Check for updates

Reading Research Quarterly

INTERNATIONAL LITERACY ASSOCIATION

ARTICLE OPEN ACCESS

Promoting Reading Accuracy and Fluency Outcomes With Complex Texts: A Grade 3 Intervention Comparison

Jake Downs[1] | Elfrieda Hiebert[2] | Kristin Conradi Smith[3] | Katie Martz[4]

[1]School of Teacher Education and Leadership, Utah State University, Logan, Utah, USA | [2]TextProject, Santa Cruz, California, USA | [3]School of Education, William & Mary, Williamsburg, Virginia, USA | [4]School of Education, University of Mississippi, Oxford, Mississippi, USA

Correspondence: Jake Downs (jake.downs@usu.edu)

Received: 26 August 2025 | **Revised:** 20 February 2026 | **Accepted:** 30 March 2026

ABSTRACT

This study examined whether third-grade readers identified for intervention could achieve better outcomes with challenging, content-rich texts supported by explicit multisyllabic word instruction compared to a more traditional intervention. Using a matched-sample quasi-experimental design, 110 third-grade students scoring below the 40th percentile on reading assessments were assigned to either the Read Like Us intervention ($n = 55$) or an active control condition using Corrective Reading curriculum ($n = 55$). The Read Like Us condition featured informational texts from 12 coherent topic areas at Grades 4–5 complexity levels, with systematic pre-teaching of multisyllabic vocabulary using a "peel-off" strategy. The active control used narrative texts with controlled vocabulary following a traditional scope-and-sequence approach. Both interventions lasted 60 instructional sessions. Comprehensive text complexity analyses were conducted using psycholinguistic databases and Coh-Metrix tools. Reading outcomes were assessed using Acadience Grade 3 measures. Despite encountering texts with substantially higher complexity (effect sizes exceeding $d = 1.50$ on most measures), Read Like Us students achieved significantly greater reading accuracy gains ($g = 0.47$, $p = 0.04$) and were 2.79 times more likely to reach the critical 98% accuracy threshold. Notably, students with the lowest initial accuracy showed the greatest gains in the challenging text condition, reversing the typical Matthew effect pattern. ORF outcomes were equivalent between conditions ($g = 0.03$, $p = 0.90$). These findings challenge conventional practices of placing striving readers in simplified texts, demonstrating that appropriately scaffolded complex texts can accelerate rather than hinder foundational skill development while also enriching students' exposure to academic language and disciplinary knowledge.

Third grade represents a watershed transition in students' academic lives—the point where curriculum demands undergo a substantial shift. The mandate of the Common Core State Standards—that 50% of Grade 3 texts be informational—fundamentally reshapes what students encounter in their daily reading (National Governors Association Center for Best Practices and Council of Chief State School Officers 2010). Whereas primary grade curricula traditionally emphasize familiar narrative structures and high-frequency vocabulary, third graders now face texts full of complex ideas, abstract concepts, and domain-specific terminology. Moreover, third-grade texts include a substantial portion of novel multisyllabic words (Kearns and Hiebert 2022), further exacerbating the challenge for readers. This increase occurs alongside escalating syntactic complexity, cohesive demands, and discourse sophistication across both informational and narrative genres (Mesmer 2025).

Taken together, the combination can create a formidable challenge for readers still mastering foundational reading proficiencies. Specifically, students who have yet to master the threshold of approximately 98% word-reading accuracy in grade-level text (Schmitt et al. 2011) may find themselves facing a widening gap between their skills and the curriculum demands. Dysfluent word recognition skills compromise meaningful engagement with complex texts. These challenges underscore the need

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2026 The Author(s). *Reading Research Quarterly* published by Wiley Periodicals LLC on behalf of International Literacy Association.

for research into effective intervention designs that not only strengthen students' decoding skills but also ensure access to complex, knowledge-rich texts.

The present study directly addresses this challenge by comparing two reading interventions. Both incorporate word study and repeated reading components, but differ in the types of words targeted for instruction, the word study strategies employed, and the complexity of the texts students encounter. These distinctions allow us to investigate whether targeted improvements in students' word recognition can occur alongside meaningful engagement with more complex texts.

1 | Review of Research: Interventions in Word Development and Repeated Reading of Texts

Effective reading interventions typically rely on two core components: systematic word-level strategies and fluency development through repeated engagement with connected text (Al Otaiba et al. 2023; Gersten et al. 2020; Hall et al. 2023; Neitzel et al. 2022). While these elements appear consistently across interventions, their forms can vary substantially.

1.1 | Word-Level Content and Strategies

The systematic, explicit instructional model that is prominent in many reading interventions in the U.S. can be traced to Project Follow Through (1968–1977). Of the 11 instructional models that were part of the project, Stebbins (1977) found that only one intervention consistently produced significant positive effects across all academic domains: the Direct Instruction (DISTAR) model. The DISTAR findings became a cornerstone for evidence-based reading instruction because they demonstrated that systematic, explicit instruction could close achievement gaps even in challenging circumstances.

Because DISTAR demonstrated consistent effectiveness within Project Follow Through's comparisons, its findings have influenced curricular designs across numerous reading interventions that emphasize systematic phonics, explicit instruction, and carefully sequenced skill progression with primary-level students (e.g., Lane et al. 2025). Although the DISTAR model was originally implemented with early elementary decoding, its influence is also evident in interventions for older striving readers, such as Corrective Reading, which repackaged direct instruction principles to Grades 3–12.

Systematic, explicit instruction models designed for primary-grade foundational skills—rooted in the Direct Instruction approach of the 1960s—prioritized phonetic decoding of monosyllabic words. Multisyllabic word decoding was not a central instructional focus because it was not seen as a primary-grade concern. However, recent research has increasingly documented that students still mastering foundational reading after the primary grades face a different challenge: accurate and automatic decoding of multisyllabic words (Hiebert et al. 2024; Kearns and Hiebert 2022; Tortorelli et al. 2024). This challenge reflects morphological complexity as a primary factor, where vowel pronunciation can shift across derived forms of a root word. This requires students to apply phonetic knowledge flexibly rather than mechanically—a cognitive demand that differs fundamentally from the systematic, sequential decoding of short, orthographically regular words.

To address this challenge, researchers have developed several strategies that combine syllabic and morphemic aspects. These hybrid approaches—such as the 'peel-off' strategy where students identify known affixes, locate remaining vowels, and blend them—represent current multisyllabic word interventions including affix instruction, syllable division strategies, morphemic analysis, and encoding practice (Kearns and Whaley 2019; Lovett et al. 2012; Traga Philippakos et al. 2025a; Vaughn et al. 2022).

Research examining these hybrid approaches demonstrates their effectiveness on isolated word reading. Toste et al. (2017, 2019) and Filderman and Toste (2022) found small to large effects on decoding unfamiliar words (ES = 0.30–0.87) among third through fifth graders, though gains in sight word recognition were more modest (ES = −0.10 to 0.29). Traga Philippakos et al. (2025a, 2025b) similarly found decoding gains (ES = 0.18–0.80) for fourth- and fifth-grade students, and additionally reported moderate to large improvements in oral reading rate (0.63–1.24).

Improved word reading alone hardly ensures reading proficiency, however; rather, students need to be able to decode connected texts with a high degree of accuracy. Although studies demonstrate consistent improvements in isolated word measures, they rarely report distal word reading measures such as passage-level accuracy and fluency, and when they do, results are mixed. Passage level accuracy is reported in single-group effect sizes (Accuracy ES = 0.20–0.91, Oral Reading Rate ES = 0.74–0.92; Traga Philippakos et al. 2025a) limiting the ability to attribute changes solely to the intervention. Further, Filderman and Toste (2022) found that business-as-usual conditions outperformed both treatment conditions on oral reading rate (ES = −0.57 to −0.41) and silent reading efficiency (ES = −1.21 to −0.82). Traga Philippakos et al. (2025b) reported two measures of silent reading efficiency with mixed results (ES = −0.81 to 0.21). Taken together, these studies demonstrate that multisyllabic instruction improves isolated word reading accuracy, but evidence for transfer to connected text—in the form of passage accuracy—remains weak, presenting a critical limitation of our current evidence base.

1.2 | The Effectiveness of Repeated Reading Interventions

Repeated reading is one of the most commonly used methods for improving reading fluency of texts. Since Samuels (1979) introduced repeated reading as a strategy for supporting fluency, numerous studies have used this approach with diverse populations and settings. Although protocols and text types vary across studies, Rasinski's Fluency Development Lesson (FDL; Rasinski et al. 1994; Zimmerman et al. 2019) illustrates typical repeated reading procedures. The FDL uses brief texts—narrative passages, speeches, or poetry—read to mastery in a single session. The protocol progresses from instructor modeling and

19362722, 2026, 2, Downloaded from https://ila.onlinelibrary.wiley.com/doi/10.1002/rrq.70117, Wiley Online Library on [07/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

discussion through choral and partner reading to a final performance for classmates.

Three recent comprehensive reviews provide compelling evidence for repeated reading's effectiveness while also revealing methodological gaps (Hudson et al. 2020; Maki and Hammerschmidt-Snidarich 2022; Padeliadu and Giazitzidou 2018). In the first, Padeliadu and Giazitzidou (2018) reported on the efficacy of interventions of repeated reading on fluency by synthesizing the results of eight meta-analyses. Effects were strong for elementary students (ES = 1.63), while effects with secondary students were strong but more moderate (ES = 0.86). The interventions produced moderate effects on word recognition (ES = 0.55), variable effects on reading rate (ES = 0.30–1.63), and smaller but meaningful gains in comprehension (ES = 0.35). The most substantial effects emerged when interventions targeted both speed and comprehension simultaneously (ES = 0.94).

In a subsequent review of 33 studies, Maki and Hammerschmidt-Snidarich (2022) found an overall fluency effect of ES = 0.46, with single-case designs producing larger effects (ES = 0.75) than group designs (ES = 0.41). Notably, overall intervention duration (time spread across days) predicted student growth better than total minutes, suggesting dosage should be reconceptualized as temporal span rather than cumulative time.

In the third analysis, Hudson et al. (2020) reviewed 16 studies involving 1000 elementary students with reading difficulties and found that repeated reading dominated interventions in their sample (86.5%). Effect sizes for fluency ranged widely (0.01–1.18), as did comprehension outcomes, though 75% of interventions produced meaningful fluency gains.

These reviews demonstrate that repeated reading consistently yields moderate to large effects across diverse populations. However, they also expose methodological gaps that constrain understanding of what makes repeated reading effective. One is the typical focus on time in intervention rather than time spent reading or words read (Hudson et al. 2020). Another is insufficient attention to text characteristics (Padeliadu and Giazitzidou 2018): repeated reading proved more effective with 3–4 repetitions with familiar texts relative to four or more with unfamiliar texts (Therrien 2004). Yet the field lacks systematic examination of how text characteristics influence reading fluency outcomes.

1.3 | The Role of Text Complexity in Oral Reading Fluency

Despite calls for greater attention to text characteristics and expectations that students engage with grade-level and complex texts, the role of texts remains underexamined in reading intervention research. This oversight is not recent. Two decades ago, an analysis of the texts reviewed by the National Reading Panel (NRP 2000) on fluency showed that controlled vocabulary texts averaged 92% high/medium frequency words compared to 83% for literature (Hiebert and Fisher 2005). Yet 74% of studies used controlled vocabulary texts, while only four used literature—with only one reporting fluency outcomes, which were nonsignificant. Because the NRP's positive fluency effects derived largely from controlled vocabulary texts, this finding raises a fundamental question: what do we know about fluency development with complex texts at the elementary level?

Following that 2005 analysis, research on text feature effects remains limited. A notable exception is Barth et al. (2014), who found that in middle school students, both student characteristics and text features uniquely contributed to oral reading fluency (ORF). Student factors accounted for 80% of ORF variance and text characteristics explained 55% of within-student variance.

The finding—that text characteristics might explain variation in student performance—was further supported in a systematic review of research on text complexity at the elementary level. An analysis of 26 studies revealed a negative correlation between increased text difficulty and accuracy, fluency, and comprehension outcomes (Amendum et al. 2018). Nearly all (92%, $n = 24$) studies showed decreased reading accuracy with increased text difficulty, 73% ($n = 19$) showed decreased reading rate, and 54% ($n = 14$) demonstrated negative effects on comprehension. However, reader skill significantly moderated these relationships—less skilled and younger readers showed consistent declines in accuracy, fluency, and comprehension with difficult texts, while skilled readers were often unaffected. Notably, the studies included in the review mostly consisted of students reading without any support.

Moving beyond correlational research, several longitudinal studies have examined the use of challenging text in different instructional configurations. In a 2-year study, Stahl and Heubach (2005) reported that second-grade students' oral reading proficiency benefited from practices such as teacher model reading, repeated reading, and comprehension work with difficult text. Similarly, O'Connor et al. (2002) found that teacher support with difficult text benefited elementary students reading at very low rates (55 words correct per minute). In a subsequent study, O'Connor et al. (2010) examined text difficulty where students read 89%–90% accurate and found similar benefits with instructor support.

Several recent quasi-experimental studies have produced encouraging results in fluency (see detailed results in Supporting Information 1). These studies have consistently reported nonsignificant differences in fluency between students reading orally in pairs with texts at grade level and those reading texts two to four grades above their designated levels (Brown et al. 2018; Downs et al. 2020). Findings have been similar in a comparison of a small group intervention where students read texts at a Flesch–Kincaid level of 6.27 and students in the control group read texts at a 1.59 level (Downs and Young 2025). The effect size for fluency in this study was modest (ES = 0.26), but students who read complex text achieved weekly target increases of 1.5 words correct per minute (WCPM), which represent appropriate intervention benchmarks (Jenkins and Terjeson 2011). Critically, this group of studies showed that students in the complex text conditions were not disadvantaged by reading more rigorous texts.

19362722, 2026, 2, Downloaded from https://ila.onlinelibrary.wiley.com/doi/10.1002/rrq.70117, Wiley Online Library on [07/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

1.4 | The Role of Text Complexity in Oral Reading Accuracy

While these findings on fluency are encouraging, they mask a troubling pattern: oral reading accuracy has been largely neglected in complex text interventions, and the studies that do address it typically reveal negative effects. As described earlier, almost all studies in the systematic review of text complexity in the elementary grades (Amendum et al. 2018) reported decreased reading accuracy (92%, $n = 24$)—more pronounced than effects on fluency and comprehension. Further, the results of the studies reported in Supporting Information 1—which did not integrate word recognition scaffolds for striving readers—essentially mirrored the findings of the systematic review: reading accuracy effects from these studies ranged from null (ES = 0.02 to −0.05) to negative (ES = −0.26 to −0.38).

These data present an instructional trade-off: desirable increases in oral reading rate appear possible when scaffolding oral reading of complex text, but these gains may come at the expense of oral reading accuracy. This rate-accuracy trade-off is particularly problematic for students who need to develop strong decoding skills.

However, O'Connor and colleagues demonstrated that this trade-off is not inevitable. In their initial study (O'Connor et al. 2002), they extracted multisyllabic words from daily texts and taught students to blend, segment, and spell these words. Students in this condition outperformed controls in word identification, word attack, and spelling, and performed similarly to a comparison group receiving word reading scaffolds with less complex text. However, in a later study by the same research team (O'Connor et al. 2010), when the word study component was omitted, students showed null results on decoding outcomes. These findings suggest that attention to word study in complex text interventions is essential if the goal is to improve word recognition accuracy and comprehension.

Students cannot comprehend texts they cannot accurately read—and they cannot read accurately texts they do not understand. Words are more likely to be read inaccurately when they are rare, contain more phonemes, and exhibit less phonetic regularity (Kearns 2020; Tortorelli et al. 2024)—all characteristics of complex text (Hiebert 2025). This bidirectional relationship between recognition and meaning underscores the need for explicit multisyllabic word study that facilitates access to meaning-rich text. The present study directly addresses this gap by examining how systematic vocabulary scaffolding—attending to both word recognition and content-rich reading—influences student performance with complex texts.

2 | Overview of the Current Study

Given the need for more research that attends to accuracy, reading rate, and text characteristics, the present investigation was designed. This study centers on the Read Like Us (RLU) intervention, which combines explicit multisyllabic word instruction with repeated reading of complex texts. By comparing this integrated approach against an active control condition, we contribute essential evidence to the emerging research on challenging texts in fluency interventions while also establishing a more rigorous methodological framework for text analysis in repeated reading research.

Specifically, the study is designed to address two questions:

Research Question #1: To what extent did complexity metrics compare for the word study and connected text portions of each intervention?

Research Question #2: To what extent did participation in the RLU versus the Active Control Condition influence students' reading accuracy, rate, and comprehension?

3 | Method

3.1 | Study Design and Sample

This study employed a matched-sample quasi-experimental design to investigate the reading outcomes between two conditions over 60 instructional sessions. Condition balancing methods were statistically employed in this study to reduce the potential bias present in the non-random student design. The initial unmatched sample used in this study included 283 third-grade students in 17 elementary schools who received word reading intervention within the partnering school district. The sample was 48.4% female ($n = 137$) and 12.4% ($n = 35$) multilingual, with 79.2% ($n = 224$) participating in the active control condition. Due to the unbalanced and nonrandom sampling, Propensity Score Matching (Lane et al. 2012; see Section 3.5) was used to statistically balance the groups. The matched sample used for all statistical analyses in this study included 55 students in each group ($n = 110$), that was 43.6% female ($n = 48$) and 4.5% multilingual ($n = 5$). The sample demographics for the matched and unmatched data are presented in Table 1.

3.2 | Instrumentation

The study used the Acadience Grade 3 Reading assessment (Good et al. 2019), which was administered as part of the partnering district's assessment protocol. Acadience is a curriculum-based measure that includes three passages, each

TABLE 1 | Participant demographics.

	Unmatched		Matched	
Demographics	**Active control**	**Pilot**	**Active control**	**Pilot**
Gender				
Female	112	25	24	24
Male	112	34	31	31
Multilingual learner				
Yes	32	3	2	3
No	192	56	53	52
Total	224	59	55	55

19362722, 2026, 2, Downloaded from https://ila.onlinelibrary.wiley.com/doi/10.1002/rrq.70117, Wiley Online Library on [07/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

read for 1 min. From these passages, the median score for words correct per minute and percentage of correct words is reported. Students also provide a retelling of each passage from which a quantitative and qualitative metric are derived. The quantitative metric ("Retell") awards one point for each word the student states that is related to the passage. The qualitative metric ("Retell Quality") ranks the retell score on a four-point likert scale. To attain the highest rank, the student must provide at least three details from the text in a logical sequence that also captures the main idea. Consistent with the WCPM and accuracy data, the median Retell and Retell Quality scores across the three passages are recorded. The Acadience suite also includes a three-minute maze assessment where students read a passage and must select the correct word among two distractors at intervals of approximately every seventh word. The WCPM, accuracy, retell, and maze results are then given varying weights and used to calculate a composite measure, fully reported by Acadience in the assessment manual (Good et al. 2019). These assessments and the composite measure have excellent technical and reliability properties. The Acadience Technical manual reports Grade Three alternate-form reliability ranging from 0.81 to 0.97, three-form alternate reliability ranging from 0.80 to 0.97, and inter-rater reliability ranging from 0.85 to 0.99 (Good et al. 2019).

3.3 | Conditions

3.3.1 | Read Like us

3.3.1.1 | Selection of Instructional Texts. The first author curated 60 texts from ReadWorks.org, a platform with open-source texts for classroom use. To ensure sufficient challenge for third-grade readers, texts were drawn from the Grades 4–5 Lexile band and sequenced to gradually increase in complexity (Kuhn and Schwanenflugel 2019). The final corpus included 12 informational topic areas, reported in detail with word count in Appendix A. Texts for each condition were also evaluated with the Flesch–Kincaid grade-level measure. Figure 1 displays the Lexile progression of RLU texts and the Flesch–Kincaid progression by condition. After selecting texts, target words were identified for multisyllabic instruction.

3.3.1.2 | Word Study Curation. The RLU word study instruction aimed to improve students' accuracy in reading multisyllabic words, focusing on a curated set of vocabulary. From each intervention text, three words of conceptual importance or perceived decoding difficulty were identified and selected by the first author for multisyllabic instruction. Examples of words chosen for instruction for one text set, related to music and performing arts, appear in Appendix B. These 15 words are illustrative of the full corpus of words chosen for instruction in that all 180 were multisyllabic with 87% of the words three syllables or longer ($n = 157$).

The target words for decoding practice in RLU had additional features that have been shown to enhance word learning. Rather than focusing on rare vocabulary, 70% of the 180 taught words appeared within the 2500 most-frequent morphological families (Hiebert et al. 2018), with only one rare word (*improvisation*). Most critically, 52.2% of the words appeared on the Academic Word List (AWL; Gardner and Davies 2014)—a characteristic of key importance for students' academic success. All chosen words directly connected to text content, ensuring meaningful context for learning. Moreover, the words were curated from the text first and only later triangulated with common morphological families and the AWL.

As is characteristic of target vocabulary in informational texts (Hiebert and Cervetti 2012), the selected terms were also repeated within texts. Overall, there was an average of 5.3 repetitions for the 180 target words achieved within passages and also across morphological families (e.g., *skateboard, skateboarding, skateboarded, skateboards, skateboarder*). This design created conceptual coherence across passages. While specific words might not repeat across all five topic texts, related concepts did. For example, in the music and performing arts set of texts, *instruments* appeared in one passage and then was connected to *bass* (an instrument) in a subsequent orchestra-focused text.

3.3.1.3 | Daily RLU Protocol. Each session contained two main segments: Word Study and Connected Text Reading, displayed in Figure 2. The Word Study portion used a peel-off strategy (Lovett et al. 2012) where students were explicitly taught to identify known prefixes and suffixes in the target word, identify the remaining vowels, and blend the word's parts together (Vaughn et al. 2022). Each word was practiced three times: initially modeled by the group instructor, followed by group practice, and concluding with individual student practice. After applying the peel-off strategy with target words, students read the phrase and full sentence containing that word three times using model, choral, and independent whisper reads. This 5–7 min procedure provided 27 total exposures to each target word prior to reading the text. Although the intervention utilized a relatively limited set of words ($n = 180$), the instructional design of RLU ensured a high volume of practice with each word. In total, each target word was encountered approximately 50 times: 27 initial exposures during word study followed by approximately 25 exposures in connected text (approximately five repetitions per word within the text and each text was read five times, see next paragraph). This high instructional dosage–spaced across the practice attempts and subsequent readings–prioritized the application of reading complex multisyllabic structures over the broader, less frequent exposures found in the active control.

In the Connected Text Reading segment of the lesson, the group engaged in a repeated reading of the target text. The reading protocol followed the procedure described by Downs and colleagues (Downs and Young 2025; Downs et al. 2025). First, the group engaged in a listening passage preview where the group instructor would model-read the text with students following along, a component shown to significantly enhance fluency outcomes for students with reading difficulties (Lee and Yoon 2017). Next, the group echo- and choral- read the text as the second and third reads. Finally, partners synchronously read the text and students concluded by independently whisper-reading the text. The result was five reads of the challenging text progressing from a passage preview by the group to an independent read by each student.

19362722, 2026, 2, Downloaded from https://ila.onlinelibrary.wiley.com/doi/10.1002/rrq.70117, Wiley Online Library on [07/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

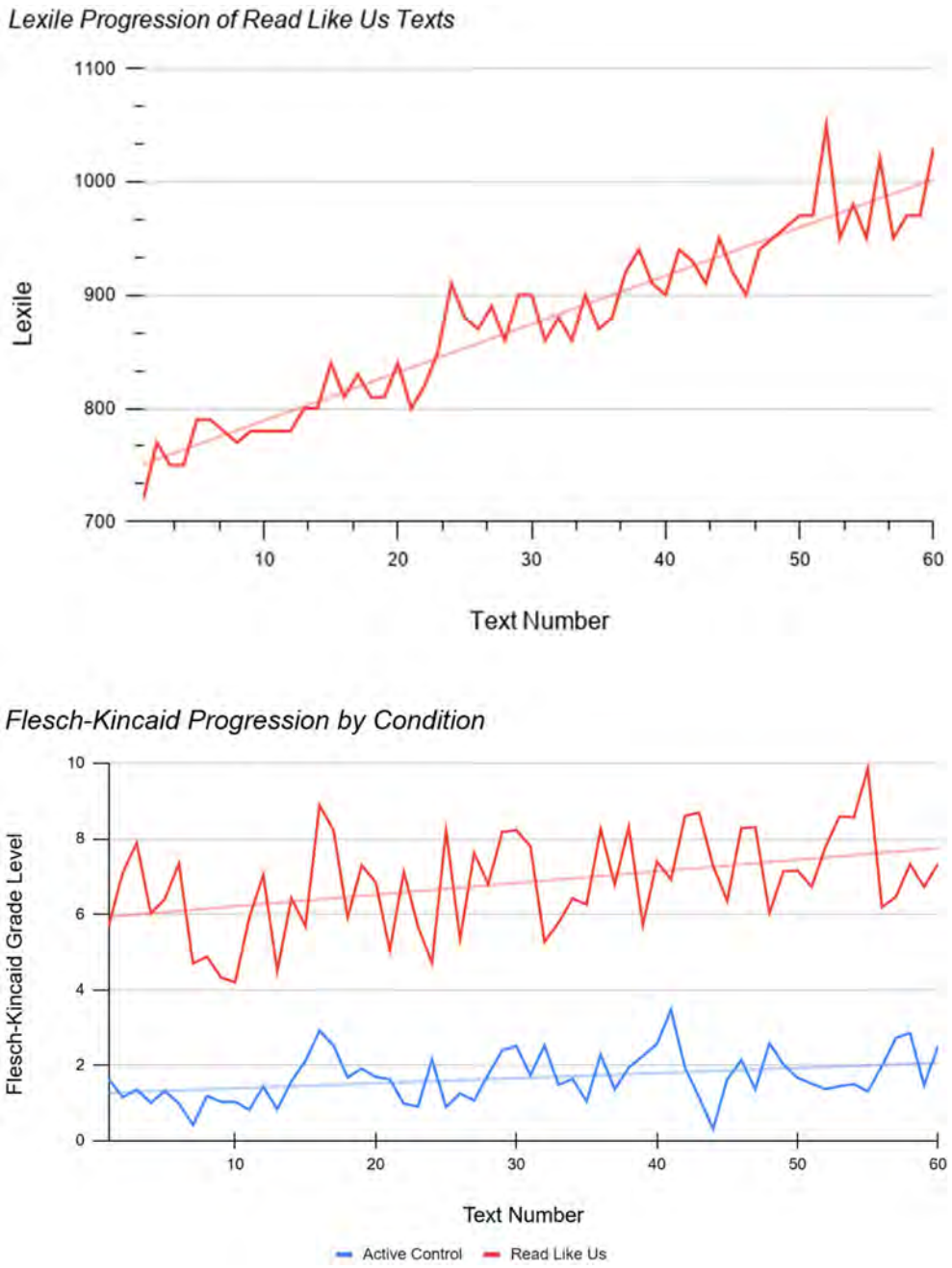


FIGURE 1 | This figure illustrates Lexile and Flesch–Kincaid measures for RLU and Active Control conditions.

3.3.2 | Active Control

This study utilized an "Active Control" condition, which is a comparison group that participates in an alternative intervention or activity rather than receiving no treatment at all or a low-rigor business-as-usual condition. The Active Control condition, Corrective Reading, has a long history as an intervention for students in Grades 3–12, nationally, as well as in the district where the study was conducted. The curriculum uses Direct Instruction principles (Engelmann and Carnine 2016) to provide scripted, sequenced lessons that promote decoding and fluency for striving readers. Corrective Reading is rated as "Promising" by the What Works Clearinghouse (2007) and "Strong" by the Center for Research and Reform in Education (n.d.), and it has demonstrated effectiveness in multiple studies (Przychodzin-Havis et al. 2005). One key study was a large-scale U.S. Department of Education trial that found significant gains in phonemic decoding and fluency among third and fifth grade striving readers (Torgesen et al. 2007). The partnering school district adopted Corrective Reading during the No Child Left Behind era and has regularly implemented it since for students scoring below the 40th percentile on word reading assessments, making it an appropriate active control condition for this study.

3.3.2.1 | **Active Control Word Study.** The Active Control condition also participated in a word study component of roughly equal time as the RLU condition. However, the curricular design of the word study was different: the word study followed a curricular scope and sequence that included common

19362722, 2026, 2, Downloaded from https://ila.onlinelibrary.wiley.com/doi/10.1002/rrq.70117, Wiley Online Library on [07/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

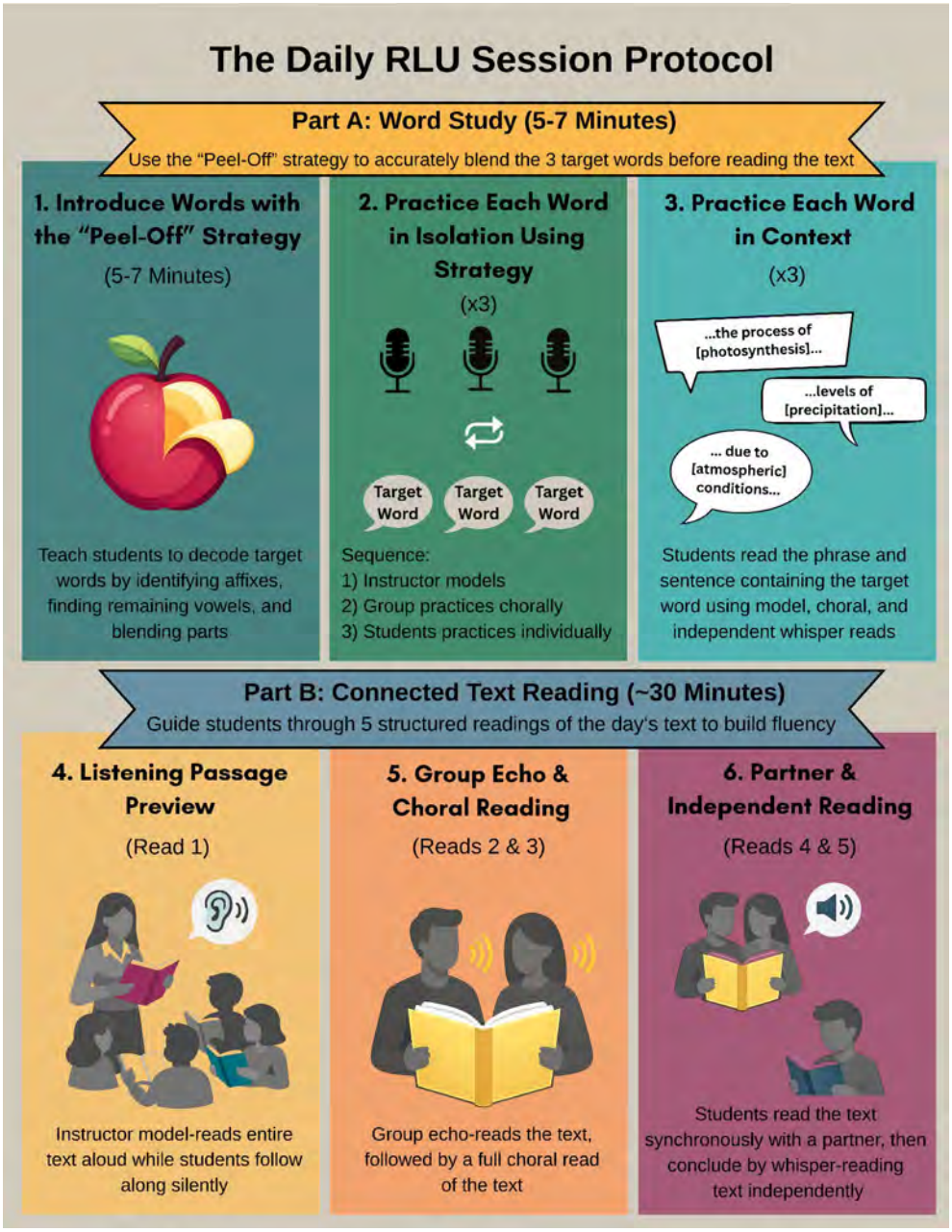


FIGURE 2 | The daily Read Like Us protocol.

letter combinations, regularly spelled words, irregularly spelled words, words with suffixes, silent-e words, and compound words. Typical components included practice with specific letter sound combinations (e.g., "ir," "ur," "er"), followed by practice reading and orally spelling words containing the target letter sound combination (e.g., *girl, turn, jerk*). The Active Control word study exhibited a much higher volume of individual words ($n = 3787$) than the RLU Condition ($n = 180$); however, only 42.7% ($n = 1617$) were multisyllabic and 7.0% ($n = 265$) were more than three syllables in the Active Control condition.

Although instructors used direct instruction techniques to teach the words, many did not appear in the subsequent reading, which limited overall volume of practice. For example, in a story about a baseball pitcher named Art, 36 of the 64 practice words were high-frequency, single syllable words (e.g., *swing, front, two*). Of the 64 words, only 33 appeared in the corresponding lesson text. Lessons also demonstrated a lack of connection between target letter-sound correspondences and practice. The lesson (#30 in the sequence) began with "oi" sound instruction, followed by practice with words like *greet, starved*, and *complained*, rather than words with the taught sound pattern. While many practice words had appeared in previous lessons, this connection was neither made explicit in the teacher's manual nor were students told which words would appear in the text for the lesson.

3.3.2.2 | **Active Control Connected Text.** Each lesson in the Active Control included connected text. These texts were narrative and included six main stories, divided into daily readings, with each daily reading subdivided into six sections. The narrative titles and number of texts associated with each

19362722, 2026, 2, Downloaded from https://ila.onlinelibrary.wiley.com/doi/10.1002/rrq.70117, Wiley Online Library on [07/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

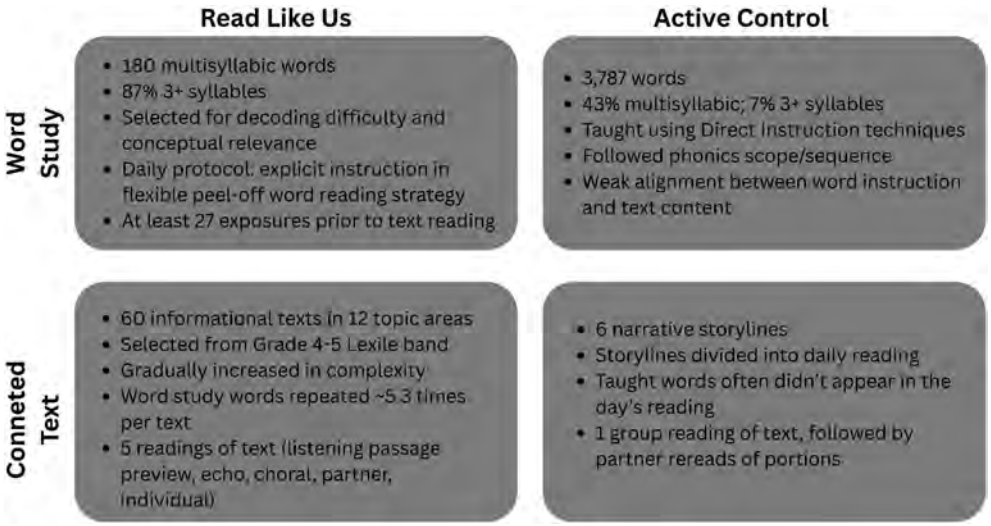


FIGURE 3 | Summary of RLU and Active Control conditions. Refer to Tables 3–7 for statistical differences among word study and text variables.

storyline are reported in Appendix C. The curricular procedure indicates for each text to be read once as a group with an additional portion of the text reread by students in partners. For the initial reading, the instructor had students take turns reading one or two sentences until the selection was finished. The instructor would tally errors for each section and if the group made two errors or less, they would continue to the next section. This process was repeated for six sections of the daily text. Then partners would individually re-read the first section of the daily text, and a portion of the previous day's text with a target words-correct-per-minute goal provided by the curriculum. Figure 3 summarizes the key features of each condition. Supporting Information 2 contains a sample vignette of the texts used during session 30 of each condition.

3.4 | Procedures

At the partnering school district, school-level reading interventionists were informed of the project, its aims, and the time commitment required for participation. These interventionists then determined which—if any—extant small groups in their school would participate in the RLU condition rather than the Active Control curriculum. Students who participated in the Active Control curriculum formed the comparison group.

The lead author conducted a 1 h training for the paraprofessionals who taught the RLU condition. This training provided an overview of each RLU segment as well as modeling and practice implementing each segment. Paraprofessionals who taught the Active Control curriculum had been previously trained in that curriculum. Each paraprofessional was observed in the two conditions at least once by the trained school interventionists. Observations in the treatment condition were conducted using a fidelity rubric created by the first author. Observations in the Active Control were conducted using a fidelity rubric designed by the National Institute for Direct Instruction (NIFDI 2014).

Acadience assessments were administered before and after the pilot by a trained assessment team. After the intake measures, the intervention occurred for approximately 40 min daily for 60 instructional days. Instruction occurred as part of the district's Response to Intervention block and was overseen by a school-level interventionist and the first author. Group instructors in both conditions maintained daily logs, and a review of those logs indicated that instructional time was allocated as prescribed with no cross-group contamination. After collecting initial data and implementing the intervention, the next step was to understand the instructional materials themselves. This involved a detailed analysis of the words and texts students encountered in each condition.

3.5 | Data Analysis

3.5.1 | Comparing Condition Words and Texts

The first phase of data analysis quantified the text complexity in each condition. Specifically, the stimuli words in the word study portion of each condition were analyzed for number of phonemes, letters, syllables, and morphemes, as well as age of acquisition, cumulative frequency, and trajectory. These variables were analyzed using the South Carolina Psycholinguistic Metabase (SCOPE; Gao et al. 2022). To reduce bias in the analysis, words that were monosyllabic and monomorphemic were removed. Thus, the final analysis compared only the words with multiple syllables and/or multiple morphemes in each condition. In addition, the multisyllabic and polymorphemic words from the Acadience Grade 3 exit assessment were also analyzed using these same variables as an indicator of grade level text.

Following the framework established by Mesmer et al. (2012), the connected texts used in each condition were analyzed for word, sentence, and discourse level differences using Coh-Metrix 2.1 (McNamara et al. 2014). Word variables captured at the text level included average word length in letters, average word length in syllables, age of acquisition for content words, word frequency, and word trajectory. Sentence level variables included the average number of words per sentence, average number of words before the main sentence verb (left embeddedness), and the number of modifiers per noun phrase. Finally, discourse level variables included syntactic similarity across sentences, frequency of causal

19362722, 2026, 2, Downloaded from https://ila.onlinelibrary.wiley.com/doi/10.1002/rrq.70117, Wiley Online Library on [07/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

connectives, narrativity, and Flesch–Kincaid grade level. The Acadience Grade 3 exit assessments were also analyzed using the same word, sentence, and discourse variables.

3.5.2 | Comparing Reading Outcomes by Condition

Our quasi-experimental study compared students in the RLU program against all other third-grade students in the district who were receiving the Active Control curriculum at the time of the study. To compare these groups, we used Propensity Score Matching (PSM) to account for the lack of random assignment in the study design. Using logistic regression, students were matched based on demographic variables (gender, English language learner status) and all intake Acadience reading subtests. Following common PSM guidelines, the caliper was set at 0.15 standard deviations (Lane et al. 2012). Before matching, we found that the two groups were heterogeneous—all covariates fell outside the 0.15 standard deviation caliper. After matching, the largest standardized difference was 0.10, indicating much better balance between conditions. Table 2 displays the matched and unmatched data by covariate. The final sample consisted of 55 students in each condition ($n = 55$). The PSM analysis was conducted in R using the MatchIt package (Ho et al. 2025).

To analyze student progress within each condition, we first calculated change scores for each student by subtracting their initial scores from their final scores on all Acadience measures (ORF, Accuracy, Retell, Retell Quality, Maze, and Acadience Composite). Given the nature of students nested within schools in the data, we used multilevel modeling to address the remaining research questions. We first fit an intercept-only model to understand the distribution of variance between schools and between students within the same school. This null model showed an intraclass correlation (ICC) of 11.9%, warranting a multilevel analysis.

Next, to answer Research Question #2, we fit two-level models for each reading outcome. Each model included condition status (whether a student participated in the RLU or Active Control) as a fixed effect, while accounting for students nested within schools. Given the rarity of word study embedded within complex text studies, we decided to thoroughly analyze the word accuracy outcomes. First, we determined 98% accuracy as a desirable outcome for students to attain over the course of the study. This threshold represents the 40%–50% ile for the end of third grade and is consistent with research indicating that exceptionally high levels of accuracy are required for adequate text comprehension (Schmitt et al. 2011). We then used logistic regression to analyze the likelihood of students in each condition attaining 98% or better accuracy by the end of the intervention. After, we utilized a bottom-up model building approach (Hox et al. 2017) to understand how condition status (RLU or Active Control), demographic variables (gender and English language learner status), as well as the Acadience reading variables, influenced reading accuracy outcomes. In this approach, the reading accuracy change score was included as an outcome variable in a two-level regression model with condition status as a fixed effect. Then demographic variables and intake variables were included one at a time as an additional fixed effect. Each time we added a variable, we compared the new model to the previous model using an omnibus likelihood ratio test. This test allowed us to determine if the additional variables significantly improved the model's ability to explain the variance in student growth. Models that accounted for significantly more variance were retained, allowing us to identify the most relevant predictors of student reading accuracy growth. The multilevel analysis for this study was conducted in R using the lme 4 package (Bates et al. 2025).

4 | Results

This matched-sample quasi-experimental study examined the reading outcomes of third-grade students who began the study reading below the 40th percentile on the Acadience Grade 3 reading assessment. Both groups received small-group instruction, delivered by a paraprofessional, that included word study and connected text reading. However, the RLU condition focused on challenging informational texts across 12 topics (totaling 60 texts), with multisyllabic word reading support using words from those texts. In contrast, the Active Control condition used a popular commercial decoding curriculum marketed for students in Grades 3–12 and designed for students with decoding difficulties. This condition consisted of direct instruction of words following a curricular scope and sequence, followed by multiple readings of narrative, controlled texts. To quantify the differences in reading complexity for students, we also analyzed the word study and text characteristics used in each condition.

Research Question #1: To what extent did complexity metrics compare for the word study and connected text portions of each intervention?

4.1 | Word Study Orthographic Properties

Words from the Word Study portion of each condition and the Acadience Grade 3 exit assessment were analyzed to quantify key differences in target words. Descriptive statistics for the condition words are displayed in Table 3. These data indicate that mean scores on most variables varied widely across the two conditions and exit assessment texts. Information provided by effect size comparisons in Table 4 reveals the target words for decoding instruction were significantly more complex in the Read Like Us condition over the Active Control condition. Indeed, the greatest difference between the two conditions was in the number of phonemes per target word ($d = 2.28$, $p < 0.0001$), with the number of morphemes per target word the smallest difference ($d = 1.17$, $p < 0.0001$). On phoneme, letter, syllable, and morpheme measures, the Read Like Us target words were also more complex than the complex words contained in the Acadience Reading Grade 3 assessment ($d = 1.29$–2.60, $p < 0.0001$). Finally, differences between the Active Control target words and the Acadience Reading complex words were much smaller and nonsignificant ($d = -0.20$ to 0.57, $p > 0.05$).

4.2 | Word Study Lexical Properties

Age of acquisition, word frequency, and word trajectory data were also analyzed. Read Like Us words were acquired later,

19362722, 2026, 2, Downloaded from https://ila.onlinelibrary.wiley.com/doi/10.1002/rrq.70117, Wiley Online Library on [07/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 2 | Propensity score matching summary of balance.

	Unmatched data ($n = 283$)			Matched data ($n = 110$)		
Measure	**Treated *M***	**Control *M***	***d***	**Treated *M***	**Control *M***	***d***
Text accuracy	95.33	91.69	0.94	95.21	95.80	−0.17
Correct words per minute	82.64	65.82	1.25	81.35	80.09	0.09
Retell score	36.95	31.63	0.29	36.72	38.93	−0.12
Retell quality	2.81	2.59	0.24	2.84	2.84	0.00
Maze	9.44	7.40	0.40	9.54	9.31	0.05
Acadience composite	277.29	223.03	1.00	275.02	281.58	−0.12
Gender—Female	42.37	49.56	−0.15	43.64	43.64	0.00
Gender—Male	57.63	50.44	0.15	56.36	56.36	0.00
EL status yes	5.08	13.84	−0.42	5.45	3.64	0.08
EL status no	94.92	86.16	0.42	94.55	96.36	−0.08

TABLE 3 | Descriptive statistics, condition words.

Measure	Condition	Min	Max	Mean	SD	Skew	Kurtosis
Number of phonemes	Acadience	3	10	5.15	1.42	1.26	5.13
	Corrective Reading	2	12	5.21	1.3	1	4.97
	Read Like Us	5	14	8.29	1.7	0.59	3.68
Number of letters	Acadience	3	12	6.39	1.7	0.92	4.55
	Corrective Reading	3	14	6.66	1.4	0.79	4.55
	Read Like Us	6	15	9.52	1.72	0.68	3.3
Number of syllables	Acadience	1	4	2.02	0.7	0.8	4.3
	Corrective Reading	1	6	1.88	0.67	0.75	5.21
	Read Like Us	2	6	3.26	0.78	0.78	4
Number of morphemes	Acadience	1	3	1.75	0.61	0.19	2.42
	Corrective Reading	1	4	1.81	0.53	−0.05	3.39
	Read Like Us	1	4	2.48	0.73	−0.04	2.72
Age of acquisition	Acadience	3.52	9.11	6.45	1.45	−0.16	2.32
	Corrective Reading	2.37	12.68	6.27	1.72	0.71	3.73
	Read Like Us	4.11	13.44	8.21	2.16	0.35	2.37
Frequency	Acadience	−1.37	7.39	3.07	2.45	0.17	3.21
	Corrective Reading	−2.15	7.48	2.21	1.97	0	2.89
	Read Like Us	−1.85	5.9	1.13	1.75	0.78	2.94
Trajectory	Acadience	−2.09	2.6	−0.23	0.89	0.35	3.2
	Corrective Reading	−2.79	2.12	−0.58	0.84	0.24	2.89
	Read Like Us	−1.57	2.91	0.93	0.8	−0.28	2.94

were less frequent, and showed steeper learning trajectories than Active Control words ($d = -0.55$ to 1.80, all $p < 0.0001$). Similar patterns emerged when comparing Read Like Us to Acadience Grade 3 words ($d = -1.45$ to 1.19, all $p < 0.0001$). Active Control words were less frequent and showed flatter trajectories than Acadience words (frequency: $d = -0.91$, $p < 0.0001$; trajectory: $d = -0.61$, $p < 0.0001$), with no meaningful difference in acquisition age ($d = -0.07$, $p = 0.75$).

19362722, 2026, 2, Downloaded from https://ila.onlinelibrary.wiley.com/doi/10.1002/rrq.70117, Wiley Online Library on [07/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 4 | Effect size comparisons, word variables.

Measure	Condition	*d*
Number of phonemes	RLU—CR	2.28***
	RLU—AG3	2.60***
	CR—AG3	0.32
Number of letters	RLU—CR	1.96***
	RLU—AG3	2.53***
	CR—AG3	0.57
Number of syllables	RLU—CR	2.02***
	RLU—AG3	1.82***
	CR—AG3	−0.2
Number of morphemes	RLU—CR	1.17***
	RLU—AG3	1.29***
	CR—AG3	0.11
Age of acquisition	RLU—CR	1.10***
	RLU—AG3	1.03***
	CR—AG3	−0.07
Frequency	RLU—CR	−0.55***
	RLU—AG3	−1.45***
	CR—AG3	−0.91***
Trajectory	RLU—CR	1.80***
	RLU—AG3	1.19***
	CR—AG3	−0.61**

Abbreviations: RLU = Read Like Us; CR = Corrective Reading; AG3 = Acadience Reading Grade 3.
***$p < 0.0001$.
**$p < 0.001$.

TABLE 5 | Frequency distribution of text words by percentage in each condition.

Condition	# Words	High	Medium	Low	Rare
RLU	17,124	74	17	6	3
Active control	46,288	80	15.5	3	1.5

Note: Word distributions derived from Hiebert et al. (2018).

4.3 | Connected Text Properties

The frequency distribution of words reported in Table 5 indicates that the Active Control texts contained a higher distribution of high- and medium-frequency words than the RLU texts and half as many low and rare frequency words. Descriptive statistics for text level data are displayed in Table 6 for the three levels of analyses: word-, sentence-, and discourse-level variables. The RLU texts exhibited more letters and syllables per word, with a higher age of acquisition and lower frequency than either Active Control or Acadience Grade 3 assessment texts. Sentence-level results indicate the RLU condition contained longer sentences, more left embeddedness and modifiers per noun phrase, with less syntactic similarity in relation to the Active Control and Acadience Assessment texts. Finally, discourse-level variables report less narrativity, more causal connectives, and a higher Flesch–Kincaid grade level rating for the RLU intervention texts.

Table 7 indicates all measured differences between the Read Like Us texts and the Active Control texts were statistically significant, with average syllables per word exhibiting the largest effect size ($d = 4.28$). Read Like Us texts were also significantly different than the Acadience text on six measures (Syllables Per Word, $d = 2.37$; Letters Per Word, $d = 2.54$; Age of Acquisition of Content Words, $d = 1.39$; Left Embeddedness, $d = 1.47$; Narrativity, $d = -2.13$; and Flesch–Kincaid, $d = 2.22$). The Active Control texts also differed significantly on six measures from the Acadience texts (Syllables Per Word, $d = -1.91$; Letters Per Word, $d = -1.90$; Sentence Length $d = -2.51$; Modifyers Per Noun Phrase, $d = -0.39$; Syntactic Similarity, $d = 3.16$; Narrativity, $d = 1.23$; Flesch–Kincaid, $d = -2.87$).

Research Question #2: To what extent did participation in the RLU versus the Active Control Condition influence students' reading accuracy, rate, and comprehension?

Table 8 reports the descriptive statistics for reading outcomes in each condition. At intake, matched groups were comparable on all measures; for example, ORF means were nearly identical (Active Control $M = 80.1$, SD = 15.9; RLU $M = 81.3$, SD = 12.9) and both groups began with approximately 95% reading accuracy.

4.4 | Oral Reading Accuracy

By the end of the study, reading accuracy results diverged by condition status. While the Active Control condition remained stagnant (Pre $M = 95.8\%$; Post $M = 95.7\%$), the RLU condition improved to a mean of 96.8%. The multilevel change score model reported in Table 9 revealed this outcome to be statistically significant ($b = 1.62$, SE = 0.77, $p = 0.04$) with a moderate effect ($g = 0.47$). These results suggest that students in the RLU condition achieved significantly greater gains in oral reading accuracy of connected text than their matched peers in the Active Control.

4.5 | Attainment of 98% Accuracy by Condition

Table 10 reports the logistic regression outcomes predicting end of year accuracy at 98% or higher, based on treatment group. These data report a statistically significant outcome in favor of the RLU condition ($b = 1.02$, SE = 0.52, $p = 0.05$), indicating that students in the RLU condition were 2.79 times as likely to reach 98% accuracy by the end of the intervention than the active control condition.

Table 11 reports the fixed and random effects of the change score multilevel models in accuracy and Table 12 reports the likelihood ratio results. Models including demographic variables indicated that English Learning status accounted for significantly more variance, while gender did not. Among all intake reading variables, only reading accuracy accounted for significantly more variance—with lower intake accuracy associated with more growth in reading accuracy. The final model—Model 5—reports an interaction between intake

19362722, 2026, 2, Downloaded from https://ila.onlinelibrary.wiley.com/doi/10.1002/rrq.70117, Wiley Online Library on [07/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 6 | Descriptive statistics, condition texts.

Measure	Condition	Min	Max	Mean	SD	Skew	Kurtosis
Average syllables per word	Acadience	1.31	1.34	1.32	0.02	0.53	1.50
	Corrective Reading	1.1	1.28	1.18	0.04	0.21	2.31
	Read Like Us	1.29	1.71	1.49	0.1	0.09	2.66
Average letters per word	Acadience	3.94	4.32	4.13	0.19	0.00	1.50
	Corrective Reading	3.41	4.08	3.75	0.12	−0.35	3.14
	Read Like Us	4	5.15	4.63	0.35	−0.29	2.93
AoA content words	Acadience	254.39	296.09	275.75	20.88	−0.09	1.50
	Corrective Reading	223.56	329.57	256.67	24.47	1.28	4.67
	Read Like Us	235	442.74	329.57	49.92	0.46	3.08
Word frequency	Acadience	1.11	1.37	1.28	0.15	−0.70	1.50
	Corrective Reading	1.06	1.9	1.52	0.18	0.04	2.56
	Read Like Us	0.52	1.96	1.19	0.29	0.40	3.14
Sentence length	Acadience	10.64	11.89	11.86	1.35	0.32	1.50
	Corrective Reading	5.76	8.7	8.7	1.01	0.25	3.65
	Read Like Us	10.14	12.44	12.44	1.47	0.90	3.70
Left embeddedness	Acadience	1.77	3.32	2.53	0.77	0.07	1.50
	Corrective Reading	1.24	3.25	2.29	0.35	−0.11	3.80
	Read Like Us	2.31	6.13	3.41	0.78	0.92	4.21
Modifyers per noun phrase	Acadience	0.65	0.81	0.75	0.09	−0.69	1.50
	Corrective Reading	0.37	0.82	0.59	0.12	0.23	2.21
	Read Like Us	0.6	1.24	0.91	0.15	−0.11	2.38
Syntax similarity all sentences	Acadience	0.1	0.12	0.11	0.01	0.15	1.50
	Corrective Reading	0.13	0.21	0.17	0.02	0.02	2.39
	Read Like Us	0.05	0.16	0.12	0.02	−0.58	4.21
Narrativity	Acadience	−0.14	1.11	0.57	0.64	−0.45	1.50
	Corrective Reading	0.27	2.22	1.13	0.47	0.39	2.43
	Read Like Us	−1.38	0.99	−0.42	0.45	0.22	3.49
Causal connectives	Acadience	3.36	27.3	17.95	0.64	−0.63	1.50
	Corrective Reading	4.45	34.38	17.82	0.47	0.27	2.70
	Read Like Us	6.08	48.23	25.32	0.45	0.08	2.25
Flesch–Kincaid	Acadience	4.36	5.03	4.61	0.36	0.65	1.50
	Corrective Reading	0.32	3.47	1.73	0.66	0.26	2.52
	Read Like Us	4.2	9.85	6.84	1.29	−0.11	2.39

accuracy and condition status, indicating that students in the RLU condition experienced greater growth in accuracy overall, and especially for students who started with lower intake accuracy. These findings in word reading accuracy are notable because the word study in the RLU condition contained far fewer words than the active control; however, the words were of much greater complexity.

4.6 | Oral Reading Fluency

Both conditions exhibited growth in oral reading rate. The treatment and active control average increased by 14 WCPM and 13.5 WCPM, respectively, with a Hedge's *g* difference of 0.03. Despite the marginally superior outcomes by the RLU condition on oral reading rate, the multilevel model indicates that these differences

19362722, 2026, 2, Downloaded from https://ila.onlinelibrary.wiley.com/doi/10.1002/rrq.70117, Wiley Online Library on [07/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 7 | Effect size comparisons, text variables.

Measure	Condition	*d*
Average syllables per word	RLU—CR	4.28***
	RLU—AG3	2.37**
	CR—AG3	−1.91**
Average letters per word	RLU—CR	4.44***
	RLU—AG3	2.54**
	CR—AG3	−1.90**
AoA content words	RLU—CR	1.89***
	RLU—AG3	1.39
	CR—AG3	−0.49
Word frequency	RLU—CR	−1.70***
	RLU—AG3	−0.68
	CR—AG3	1.03
Sentence length	RLU—CR	2.97***
	RLU—AG3	0.46
	CR—AG3	−2.51***
Left embeddedness	RLU—CR	1.86***
	RLU—AG3	1.47*
	CR—AG3	−0.39
Modifyers per noun phrase	RLU—CR	1.86***
	RLU—AG3	1.47
	CR—AG3	−0.39
Syntax similarity all sentences	RLU—CR	−2.67***
	RLU—AG3	0.49
	CR—AG3	3.16***
Narrativity	RLU—CR	−3.36***
	RLU—AG3	−2.13**
	CR—AG3	1.23
Causal connectives	RLU—CR	0.82***
	RLU—AG3	0.8
	CR—AG3	−0.01
Flesch–Kincaid	RLU—CR	5.09***
	RLU—AG3	2.22**
	CR—AG3	−2.87***

***$p < 0.0001$.
**$p < 0.001$.
*$p < 0.01$, $p < 0.05$.

were not statistically significant ($b = 0.34$, $p = 0.90$). This finding of equivalence is notable given the disparity in text complexity; despite reading texts that were dramatically more complex (Flesch–Kincaid ES = 5.09), RLU students increased in oral reading rate at the same rate as the active control condition.

4.7 | Basic Comprehension and Composite Measures

Performance on the retell, retell quality, and maze measures were statistically equivalent between groups. Students in the RLU condition increased more than the active control with the number of words they could say about the text ($b = 8.26$, SE = 2.88), but this difference was pragmatically and significantly null ($g = 0.05$, $p = 0.80$). Similar results were also found in the retell quality outcomes ($b = -0.05$, SE = 0.20, $g = -0.05$, $p = 0.78$) and maze outcomes ($b = -0.32$, SE = 1.10, $g = -0.04$, $p = 0.77$). Outcomes on the Acadience Composite measure favored the RLU condition, however the results were not statistically significant ($b = 12.22$, SE = 12.98, $g = 0.24$, $p = 0.24$).

4.7.1 | Results Summary

This study compared the outcomes of two interventions on striving readers' word reading accuracy and oral reading fluency. After establishing that the words and texts used in the RLU condition were significantly more complex than those in the Active Control, we next set out to examine students' growth. The main finding of this study is that students in the RLU intervention performed as well as—and in some cases better—than students in the Active Control intervention. Despite significantly greater word and text complexity, the Read Like Us intervention yielded equal or superior reading outcomes compared to students in the Active Control condition. Outcomes were particularly favorable for accuracy and for students who started the intervention with lower initial accuracy. We make sense of these findings below.

5 | Discussion

Despite encountering substantially more complex words and texts, students in the RLU intervention matched or exceeded the reading gains of students in the Active Control condition. The gains proved most dramatic for word reading accuracy—the very dimension of reading fluency that has often proven most resistant to intervention. This finding carries particular weight in that students who entered the intervention condition with lower initial accuracy showed the strongest growth, suggesting that the intervention succeeded precisely where struggle can be most entrenched.

5.1 | Word Reading Accuracy: The Core Finding

The contrasting approaches to word instruction produced different accuracy outcomes. While Active Control students received direct instruction on high-volume, lower-complexity words—only half of which appeared in subsequent connected text—RLU students engaged in explicit multisyllabic word reading strategy targeting fewer but more complex words. Critically, 100% of RLU target words appeared in the texts students read, with each word repeated an average of 5.2 times in authentic contexts. The words themselves were substantially more challenging: RLU students practiced words averaging 3.3 syllables, while the words that made up practice in the Active Control condition averaged 1.9 syllables.

19362722, 2026, 2, Downloaded from https://ila.onlinelibrary.wiley.com/doi/10.1002/rrq.70117, Wiley Online Library on [07/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 8 | Descriptive statistics, reading outcomes.

Condition	Accuracy	ORF	Retell	Retell quality	Maze	Acadience composite
Corrective reading pre	95.8 (3.4)	80.1 (15.9)	38.9 (17.8)	2.8 (1.0)	9.3 (4.4)	281.6 (70.1)
Corrective reading post	95.7 (3.3)	93.6 (20.2)	47.2 (18.6)	3.3 (1.0)	15.7 (4.8)	336.7 (74.4)
Read like us pre	95.2 (3.5)	81.3 (12.9)	36.7 (16.8)	2.8 (0.9)	9.5 (5.1)	275.0 (51.7)
Read like us post	96.8 (2.7)	95.3 (15.3)	46.0 (23.2)	3.2 (1.0)	15.8 (4.2)	345.0 (64.7)

TABLE 9 | Results of multilevel model for fixed and random effects, research question 1.

	Fixed effects						Random effects			
Measure	**Coefficient**	**SE**	***t***	**R2M**	***p***	***g***	**VC**	**σϵ2**	**R2C**	**ICC**
Oral reading accuracy										
Intercept	0.01	0.55	0.02	0.05	0.99		0.87	12.78	0.11	0.06
Condition	1.62	0.77	2.12		0.04	0.47				
Words correct per minute										
Intercept	13.50	2.07	24.74	0.00	<0.01		17.41	156.51	0.10	0.10
Condition	0.34	2.81	0.12		0.90	0.03				
Maze										
Intercept	6.42	0.78	8.20	0.00	<0.01		1.25	27.86	0.04	0.04
Condition	−0.32	1.10	−0.29		0.77	−0.04				
Retell										
Intercept	8.26	2.88	2.86	0.00	<0.01		0.0	455.70	0.00	0
Condition	1.02	4.07	0.25		0.80	0.05				
Retell quality										
Intercept	0.44	0.14	3.12	0.00	<0.01		0.00	1.08	0.00	0
Condition	−0.05	0.20	−0.28		0.78	−0.05				
Acadience composite										
Intercept	56.65	9.53	5.95	0.01	<0.01		372.8	3322.8	0.11	0.10
Condition	12.22	12.98	0.94		0.35	0.24				

This pedagogical choice yielded a statistically significant moderate effect favoring RLU (0.47), with RLU students 2.79 times more likely to reach the critical proficiency threshold of 98% accuracy by study's end. Most tellingly, multilevel analyses revealed an interaction by initial accuracy: students with the lowest initial word reading accuracy demonstrated significantly greater growth in the RLU condition—a reversal of the typical Matthew effect (Stanovich 1986). Rather than benefiting most from simplified word practice, struggling readers showed accelerated progress when practicing challenging, content-aligned vocabulary.

The assessment findings underscore the robustness of this effect. Although the test passages contained words with a lower mean syllable count (2.0) than RLU's taught words (3.3), they included 71 multisyllabic words with three or more syllables. Of these 71 words, only 10 had been explicitly taught in the RLU intervention—meaning RLU students demonstrated superior accuracy on the vast majority of multisyllabic words in the assessments that they had not encountered during instruction. This pattern demonstrates genuine transfer: students who practiced complex word reading strategies generalized those strategies to untrained polysyllabic words in grade-level text.

These accelerated accuracy outcomes are particularly significant given that recent challenge-text studies report null and even negative outcomes on reading accuracy effects (Brown et al. 2018; Downs et al. 2020, 2025). The contrast is instructive. RLU did not simply expose students to challenging text; instead, it combined explicit strategy instruction targeting polysyllabic words with systematic exposure to those words in authentic contexts before encountering them in passages. This integration of targeted word study, strategic scaffolding, and contextual embedding appears to

19362722, 2026, 2, Downloaded from https://ila.onlinelibrary.wiley.com/doi/10.1002/rrq.70117, Wiley Online Library on [07/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

distinguish RLU's positive accuracy effects from challenge-text approaches that rely primarily on exposure to text of elevated complexity without corresponding word-level instruction.

TABLE 10 | Logistic regression predicting end-of-year accuracy threshold attainment (EOY) among students below threshold at midyear (MOY).

Predictor	*B*	SE	Wald *z*	*p*	Odds ratio
Intercept	−1.18	0.4	−2.92	0.004**	0.31
Group (RLU)	1.02	0.52	1.98	0.047*	2.79

Abbreviation: RLU = Read Like Us.
**$p < 0.001$.
*$p < 0.01$.

5.2 | Oral Reading Fluency: Complexity Without Cost

While accuracy outcomes revealed a striking advantage for RLU, oral reading fluency (ORF) outcomes demonstrated something equally important: equivalence in the face of substantially greater text complexity. Across the 60-day intervention, RLU students achieved an increase of approximately 14 words per minute—0.23 WCPM per session—comparable to benchmarks for effective Grade 3 fluency interventions (Fuchs et al. 1993) and matching the growth of Active Control students despite reading substantially more challenging texts.

Critically, students in the RLU condition experienced no cost in fluency development from exposure to substantially more challenging texts, greater vocabulary complexity, or more

TABLE 11 | Results of multilevel model for fixed and random effects for reading accuracy outcomes 3.

		Fixed effects				Random effects			
Measure	**Coefficient**	**SE**	***t***	**R2m**	***p***	**VC**	**σε2**	**r2c**	**ICC**
Model 1									
Intercept	0.01	0.55	0.02	0.05	0.99	0.87	12.78	0.11	0.06
Condition	1.62	0.77	2.12		0.04				
Model 2									
Intercept	1.36	1.71	0.8	0.05	0.43	0.79	12.87	0.11	0.06
Condition	1.62	0.76	2.12		0.04				
Language status	−1.42	1.69	−0.84		0.4				
Model 3									
Intercept	67.23	7.7	8.73	0.44	< 0.0001	0.02	7.99	0.45	0
Condition	1.3	0.054	2.4		0.02				
Language status	−1.51	1.3	−1.16		0.25				
Intake accuracy	−0.69	0.08	−8.68		< 0.0001				
Model 4									
Intercept	57.57	6.9	8.33	0.58	< 0.0001	0	6.08	0	0
Condition	1.01	0.48	2.12		0.04				
Language status	−0.63	1.14	−0.55		0.58				
Intake accuracy	−0.61	0.07	−8.67		< 0.0001				
Acadience composite change	0.02	0	5.88		< 0.0001				
Model 5									
Intercept	38.17	9.23	4.13	0.61	< 0.0001	0	5.64	0	0
Condition	39.57	12.72	3.11		0.002				
Language status	−0.68	1.1	−0.62		< 0.0001				
Intake accuracy	−0.41	0.1	−4.26		< 0.0001				
Acadience composite change	0.02	0	5.98		< 0.0001				
Interaction: Condition × Intake accuracy	−0.4	0.13	−3.03		0.003				

19362722, 2026, 2, Downloaded from https://ila.onlinelibrary.wiley.com/doi/10.1002/rrq.70117, Wiley Online Library on [07/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 12 | Results of logliklihood ratio omnibus tests.

Model comparison	LR statistic	df	*p*
Model 1 vs. Model 2	3.59	1	0.05
Model 2 vs. Model 3	53.32	1	<0.0001
Model 3 vs. Model 4	20.97	1	<0.0001
Model 4 vs. Model 5	6.7	1	0.009

substantive content demands. This finding challenges long-standing assumptions in intervention design that striving readers require constrained vocabulary and simplified texts to make progress. The Active Control condition represents precisely this approach: carefully controlled word complexity organized around orthographic patterns, with narrative texts simplified for accessibility. Yet RLU students achieved equivalent fluency gains while engaging with complex informational texts from content domains and multisyllabic vocabulary guidance. Equivalence to an established, widely-adopted intervention—particularly one explicitly marketed for students with decoding difficulties—suggests we may need to reconsider whether orthographic constraint and textual simplification represent optimal conditions for supporting striving readers' foundational skill development beyond the early elementary years.

These fluency gains align with broader research demonstrating ORF growth even with the use of challenging texts (see Supporting Informations 1 and 2). When compared to other recent challenge-text studies employing Acadience (DIBELS) as the ORF measure, RLU yielded comparable growth rates to previous research (Downs et al. 2020, 2025) and higher growth than documented by others (Brown et al. 2018). This convergence of fluency adequacy and accuracy superiority suggests that the tension traditionally posited between challenge and foundational skill development may be false—or, at minimum, resolvable through careful integration of word-level instruction with appropriately complex texts.

5.3 | Mechanisms Underlying RLU Effectiveness

The RLU intervention and Active Control condition differed substantively across three dimensions: text complexity, vocabulary exposure density, and content/background knowledge demands. RLU texts were substantially more complex than Active Control texts across word, sentence, and discourse levels, with effect sizes for complexity differences exceeding 1.5 for most measures and 5.09 on Flesch–Kincaid grade-level. RLU students encountered grade-level informational texts addressing science and history topics, with 27 systematic exposures to target multisyllabic words in texts before reading the texts. By contrast, Active Control students read below-grade-level narrative texts centered on fictional scenarios with substantially fewer structured word study opportunities.

Because the three factors of text complexity, vocabulary exposure, and text content were not independently manipulated, we cannot definitively isolate which feature—or combination of features—accounts for RLU's equivalent or superior outcomes. However, this limitation itself illuminates an important constraint inherent in reading instruction: the defining elements of the Active Control intervention are not theoretically or practically separable: Orthographic scope constrains conceptual territory; simplified discourse supports predictable narrative structures. Future research with more granular experimental designs could refine specific effects, but attempting to disentangle these dimensions in the Active Control context may be neither feasible nor theoretically meaningful.

Rather than speculate about inseparable mechanisms, we focus our analysis on two dimensions where the evidence base and theoretical grounding allow for stronger claims: the role of repetition and exposure density in vocabulary learning and the time allocation within interventions to vocabulary practice and text application.

5.3.1 | Repetition and Opportunities for Word Learning

RLU was intentionally designed with brief but consistent exposure to complex, contextually relevant words and extended exposure to thematically related texts. While RLU contained minimal explicit vocabulary instruction—explanations were relegated to brief definitions of specific affixes and words—the repetition that facilitated word accuracy and fluency also likely increased the probability for incidental word learning.

Incidental word acquisition is enhanced by repeated exposures within a short span, morphological analysis, and encounters embedded within and across texts (Bowers and Kirby 2010; Swanborn and de Glopper 1999; Uchihara et al. 2019; Wang et al. 2025). The RLU design operationalized these facets by incorporating the peel-off morphological strategy, dozens of repetitions per target word, immediately reinforcing each decoding attempt by application in connected text, and providing exposure to semantic networks of related words across thematic texts. In sharp contrast, the Active Control lacked these conditions. Target words averaged just 1.2 encounters, appeared in the subsequent daily text only 50% of the time, and were presented as a decontextualized list based on orthographic pattern. Ultimately, the Active Control lacked the design to provide contextually relevant exposures necessary for increasing word accuracy and word learning. Given that the majority of stadolescent readers exhibit concurrent difficulties in fluency and vocabulary (Clemens et al. 2017), designing pragmatic approaches to support each interdependently is more than a pedagogical preference—it is likely a design necessity. Although word learning was not measured in this study, the reading outcomes combined with the favorable word learning affordances of RLU suggest decoding automaticity and vocabulary can be effectively targeted within a single instructional framework.

5.3.2 | Instructional Time Allocation

RLU was intentionally designed to allocate approximately 20% of lesson time to explicit, systematic word study and the remaining 80% devoted to text reading. This allocation reflects a deliberate choice to provide targeted but modest intervention

19362722, 2026, 2, Downloaded from https://ila.onlinelibrary.wiley.com/doi/10.1002/rrq.70117, Wiley Online Library on [07/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

intensity—brief, explicit exposure to challenging words embedded within authentic reading contexts—rather than expand the scope of multisyllabic instruction to include components such as extended morphological analysis or rapid drilling of isolated affixes. The rationale is straightforward: given constrained instructional time where teachers must address multiple facets of literacy, the critical variable is not simply how much time is spent on word study, but what students do during that time and how it serves their reading of meaningful texts.

Several recent multisyllabic interventions have adopted different allocations where approximately 80% of lesson time is devoted to word study activities and 20% to connected text reading (Toste et al. 2017, 2019; Traga Philippakos et al. 2025a, 2025b). This inverted ratio reflects an assumption that intensive, isolated word-level work is necessary to build competence with complex words. An economically allocated approach—providing systematic, explicit exposure to multisyllabic words positioned directly in service of reading grade-level connected text that addresses content and other reading goals—appears promising. The question becomes not merely one of time allocation, but of task authenticity: whether brief, purposeful word study embedded within meaningful reading contexts can rival or exceed the effects of extended isolated word work. The results presented here suggest it can.

5.4 | Implications

These findings—that striving readers benefit from strategic engagement with appropriately scaffolded challenging words embedded within complex texts—carry substantial weight precisely because they contradict long-standing instructional assumptions. But findings alone do not direct instructional practice. What remains is a pressing question: Given that this particular configuration of instruction yielded superior outcomes, how should teachers move this evidence into their daily practice? The present section addresses two complementary dimensions of implementation that emerge from the RLU design: pre-teaching of words and selecting connected and meaningful texts.

5.5 | Strategic Decoding of Critical Vocabulary

The interrelationships of word complexity, word frequency, and age of acquisition within instructional text present a cruel trade-off for inaccurate readers: the words that are the hardest to read in a text tend to convey the most meaning. In any complex text polysyllabic words are not limited to a handful of words. Consider the text "Bass" within the Music topic of the RLU intervention. Of the 182 unique words in the text, three—*orchestras*, *composition*, and *experimenting*—were selected for explicit instruction and practice. Yet among the remaining 30 polysyllabic words in that passage, students encountered complex vocabulary (e.g., *repertoire*, *virtuoso*) along with general academic words (e.g., *requires, positions*) and potentially unfamiliar two-syllable words (e.g., *octave*, *rhythmic*).

Rather than attempting to explicitly scaffold all multisyllabic words—an impossibility in dense, content-rich texts—RLU employed a dual strategy. First, for each instructional text, three words were carefully chosen based on conceptual centrality to the text and decoding difficulty. Across texts, the 180 target words taught over the 60 sessions averaged 5.3 repetitions within texts. Second, students were taught a morphological strategy called "peel-off," in which they systematically analyzed word roots, affixes, inflected endings, and vowels to decode polysyllabic words. This approach meant that when students confronted previously unencountered words in texts, they could potentially apply morphological understanding and semantic context—skills developed through explicit instruction on the focal three words—to successfully decode and hopefully comprehend untaught words.

Thus, strategic decoding of critical vocabulary addresses a fundamental tension in literacy instruction: the impossibility of teaching every challenging word in a content-rich text, balanced against the necessity of building students' word knowledge and decoding strategies. By deliberately limiting explicit instruction and blending support to three conceptually central words per text while simultaneously equipping students with morphological tools to tackle untaught polysyllabic words independently, RLU creates a scalable model for classroom practice.

5.6 | Strategic Text Selection

The results of this study point to another implication: educators must attend carefully to their selection of texts and topics. The domains represented in Reading Like Us were not arbitrary selections but rather represent domains students will encounter repeatedly in school and beyond. Recent research supports this principled approach: Kim and Cao's (2025) meta-analysis confirms that students with relevant background knowledge comprehend texts more deeply, retain information more durably, and transfer their learning to new reading experiences.

Our intervention provided students with exposure to substantive, coherent bodies of knowledge organized around genuine curricular domains—a marked contrast to the Active Control condition. There, students encountered unengaging narratives with implausible themes such as a woman who invents invisible paint or stinkbugs engaged in competition. These novelty topics lack curricular relevance, are unlikely to appear in future texts or real-world contexts, and function as isolated narratives that build no transferable knowledge.

A critical question is whether this structural difference generated measurable knowledge gains. We cannot directly answer this because the study included no measure of background knowledge. However, converging evidence from the literature provides grounds for optimism. Research demonstrates that background knowledge is foundational to reading comprehension (Ahmed et al. 2016; Kim and Cao 2025) and that domain knowledge accumulates from exposure to related content over time (Kim et al. 2024; Huxley 2006).

Taken together, these findings underscore that a reading intervention such as RLU addresses far more than fluency and word

19362722, 2026, 2, Downloaded from https://ila.onlinelibrary.wiley.com/doi/10.1002/rrq.70117, Wiley Online Library on [07/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

accuracy. By organizing texts around substantive, curricular domains rather than isolated novelty topics, educators create conditions where novel words and background knowledge accumulate meaningfully over time, supporting stronger comprehension and fostering the development of interconnected knowledge structures that students can draw upon throughout their academic careers and beyond.

5.7 | Limitations

Several limitations inform the interpretations of these data. While the propensity matching identified students in the existing intervention curriculum whose data profile most closely matched students in the RLU condition, randomization was not utilized in this study design. Further, due to the contrasting approaches, it is difficult to isolate any outcomes to specific aspects of the condition. The word accuracy outcomes have been attributed primarily to the explicit instruction of stimuli words and the ORF outcomes are due primarily to the connected text reading. However, given the interrelationships of accuracy and fluency, how much variance in outcomes can be attributed to specific condition aspects remains unknown.

Similarly, word complexity was not tightly controlled in this study. These data do not suggest an optimal degree of complexity for words or texts for promoting word accuracy and text automaticity outcomes. Rather, it appears that practice with difficult words extracted from text benefitted students' accuracy more than with more common, less difficult words, and that repeated practice with more challenging texts did not hinder students' fluency development.

Although a common fluency measure with high technical adequacy was utilized, additional fluency measures, such as a prosody measure, may have provided more robust understandings of the RLU outcomes.

Last, this study compared RLU with a commercial curriculum marketed toward students in Grades 3–12. Although each approach contained similar features such as word study and repeated reading, it is unknown how RLU would compare against other repeated reading protocols reported in the literature (i.e., Fluency Development Lesson, Fluency Oriented Reading Instruction).

5.8 | Future Research

The findings from this study suggest that practices such as explicit multisyllabic word reading practice and repeated reading may play a beneficial role in promoting elementary students' reading outcomes in complex text. An underexplored yet promising feature of the present study is RLU's use of coherent, thematically aligned texts—a shift from disconnected passages often used in interventions. By using coherent, content-aligned texts, RLU affords opportunities both to integrate and advance decoding, fluency, vocabulary, and comprehension. This represents a marked shift from interventions that move arbitrarily from topic to topic in pursuit of specific orthographic patterns and warrants future study. We recommend that future research attend further to *how* content coherence, delivered within the context of an intervention, might contribute to gains while also fostering domain expertise that makes increasingly complex material more accessible.

Moreover, such approaches might prove as a buffer against further widening gaps for students who are behind in reading. Specifically, research should examine whether explicit instruction could enhance comprehension outcomes by helping students synthesize knowledge across complex texts, connect new vocabulary to existing conceptual frameworks, and build coherent mental models that integrate word-level learning with textual meaning. Such instruction might simultaneously support the rate of automatic word recognition.

Future research must also continue to investigate effects of dosages of long word instruction. The strategy used in this study can best be described as economical: Students were engaged in the explicit teaching, guided practice, and independent practice of a peel-off approach, consuming less than ten minutes of instructional time. While this approach yielded statistically significant results, it is unknown how much added value other recommended approaches to long word reading instruction would have provided. More extensive instruction in syllabication strategies may have yielded better results; however, this comes at the sacrifice of more time devoted to reading and knowledge development. Given the finitude of instructional minutes in school contexts, future studies should research the most productive strategies to promote student threshold of accuracy across a variety of texts.

6 | Conclusion

Our findings demonstrate that when third-grade students engaged with challenging, content-rich informational texts supported by strategic multisyllabic word instruction, they not only maintained their reading development but accelerated their progress in foundational accuracy skills. Perhaps most remarkably, this intervention reversed the typical "Matthew Effect," with the lowest-performing students showing the greatest gains when exposed to complex texts rather than being relegated to oversimplified materials. By organizing instruction around coherent content domains rather than disconnected skill-building exercises, we can simultaneously develop foundational reading skills and build the rich background knowledge that distinguishes skilled comprehenders from their striving peers.

Acknowledgments

The authors gratefully acknowledge Devin Kearns for his input on multisyllabic blending, the Center for the School of the Future at Utah State University for their support in preparing this article, and the school-based colleagues who partnered in conducting this research.

Funding

The authors have nothing to report.

Ethics Statement

This research was approved by USU IRB # 14231.

19362722, 2026, 2, Downloaded from https://ila.onlinelibrary.wiley.com/doi/10.1002/rrq.70117, Wiley Online Library on [07/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Word- and text-level data are available upon request from the first author. Student-level data are not available due to restrictions outlined in the Institutional Review Board protocol.

References

Ahmed, Y., D. J. Francis, M. York, J. M. Fletcher, M. Barnes, and P. Kulesz. 2016. "Validation of the Direct and Inferential Mediation (DIME) Model of Reading Comprehension in Grades 7 Through 12." *Contemporary Educational Psychology* 44: 68–82.

Al Otaiba, S., K. McMaster, J. Wanzek, and M. W. Zaru. 2023. "What We Know and Need to Know About Literacy Interventions for Elementary Students With Reading Difficulties and Disabilities, Including Dyslexia." *Reading Research Quarterly* 58, no. 2: 313–332.

Amendum, S. J., K. Conradi, and E. Hiebert. 2018. "Does Text Complexity Matter in the Elementary Grades? A Research Synthesis of Text Difficulty and Elementary Students' Reading Fluency and Comprehension." *Educational Psychology Review* 30, no. 1: 121–151. https://doi.org/10.1007/s10648-017-9398-2.

Barth, A. E., T. D. Tolar, J. M. Fletcher, and D. Francis. 2014. "The Effects of Student and Text Characteristics on the Oral Reading Fluency of Middle-Grade Students." *Journal of Educational Psychology* 106, no. 1: 162–180.

Bates, D., M. Maechler, B. Bolker, et al. 2025. "lme4: Linear Mixed-Effects Models Using 'Eigen' and S4 (Version 1.1-37) [R Package]." https://github.com/lme4/lme4/.

Bowers, P. N., and J. R. Kirby. 2010. "Effects of Morphological Instruction on Vocabulary Acquisition." *Reading and Writing* 23, no. 5: 515–537. https://doi.org/10.1007/s11145-009-9172-z.

Brown, L. T., K. A. J. Mohr, B. R. Wilcox, and T. S. Barrett. 2018. "The Effects of Dyad Reading and Text Difficulty on Third-Graders' Reading Achievement." *Journal of Educational Research* 111, no. 5: 541–553. https://doi.org/10.1080/00220671.2017.1310711.

Center for Research and Reform in Education. n.d. "Corrective Reading–Elementary. Evidence for ESSA." https://www.evidenceforessa.org/program/corrective-reading-elementary/.

Clemens, N. H., D. Simmons, L. E. Simmons, H. Wang, and O. Kwok. 2017. "The Prevalence of Reading Fluency and Vocabulary Difficulties Among Adolescents Struggling With Reading Comprehension." *Journal of Psychoeducational Assessment* 35, no. 8: 785–798. https://doi.org/10.1177/0734282916662120.

Downs, J. D., K. A. J. Mohr, and T. S. Barrett. 2020. "Determining the Academic and Affective Outcomes of Dyad Reading Among Third Graders." *Journal of Educational Research* 113, no. 2: 120–132.

Downs, J., and C. Young. 2025. "Complex Text and Fluency: Evaluating the Read Like Us Protocol With Third and Fourth Grade Students." *Journal of Educational Research* 118, no. 1: 51–62. https://doi.org/10.1080/00220671.2024.2436875.

Downs, J., C. Young, and A. Cole. 2025. "Promoting Fluency Through Challenge: Repeated Reading With Texts of Varying Complexity." *Reading Teacher* 79, no. 3: e70024. https://doi.org/10.1002/trtr.70024.

Engelmann, S., and D. Carnine. 2016. *Theory of Instruction: Principles and Applications*. Rev. ed. NIFDI Press.

Filderman, M. J., and J. R. Toste. 2022. "Effects of Varying Levels of Data Use to Intensify a Multisyllabic Word Reading Intervention for Upper Elementary Students With or at Risk for Reading Disabilities." *Journal of Learning Disabilities* 55, no. 5: 393–407. https://doi.org/10.1177/00222194211048405.

Fuchs, L. S., D. Fuchs, C. L. Hamlett, L. Walz, and G. Germann. 1993. "Formative Evaluation of Academic Progress: How Much Growth Can We Expect?" *School Psychology Review* 22, no. 1: 27–48. https://doi.org/10.1080/02796015.1993.12085636.

Gao, C., S. V. Shinkareva, and R. H. Desai. 2022. "SCOPE: The South Carolina Psycholinguistic Metabase." *Behavior Research Methods* 55: 2853–2884. https://doi.org/10.3758/s13428-022-01934-0.

Gardner, D., and M. Davies. 2014. "A New Academic Vocabulary List." *Applied Linguistics* 35, no. 3: 305–327.

Gersten, R., K. Haymond, R. Newman-Gonchar, J. Dimino, and M. Jayanthi. 2020. "Meta-Analysis of the Impact of Reading Interventions for Students in the Primary Grades." *Journal of Research on Educational Effectiveness* 13, no. 2: 401–427.

Good, R. H., R. A. Kaminski, E. N. Dewey, J. Wallin, K. A. Powell-Smith, and R. Latimer. 2019. *Acadience Reading K-6 Technical Manual*. Acadience Learning Inc.

Hall, C., K. Dahl-Leonard, E. Cho, et al. 2023. "Forty Years of Reading Intervention Research for Elementary Students With or at Risk for Dyslexia: A Systematic Review and Meta-Analysis." *Reading Research Quarterly* 58, no. 2: 285–312.

Hiebert, E. H. 2025. "Unpacking Automaticity: Scaffolded Texts and Comprehension." *Journal of Adolescent & Adult Literacy* 68, no. 4: 369–379. https://doi.org/10.1002/jaal.1394.

Hiebert, E. H., and G. N. Cervetti. 2012. "What Differences in Narrative and Informational Texts Mean for the Learning and Instruction of Vocabulary." In *Vocabulary Instruction: Research to Practice*, edited by J. Baumann and E. Kame'enui, 2nd ed., 322–344. Guilford Press.

Hiebert, E. H., and C. W. Fisher. 2005. "A Review of the National Reading Panel's Studies on Fluency: The Role of Text." *Elementary School Journal* 105, no. 5: 443–460. https://doi.org/10.1086/431888.

Hiebert, E. H., A. P. Goodwin, and G. N. Cervetti. 2018. "Core Vocabulary: Its Morphological Content and Presence in Exemplar Texts." *Reading Research Quarterly* 53, no. 1: 29–49. https://doi.org/10.1002/rrq.183.

Hiebert, E. H., A. Pugh, and D. M. Kearns. 2024. "The Presence and Progression of Rare Vocabulary in Texts Across Elementary Grades and Between Genres." *Education in Science* 14, no. 12: 1314. https://doi.org/10.3390/educsci14121314.

Ho, D., K. Imai, G. King, E. A. Stuart, and N. Greifer. 2025. "MatchIt: Nonparametric Preprocessing for Parametric Causal Inference (Version 4.7.2) [R Package]." https://kosukeimai.github.io/MatchIt/.

Hox, J., M. Moerbeek, and R. Van de Schoot. 2017. *Multilevel Analysis: Techniques and Applications*. Routledge.

Hudson, A., P. W. Koh, K. A. Moore, and E. Binks-Cantrell. 2020. "Fluency Interventions for Elementary Students With Reading Difficulties: A Synthesis of Research From 2000–2019." *Education in Science* 10, no. 3: 52. https://doi.org/10.3390/educsci10030052.

Huxley, A. B. P. 2006. *A Text-Based Intervention of Reading Fluency, Comprehension, and Content Knowledge*. University of Michigan.

Jenkins, J., and K. J. Terjeson. 2011. "Monitoring Reading Growth: Goal Setting, Measurement Frequency, and Methods of Evaluation." *Learning Disabilities Research & Practice* 26, no. 1: 28–35. https://doi.org/10.1111/j.1540-5826.2010.00322.x.

Kearns, D. M. 2020. "Does English Have Useful Syllable Division Patterns?" *Reading Research Quarterly* 55, no. S1: S145–S160. https://doi.org/10.1002/rrq.342.

Kearns, D. M., and E. H. Hiebert. 2022. "The Word Complexity of Primary-Level Texts: Differences Between First and Third Grade in

19362722, 2026, 2, Downloaded from https://ila.onlinelibrary.wiley.com/doi/10.1002/rrq.70117, Wiley Online Library on [07/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Widely Used Curricula." *Reading Research Quarterly* 57, no. 1: 255–285. https://doi.org/10.1002/rrq.429.

Kearns, D. M., and V. M. Whaley. 2019. "Helping Students With Dyslexia Read Long Words: Using Syllables and Morphemes." *Teaching Exceptional Children* 51, no. 3: 212–225. https://doi.org/10.1177/0040059918810010.

Kim, J. S., J. B. Gilbert, J. E. Relyea, et al. 2024. "Time to Transfer: Long-Term Effects of a Sustained and Spiraled Content Literacy Intervention in the Elementary Grades." *Developmental Psychology* 60, no. 7: 1279–1297. https://doi.org/10.1037/dev0001710.

Kim, Y. S. G., and Y. Cao. 2025. "Content Knowledge and Comprehension: A Meta-Analytic Review of Correlational and Causal Associations." *Psychological Bulletin* 151, no. 10: 1219–1244.

Kuhn, M. R., and P. J. Schwanenflugel. 2019. "Prosody, Pacing, and Situational Fluency (Or Why Fluency Matters for Older Readers)." *Journal of Adolescent & Adult Literacy* 62, no. 4: 363–368. https://doi.org/10.1002/jaal.867.

Lane, F. C., Y. M. To, K. Shelley, and R. K. Henson. 2012. "An Illustrative Example of Propensity Score Matching With Education Research: Career and Technical Education Research." *Career & Technical Education Research* 37, no. 3: 187–212. https://doi.org/10.5328/cter37.3.187.

Lane, H. B., V. A. Contesse, N. A. Gage, and M. K. Burns. 2025. "Effect of an Instructional Program in Foundational Reading Skills on Early Literacy Development of Students in Kindergarten and First Grade." *Reading Research Quarterly* 60, no. 1: e607. https://doi.org/10.1002/rrq.607.

Lee, J., and S. Y. Yoon. 2017. "The Effects of Repeated Reading on Reading Fluency for Students With Reading Disabilities: A Meta-Analysis." *Journal of Learning Disabilities* 50, no. 2: 213–224. https://doi.org/10.1177/0022219415605194.

Lovett, M. W., L. Lacerenza, M. De Palma, and J. C. Frijters. 2012. "Evaluating the Efficacy of Remediation for Struggling Readers in High School." *Journal of Learning Disabilities* 45, no. 2: 151–169.

Maki, K. E., and S. Hammerschmidt-Snidarich. 2022. "Reading Fluency Intervention Dosage: A Novel Meta-Analysis and Research Synthesis." *Journal of School Psychology* 92: 148–165. https://doi.org/10.1016/j.jsp.2022.03.008.

McNamara, D. S., A. C. Graesser, P. M. McCarthy, and Z. Cai. 2014. *Automated Evaluation of Text and Discourse With Coh-Metrix*. Cambridge University Press.

Mesmer, H. A. E. 2025. "Exploring the Frequency Contours in Close Reading Texts." *Reading and Writing* 38, no. 1: 1–35. https://doi.org/10.1007/s11145-023-10493-5.

Mesmer, H. A., J. W. Cunningham, and E. H. Hiebert. 2012. "Toward a Theoretical Model of Text Complexity for the Early Grades: Learning From the Past, Anticipating the Future." *Reading Research Quarterly* 47, no. 3: 235–258. https://doi.org/10.1002/rrq.019.

National Governors Association Center for Best Practices, & Council of Chief State School Officers. 2010. "Common Core State Standards for English Language Arts and Literacy in History/Social Studies, Science, and Technical Subjects; Appendix A: Research Supporting Key Elements of the Standards, Glossary of Key Terms." Author.

National Institute for Direct Instruction (NIFDI). 2014. *Achieving Success for Every Student With Direct Instruction*. NIFDI Press. https://www.nifdi.org/docman/new-to-di/new-school-handouts/528-achieving-success-for-every-student-with-direct-instruction/file.html.

National Reading Panel. 2000. *Teaching Children to Read: An Evidence-Based Assessment of the Scientific Research Literature on Reading and Its Implications for Reading Instruction (NIH Publication No. 00-4769)*. U.S. Government Printing Office. https://www.nichd.nih.gov/sites/default/files/publications/pubs/nrp/Documents/report.pdf.

Neitzel, A. J., C. Lake, M. Pellegrini, and R. E. Slavin. 2022. "A Synthesis of Quantitative Research on Programs for Struggling Readers in Elementary Schools." *Reading Research Quarterly* 57, no. 1: 149–179.

O'Connor, R. E., K. M. Bell, K. R. Harty, L. K. Larkin, S. M. Sackor, and N. Zigmond. 2002. "Teaching Reading to Poor Readers in the Intermediate Grades: A Comparison of Text Difficulty." *Journal of Educational Psychology* 94, no. 3: 474–485. https://doi.org/10.1037/0022-0663.94.3.474.

O'Connor, R. E., H. L. Swanson, and C. Geraghty. 2010. "Improvement in Reading Rate Under Independent and Difficult Text Levels: Influences on Word and Comprehension Skills." *Journal of Educational Psychology* 102, no. 1: 1–19. https://doi.org/10.1037/a0017488.

Padeliadu, S., and S. Giazitzidou. 2018. "A Synthesis of Research on Reading Fluency Development: Study of Eight Meta-Analyses." *European Journal of Special Education Research* 3, no. 4: 232–256. https://doi.org/10.5281/ZENODO.1477124.

Przychodzin-Havis, A. M., N. E. Marchand-Martella, R. C. Martella, et al. 2005. "An Analysis of Corrective Reading Research." *Journal of Direct Instruction* 5, no. 1: 37–65.

Rasinski, T. V., N. Padak, W. Linek, and E. Sturtevant. 1994. "Effects of Fluency Development on Urban Second-Grade Readers." *Journal of Educational Research* 87, no. 3: 158–165. https://doi.org/10.1080/00220671.1994.9941237.

Samuels, S. J. 1979. "The Method of Repeated Readings." *Reading Teacher* 32, no. 4: 403–408.

Schmitt, N., X. Jiang, and W. Grabe. 2011. "The Percentage of Words Known in a Text and Reading Comprehension." *Modern Language Journal* 95, no. 1: 26–43. https://doi.org/10.1111/j.1540-4781.2011.01146.x.

Stahl, S. A., and K. M. Heubach. 2005. "Fluency-Oriented Reading Instruction." *Journal of Literacy Research* 37, no. 1: 25–60. https://doi.org/10.1207/s15548430jlr3701_2.

Stanovich, K. E. 1986. "Matthew Effects in Reading: Some Consequences of Individual Differences in the Acquisition of Literacy." *Reading Research Quarterly* 21, no. 4: 360–407. https://doi.org/10.1598/RRQ.21.4.1.

Stebbins, L. B. 1977. *Education as Experimentation: A Planned Variation Model*. Vol. 4. Bloomsbury Publishing PLC.

Swanborn, M. S. L., and K. de Glopper. 1999. "Incidental Word Learning While Reading: A Meta-Analysis." *Review of Educational Research* 69, no. 3: 261–285. https://doi.org/10.3102/00346543069003261.

Therrien, W. J. 2004. "Fluency and Comprehension Gains as a Result of Repeated Reading: A Meta-Analysis." *Remedial and Special Education* 25, no. 4: 252–261. https://doi.org/10.1177/07419325040250040801.

Torgesen, J. K., A. Schirm, L. Castner, et al. 2007. "National Assessment of Title I: Final Report." In *Volume II: Closing the Reading Gap: Findings From a Randomized Trial of Four Reading Interventions for Striving Readers (NCEE 2008–4013)*. U.S. Department of Education, Institute of Education Sciences. https://ies.ed.gov/ncee/pubs/20084013/.

Tortorelli, L. S., J. Z. Strong, and B. E. Anderson. 2024. "Multisyllabic Decoding Achievement and Relation to Vocabulary at the End of Elementary School." *Journal of Experimental Child Psychology* 246: 106018. https://doi.org/10.1016/j.jecp.2024.106018.

Toste, J. R., P. Capin, S. Vaughn, G. J. Roberts, and D. M. Kearns. 2017. "Multisyllabic Word-Reading Instruction With and Without Motivational Beliefs Training for Struggling Readers in the Upper Elementary Grades: A Pilot Investigation." *Elementary School Journal* 117, no. 4: 593–615. https://doi.org/10.1086/691684.

Toste, J. R., P. Capin, K. J. Williams, E. Cho, and S. Vaughn. 2019. "Replication of an Experimental Study Investigating the Efficacy of a Multisyllabic Word Reading Intervention With and Without

19362722, 2026, 2, Downloaded from https://ila.onlinelibrary.wiley.com/doi/10.1002/rrq.70117, Wiley Online Library on [07/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Motivational Beliefs Training for Struggling Readers." *Journal of Learning Disabilities* 52, no. 1: 45–58. https://doi.org/10.1177/00222194187751114.

Traga Philippakos, Z. A., M. F. Quinn, and L. M. Rocconi. 2025b. "Developing Multisyllabic Decoding and Encoding Skills With Upper Elementary Learners: Reporting Two Cycles of Design-Based Research." *Reading & Writing Quarterly* 41, no. 3: 193–217. https://doi.org/10.1080/10573569.2024.2406005.

Traga Philippakos, Z. A., M. F. Quinn, and L. Rocconi. 2025a. "Multisyllabic Instruction for Upper Elementary Learners: A Pilot Study Exploring Feasibility and Effectiveness." *Journal of Research in Childhood Education* 40: 1–24. https://doi.org/10.1080/02568543.2025.2465585.

Uchihara, T., S. Webb, and A. Yanagisawa. 2019. "The Effects of Repetition on Incidental Vocabulary Learning: A Meta-Analysis of Correlational Studies." *Language Learning* 69, no. 3: 559–599. https://doi.org/10.1111/lang.12343.

Vaughn, S., R. Gersten, J. Dimino, et al. 2022. *Providing Reading Interventions for Students in Grades 4–9 (WWC 2022007).* National Center for Education Evaluation and Regional Assistance (NCEE), Institute of Education Sciences, U.S. Department of Education. https://whatworks.ed.gov/.

Wang, Z., T. O'Reilly, M. Flor, B. Beigman Klebanov, and K. Bruce. 2025. "How Much Should I Read? An Analysis of Word Learning Opportunities in Children's Novels." *Reading Research Quarterly* 60, no. 3: e70025. https://doi.org/10.1002/rrq.70025.

What Works Clearinghouse. 2007. *WWC Intervention Report: Corrective Reading.* U.S. Department of Education, Institute of Education Sciences. https://ies.ed.gov/ncee/WWC/Docs/InterventionReports/WWC_Corrective_Reading_070207.pdf.

Zimmerman, B. S., T. V. Rasinski, C. A. Was, et al. 2019. "Enhancing Outcomes for Struggling Readers: Empirical Analysis of the Fluency Development Lesson." *Reading Psychology* 40, no. 1: 70–94. https://doi.org/10.1080/02702711.2018.1555365.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Data S1:** rrq70117-sup-0001-Supinfo.docx.

19362722, 2026, 2, Downloaded from https://ila.onlinelibrary.wiley.com/doi/10.1002/rrq.70117, Wiley Online Library on [07/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Appendix A

Text content description, read like us texts.

Topic categories	*N*	Word count *M* (SD)
Civics & government	5	312.2 (94.49)
Earth & space Science	5	305.0 (19.57)
Geography, societies & culture	5	288.8 (62.58)
Life Science	5	302.6 (26.54)
Music & performing arts	5	269.8 (47.93)
Physical science	5	294 (68.0)
School & family life	5	236.2 (78.74)
Sports, health & safety	5	321.6 (19.72)
Technology & engineering	5	380.0 (68.15)
U.S. history	5	295.2 (52.29)
Visual arts	5	251.4 (57.15)
World history	5	321.8 (47.05)
All topic categories	60	298.2 (36.94)

Appendix B

Music and performing arts topic category, read like us condition.

Text title	Explicitly taught words
The Woodwind Family	Provided
	Encircled
	Included
Friends and Relatives (Strings)	Horizontally
	Distinctive
	Relatives
Musical Instruments: Brass Instruments	Instruments
	Musicians
	Examples
Bass	Orchestras
	Composition
	Experimenting
Jazz	Characteristic
	Improvisation
	Innovative

Appendix C

Text description, corrective reading texts

Topic categories	*N*	Word count *M* (SD)
Five Stink Bugs Have a Contest	2	623.5 (68.5)
Art's Fastball	12	693.83 (44.01)
The President and the Con Man	5	662.4 (16.31)
Hurn the Wolf	10	706.4 (52.27)

19362722, 2026, 2, Downloaded from https://ila.onlinelibrary.wiley.com/doi/10.1002/rrq.70117, Wiley Online Library on [07/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Topic categories	*N*	Word count *M* (SD)
Irma's Invisible Paint	12	697.0 (62.05)
Old Salt, The Retired Sailor	24	808.54 (81.31)

19362722, 2026, 2, Downloaded from https://ila.onlinelibrary.wiley.com/doi/10.1002/rrq.70117, Wiley Online Library on [07/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License