the progression for phonological awareness were too easy for vendors to meet.

A minority of reviewers speculated that some vendors did not seem to understand the rubric, and that it was easier to score when vendors exhibited an understanding of the rubric. Overall, most reviewers agreed that the two-phase process was needed and that both rubrics helped to make accurate selections of programs for the advisory list.

“ Let’s see. Usability was super easy to do, but I wouldn’t cut it, because I think there is great value in providing materials that have ease of use for teachers. So that was an easy part, but I wouldn’t cut it.”
— Reviewer about ease of rubric criteria

Reviewer Recruitment and Selection

This section describes CDE’s recruitment for reviewer applications, the content of reviewer applications, reviewer selection, and similarities and differences between the 2020 and 2022 review cycles. The evaluation team interviewed CDE staff and reviewers and analyzed outreach materials (website and emails), the reviewer application, and submitted applications for each of the two rounds of review (2020 and 2022).

CDE Recruitment of Reviewers

CDE reported it was looking for reviewers with an understanding of the science of reading, who had district leadership roles and expertise in the field, and who had ideally done work in early reading instruction in the past. CDE also targeted “typical areas” in the state, like districts with a high English learner population to help find reviewers to review Spanish programs as well as English programs. Important to the CDE leadership team was ensuring that candidates selected as reviewers reflected the diversity of districts across the state including rural, urban, and suburban districts.



3. CDE's Processes for Reviewing Instructional Materials

CDE reached out to prospective reviewers through direct email, a listserv, and CDE's newsletter "The Scoop." CDE leaders shared that the listserv was specific to the READ Act and already included interested stakeholders from advocacy groups, districts, past CDE employees, consultants, and others in the field interested in the READ Act. The newsletter reached those inside and outside of Colorado who signed up for it.

Interviewees confirmed CDE's process, saying that they received an email from CDE inviting participation, found the opportunity on the CDE website, or saw the opportunity in the newsletter. CDE leaders shared that they "had a really long window of time to request reviewers," sending out requests multiple times, knowing that messages can "get lost in an inbox [and] people don't necessarily know their availability right away."

CDE's recruitment process was the same for both the 2020 and 2022 review cycles.

Reviewer Applications

For the 2020 and 2022 review cycles, the reviewer application asked why applicants were interested in participating in the review; for a description of their education and continued professional development in literacy; what they considered the most and least important characteristics of instructional programs; what they considered the most important qualities in a reviewer; whether they could review English programs, Spanish programs, or both; dates they were available; and for a resume.

The 2022 application added three questions about applicants' knowledge and experiences: education and continued professional development in

"We [looked] for somebody that really has had some expertise in the field. And we've gathered a pool of people that we know have been engaged and done this work in the past. So we... have a known group of people that we'll pull from. But then also we're always looking for those... that have representative populations and [rural and urban] representation." — CDE leader on reviewer recruitment



3. CDE's Processes for Reviewing Instructional Materials

evidence-based and scientifically based reading practices; experience in reviewing instructional programming materials in K–3 reading (e.g., district curriculum review committees, state curriculum reviews, national curriculum reviews); and understanding of tiered ESSA evidence levels as they relate to early literacy. The lack of understanding of ESSA evidence levels did not disqualify prospective reviewers.

Questions in 2022 also clarified what it meant to review programs in languages other than English. Reviewers for additional languages had to be able to read, speak, and have knowledge on reading instruction for the language indicated.

Most interviewees reported that the application process was thorough or had no feedback about the application. One interviewee described the process as “long and involved.”

Interviewees reported on reasons for applying, which included inputs into the process like representing an underrepresented region in the state, “giving back to the profession,” and advocating for readers in their community, or to gain understanding of the review process for communication about the process of program selection with their districts.

Reviewer Selection

CDE engaged a small internal application review team in carefully vetting whether applicants demonstrated an understanding of the science of reading and the developmental progression of early reading skills from kindergarten to 3rd grade. Both CDE and external interviewees emphasized the importance of reviewers having a science of reading background. CDE viewed this background as necessary for reviewers to successfully apply the rubric criteria to vendors' evidence. As one CDE leader explained,

“We know that our universities don't always know, our teachers don't always know... We don't want to start from the place of having to train somebody up on what the science of reading is. We need to know that



3. CDE’s Processes for Reviewing Instructional Materials

they’re already doing it, living it, and implementing [it]. So just that team of experts is something loosely defined, and if that means it might not be a PhD literacy researcher but a kindergarten teacher that’s been doing this for 20 years and understands how to teach phonics, or has lots of different experience with lots of different curricula.... I think we often get asked, like, ‘Are you using reading researchers or scientists?’ And it’s, like, no, we’re not. We’re using people that actually understand the science of reading and that we know what they do.”

In 2020, 44 reviewers participated in the review cycle, while 38 participated in 2022. Exhibit 3.11 shows external applications submitted, external reviewers selected, CDE internal reviewers invited and accepted, and total number of reviewers. The CDE leadership team reported that three 2022 applicants were affiliated with vendors who sell reading programs and were therefore not accepted as reviewers due to a potential conflict of interest.

Exhibit 3.11. Reviewer Application and Selection

Year	External Applications Submitted	External Reviewers Selected	CDE Internal Reviewers Invited and Accepted	Total Reviewers
2020	33	33 (2 did not participate)	13	44
2022	33	30 (7 did not participate)	15	38

In reporting challenges in reviewer selection, CDE said that although it had a pool of candidates and recruited across the state, it still struggled with finding an ideal pool of review candidates: “We’re always looking for those that have representative populations and district representation. That said, we always struggle. Always, always struggle with finding folks that really can accurately understand curricula in Spanish and English. So basically we try really hard in those areas, but there is a significant lack of expertise [and] a dance between who actually has the time to complete the reviews.”



3. CDE's Processes for Reviewing Instructional Materials

Reviewer Training Process and Content

CDE's training for reviewers included training materials to be used for reviews, an introductory webinar, and a training session the morning that the reviews began.

This section describes CDE's training process and content, including reviewer feedback about the training. Evaluators interviewed CDE staff and reviewers and analyzed the training timeline and training materials (slides and guidance document for reviewers).

Session 1

CDE offered two opportunities for Session 1 training for reviewers. The first was a webinar that CDE recorded and sent to reviewers along with access to materials. CDE offered the same webinar live for interested reviewers 2 days later. In the webinar, CDE first presented the READ Act Statute and Rule guiding the review and reviewer responsibilities of commitment to a fair and unbiased process, maintaining confidentiality, and attesting to not having a conflict of interest. Second, CDE described core, supplemental, and intervention programming along with the vendor application process. Finally, CDE spent most of the training on introducing materials for the review process which included locations of the Part II Program Review vendor applications, three rubrics (core, supplemental, and intervention), and scoring guides for the rubrics.

CDE sent reviewers access to the materials along with the recorded webinar then walked through how to access materials in the live webinar session. These materials included rubrics for core, supplemental, and intervention programs. CDE also provided three scoring guides. The scoring guide for the rubric provided reviewers with a description of Phase 1 and Phase 2 reviews as well as criteria for assignment of points for full versus partial recommendation in each phase. The scoring guide for core and supplemental programs listed instructions for reviewers, Phase 1 required features of evidence-based or scientifically based core reading programs, and Phase 2 required instructional



3. CDE's Processes for Reviewing Instructional Materials

practices for teaching essential early literacy skills in each of kindergarten through 3rd grade. The scoring guide for the intervention program review described characteristics of intervention instruction, intervention design and delivery, the degree of instructional intensity in each of the five reading components, how to score programs using the rubric, and scoring steps, from reading the cover page to rating programs and submitting a final scoring summary.

Session 2

CDE planned the second training for the first 90 minutes of a full-day schedule from 8:30 a.m. to 4:00 p.m., with the rest of the day set aside for teams to start reviewing instructional programs. First, CDE presented an overview of the process, which included the scope of work for reviewers. CDE explained that when programs met expectations for Phase 1 then they moved to a Phase 2 review. CDE also explained how programs were assigned to review teams and the submission procedure when teams completed each rubric. Next, CDE presented instructions for reviewers along with a practice review. Instructions included the four items needed to start reviews: vendor application, appropriate rubric, scoring guide, and folder for the program being reviewed. During practice, reviewers walked through each of the five review steps:

1. Reviewing the application for type of program, target audience, reading components being reviewed if the program type was supplemental or intervention, and required verification.
2. Identifying sections and evaluation criteria, which included the review of the rubric and scoring guide and expectations in Phase 1 criteria.
3. Rating of each indicator, which involved assigning points in each row within a section, using the Scoring Guidance Document, providing evidence or feedback for all items marked "partially meets" or "doesn't meet," and citing concrete examples.



3. CDE's Processes for Reviewing Instructional Materials

4. Determining the score for each section, which involved identifying where totals from each section autopopulated in the rubric, checking scores, and adding general comments where applicable.
5. Completing the final summary, which included documentation of the program name, publisher, and year; type of program; review team identifiers; Phase 1 and Phase 2 areas that autopopulate and what rubric areas CDE completed after team submission.

“ Within the training, we do try to do a little bit of [practice] on what an objective versus a subjective statement might be so that we can keep our reviewers objective rather than subjective. So we spend a lot of time on how to write an objective statement. What does that look like?”

— CDE trainer

Finally, the CDE team presented reminders to review teams about scoring guides, deletion of vendor materials, and what happens with CDE review tools when the process is complete. CDE placed all training attendees into assigned teams to plan team logistics for the rest of the day and for subsequent review meetings over the next 10 days.

Both CDE and reviewers reported that training focused on how to score applications using the rubrics. CDE's training materials described that Phase 1 rubrics identified whether the program is aligned with evidence-based and scientifically based practices and that Phase 2 reviewed for program rigor and ability to meet the needs of students with significant reading deficiencies. The training did not focus on the science of reading background because CDE only selected reviewers with a background in the science of reading.

Most external reviewers reported that the training walked through the rubric and examples of evidence to meet criteria were helpful.¹¹ About half of reviewers reported that the 2-hour training on the rubric and scoring guide was

¹¹ The interviews occurred nearly 11 months after reviewers participated in the training. Most provided general impressions rather than detailed accounts of the training.



3. CDE's Processes for Reviewing Instructional Materials

“just right.” One respondent indicated that outlining the big picture during the large group presentation was helpful. The rest expressed a desire to have attended a longer, more in-depth large group training. An experienced reviewer commented that experience of the reviewer mattered and that less experienced reviewers may have needed more training time.

Some external reviewers also viewed program review activities as training because of the coaching support of CDE staff. Several respondents stated that the CDE representative in reviews supported both the review process and learning during reviews. One reviewer called the reviews “training by doing.”

The Program Review Structure and Processes for Assigning Points

Review teams consisted of three to four members, which included a CDE representative who coached and supported the team. Each team reviewed between three and seven programs and used consensus decision-making to resolve disagreements about points. Reviewers resolved disagreements on borderline cases, and most agreed that quality and quantity of vendor materials varied. CDE coaches met separately to share team processes and problem-solve disagreements about scores that emerged during team consensus decision-making.

This section describes review team structures and how reviewers assigned points. Evaluators reviewed CDE's program review materials and interviewed CDE staff and reviewers for information about the program review process.

Review Team Structure

CDE assigned teams between three and seven programs to review, and review teams began reviewing programs shortly after the training ended. Teams reviewing core programs had fewer (three to four) programs to review because, on average, core programs took more time to review because teams reviewed for each of four grade levels as well as all five components of scientifically based



3. CDE's Processes for Reviewing Instructional Materials

reading. Supplemental and intervention programs did not take as long to review because of the targeting of one or more components of scientifically based reading, and intervention programs focused across grade levels rather than on each grade level.

CDE organized program review teams with three to four members, with each team including a reviewer from CDE to “always have staff involved [and reviewers to] have support during the process” per one CDE leader. In some groups, those with more experience (e.g., veteran curriculum directors and those who participated in the previous review cycle) led the team and consulted with the CDE reviewer when needed. CDE leaders reported that the CDE staff member “intentionally scaffolded” most groups as a coach. In some cases, coaches facilitated a gradual release of responsibility where the team worked on a program review together with the CDE lead, then worked in pairs and individually before returning to the team to report scores. In other cases, reviewers needed more support. One CDE lead commented on providing more focused support: “If this person is feeling like they really can’t finish this review we’re going to partner with them versus gradual release for those that can do it on their own.”

One CDE leader said that CDE staff met separately not only on consensus and problem-solving but also to review team processes: “We talk about how each of the small groups are going. Is this person not getting it? What are you doing for it? And we try to stay reliable with each other in the way that we’re coaching those small teams so that we constantly have an eye on the support that those reviewers are getting.” This internal process, they reported, helped with reliability of scoring across teams.

Consistent with CDE’s planned design, reviewers reported in interviews that they received guidance, support, and feedback as a group during the review process and that each team included at least one CDE staff member. Specifically, they reported that the CDE staff member in their group answered questions or quickly sought answers from the internal CDE team when answers



3. CDE's Processes for Reviewing Instructional Materials

were unknown, for example to clarify definitions and locate certain evidence in vendor materials.

Reviewers reported that all teams used CDE's program review rubrics and vendor-submitted materials to find evidence for rubric criteria and assign points in the program review. All reviewers reported that they also evaluated programs with Phase 1 of the rubric, then moved to Phase 2 when vendors received at least minimum points in Phase 1. There was one exception for supplemental programs. If a supplemental program was part of a core program that a different set of reviewers rated for Phase 1, a set of reviewers who were charged with reviewing a supplemental program rated it for Phase 2.

Interviewees said that reviewing digitally, for the most part, worked well. Seven reviewers interviewed reviewed both paper (2020 review cycle) and digital (2022 review cycle) formats, and some described similarities and differences in the formats. Similarities were the receipt of packets of materials from vendors and the decision-making process for whether submitted materials met the criteria. A difference was that reviewers were unable to manipulate print materials and look for examples beyond what vendors submitted for an item, while sometimes they could with digital submissions, although access was still limited to what the vendor provided for the rubric item.

Review Team Process of Assigning Points

Reviewers assigned points for each item within rubric sections. They described looking for evidence within vendor applications to determine the rubric rating. In most teams, reviewers looked at evidence together. Work varied in some teams, like being divided by grade level, or all reviewers rated each item and came to a consensus before moving to the next item. Several reviewers described writing extensive comments to justify scores, especially for items that did not receive the highest ratings.

Consensus process. CDE leaders explained that there was no formal process of interrater reliability when assigning points. Rather, teams came to



3. CDE's Processes for Reviewing Instructional Materials

consensus and, as a quality check, CDE staff from each team provided regular team reports to other CDE staff on the internal leadership team. The CDE team also completed a final check of review teams' scoring sheets, evidence indicators, and comments written by reviewers. One leader described expecting that a review team would "come together and meet and go off on [its] own," but one especially "tricky" reviewer needed a lot of support, so the leader brought the problem to the internal team to collaboratively problem-solve.

One CDE leader explained that the process of internal checks on the leadership team was one CDE created but had not yet codified but helped with resolving disagreements during the consensus process. Another stated, "I'm going to just say that that is an area that we need to evolve. We don't have a scored interrater reliability process. We just don't. We have the training, the work with CDE staff, the consistent reviewing from outside of the review. But we don't have a formal reliability process."

Reviewers concurred that they did not get trained on a formal calibration process and instead used a process of consensus among team members during reviews. When disagreements arose, reviewers noted two strategies for resolving them. The first strategy was to talk through disagreements by rereading the rubric and looking at the evidence that vendors provided. Most reviewers found conversations to be adequate for resolving disagreements. The second strategy was to ask the CDE team member for support. If a CDE team member was unable to resolve the disagreement, the team turned to the internal CDE leadership team to discuss the rubric and evidence for resolution.

One reviewer said that the expectation was to come to a consensus so "it wasn't based on individual scoring." Another reviewer talked about consensus steps:

"We talked through and shared our rationale of why we believed something should or should not have a point.... Everyone on the team was experienced when it came to literacy curriculum and instruction. So we never really ran into a point where we said, 'I can't give that point' or 'I



3. CDE's Processes for Reviewing Instructional Materials

must give that point.' It was all very collaborative and we could reach consensus well. We worked well together as a team."

Another reviewer explained that when questions arose, they could "work their way through the notes" and that there were individuals on the team with more experience who could guide and say, "in the past we've done this," which helped with consensus building. The reviewer commented that the rubric helped in the consensus discussions because it was "thorough."

A reviewer of supplemental programs said that the team's consensus process was to "take time away" before returning to the team to discuss discrepant cases: "We were able to step away, work individually and then come back together and share, compare, and then again look at one another's work and [ask], "Why did you give that point?" ... Or, "I give that [1] point. Why didn't you?" [then] share some of the evidence.... that helped our workflow [to] work individually and come back together and then reach consensus." This reviewer added that this "step away" process for consensus may have been easier for the team because the focus was supplemental programs rather than core. "One of them," the reviewer said, "was just a vocabulary program... It didn't take as long. I would imagine that if we were doing a core program, it would require a lot more time." Interviewees who reviewed both core and supplemental or intervention programs confirmed that core program reviews were more in-depth because submissions were more comprehensive than for supplemental or intervention programs.

Determination of points each section was worth. Most reviewers who described assigning points at the section level explained that scores for sections were calculated automatically in the score sheet. This is consistent with the explanation that CDE provided during the reviewer training. Reviewers reported viewing the automatically populated score then, if the section rating did not align with the team's perceptions of whether a program met, partially met, or did not meet expectations, the team reexamined then adjusted individual item scores by consensus. Another reviewer described a situation in which the team was



3. CDE's Processes for Reviewing Instructional Materials

“constantly dissenting” on some individual items, where the team conducted rounds of reading the rubric element aloud and going back into the materials to show fellow reviewers where the element could or could not be found in the materials. Two reviewers perceived that rating core programs grade by grade on all items led to inconsistent results across grade levels.

Ratings for borderline cases. About half of reviewers believed that the application of the rubric was straightforward and that decisions for assigning points were clear. The other half of reviewers noted that borderline cases did present some challenges. One reviewer noted that the team spent more time discussing borderline cases than other cases and another noted that “if an item was not present” a submission was considered borderline. One reviewer noted simply that “borderline cases were difficult.” Two reviewers said that borderline cases resulted in a lack of consensus. Another noted that the team sent two programs forward even though there was disagreement about one of them.

Sufficient evidence from vendors for scoring. Reviewers reported that both the quality and the quantity of what vendors submitted varied. Reviewers believed some programs but not others provided sufficient evidence. Reviewers reported that they both looked at vendors’ answers to questions and verified those answers with program materials provided by vendors.

Most reviewers reported that they would have preferred to have access to all materials associated with the programs. They took their task of verifying vendor statements with actual examples from the programs very seriously. One reviewer expressed that “given the stakes of the review, access to more extensive materials is warranted.”

One reviewer said that if a vendor stated that there was evidence in a section for an item, the reviewer would go to that section to verify the evidence. The reviewer expressed being “very cautious” not to say “meets expectations” if the evidence was not entirely there. Sometimes an application would indicate that the program provided evidence but reviewers did not have access to the resources. In cases with limited evidence or access, one reviewer stated, “our



3. CDE's Processes for Reviewing Instructional Materials

scores went to the 'partially [meets].' And then the note was the program indicates that there are several resources and we have access to very few, or something like that." One interviewee suggested that vendors may not have their content experts completing the applications, which may result in vendors not submitting the appropriate materials, adding, "[the vendors] aren't deeply steeped in the science of reading, which is a problem."

Summary of Findings

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. Reviewers were well qualified and well trained. 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. Challenges included vendor material not always matching with evidence needed and reviewers needing to search for evidence, a lack of indicators to provide evidence for English learners and cultural responsiveness, and a need for a consistent consensus approach across review teams.



4

CDE's Processes for Reviewing Spanish-Language Instructional Programs

CDE reviews Spanish-language core, supplemental, and intervention programs. Chapter 2 presented findings of the secondary review of programs added to Advisory Lists as well as those that passed Phase 1, but not Phase 2, of CDE's review.

This chapter presents an analysis of CDE's process for identifying Spanish-language instructional programs for its READ Act Advisory Lists.

Key findings

- In 2020 and 2022, CDE used the same processes and rubrics for identifying Spanish-language instructional programs as they did for English-language instructional programs.
- CDE's processes are aligned with the READ Act.
 - Phase 1 of CDE's selection process can reasonably be applied to both Spanish- and English-language programs.
 - Phase 2 of CDE's selection process includes criteria that are not appropriate for teaching foundational reading skills in Spanish.



4. Spanish-Language Instructional Programs

Introduction

This section summarizes findings related to CDE’s selection of Spanish-language instructional programs. This analysis was conducted by examining the following review rubrics: core program Part II, supplemental program Part II, and intervention program Part II. The evaluation is also based on interviews conducted with reviewers who have specialized backgrounds in Spanish-language reading development and instruction as well as deep grounding in scientifically based reading instruction as defined in the READ Act. The evaluation team’s analysis for this section focused primarily on reviewers’ comments about the core program review rubric and their experience using this rubric to review Spanish-language literacy materials.

Summary of General Issues in Applicability of Rubrics to Spanish Literacy Materials

The writing systems of English and Spanish are both based on the Roman alphabet, meaning sounds are associated with particular letters or combinations of letters; thus, early literacy instruction in alphabets is similar for both languages. In particular, letter recognition, sound-symbol correspondence, and phonemic awareness form the initial building blocks for decoding in both languages. Research has shown that, while Spanish- and English-speaking children acquire these skills in similar ways (e.g., Anthony et al., 2011; Branum-Martin et al., 2006; Pollard-Durodola, 2009), differences between the two languages and their alphabets have important implications for how children learn to read (Goldenberg et al., 2014; Jiménez González & Ortíz González, 2000).

Consider the relationship between reading fluency and reading comprehension. For students learning English, reading fluency is widely recognized as being predictive of reading comprehension (e.g., Fuchs et al., 2001; Jenkins et al., 2009; Kim et al., 2010). For students learning Spanish, however, research has shown that reading accuracy (a crucial component of



4. Spanish-Language Instructional Programs

fluency) is a poor indicator of reading comprehension (López-Escribano et al., 2013). The reason for this important difference is related to the nature of the languages themselves. English has 26 letters in the alphabet and 44 phonemes or sounds, whereas Spanish has 27 letters and only 22–24 phonemes. In general, the sound-symbol correspondences in Spanish are consistent, meaning that once students master those, they can decode most words fairly well. As a result, students learning Spanish can often decode text well beyond the level at which they have good comprehension of what they are reading, a pattern not seen in English literacy development where sound-symbol correspondences are not always predictable and decoding is therefore more challenging. Linguistic nuances, as in this example, necessitate a different focus on certain elements of early literacy instruction in Spanish, such as sound-symbol correspondences, word learning, and early reading practices.

Due to the nature of these differences, a rubric designed to evaluate materials for early literacy instruction in English would need to be adapted to effectively evaluate materials for early literacy instruction in Spanish. An overview of these rubric elements and the related applicability issues is provided below, with related recommendations included in Chapter 5.

Overview of Applicability Issues in Rubrics

Exhibit 4.1 provides an overview of issues related to applicability of an early literacy rubric based on instruction in English to materials designed for early literacy instruction in Spanish. Specific criteria (rows) affected by these issues are highlighted in attached versions of each rubric. (Note that these may not be the only criteria that may be affected—see the Recommendations in Chapter 5 on conducting a review of rubrics by specialists in early literacy instruction in Spanish.)



4. Spanish-Language Instructional Programs

Exhibit 4.1. Overview of Potential Applicability Issues of English-language Review Rubrics¹²

Grade Levels	Section of Core Program Rubric	Comments
K-1	Phonological and phonemic awareness	<ul style="list-style-type: none">• Due to the phonemic nature of the Spanish alphabet, and the syllable and word structure of Spanish, these early literacy skills in Spanish are taught at the syllable level rather than the individual phoneme level; students learn to isolate and manipulate phonemes by manipulating sounds in syllables rather than blending and segmenting individual sounds in words, as in English literacy instruction.• Since sound-symbol correspondences in Spanish are very regular and phonemic awareness is taught at the syllable level, the scope and sequence of phonological and phonemic skills in Spanish literacy instruction is different from that in English.
K-3	Phonics and word study	<ul style="list-style-type: none">• Vowel sounds and syllable patterns in Spanish are different from those in English.• High-utility letters, spelling patterns, and words may be different in Spanish and English.• Easily confused letters, letter-sounds, and words may be different in Spanish and English.• There are fewer irregular words with unpredictable sound-letter correspondences in Spanish than in English, so skills related to recognizing irregular words can be less of a focus in Spanish.• As noted in the Phonological and Phonemic Awareness section, these early literacy skills in Spanish are taught at the syllable level rather than the individual phoneme level.

¹² This table draws on information provided by the following resources: A comparison of initial literacy development in Spanish and in English (<https://www.teachingforbilitery.com/wp-content/uploads/2014/09/Comparing-initial-literacy-in-English-and-Spanish.pdf>); Early Literacy Instruction in Spanish: Teaching the Beginning Reader (<https://www.colorincolorado.org/article/early-literacy-instruction-spanish-teaching-beginning-reader>).



4. Spanish-Language Instructional Programs

Grade Levels	Section of Core Program Rubric	Comments
K–3	Vocabulary	<ul style="list-style-type: none">The vocabulary criteria are general enough to apply to both Spanish and English instruction.
K	Listening comprehension	<ul style="list-style-type: none">The listening comprehension criteria are general enough to apply to both Spanish and English instruction.
1–3	Text reading and fluency	<ul style="list-style-type: none">Syllable patterns in Spanish are different from those in English (relevant to criteria in 1st Grade only).
1	Listening and reading comprehension	<ul style="list-style-type: none">Syllable patterns, sounds, spelling patterns, and high-utility words in Spanish are different from those in English.
2–3	Reading comprehension	<ul style="list-style-type: none">The reading comprehension criteria are general enough to apply to both Spanish and English instruction.



5

Conclusion and Recommendations

This chapter summarizes findings about the alignment of materials on the Advisory Lists of Professional Development and Instructional Programming and CDE processes for selecting materials for these lists. It also outlines recommendations for further improving CDE's selection processes.

Key findings

- **Materials CDE approved for use with READ Act funds meet minimum requirements of SB 19-199.**
- **In an overall summary rating, 22 instructional programs received a rating of “fully met,” 13 received a rating of “largely met,” and 1 received a rating of “partially met.”**
- **CDE’s rubrics and processes for vetting instructional programs are aligned with the provisions of the READ Act and are grounded in evidence-based practices for teaching reading.**
 - Phase 1 of CDE’s selection process can reasonably be applied to both Spanish- and English-language programs.
 - Phase 2 of CDE’s selection process includes criteria that are not appropriate for teaching foundational reading skills in Spanish.



Conclusions and Recommendations

Summary of Findings

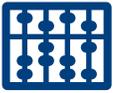
By and large, **instructional programs on the advisory list meet minimum requirements of SB 19-199** (Exhibit 5.1).

Exhibit 5.1. Summary Rating for Instructional Programs

Program	Rating			
	Fully meets	Largely meets	Partially meets	Does not meet / not rated
Instructional programs	22	13	1	0

CDE’s two-phase process for reviewing instructional programs is grounded in the science of reading and provides a rigorous and reliable way to vet these programs. In the evaluation team’s review of approved instructional programs, all but two programs met the minimum threshold for evidence—a clear logic model rooted in the science of reading; this suggests that all programs have the potential to make a positive impact on students’ reading outcomes. All but one program demonstrated the presence of skill development in phonemic awareness, phonics, reading fluency, and reading comprehension (as applicable), with 36 programs meeting criteria for explicit and systematic skill development. All core programs met 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.

In its review of the Advisory List of Instructional Programming, the evaluation team identified several cross-cutting findings. First, as was true with the 2021 evaluation, **the advisory list offered clear guidance about explicit and systematic instruction in the elements of scientifically based reading instruction.** Maintaining a clear and consistent focus on these elements over time, while continuing to integrate new empirical research about K–3 reading development, promises to support educators in creating classroom environments that reflect the most up-to-date science of reading.



Conclusions and Recommendations

Second, again consistent with the 2021 evaluation, **vendors submitted empirical evidence or promise of effectiveness to CDE that varied significantly in type and quality.** Only a handful of instructional program vendors submitted high-quality randomized controlled trials (ESSA Tier 1) and quasi-experimental studies (ESSA Tier 2). In contrast, some vendors that submitted evidence to claim ESSA Tier 3 included research summaries or data tables that lacked contextual information or interpretation, case studies from a small number of observations, or research that relied on poorly designed empirical studies.

Third, **the instructional program review process reflected requirements of READ Act legislation.** Each step of CDE's vetting process emphasized the components of scientifically based reading instruction. Vendors at both the Letter of Intent (Part I) and Application (Part II) stages were required to submit evidence that their approaches to reading instruction were aligned with the five components of reading instruction named in the READ Act. The Phase 1 and Phase 2 rubric criteria for each grade level were rooted in evidence, were clearly specified, and reflected both historical and current understandings of how the science of reading could be applied effectively in classroom practice. Reviewers in both years were well qualified, demonstrating an understanding of and educational experience implementing scientifically based reading instruction, and well trained to apply CDE's explicit criteria focused on scientifically based reading instruction. 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.

Fourth, **while the instructional review process was well aligned with the READ Act and reviewers found most aspects of the process to be well run and supportive, some reviewers did note a few challenges.**

- The training process involved a webinar, with reviewers receiving a recording of the webinar and all materials prior to trainings, as well as the opportunity to attend a live session of the same webinar. Most reviewers said that the



Conclusions and Recommendations

training was “just right,” but a few wished it had been longer. CDE leaders discussed different supports needed for reviewers, and experienced reviewers suggested that those not as experienced may need more support.

- Some reviewers found it challenging to find the right evidence for each area, with one suggesting that vendors may not have content experts supporting application submission.
- Some reviewers believed that their colleagues in districts did not understand processes that CDE uses to vet and recommend instructional materials.
- Some reviewers noted that the Part II, Phase 1 criterion “Differentiation and support are provided for supporting English learners, students who are struggling, and those who need acceleration” was difficult to score because it combines supports for students with disparate needs.

Fifth, **CDE was successful in identifying Spanish-language materials.**

The materials selected in 2022 included the required components of scientifically based reading that met the evaluation’s thresholds. However, the evaluation identified a significant challenge for reviewing Spanish-language programs: reviewers were required to use rubrics and scoring guides for English-language programs. Several Phase 2 criteria were defined in ways specific to English-language reading instruction and were not applicable to Spanish-language reading instruction.

Finally, in 2020, **the evaluation revealed that CDE was less successful in identifying instructional materials in English that offer differentiated support for English learners; this was especially true of intervention and supplemental programs.** While 2022 programs were not reviewed for this criterion, several reviewers from outside CDE mentioned that one of the most difficult Phase 1 criteria to score was the one on differentiation. These reviewers noted that English learners, students with learning differences, and students who require advanced learning opportunities have different needs.



Recommendations

Recommendation 1: Keep rubrics and the two-phase process for reviewing English-language programs the same. CDE developed its rubrics based on the current science of reading literature, and updated literature also supports items and sections within rubrics. Almost all reviewers stated that both Phase 1 and Phase 2 rubrics were needed. In Phase 1, reviewers evaluated programs on key elements and features of scientifically based reading instruction. If the vendor met criteria in Phase 1, reviewers then used the Phase 2 rubric to evaluate the extent to which programs implemented effective instructional practices for teaching the essential early literacy skills relevant to each grade level (K–3). The rubrics and two-phase review process are sound. They should be kept the same.

Recommendation 2: Develop and list separate rubric indicators in the Phase 1 rubric for meeting the needs of English learners and consider adding indicators for cultural representativeness. Research on both students who are struggling with English as their first language and English learners show similar difficulties with phonological awareness, word reading, and spelling (August & Shanahan, 2006). At the same time, there is strong evidence for two approaches to assisting English learners: explicit teaching of vocabulary words using a variety of instructional activities intensively over several days and the integration of oral and written language in content areas (Baker et al., 2014). English learners need additional instructional supports to meet grade-level proficiency levels in reading as well as their academic subjects; few programs offered the kinds of differentiated support necessary to help English learners successfully learn to read. A separate indicator of English learners would help vendors provide evidence for this need. CDE should enlist a panel of experts to support the development of such criteria. We recommend that experts include people with the following qualifications:



Conclusions and Recommendations

- Knowledge of theory and practice of teaching English early literacy skills to English learners in the United States.
- Experience teaching English early literacy skills and/or researching and writing about early literacy skill development, especially for English learners in the U.S. Research and writing could include: conducting original research or research syntheses or publishing findings in a peer-reviewed journal or in a book by a trusted publisher (e.g., National Academy of Sciences, National Literacy Panel).
- Experience developing and/or selecting evidence-based curriculum materials for English early literacy instruction, especially for English learners in the United States.

A separate indicator for cultural responsiveness would also help vendors meet current research needs for providing texts that represent student backgrounds and communities. Representation positively contributes to child development (Hughes-Hassell & Cox, 2010). Children interact with text features while reading, and when texts are diverse, children engage in acceptance, exposure to different cultures, and perspectives that challenge their worldviews (Okoye-Johnson, 2011; Shachar, 2012). However, texts are lacking in vendor programs. The evaluation team further recommends that CDE consider guidance for vendors that instructional materials have diverse authorship and that characters are intentionally diverse, minimally at or above the current research-based percentages discussed in the Year 1 report.

These recommendations align with the READ Act provision that instructional programming include 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))*.

Recommendation 3: 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-



Conclusions and Recommendations

based early literacy instruction in Spanish. CDE currently uses the same process and rubrics for both English- and Spanish-language programs. The overall process works well for both types of programs. Phase 1 rubrics, because of their general nature, also work well for both types of programs.

The main recommendation is that Phase 2 rubrics be reviewed by experts in early literacy learning and instruction in Spanish for Spanish speakers and then be revised to reflect best practices of evidence-based early literacy instruction in Spanish. We recommend that experts include people with the following qualifications:

- Knowledge of theory and practice of teaching early literacy skills in English, especially to Spanish-speaking English learners in the United States.
- Knowledge of theory and practice of teaching early literacy skills in Spanish, especially to Spanish-speaking English learners in the United States.
- Experience teaching early literacy skills in Spanish and/or researching and writing about early literacy skill development and instruction in Spanish.
- Experience developing and/or selecting evidence-based curriculum materials for early literacy instruction in Spanish, especially for Spanish-speaking English learners in the United States.

There are three options for revising rubrics:

1. Revise relevant individual criteria of rubrics to be more general so that criteria apply to both Spanish and English early literacy instruction.
2. Revise relevant individual criteria of rubrics to include practices and examples for both Spanish and English.
3. Develop a separate version of each rubric to align with evidence-based early literacy scope and sequence and instructional practices for Spanish,



Conclusions and Recommendations

including examples relevant to the Spanish language (and its various regional dialects) and the Spanish writing system.

Recommendation 4: Refine the process for reviewing English-language programs in three ways: a) add differentiated supports for less-experienced reviewers; b) codify the consensus-based process that is currently used for program selection; and c) provide vendors with additional guidance on submitting materials that align with rubric criteria.

Recommendation 4a: Consider differentiation of the training process for program reviewers. CDE should offer a tiered training process based on its level of prior experience and depth of background knowledge about using instructional programming grounded in scientifically based reading instruction. Such differentiated training may include additional guided practice and coaching with sample programs, prior to reviewing programs currently under review. Differentiated support could also include additional advance preparation. As part of the training and review process, CDE could also offer more detailed feedback on scoring for less experienced reviewers.

Recommendation 4b: Further codify the consensus decision-making process used by reviewers. Overall, CDE's consensus-based review process worked well. Reviewers described that when there were disagreements, they discussed, located evidence, and came to a consensus. When they were unable to come to a consensus, they consulted their CDE team member to resolve disagreements. When CDE team members needed assistance with problem-solving, they consulted the internal CDE leadership team. This consensus and problem-solving process was informal and not written down, so operationalizing and formalizing these processes could further contribute to consistent scoring.

Recommendation 4c: Provide vendors with additional guidance on submitting materials that align with rubric criteria. One way to guide vendors would be to suggest their content expert assist in locating and describing the evidence. Another way is to provide the rubric along with application materials or



Conclusions and Recommendations

a sample of a quality submission. Several reviewers also noted that they would have preferred to have a full set of program materials to assist with review.

Recommendation 5: Clearly identify the degree to which vendors' evidence meets ESSA evidence levels, make studies available, and provide guidance for how districts and schools might consider this research when selecting instructional programs. Meeting ESSA Tiers 1 and 2 requires vendors to conduct rigorous research to test whether their theories of action and logic models actually result in impact on student outcomes. We recommend that CDE clearly identify the degree to which evidence meets ESSA evidence levels, make studies available, and provide guidance for how districts and schools might consider this research when selecting instructional programs. (Aligns with *Is evidence-based (22-7-1209 (2) (b) (I))*)

Recommendation 6: Provide districts with streamlined and accessible information about how programs are vetted. Reviewers recommended increasing communications and understanding about processes and selection criteria could support curricular adoption and policy change within districts. One reviewer suggested creating documents for districts that explain the advisory list selection process, in addition to the documentation already included on CDE's website.



References

- Adams, M. J. (1990). *Beginning to read: Thinking and learning about print*. MIT Press.
- Adams, M. J. (1998). The three-cueing system. In J. Osborn & F. Lehr (Eds.), *Literacy for all: Issues in teaching and learning* (pp. 73–99). The Guilford Press.
- Anthony, J. L., Williams, J., Duran, L. K., Gillam, S., Liang, L., Landry, S. (2011). Spanish phonological awareness: Dimensionality and sequence of development during the preschool and kindergarten years. *Journal of Educational Psychology*, 103(4), 857–876. <https://doi.org/10.1037/a0025024>
- August, D., & Shanahan, T. (Eds.). (2006). *Developing literacy in second-language learners: Report of the National Literacy Panel on Language-Minority Children and Youth*. Lawrence Erlbaum Associates Publishers.
- Baker, S., Mercier Smith, J., Fien, H., Otterstedt, J., Katz, R., Luft Baker, D., & Thomas Beck, C. (2007). *Three-year report on Oregon Reading First: Impact and implementation*. Oregon Reading First Center.
- Baker, S., Lesaux, N., Jayanthi, M., Dimino, J., Proctor, C. P., Morris, J., Gersten, R., Haymond, K., Kieffer, M. J., Linan-Thompson, S., & Newman-Gonchar, R. (2014). *Teaching academic content and literacy to English learners in elementary and middle school* (NCEE 2014-4012). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. http://ies.ed.gov/ncee/wwc/publications_reviews.aspx
- Beeman, K., & Urow, C. (n.d.). *A comparison of initial literacy development in Spanish and in English*. Center for Teaching for Biliteracy. <http://www.teachingforbiliteracy.com/wp-content/uploads/2014/09/Comparing-initial-literacy-in-English-and-Spanish.pdf>
- Berkeley, S., Scruggs, T. E., & Mastropieri, M. A. (2010). Reading comprehension instruction for students with learning disabilities, 1995–2006: A meta-analysis. *Remedial and Special Education*, 31(6), 423–436. <https://doi.org/10.1177/0741932509355988>
- Borman, G. D., & D'Agostino, J. V. (1996). Title I and student achievement: A meta-analysis of federal evaluation results. *Educational Evaluation and Policy Analysis*, 18(4), 309–326. <https://doi.org/10.3102/01623737018004309>
- Branum-Martin, L., Mehta, P. D., Fletcher, J. M., Carlson, C. D., Ortiz, A. A., Carlo, M. S., & Francis, D. J. (2006). Bilingual phonological awareness: Multilevel construct validation among Spanish-speaking kindergarteners in transitional bilingual education classrooms. *Journal of Educational Psychology*, 98(1), 170–181. <https://doi.org/10.1037/0022-0663.98.1.170>
- Brown, K. J., Patrick, K. C., Fields, M. K., & Craig, G. T. (2021). Phonological Awareness materials in Utah kindergartens: A case study in the science of reading. *Reading Research Quarterly*, 56, S249–S272. <https://doi.org/10.1002/rrq.386>
- 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>



References

- Center on Enhancing Early Learning Outcomes & Council of Chief State School Officers. (2019). *Third grade reading laws: Implementation and impact*. <https://ccsso.org/sites/default/files/2019-08/CCSSO%20CEELO%20third%20grade%20reading.pdf>
- 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>
- Clemens, N. H., Solari, E., Kearns, D. M., Fien, H., Nelson, N. J., Stelega, M., Burns, M., St. Martin, K. & Hoeft, F. (2021). They say you can do phonemic awareness instruction “in the dark”, but should you? A critical evaluation of the trend toward advanced phonemic awareness training. <https://psyarxiv.com/ajxbv/download>
- Connor, C. M., Alberto, P. A., Compton, D. L., & O’connor, R. E. (2014). *Improving reading outcomes for students with or at risk for reading disabilities: A synthesis of the contributions from the Institute of Education Sciences research centers*. <https://ies.ed.gov/ncser/pubs/20143000>
- Dahlen, S. P. (2020). “We need diverse books”: Diversity, activism, and children’s literature. *Literary cultures and twenty-first-century childhoods, 83–108*. https://doi.org/10.1007/978-3-030-32146-8_5
- Ehri, L. C., & McCormick, S. (1998). Phases of word learning: Implications for instruction with delayed and disabled readers. *Reading & Writing Quarterly: Overcoming Learning Difficulties, 14*(2), 135–163. <https://doi.org/10.1080/1057356980140202>
- 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*, 393–447. <https://doi.org/10.3102/00346543071003393>
- 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>
- 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>
- Foorman, B., Beyler, N., Borradaile, K., Coyne, M., Denton, C. A., Dimino, J., Furgeson, J., Hayes, L., Henke, J., Justice, L., Keating, B., Lewis, W., Sattar, S., Streke, A., Wagner, R., & Wissel, S. (2016). *Foundational skills to support reading for understanding in kindergarten through 3rd grade* (NCEE 2016-4008). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. https://ies.ed.gov/ncee/wwc/Docs/PracticeGuide/wwc_foundationalreading_040717.pdf



References

- Foorman, B. R., Smith, K. G., & Kosanovich, M. L. (2017). Rubric for evaluating reading/language arts instructional materials for kindergarten to grade 5 (REL 2017–219). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance; Regional Educational Laboratory Southeast.
<https://ies.ed.gov/pubsearch/pubsinfo.asp?pubid=REL2017219>
- Ford, K., & Palacios, R. (2015). Early literacy instruction in Spanish: Teaching the beginning reader. *Colorín Colorado*. <https://www.colorincolorado.org/article/early-literacy-instruction-spanish-teaching-beginning-reader>
- Fuchs, L. S., Fuchs, D., Hosp, M. K., & Jenkins, J. R. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading*, 5, 239–256.
https://doi.org/10.1207/S1532799XSSR0503_3
- Gamse, B. C., Jacob, R. T., Horst, M., Boulay, B., & Unlu, F. (2008). *Reading first impact study final report* (NCEE 2009-4038). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. <https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=NCEE20094038>
- 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>
- Goldenberg, C., Tolar, T. D., Reese, L., Francis, D. J., Bazan, A. R., & Mejía-Arauz, R. (2014). How important is teaching phonemic awareness to children learning to read in Spanish? *American Education Research Journal*, 51(3), 604–633.
<https://doi.org/10.3102/0002831214529082>
- Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, 7(1).
<https://doi.org/10.1177/074193258600700104>
- Greene, J. P. (1998). *A meta-analysis of the effectiveness of bilingual education*. Tomas Rivera Policy Institute; University of Texas at Austin, Public Policy Clinic of the Department of Government; Program on Education Policy and Governance at Harvard University.
- Hall, C., Dahl-Leonard, K., Cho, E., Solari, E. J., Capin, P., Conner, C. L., Henry, A. R., Cook, L., Hayes, L., Vargas, I., Richmond, C. L., & Kehoe, K. F. (2022). Forty years of reading intervention research for elementary students with or at risk for dyslexia: A systematic review and meta-analysis. *Reading Research Quarterly*.
<https://doi.org/10.1002/rrq.477>
- Herlihy, C., Kemple, J., Bloom, H., Zhu, P., & Berlin, G. (2009). *Understanding Reading First: What we know, what we don't, and what's next*. MDRC.
https://www.mdrc.org/sites/default/files/understanding_reading_first.pdf



References

- Hernandez, D. J. (2012). *Double jeopardy: How third-grade reading skills and poverty influence high school graduation*. Annie E. Casey Foundation.
- Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal*, 2(2), 127–160. <https://doi.org/10.1007/BF00401799>
- Hughes-Hassell, S., & Cox, E. J. (2010). Inside board books: Representation of people of color. *Library Quarterly*, 80(3), 211–230. <https://doi.org/10.1086/652873>
- Jenkins, J. R., Graff, J. J., & Miglioretti, D. L. (2009). Estimating reading growth using intermittent CBM progress monitoring. *Exceptional Children*, 75, 151–163. <https://doi.org/10.1177/001440290907500202>
- 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>
- Jiménez González, J. E. & Ortiz González, M. R. (2000). Metalinguistic awareness and reading acquisition in the Spanish language. *The Spanish Journal of Psychology* 3(1), 37–46. <https://doi.org/10.1017/S1138741600005527>
- 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>
- 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.
- Kilpatrick, D. A., & O'Brien, S. (2019). Effective prevention and intervention for word-level reading difficulties. In D. K. Kilpatrick, R. M. Joshi, & R. K. Wagner (Eds.) *Reading development and difficulties* (pp. 179–210). Springer.
- Kim, Y.-S., Petscher, Y., Schatschneider, C., & Foorman, B. (2010). Does growth rate in oral reading fluency matter in predicting reading comprehension achievement? *Journal of Educational Psychology*, 102(3), 652–667. <https://doi.org/10.1037/a0019643>
- Langenberg, D. N., Corroero, G., Ferguson, G., Kamil, M. L., Samuels, S. J., Shaywitz, S. E., Williams, J., Yatvin, J., Ehri, L., Garza, N., Marrett, C. B., Shanahan, T., Trabasso, T., & Willows, D. (2000). *Report of the National Reading Panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. U.S. Department of Health and Human Services, National Institutes of Health, National Institute of Child Health and Human Development. <https://www.nichd.nih.gov/sites/default/files/publications/pubs/nrp/Documents/report.pdf>
- Liberman, I., Shankweiler, D., Fischer, F., & Carter, B. (1974). Explicit syllable and phoneme segmentation in the young child. *Experimental Child Psychology*, 18, 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>



References

- López-Escribano, C. L., Elosúa, M. R., Gómez-Veiga, I., & García Madruga, J. A. (2013). A predictive study of reading comprehension in third-grade Spanish students. *Psicothema*, 25, 199–205. <https://doi.org/10.7334/psicothema2012.175>
- 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*, 2018(11). <https://doi.org/10.1002/14651858.CD009115.pub3>
- National Center for Education Statistics. (2019). National 4th grade reading scores. In *National Assessment of Educational Progress*.
- National Center for Education Statistics. (2022). Racial/ethnic enrollment in public schools. In *Condition of Education*. U.S. Department of Education, Institute of Education Sciences.
- Okoye-Johnson, O. (2011). Does multicultural education improve students' racial attitudes? Implications for closing the achievement gap. *Journal of Black Studies*, 42(8), 1252–1274. <https://doi.org/10.1177/0021934711408901>
- Pollard-Durodola, S. D., & Simmons, D. C. (2009). The role of explicit instruction and instructional design in promoting phonemic awareness development and transfer from Spanish to English. *Reading & Writing Quarterly*, 25, 139–161. <https://doi.org/10.1080/10573560802683531>
- 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.
- Shachar, R. (2012). Structuring of gender in Israeli society through children's reading and textbooks: Where is mom's apron? *Journal of Research on Childhood Education*, 26, 249–263. <https://doi.org/10.1080/02568543.2012.685124>
- Shanahan, T., Callison, K., Carriere, C., Duke, N. K., Pearson, P. D., Schatschneider, C., & Torgesen, J. (2010). Improving reading comprehension in kindergarten through third grade: A practice guide. U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.
- 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>
- Slavin, R. E., & Cheung, A. (2005). A synthesis of research on language of reading instruction for English language learners, *Review of Educational Research* 75(2) 247–284). <https://doi.org/10.3102/00346543075002247>
- 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>



References

- 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>
- U.S. Department of Education. (2015). *Reading First funding status*. <https://www2.ed.gov/programs/readingfirst/funding.html>
- Vaughn, S., & Wanzek, J. (2014). Intensive interventions in reading for students with reading disabilities: Meaningful impacts. *Learning Disabilities Research & Practice, 29*(2), 46–53. <https://doi.org/10.1111/ldrp.12031>
- Weiss, S., Stallings, D. T., & Porter, S. (2018). *Is Read to Achieve making the grade? An assessment of North Carolina's elementary reading proficiency initiative*. NC State University, Friday Institute for Educational Innovation.
- 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. (2017). 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. Advisory List of Instructional Programs

- B. Summary of individual program ratings
 - B.1. Advisory List of Instructional Programs



Appendices
