Research Article

Hammill Institute on Disabilities

Journal of Learning Disabilities
2022, Vol. 55(5) 393–407

DOI: 10.1177/00222194211048405
journaloflearningdisabilities.sagepub.com
SAGE

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

Marissa J. Filderman, PhD[1], and Jessica R. Toste, PhD[2]

Abstract

For students who show inadequate response to research-based intervention, intensification of intervention using data-based decision making (DBDM) is recommended. There is a paucity of research on upper elementary students related to the efficacy of (a) word reading interventions and (b) DBDM procedures. This randomized controlled trial examined the differential effects of data use at two timepoints to intensify a multisyllabic word reading intervention for Grades 4 and 5 students with or at risk for reading disabilities (RD). Eighty-eight students were randomly assigned to one of three conditions: initial customization of the intervention (IC-only), initial customization with DBDM (IC + DBDM), or a business-as-usual comparison condition. Results indicated that (a) students in both treatment conditions outperformed the comparison condition on multisyllabic word reading and (b) students in the IC + DBDM condition also outperformed comparison students on decoding. Implications, including the use of DBDM in addition to IC-only as well as methods for decision-making within a small-group context, are discussed.

Keywords

data-based decision making, multisyllabic word reading, intervention, upper elementary

Despite ongoing efforts to improve reading achievement, data collected by the National Center for Education Statistics (NCES) indicate that in the United States 34% of fourth-grade students, and 70% of fourth-grade students with disabilities (SWD), are performing below basic levels in reading (NCES, 2019). Many students who fail to meet basic levels of reading performance in upper elementary grades struggle particularly with word reading (Leach et al., 2003). Students with word reading difficulties in upper elementary grades (i.e., Grades 4 and 5) face unique challenges, as word reading instruction is rarely provided (Leach et al., 2003) and the ability to read independently to learn new material is expected (Council of Chief State School Officers, 2013). In addition, students are exposed to an increasing number of multisyllabic words as they move through the elementary grades (Kearns et al., 2016), many of which carry the meaning of the text (Archer et al., 2003). As many students tend to read multisyllabic words less fluently (Yap & Balota, 2009), this presents a potential barrier to gauging meaning from text, particularly for struggling readers (Kearns et al., 2016).

To support fluent multisyllabic word reading, some students require explicit multisyllabic word reading (MWR) intervention (Toste et al., 2019). Although MWR interventions have been demonstrated to be effective for upper elementary readers (e.g., Toste et al., 2017a, 2019), it has been suggested that 10% to 50% of students do not respond adequately to evidence-based reading interventions in general and require more intensive intervention (D. Fuchs & Fuchs, 2015). One method to intensify intervention is by using student data to adjust instruction over time to meet student needs, a process known as data-based decision making (DBDM; Deno & Mirkin, 1977). There is presently little research on the use of DBDM in reading intervention for upper elementary students (Filderman et al., 2018). The present study tested the efficacy of an evidence-based MWR intervention paired with DBDM for upper

[1]The University of Alabama, Tuscaloosa, USA
[2]The University of Texas at Austin, USA

Corresponding Author:
Marissa J. Filderman, PhD, Department of Special Education and Multiple Abilities, College of Education, The University of Alabama, 902 University Blvd., 201 Graves Hall Box 870232, Tuscaloosa, AL 35487, USA.
Email: mjfilderman@ua.edu

elementary students with or at risk for reading disabilities (RD).

MWR Intervention in Upper Elementary Grades

MWR relies on more complex processes compared with basic word reading. For instance, MWR requires using grapheme-phoneme correspondences to consolidate letters into larger units (Ehri, 2005) and syllables (Perry et al., 2010), stressing syllables (Seva et al., 2009), and knowledge of complex morphology (Nagy et al., 2006). Due to these unique processes, many studies of MWR have been conducted at the secondary level, with evidence suggesting that MWR interventions, in general, are effective in improving reading outcomes for struggling readers at this level (e.g., Bhattacharya & Ehri, 2004; Diliberto et al., 2008; Lenz & Hughes, 1990; Penney, 2002). Although less research has been conducted on students in elementary grades, some evidence exists.

Das-Smaal and colleagues (1996) trained upper elementary students to recognize multi-letter units in words. The treatment resulted in improved fluency (d = 1.42) and accuracy (d = 0.98) in ability to recognize these units in words. Subsequent work extended teaching multi-letter units within the context of structural analysis (i.e., breaking words apart into recognized parts: affixes, roots, syllables). Vadasy and colleagues (2006) conducted two studies with Grade 2 and 3 struggling readers that focused on teaching structural analysis. The first half of the intervention consisted of reading and spelling of monosyllabic words with complex spelling patterns, and practice reading and spelling words with inflectional endings. After this, tutors explicitly taught students how to chunk words into syllables orally and in writing by noticing vowels, identifying and reading syllables, and then putting the parts together. Students then practiced reading orally from grade-level texts. Students in the intervention group outperformed control peers on standardized measures of decoding efficiency (d = 0.86 Study 1; d = 0.70 Study 2), reading fluency (d = 0.82; 1.09), and accuracy (d = 0.71; 1.06).

Recent research has focused on combining these proven structural analysis methods with repeated exposure to words, as repeated exposure to multisyllabic words has been posited to allow for more application of phonics-based rules in context and increase morpheme recognition to reduce the cognitive load required to read multisyllabic words (Heggie & Wade-Woolley, 2017a). To test this theory, two studies were conducted by Toste and colleagues (2017, 2019). Results of the 2017 study with Grade 3 and 4 students indicated significant positive effects compared with a control condition for word reading fluency (g = .73), phonemic decoding efficiency (g = .31), letter-word identification (g = .29), and untimed decoding (g = .30). The 2019 study with fourth and fifth graders indicated significant positive effects on standardized measures of decoding (g = .43), word recognition (g = .29), comprehension (g =.26), and MWR (g = .90). Although studies of MWR interventions demonstrate the potential in improving word reading and fluency outcomes, as previously noted, many students continue to struggle despite receiving such intervention (D. Fuchs & Fuchs, 2015) and require interventions to be intensified.

DBDM in Reading Intervention

To intensify intervention, a widely recommended practice suggests using student data to adjust intervention, a process known as DBDM (National Center on Intensive Intervention [NCII], 2013). To implement DBDM, teachers must first select an intervention that (a) is aligned with student needs, (b) has demonstrated effects on student achievement, (c) targets skills that can be applied across contexts, and (d) follows the principles of explicit, systematic instruction (L. S. Fuchs et al., 2017).

Teachers then monitor progress using general outcome measures (GOM) and/or specific subskill mastery measures (SSMM; Hintze et al., 2006). GOM evaluate student progress toward a long-term skill over time, thereby providing valuable information on retention and generalization of skills (Hintze et al., 2006). Curriculum-based measurement (CBM) is a widely used GOM developed to provide quantifiable information on student progress toward long-term goals of the curriculum (Deno, 1985); the CBM of oral reading fluency (CBM-R) is the most frequently used CBM (Ardoin et al., 2013). SSMM focus on mastery of short-term objectives and are more closely aligned with instruction (Hintze et al., 2006). To promote the utility of both types of measurement, it is recommended that they be used in tandem, each providing valuable information that can be used to adjust instruction (VanDerHeyden & Burns, 2018).

Once a tool is selected, teachers must set a goal by determining an appropriate expected rate of improvement (ROI). Expected ROI on CBM-R of 1.39 (Deno et al., 2001) to 1.67 (Jenkins & Terjeson, 2011) words correct per minute per week have been reported for students in Grades 2 through 6 receiving evidence-based reading intervention. A systematic review found researchers identified the most statistically accurate ROI after collection of data ranging from 10 to 14 weeks with one data point per week (Ardoin et al., 2013). However, Jenkins and Terjeson (2011) found that, although there may be fluctuations in data over a shorter period of time (i.e., up to 8 weeks), the instructional decisions that would be made remain the same in practice. SSMM, however, may be more sensitive to student growth as a result of being proximal to the intervention and decisions may be made over shorter durations (VanDerHeyden & Burns, 2018).

Finally, after a sufficient duration and frequency of monitoring progress, the student's observed ROI is compared with their expected ROI to determine whether sufficient progress has been made. The slope method, wherein a student's observed ROI is compared with an expected ROI, is the most accurate method to evaluate student progress (Ardoin et al., 2013). If it is determined that sufficient progress has not been made, teachers may intensify intervention by increasing expertise of the instructor; decreasing group size; varying the type, delivery, or pacing of treatment (L. S. Fuchs et al., 2017; D. Fuchs & Fuchs, 2015); increasing explicit instruction; individualizing instruction; increasing opportunities for student response and feedback; and/or adjusting the dosage of each intervention component (Vaughn et al., 2012).

Upper Elementary Students. Literature reviews indicate DBDM is effective within the context of reading intervention (Filderman et al., 2018; Jung et al., 2018); however, only one study has looked specifically at DBDM in a reading intervention for upper elementary students. Vaughn and colleagues (2016) conducted a study of Grade 4 struggling readers receiving a multicomponent reading intervention with DBDM. On a monthly basis, fluency data were used to determine risk of falling behind, whereas mastery measures were used to determine the specific skills to target instructionally. Instruction was then targeted within small groups, with the researcher choosing one student per day for whom instruction would be targeted. As the authors did not report significant effects, this study and the literature on DBDM as a whole suggest a need to better understand the use of DBDM with upper elementary students.

Students With Limited English Proficiency. There have been mixed findings on the efficacy of word reading interventions for students with limited English proficiency (LEP). For instance, students with LEP in Grades 2 through 5 who received systematic phonics instruction demonstrated significant improvement in word recognition ($d = 0.84$) but not in decoding or comprehension (Denton et al., 2004). Toste and colleagues' (2019) MWR intervention reported equivalent effects for students with and without LEP across reading outcomes with the exception of spelling. Only one study has explored the effects of DBDM for students with LEP. Students with LEP in middle school grades who received DBDM performed better than students without LEP on a standardized measure of phonemic decoding and similarly on other outcome measures (Vaughn et al., 2011). There is potential to support the learning needs of students with LEP who struggle with word reading, but this evidence is limited.

Challenges of Data Use in Today's Classrooms

Despite evidence supporting DBDM, teachers continue to report that they do not consistently and efficiently use data to inform their practice (Datnow & Hubbard, 2016; Gallagher et al., 2008). Although teachers frequently gather data on their students (Gallagher et al., 2008), many teachers report that they lack the training and time needed to successfully implement DBDM (Datnow & Hubbard, 2016). Given the restraints teachers may face associated with the use of DBDM, it is imperative to explore potential ways to make the use of student data more feasible for teachers to implement. Therefore, the DBDM aspects of the present study were designed to consider practices that may have the potential for supporting teacher implementation.

Different Levels of Data Use. DBDM relies on the ongoing use of data to evaluate the student's response and adjust intervention accordingly (NCII, 2013). Some studies have only used progress monitoring data after several weeks of intervention to make instructional decisions (e.g., Denton et al., 2013; L. S. Fuchs et al., 1989; Vaughn et al., 2016). A range of effects from negative (Vaughn et al., 2016) to small (Denton et al., 2013) to large (L. S. Fuchs et al., 1989) and a range of instructional decisions (e.g., individualization, pacing, and teaching to mastery) make it challenging to ascertain whether using data only after implementation of an intervention is effective. Alternatively, some studies of DBDM have included the use of data at an initial timepoint to provide more aligned intervention from the onset, as well as on an ongoing basis to evaluate student response and adjust intervention accordingly (e.g., Coyne et al., 2013; Vaughn et al., 2012). As these studies demonstrated strong positive effects, it is unclear whether the initial adjustment was effective in isolation or whether there was an added benefit of ongoing data use to further adjust intervention. If data need to only be used at an initial timepoint, the process would be less time-intensive for teachers and would be far more feasible to train teachers on. The present study compares two levels of data use: intervention customized at an initial timepoint only (IC-only) and initial customization plus DBDM (IC+DBDM) adjustments based on students' responsiveness to the intervention.

Decision-Making Within Small Groups. Within schools, it is often a challenge to have flexible grouping due to scheduling restraints and limited personnel. Despite this "real life" challenge, many studies of DBDM examine small-group interventions wherein students change grouping based on their response to instruction (e.g., Coyne et al., 2013; Denton et al., 2013; Mathes et al., 2005). Alternatively, some studies targeted instructional adjustments for one student each lesson (Vaughn et al., 2012), which does not consistently address individual students' needs or the needs of students within the group who are responsive to instruction. For the present study, intervention was delivered in small groups and students remained in their small groups regardless of the DBDM outcome.

Study Purpose

For students who show inadequate response to evidence-based intervention, intensification using DBDM is recommended (NCII, 2013). Although there is evidence to support the use of DBDM, there is a paucity of research for DBDM in the upper elementary grades. Moreover, DBDM remains a challenge for teachers to implement, necessitating research that explores ways to improve the feasibility of DBDM. As such, this experiment addressed four research questions.

Research Questions

Research Question 1 (RQ1): What are the effects of IC-only of an MWR intervention compared with business-as-usual (BAU) on reading outcomes for Grade 4 and 5 students with word reading difficulties?
Research Question 2 (RQ2): What are the effects of IC + DBDM of an MWR intervention compared with BAU on reading outcomes for Grade 4 and 5 students with word reading difficulties?
Research Question 3 (RQ3): What are the effects of IC + DBDM compared with IC-only on reading outcomes for Grade 4 and 5 students with word reading difficulties?
Research Question 4 (RQ4): Are the effects of either treatment moderated by student characteristics, (i.e., initial levels of word reading or LEP)?

Method

Participants

Participants were recruited from four public charter schools in Texas using a two-step screening process. Administrators were asked to nominate Grades 4 and 5 students who demonstrated difficulties with word reading per school-collected data. Nominated students were then administered a screening assessment and qualified for the study if they scored at or below the 25th percentile on at least one of the two subtests of the Test of Word Reading Efficiency–Second Edition (TOWRE-2; Torgesen et al., 2012). In all, 159 students were nominated, and 109 students returned parent consent forms.

The final sample consisted of 88 students (52% female) who met screening criteria and assented to participation: 47 Grade 4 students and 41 Grade 5 students. The average age was 10.24 years (SD = 1.27) and 55% of the sample were identified as having LEP (i.e., basic or intermediate performance on the Texas English Language Proficiency Assessment System). According to school records, 86 students were White and 2 were Black. Their ethnicity was further reported as White Hispanic (n = 81) or non-Hispanic. Nineteen students were receiving special education services, with the majority served under the specific learning disability category (n = 13). Table 1 presents demographic information disaggregated by condition.

To determine baseline equivalence, chi-square analyses were conducted that compared categorical variables between each of the conditions (i.e., LEP, SWD, ethnicity, gender). Then, we conducted one-way analysis of variance (ANOVA) to compare students' scores on each measure between conditions. No baseline differences were detected based on condition. Overall attrition was 2.2% with one student being withdrawn from each treatment condition; conservative standards indicate low attrition and no expected differential attrition (What Works Clearinghouse, 2014).

Study Procedures

The first 5 weeks of intervention were parallel for students in both the IC-only and IC + DBDM conditions, wherein both conditions received the initial customization described below. Following the fifth week of instruction, students in the IC-only condition (n = 28) continued with these customized intervention sessions, whereas students in the IC + DBDM condition (n = 31) had their intervention adjusted based on DBDM procedures (see Figure 1). Intervention sessions were delivered by a tutor in small groups of three to five students during the typical school day in classrooms and intervention spaces. The total intervention dosage consisted of 38 to 40 sessions, for 30 to 40 min per session (M = 35.42; SD = 5.18).

MWR intervention. The MWR intervention used in this study consisted of seven key instructional components as developed and tested by Toste and colleagues (2017a, 2019). The first component of each lesson was a "Warm-Up," during which students reviewed prerequisite vowel patterns (e.g., short vowels, vowel digraphs, diphthongs, r-controlled vowels) necessary to read multisyllabic words through explicit instruction and practice opportunities with the patterns isolated and in words. The next component was "Affix Bank," during which tutors explicitly taught new affixes and sample words that contained the affix, and students added the affix to their "bank" of prefixes and suffixes. Affixes were selected from a list of the most commonly used prefixes and suffixes for students in third through ninth grades (White et al., 1989). The next three components focused on structural analysis of words and

Table 1. Demographic Data for Participants in the MWR Study.

Demographic	IC-only[a] (*n*)	IC + DBDM[b] (*n*)	BAU[c] (*n*)	Total[d] *n*	Total[d] %
Gender					
Male	10	17	15	42	48
Female	17	15	14	46	52
Grade					
4	16	16	15	47	53
5	11	16	14	41	47
Ethnicity					
Hispanic	25	29	27	81	92
Black	2	2	1	5	6
White	0	1	1	2	2
Other	0	0	0	0	0
FRL	28	31	29	88	100
Identified disability	7	5	7	19	22
Limited English proficiency	14	18	16	48	55

Note. IC-only = initial customization; IC + DBDM = initial customization + data-based decision making; BAU = business-as-usual; FRL = free and reduced-price lunch.
[a]$n = 28$. [b]$n = 31$. [c]$n = 29$. [d]$N = 88$.

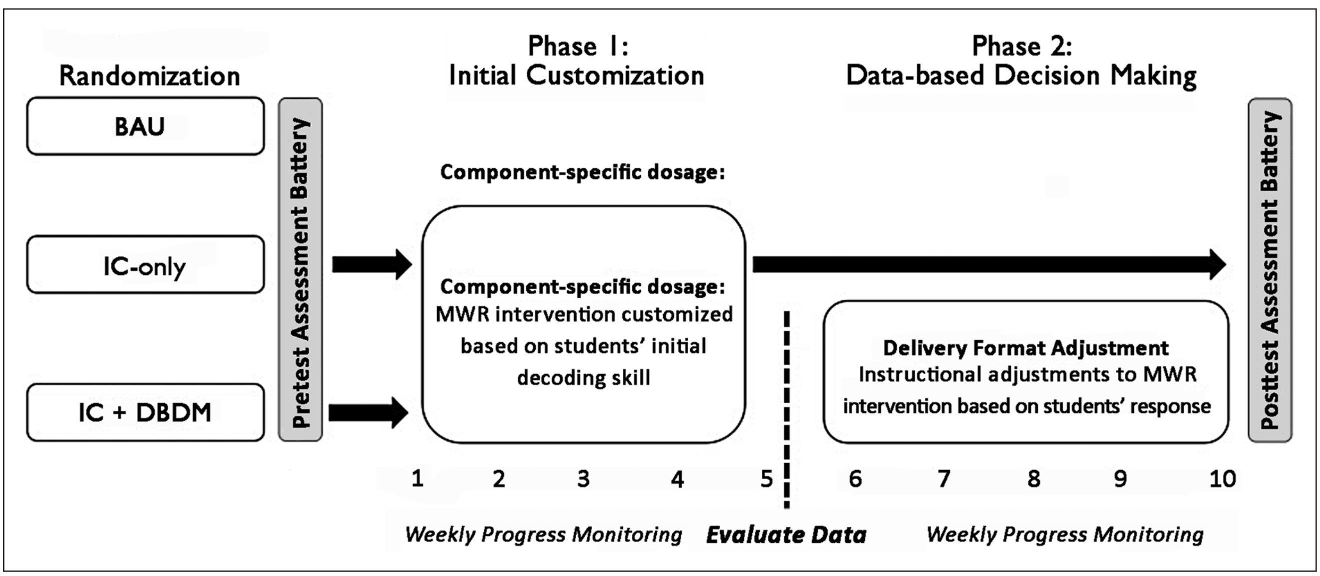

Figure 1. Research design for study on effects of data use on upper elementary reading.
Note. IC-only = initial customization; IC + DBDM = initial customization + data-based decision making; BAU = business-as-usual; MWR = multisyllabic word reading intervention.

automaticity. During "Word Play," students added affixes to targeted base words while playing games. "Beat the Clock" had students practice a peel-off strategy wherein they circled affixes, chorally read affixes in isolation, chorally read whole words, and then completed timed readings. The next component, "Write Word," had students practice encoding multisyllabic words using the targeted affix. The final two components focused on repeated exposure and fluency. During "Speedy Read," students read aloud from a bank of high frequency words that included both monosyllabic words that contained the prerequisite vowel patterns and multisyllabic words. Finally, "Text Reading" consisted of repeated reading of sentences, maze or cloze sentences, and passages. For a full description of the components of the intervention and sample materials, see Toste, et al., 2017b.

Table 2. Time Adjustments for MWR and MWR-D Lesson Customizations.

Component	Description	Time (min)	
		MWR	MWR-D
	Initial customization		
1. Warm-up	Review and practice prerequisite skills, focusing on target vowel patterns in isolation and in words.	2	7
2. Affix Bank	Explicit instruction of new affixes. Students write the affix in their Affix Bank.	3	5
3. Word Play	Alternating games that focus on combining newly learned affixes with base words.	5	6
4. Beat the Clock	Peel-off strategy wherein students circle affixes, read the affixes, read the words, and then complete timed readings of words.	9	8
5. Write Word	Students write two to three syllable words with targeted affix	5	5
6. Speedy Read	Timed reading of high-frequency words, words with prerequisite letter patterns, and multisyllabic words	8	5
7. Text Reading	Repeated readings of sentences and passages	8	4
	Initial customization + DBDM		
1. Warm-up		2	7
2. Affix Bank		3	5
3. Beat the Clock		9	7
4. Word Play		6	5
5. Speedy Read		8	4
6. Adjusted Instruction		12	12

Note. MWR = multisyllabic word reading intervention; MWR-D = multisyllabic word reading intervention with additional decoding instruction; DBDM = data-based decision making.

Differences between treatment conditions

Initial customization. Students in each condition were assigned to customized tutoring groups based on their skill profile, determined by their performance on a diagnostic assessment (Houghton Mifflin, 2001). All intervention components were included in the customized tutoring groups, with time in each component adjusted based on student need. Students who demonstrated mastery of monosyllabic decoding patterns received the base MWR intervention. Students who did not demonstrate mastery of these prerequisite skills (i.e., low decoders, n = 41 total; n = 22 IC-only; n = 19 IC + DBDM) received adjusted MWR intervention lessons, with time in each component adjusted so that there was more focus on decoding instruction (MWR-D; see Table 2).

MWR lessons. MWR lessons focused on applied practice, with less focus on explicit decoding instruction. More time with applied practice in the "Speedy Read" and "Text Reading" components allowed students to apply newly learned skills at an increased rate to promote the generalized fluent reading of multisyllabic words. The lessons followed the scope and sequence of the original intervention developed by Toste and colleagues (2017a). Warm-ups progressed from short to long vowels, r-controlled vowels, diphthongs, and vowel flexing. The latter half of intervention reviewed each of these remedial skills. Affixes targeted were grouped each day based on similarity of meaning (e.g., dis-, un-, sub) or morphological features (e.g., inflectional suffixes).

MWR-D lessons. MWR-D lessons focused on explicit instruction and decoding practice, with less focus on applied practice. Increased time in "Warm-Up" allowed for additional instruction and practice of remedial skills, increased time in "Affix Bank" allowed for increased explicit teaching of new affixes, and increased time in "Beat the Clock" allowed for more opportunities to use strategies on multisyllabic words. The scope and sequence for MWR-D lessons was aligned with the MWR lessons, with the exception of warm-up remedial skills targeted. Students started with the same progression of vowels and then, rather than review, other remedial skills were taught (i.e., common blends, digraphs, trigraphs).

Initial customization with DBDM. Following the fifth week of instruction, students in the IC + DBDM condition had their intervention adjusted based on DBDM procedures.

Identifying response. Adequate or inadequate response was assessed by the primary author using CBM-R and SSMM scores collected by tutors weekly. First, an expected ROI of 1.5 words per minute per week was chosen (Deno et al., 2001; Jenkins & Terjeson, 2011). After collecting baseline data, the following formula was used to calculate outcome goals: Outcome = Baseline + (5 weeks × 1.5 wcpm). After 5 weeks of data collection, ordinary least squares regression was used to determine a line of best fit for the data; then, the expected ROI was visually compared with the observed ROI (Ardoin et al., 2013). Students whose observed ROI was below the expected ROI on both measures, or who were below the expected ROI on SSMM but above on CBM-R, were considered inadequate responders. Students who were below the expected ROI on CBM-R but above the expected ROI on the SSMM, or who were above the slope on both measures, were considered adequate responders because the SSMM may have been more sensitive to growth (VanDerHeyden & Burns, 2018). Finally, administration of a mid-point diagnostic assessment provided information that allowed for selection of remedial skills to target during explicit instruction. The diagnostic assessment was administered at the mid-point because the decoding skills assessed were those targeted explicitly in the intervention, and thus could be expected to have shifted during the course of intervention (see "Data Collection and Measures" section).

Instructional adjustments. For 12 min of each lesson, students who demonstrated inadequate response (n = 18) received additional teacher-led explicit instruction, whereas students who demonstrated adequate response (n = 13) engaged in peer-led practice. These changes to the delivery format were in addition to the initial customizations to maintain the instructional emphases of the customized lessons; however, "Write Word" and "Text Reading" were removed from both customizations to allow time for the instructional adjustments (see Table 2). "Write Word" was removed because it was the only component that would not detract from time spent reading, while "Text Reading" was removed because students spent time with connected text reading within the instructional adjustments. Students were trained on the routines of the adjusted lessons during the two instruction blocks immediately following the decision.

Teacher-led instruction consisted of explicit instruction on skills for which students lacked proficiency based on the diagnostic assessment, as well as increased exposure to multisyllabic words with immediate, corrective feedback. First, explicit instruction of decoding rules and patterns was provided in mini-lessons, followed by modeling and practice. A scope and sequence was developed for the mini-lessons (i.e., short vowels, blends, digraphs, vowel teams, multisyllabic word reading). A targeted scope and sequence of lessons to teach from the mini-lesson book was then provided for each tutoring group, resulting in eight unique mini-lesson sequences with 18 mini-lessons each. The mini-lessons were more intensive because of the increased explicit, targeted instruction (Vaughn et al., 2012). Next, flashcards were used with multisyllabic words containing affixes that were taught in the main part of the lesson. As the words included the same affixes taught in the intervention sessions, students had an additional opportunity for review of the lesson content. Finally, students read from the connected text of the main lesson to apply skills in context. The flashcards and connected text reading followed the same sequence for the teacher-led instruction. Flashcards and connected text reading were more intensive because there was a decreased teacher–student ratio and increased immediate targeted feedback and error correction (e.g., Vaughn et al., 2012).

Peer-led practice was considered less intensive because students did not receive teacher feedback or explicit instruction during this time, but students still had guided practice as well as immediate feedback and error correction from peers (Greenwood et al., 1989). Peer practice was modeled after the "Partner Reading" component of Peer-Assisted Learning Strategies (PALS; D. Fuchs et al., 1997). Using decodable books with multisyllabic words, students identified and chunked multisyllabic words into their word parts before reading. Then, students read aloud for 4 min each, followed by retelling the story for 1 min. Their partner acted as a coach and helped the student notice and correct errors while reading.

Business-as-usual. Students in treatment conditions received the present study's intervention during their school's standard intervention block. Based on interviews with classroom teachers, during the standard intervention block comparison students participated in independent reading, small-group teacher-designed reteach of comprehension skills, small-group Leveled Literacy Intervention (Fountas & Pinnell, 2009), online independent work with IXL Learning, assignments with the teacher related to preparation for the state standardized assessment, or homework. Teachers were also asked to describe how they used data to inform instruction and intervention. They reported primarily using informal sources of data (e.g., exit tickets, mastery measures, unit tests) to determine which students were in need of small-group support, to adjust small skill-based groups, and to select skills that needed to be taught again.

Implementation Fidelity

All tutors worked with groups in each condition. All intervention sessions were audio-recorded and uploaded on a weekly basis. Upon conclusion of intervention, 30% of each tutor's audio recordings within each condition were selected at random and scored for fidelity by a trained research assistant (RA). The RA met a minimum of 90% interrater agreement with the primary investigator prior to scoring remaining audios. Fidelity was assessed with a checklist that included adherence to general and intervention-specific features for each condition. The number of items observed was divided by the total number of items in each session to calculate fidelity. In addition, a rubric was used that assessed quality of instructional pacing, error corrections, and behavioral corrections, scored on a Likert-style scale of 0 (*poor*) to 3 (*high*).

The overall mean implementation adherence score across components, conditions, and interventionists was 89.52% (SD = 7.98%). Mean implementation adherence scores for IC-only and IC + DBDM were 88.80% (SD = 8.88) and 90.02% (SD = 7.31), respectively. The overall mean quality score across conditions, sessions, and tutors was 2.98 (SD = 0.18). Across conditions, the scores were 2.96 (SD = 0.19) for pacing, 3.0 (SD = 0) for corrections, and 2.99 (SD = 0.25) for behavior management. Mean quality scores for IC-only and IC + DBDM were 3.00 (SD = 0) and 2.97 (SD = 0.17), respectively.

Data Collection and Measures

All participants completed a measure of word reading fluency during screening. Pretest and post-test assessment batteries consisted of measures of (a) word reading and (b) fluency and comprehension. During intervention, progress was monitored weekly. Finally, students were administered a diagnostic assessment prior to receiving intervention, and students in the IC + DBDM condition were administered this assessment again at the mid-point of the study. A trained team of RAs along with tutors administered assessments individually to students with the exception of the Test of Silent Reading Efficiency and Comprehension (TOSREC), which was administered in small groups. All assessments were double scored by two assessors prior to data entry.

Screening. The TOWRE-2 is a standardized, norm-referenced test of word reading that consists of two subscales, Sight Word Efficiency and Phonemic Decoding Efficiency, which measure the amount of real and pronounceable nonsense words a student can read in 45 s, respectively. Test–retest reliability ranges from .89 to .93 for students in Grades 1 through 12 (Torgesen et al., 2012).

Pre- and Post-test Assessment Battery

Word reading. Students completed two subtests from the Woodcock Reading Mastery Tests–Third Edition (WRMT-3; Woodcock, 2011). The Word Attack (WA) and Word Identification (WID) subtests are untimed measures that assess decoding and word recognition, with test–retest reliability for Grades 3 to 8 of .88 and .92, respectively (Woodcock, 2011).

Fluency and comprehension. The WRMT-3 Passage Comprehension (PC) subtest uses a cloze passage structure to measure comprehension. Test–retest reliability for this measure is reported to be .90 for Grades 3 to 8 (Woodcock, 2011). The TOSREC (Wagner et al., 2010) was also used to measure fluency and comprehension. During this 3-min assessment, students silently read sentences and determine whether the sentence is truthful or not. Students are scored based on the number of total correct responses, with incorrect responses subtracted from the total. Alternate form reliability coefficients are reported at .86 for Grade 4 and .89 for Grade 5 (Wagner et al., 2010).

Progress monitoring. The Fastbridge CBM-R measures the words students can read correctly during a 1-min reading of an instructional-level passage (Christ et al., 2014). Instructional level was determined as the highest-level passage in which a student could read 10 to 50 words with 90% accuracy (Stecker & Lembke, 2011) and was determined during administration of the pretest battery. The alternate form reliability ranges from .87 to .92 for Grades 2 through 3 level passages, and .92 to .95 for Grades 4 through 6 level passages, both administered to students across elementary grade levels (i.e., Grades 1–5). The second subtest of the Big Word Reading Test (Toste et al., 2019), which consists of 94 multisyllabic words with learned affixes, was modified for the SSMM. Internal reliability is reported to be .94 to .93 (pre to post). For the SSMM, a subset of 50 words, controlled for a number of syllables in the words, were randomly selected and ordered to create a unique word list each week. Students were scored on the number of words read correctly in 1 min.

Diagnostic assessment. The reading and decoding portion of the Houghton Mifflin (2001) was used to assess skills in need of remediation. This 5-min section measured knowledge of short vowel sounds in consonant-vowel-consonant words, digraphs, trigraphs, consonant blends, r- and l-controlled vowels, vowel diphthongs, and MWR with various syllable types (closed; silent e; open, closed; open, silent e; consonant -le; r-controlled, vowel team).

Data Analysis

To account for clustering in the data and the potential increase in Type 1 error, parameters were estimated as multilevel (Hox, 2002) using Stata statistical software (StataCorp, 2019). Benjamini–Hochberg corrections were used with a false discovery rate of .05 to account for the multiple contrasts (Benjamini & Hochberg, 1995). To determine appropriate modeling, specifically the need for models that account for two-level partial nesting (Roberts & Roberts, 2005) or cross-classification (Luo et al., 2015), unconditional models as follows were run:

$$\text{Level1} \quad Y_{ij} = \beta_{0j} + e_{ij}$$
$$\text{Level 2} \quad \beta_{0j} = \gamma_{00} + u_{0j}$$

where Y_{ij} represents an individual student's reading outcome at post-test, β_{0j} is the intercept representing the mean post-test score for all participants at post-test in classroom j, γ_{00} represents the grand mean intercept, and u_{0j} is the deviation of the mean of classroom j from the grand mean. Significant school-level variance with intraclass correlation coefficients (ICCs) ranging from .00 to .24 substantiated the need for multilevel models (Hedges & Hedberg, 2007). As tutor-level ICCs ranged from <.00 to .02, cross-classification and partial nesting were not considered within the remaining models.

Next, fully specified hypothetical models were analyzed. In the first and second models, post-test scores were modeled as a function of the student's grand mean centered pretest score, as well as the treatment students received. Specifically, the hypothetical model was as follows:

Level 1 $$Y_{ij}=\beta_{0j}+\beta_{1j}\text{IC-only}_{ij}+\beta_{2j}\,\text{IC+DBDM}_{ij}..+\beta_{3j}\,(\text{pre}_{ij-}\overline{\text{pre}}_{ij-j})+e_{ij}$$

Level 2 $$\beta_{0j}=\gamma_{00}+u_{0j}$$

$$\beta_{1j}=\gamma 10$$

$$\beta_{2j}=\gamma 20$$

$$\beta_{3j}=\gamma 30$$

where Y_{ij} represents an individual student's post-test score, β_{0j} is the expected class average, β_{1j} represents the estimated slope for students who received IC-only compared with the comparison condition, β_{2j} represents the estimated slope for students who received IC + DBDM compared with the comparison condition, and β_{3j} represents the expected pretest slope on the outcome of interest. The second model was identical to the first model but β_{1j} compared IC + DBDM with IC-only, with β_{2j} being the comparison condition. Results from the hypothetical models indicated significant variance in the intercept across classrooms for Woodcock Reading Mastery Test–WID ($\tau_{oj}=1.18$), WA ($\tau_{oj}=2.87$), and PC ($\tau_{oj}=14.70$), which substantiated the need for random effects of intercept. Although it was hypothesized that pretest may randomly vary by classroom, random slopes for pretest were non-different from zero across outcomes and thus modeled as fixed. Then, moderated effects models were run that added upon the conditional models with a full set of hypothesized explanatory variables (i.e., initial levels of word reading performance, LEP). To assess moderation, interactions were run between each treatment and each moderator of interest.

Results

The following sections describe our model tests and results for each of the research questions based on final models. Table 3 presents the descriptive statistics for pre- and post-test by condition, and Table 4 presents estimated treatment effects. Table S1 in the online supplemental materials presents pre- and post-test scores disaggregated by LEP and LD identification, as well as by MWR and MWR-D customizations.

Empirical Main Effects Models

To address RQ1, we examined the effects of IC-only compared with BAU on reading outcomes. Students in both IC-only and BAU conditions performed similarly on the WRMT-3 word reading subtests of WID, $\beta = -0.65$, $t(85) = -0.26$, $p = .81$, WA, $\beta = 2.99$, $t(85) = 0.99$, $p = .40$, and PC, $\beta = 3.29$, $t(85) = 1.06$, $p = .37$. Students in the IC-only condition performed significantly lower on the TOSREC, $\beta = -8.30$, $t(85) = -4.28$, $p = .00$, than students in BAU. Students in both conditions performed similarly on the fluency measure of CBM-R, $\beta = -6.99$, $t(85) = -1.55$, $p = .16$, but students in the IC-only condition significantly outperformed students in BAU on the SSMM, $\beta = 7.14$, $t(85) = 4.40$, $p = .00$.

Next, to address RQ2, we examined the effects of IC + DBDM compared with BAU. Students in both IC + DBDM and BAU conditions performed similarly on the WRMT-3 word reading subtest of WID, $\beta = -0.57$, $t(85) = -0.29$, $p = .81$, but IC + DBDM significantly outperformed students in BAU on WA, $\beta = 7.68$, $t(85) = 2.66$, $p = .00$. Students performed similarly on the WRMT-3 PC subtest, $\beta = 1.05$, $t(85) = .35$, $p = .80$, and students in IC + DBDM performed significantly lower than students in the comparison condition on the TOSREC, $\beta = -4.50$, $t(85) = -2.37$, $p = .03$. Although students performed similarly on the CBM-R, $\beta = -3.42$, $t(85) = -0.79$, $p = .50$, IC + DBDM significantly outperformed BAU on the SSMM, $\beta = 9.68$, $t(85) = 6.22$, $p = .00$.

RQ3 examined effects of IC + DBDM compared with IC-only. Students in both treatment conditions performed similarly on WID, $\beta = 0.08$, $t(85) = 0.04$, $p = .97$; WA, $\beta = 4.69$, $t(85) = 1.57$, $p = .16$; PC, $\beta = -2.24$, $t(85) = -0.73$, $p = .53$; TOSREC, $\beta = 3.81$, $t(85) = 1.99$, $p = .07$; CBM-R, $\beta = 3.56$, $t(85) = 0.80$, $p = .50$; and SSMM, $\beta = 2.54$, $t(85) = 1.60$, $p = .16$.

Moderated Effects Models

RQ4 considered whether the effects of treatment were moderated by student characteristics. Initial word reading significantly moderated the effects of the intervention for the IC-only condition, $\beta = -.47$, $t(85) = -2.74$, $p = .01$, for TOSREC. An adapted Johnson–Neyman technique indicated that clinical significance began for students who performed at or above −5.02 points below the mean (standard score of approximately 72). That is, the effects of the IC-only treatment were poorer for students with a standard score of 72 or above on the TOSREC. There were no other significant moderators of the effects based on initial word reading or LEP status.

Table 3. Pre- and Post-Test Descriptive Statistics for the Full Sample.

Measure	Pretest IC-only[a] M (SD)	Pretest IC + DBDM[b] M (SD)	Pretest BAU[c] M (SD)	Post-test IC-only[a] M (SD)	Post-test IC + DBDM[b] M (SD)	Post-test BAU[c] M (SD)
TOWRE	76.15 (10.60)	81.48 (8.86)	77.24 (9.55)			
WID	81.23 (11.81)	85.65 (11.94)	83.38 (15.61)	84.96 (12.81)	88.77 (9.66)	87.41 (17.83)
WA	85.04 (15.15)	88.48 (12.24)	88.14 (16.63)	89.15 (14.73)	96.32 (13.72)	88.38 (16.62)
PC	76.65 (10.69)	79.23 (10.95)	76.45 (11.63)	81.92 (15.63)	81.52 (10.88)	78.62 (14.00)
TOSREC	76.69 (9.69)	79.35 (9.90)	74.03 (9.92)	74.04 (8.40)	79.74 (10.66)	80.45 (10.51)
CBM-R	88.27 (32.60)	95.77 (29.35)	86.03 (32.24)	90.65 (32.27)	100.68 (28.61)	95.72 (34.11)
SSMM	19.46 (11.15)	20.61 (9.66)	19.59 (10.12)	32.35 (11.62)	35.87 (9.89)	25.31 (10.48)

Note. IC-only = initial customization; IC + DBDM = initial customization + data-based decision making; BAU = business-as-usual; TOWRE = Test of Word Reading Efficiency; WID = Woodcock Reading Mastery Test–Word Identification; WA = Woodcock Reading Mastery Test–Word Attack; PC = Woodcock Reading Mastery Test–Passage Comprehension; TOSREC = Test of Silent Reading Efficiency and Comprehension; CBM-R = Curriculum-Based Measurement–Reading; SSMM = specific subskill mastery measurement.
[a]$n = 26$. [b]$n = 31$. [c]$n = 29$.

Discussion

The purpose of this study was to estimate the effects of data use to customize and target an intervention for students with or at risk for RD in the upper elementary grades. Significant positive effects were detected for both treatments compared with BAU on multisyllabic word reading, and additionally for IC + DBDM on decoding. Students in both treatment conditions performed significantly worse on comprehension efficiency than BAU, with initial word reading moderating these effects for students in IC-only. There were no other significant moderators of effects based on initial word reading or LEP. Potential explanations for these findings will be discussed in detail in the following sections.

To What Extent Does Data Use Improve Reading Outcomes?

Findings related to word reading. Results indicated that students in the IC-only ($g = .47$) and IC + DBDM ($g = 1.12$) conditions both significantly outperformed their peers in the BAU condition on the researcher-designed SSMM. This finding aligns with the study of the base intervention tested by Toste and colleagues (2019) with a sample of Grades 4 and 5 students. Students in upper elementary grades are noted to be less responsive to intervention; thus, it is quite challenging to improve reading outcomes for students in these grades (Vaughn & Wanzek, 2014). This is particularly meaningful because multisyllabic words represent a majority of new words upper elementary students are exposed to (Kearns et al., 2016) and often carry the meaning of the text (Archer et al., 2003). Improving upper elementary students' ability to read multisyllabic words is no small feat and may have major impacts on text accessibility.

Toste and colleagues reported a positive effect on WID ($g = .26$), whereas our observed effects were negligible. Often, tutors conducted progress monitoring while students were engaged in sight word reading in pairs (i.e., "Speedy Read"). This intervention component is directly aligned with word recognition. Because the students worked in pairs with less teacher feedback and error correction, this may have detracted from the effects on WID. However, students in the IC + DBDM condition significantly outperformed BAU on WA ($g = .87$), whereas students in IC-only also outperformed comparison peers ($g = .26$) but this finding was not significant. This finding aligns with the findings noted by Toste and colleagues (2019; $g = .43$), with a substantially larger effect noted for students in the IC + DBDM condition. One possible explanation for the significant and larger effects associated with IC + DBDM is the targeted decoding mini-lessons within teacher-led instruction that were designed to remediate skill deficits. Intensification of an intervention using targeted, explicit instruction may, therefore, be a particularly beneficial adjustment within the DBDM process. This would lend support to the importance of making qualitative (e.g., changing content delivery method, increasing opportunities to respond) rather than quantitative (e.g., decreasing group size, increasing dosage) instructional adjustments when intensifying intervention (NCII, 2013). This is further supported by findings reported in a systematic review of long-term reading interventions for struggling readers in 4th through 12th grades, which indicated that quantitative intervention features did not improve the overall effects of reading interventions (Wanzek et al., 2013). Unlike previous studies, the present study made *only* qualitative adjustments, thus lending empirical evidence to this method for intensifying intervention with struggling upper elementary readers.

Findings related to connected text reading fluency and comprehension efficiency. Findings indicated that students

Table 4. Estimated Treatment Effects.

Fixed effect	Estimate	95% CI	*SE*	*t*	*p*-value	ES
WID	0.82		0.07			
Pretest		[0.69, 0.96]	0.07	11.87	.00	
IC-only vs. BAU	−0.65	[−5.09, 3.78]	2.26	−0.26	.81	−0.02
IC + DBDM vs. BAU	−0.57	[−4.82, 3.68]	2.17	−0.29	.81	−0.10
Constant	87.75	[84.48, 91.01]	1.67	52.66	.00	
IC + DBDM vs. IC-only	0.08	[−4.32, 4.49]	2.25	0.04	.97	−0.08
Constant	87.10	[83.65, 90.54]	1.76	49.55	.00	
WA	0.69		0.08	8.26		
Pretest		[0.52, 0.85]			.00	
IC-only vs. BAU	2.99	[−2.94, 8.92]	3.03	0.99	.40	0.26
IC + DBDM vs. BAU	7.68	[2.03, 13.33]	2.88	2.66	.00	0.87
Constant	88.19	[83.73, 92.65]	2.27	38.79	.00	
IC + DBDM vs. IC-only	4.69	[−1.16, 10.54]	2.99	1.57	.16	0.60
Constant	91.18	[86.49, 95.87]	2.39	38.09	.00	
PC	0.61		0.11	5.31		
Pretest		[0.38, 0.83]			.00	
IC-only vs. BAU	3.29	[−2.81, 9.39]	3.11	1.06	.37	0.10
IC + DBDM vs. BAU	1.05	[−4.82, 6.92]	3.00	0.35	.80	0.02
Constant	80.70	[74.92, 86.48]	2.95	27.37	.00	
IC + DBDM vs. IC-only	−2.24	[−8.28, 3.81]	3.08	−0.73	.53	−0.09
Constant	83.98	[78.02, 89.95]	3.04	27.59	.00	
TOSREC						
Pretest	0.71	[−.56, 0.87]]	0.08	8.94	.00	
IC-only vs. BAU	−8.30	[−12.11, –4.50]	1.94	−4.28	.00	−1.21
IC + DBDM vs. BAU	−4.50	[−8.21, –0.78]	1.89	−2.37	.03	−0.82
Constant	82.39	[79.75, 85.02]	1.34	61.28	.00	
IC + DBDM vs. IC-only	3.81	[0.06, 7.54]	1.91	1.99	.07	0.39
Constant	74.08	[71.34. 76.83]	1.40	52.87	.00	
CBM-R	0.86		0.06	14.70		
Pretest		[0.75, 0.97]			.00	
IC-only vs. BAU	−6.99	[−15.84, 1.85]	4.51	−1.55	.16	−0.57
IC + DBDM vs. BAU	−3.42	[−11.95, 5.11]	4.35	−0.79	.50	−0.41
Constant	99.32	[93.23, 105.42]	3.11	31.92	.00	
IC + DBDM vs. IC-only	3.57	[−5.18, 12.32]	4.46	0.80	.50	0.16
Constant	92.33	[85.91, 98.76]	3.28	28.17	.00	
SSMM	0.86		0.06	13.31		
Pretest		[0.73, 0.98]			.00	
IC-only vs. BAU	7.14	[3.96, 10.33]	1.63	4.40	.00	0.47
IC + DBDM vs. BAU	9.68	[6.63, 12.73]	1.56	6.22	.00	1.12
Constant	25.59	[23.40, 27.78]	1.12	22.90	.00	
IC + DBDM vs. IC-only	2.54	[−0.60, 5.68]	1.60	1.59	.16	0.65
Constant	32.74	[30.42, 35.05]	1.18	27.73	.00	

Note. All *p* values were corrected using Benjamini–Hochberg corrections. CI = confidence interval; WID = Woodcock Reading Mastery Test–Word Identification; IC-only = initial customization; BAU = business-as-usual; IC + DBDM = initial customization + data-based decision making; WA = Woodcock Reading Mastery Test–Word Attack; PC = Woodcock Reading Mastery Test–Passage Comprehension; TOSREC = Test of Silent Reading Efficiency and Comprehension; CBM-R = Curriculum-Based Measurement–Reading; SSMM = specific subskill mastery measurement; ES = Effect Sizes.

did improve their fluency in reading multisyllabic words in isolation; however, they did not perform as well with measures that involved connected text reading. On the TOSREC, students in IC-only ($g = -1.21$) and IC + DBDM ($g = -.82$) performed significantly worse than their peers in the BAU condition. As the TOSREC requires fluency to comprehend efficiently, it is possible that this finding was related to students' overall reading fluency. The moderate, though not significant, negative effects observed on the CBM-R when comparing IC + DBDM ($g = -.41$) and IC-only ($g = -.57$) to BAU may support this explanation.

Verbal efficiency theory may help to explain findings related to fluency and comprehension efficiency. This theory posits that dysfluent reading affects comprehension because the energy students spend decoding words clogs the working memory and thus limits the space to engage with the text for the purpose of making meaning (Perfetti, 1985). Although the intervention developed by Toste and colleagues (2017a, 2019) was designed to promote automaticity and reduce cognitive load while reading multisyllabic words, this may not have translated into fluent reading in the present study. This is because, while the base intervention included substantial time with connected text reading, there was less time spent on this segment for students who received increased decoding lessons (MWR-D). Moreover, students in BAU were often engaged in connected text reading in their school-based interventions, which compounds the negative effects that may have occurred as a result of the instructional focus. Connected text reading has been found to be highly effective for students with or at risk for RD (Swanson et al., 2009). It is, therefore, possible that the relatively reduced exposure to multisyllabic words in context could have affected fluency. Verbal efficiency theory would indicate students were expending more energy on reading accurately at the expense of fluency and comprehension. Although Toste and colleagues did not measure fluency, our findings align with theirs on PC. As PC is related to fluency, it is possible findings would have aligned on fluency, providing further support for this theory.

Potential value-added of DBDM. The central focus of this study was to compare interventions with varying levels of data use. Both conditions were customized initially using diagnostic data, but ongoing progress monitoring data were used to target instruction for students in the IC + DBDM condition. Findings indicated that students in IC + DBDM did not significantly outperform students in IC-only. However, effect sizes indicate that students in IC + DBDM had greater gains than students in IC-only on the SSMM (g = .65), WA (g = .60), and TOSREC (g = .39). Although not significant, these moderate to large effect sizes may support the use of DBDM in upper elementary grades within the context of a word reading intervention, particularly when the relatively short duration of the DBDM treatment is taken into account (Kraft, 2020).

Do Student Characteristics Influence Intervention Effects?

In line with the previous study of the base intervention, there was no significant moderation of the effects of treatment on reading outcomes based on LEP status (Toste et al., 2019). Initial word reading did moderate effects for performance on the TOSREC; specifically, the effects of IC-only treatment compared with BAU were poorer for students who performed relatively higher on initial word reading (i.e., standard score of 72 and above on TOWRE-2). In theory, students with higher levels of initial word reading would likely have been placed in MWR tutoring groups; however, based on the diagnostic assessment of untimed decoding skills, only four students were in one MWR tutoring group in this condition. In contrast, 15 students in IC-only scored above the 72 cut-off that was associated with poorer TOSREC outcomes. These students may have benefited from the additional embedded text practice of MWR lessons. The fact that initial word reading was not a significant moderator of the effects of IC + DBDM also supports the idea that the extra fluency practice after the first 5 weeks may have supported comprehension efficiency for students with higher initial word reading. Although comprehension was not the focus of this intervention, it is essential to determine how to support the scope of reading needs of struggling learners who may have different initial levels of reading.

Study Limitations and Future Directions

The first limitation of the present study is that it may have been underpowered. Power analyses called for 12 teachers, with five students per teacher; however, the four schools recruited each had only one reading teacher. We, therefore, recommend caution when interpreting results; however, the present study presents a promising preliminary investigation of this important topic and warrants further research to confirm results. A second limitation is related to duration and dosage. Due to school scheduling, students received 38 to 40 sessions, for 30 to 40 min each session. It is possible that the time during which the two treatments differed was also too brief to determine effects; thus, future research over a longer duration is needed. An additional limitation of the present study is that of generalizability; students were mostly Hispanic and LEP, which limits generalizability to other populations of students. Another limitation is that, although tutors implemented progress monitoring and instructional adjustments, data-based decisions were made by the primary author. We note, however, that the ultimate goal of this study was to better understand the effects of frequency of data-based decisions and student response to establish an evidence base for practices that may inform teacher practice. Future research is needed to determine the effects of teacher implementation.

Relatedly, future research may evaluate the effects of each of the instructional adjustments made within the DBDM condition to further support teacher implementation. Finally, the present study utilized a researcher-developed measure of MWR as there are no standardized measures of this skill. It has been previously demonstrated that smaller effects are found on standardized measures compared with researcher-developed measures, a finding that is apparent in the present study (e.g., Edmonds et al., 2009). Future research is needed to develop a standardized MWR measure to further support findings.

Implications for Practice

Findings from this study indicate that the use of data for upper elementary students, both at an initial timepoint and on an ongoing basis, improved word reading outcomes, particularly those directly associated with the intervention. In addition, the moderate effects of IC + DBDM above and beyond those observed as a result of IC-only indicate that a customized ongoing approach may support this population of students. Although there were limited findings for transfer of these skills to fluency and comprehension, this provides a promising start to understanding how to move the dial on this struggling population of students.

The current study has implications for classroom teachers regarding the implementation of data-based decisions within intervention groups. First, teachers may group students based on their performance on a diagnostic assessment—that is, teachers may initially customize intervention by creating skill-based groups for students based on their mastery of a set of criteria. Second, teachers may consider ways in which they might make progress monitoring more efficient, such as pairing students up for timed readings while working to assess individual students. In addition, teachers may divide students based on response within the small group. For example, teachers can assign responsive students to peer tutoring for part of the intervention time to allow nonresponsive students to receive targeted content. Finally, teachers should be careful when selecting components of intervention to adjust. It is possible that spending less time in connected text fluency practice compared with BAU was associated with lesser effects on fluency and comprehension outcomes. However, we note that the intervention adjustments were designed to remediate word reading and there were significant improvements in this area that may have generalized to fluency and comprehension given more time. Each of these adjustments makes the use of DBDM more feasible for teachers to implement in their classrooms.

Conclusion

The use of data to customize and target a research-based MWR intervention for students with or at risk for RD in upper elementary grades is supported by the findings of this randomized controlled trial. This is particularly important because it can be challenging to improve the outcomes of students with reading difficulties beyond the primary elementary grades, and students during this time are presented with higher expectations in reading. In addition to lending support to the evidence for DBDM in word reading interventions for students in upper elementary grades, this study also offers some ways in which data may be used within a small-group context to support teacher implementation in practice. Future research is needed to determine how to transfer the effects of intervention across reading domains to create a lasting impact for upper elementary students with or at risk for RD.

Authors' Note

These contents do not necessarily represent the policy of the funding agency, and one should not assume endorsement by the federal government.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

This work was supported in part by grant H325D150056 from the Office of Special Education Programs, U.S. Department of Education. This research was also supported, in part, by the Division of Learning Disabilities of the Council for Exceptional Children, Candace S. Bos Innovative Grant Award.

ORCID iD

Jessica R. Toste https://orcid.org/0000-0002-6327-0054

Supplemental Material

Supplementary material for this article is available on the *Journal of Learning Disabilities* website with the online version of this article.

References

Archer, A. L., Gleason, M. M., & Vachon, V. L. (2003). Decoding and fluency: Foundation skills for struggling older readers. *Learning Disability Quarterly*, *26*(2), 89–101. https://doi.org/10.2307/1593592

Ardoin, S. P., Christ, T. J., Morena, L. S., Cormier, D. C., & Klingbeil, D. A. (2013). A systematic review and summarization of the recommendations and research surrounding Curriculum-Based Measurement of oral reading fluency (CBM-R) decision rules. *Journal of School Psychology*, *51*, 1–18. https://doi.org/10.1016/j.jsp.2012.09.004

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Bhattacharya, A., & Ehri, L. C. (2004). Graphosyllabic analysis helps adolescent struggling readers read and spell words. *Journal of Learning Disabilities*, *37*, 331–348.

Christ, T. J., Arañas, Y. A., Kember, J. M., Kiss, A. J., McCarthy-Trentman, A., Monaghen, B. D., . . . White, M. J. (2014).

Formative assessment system for teachers™ technical manual: EarlyReading, CBMReading, aReading, aMath, and earlyMath. Formative Assessment System for Teachers.

Council of Chief State School Officers. (2013). *Interstate Teacher Assessment and Support Consortium (InTASC) model core teaching standards and learning progressions for teachers*. 1.0. Author.

Coyne, M., Simmons, D., Hagan-Burke, S., Simmons, L., Kwok, O., Kim, M., . . . Rawlinson, D. (2013). Adjusting beginning reading intervention based on student performance: An experimental evaluation. *Exceptional Children*, *80*(1), 25–44. https://doi.org/10.1177/001440291308000101

Das-Smaal, E. A., Klapwijk, M. J., & van der Leij, A. (1996). Training of perceptual unit processing in children with a reading disability. *Cognition and Instruction*, *14*, 221–250. https://doi.org/10.1207/s1532690xci1402_3

Datnow, A., & Hubbard, L. (2016). Teacher capacity for and beliefs about data-driven decision making: A literature review of international research. *Journal of Educational Change*, *17*(1), 7–28. https://doi.org/10.1007/s10833-015-9264-2

Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children*, *52*, 219–232. https://doi.org/10.1177/001440298505200303

Deno, S. L., Fuchs, L. S., Marston, D., & Shin, J. (2001). Using curriculum-based measurement to establish growth standards for students with learning disabilities. *School Psychology Review*, *30*, 507–525. https://doi.org/10.1080/02796015.2001.12086131

Deno, S. L., & Mirkin, P. (1977). *Data-based program modification: A manual*. Council for Exceptional Children.

Denton, C. A., Anthony, J. L., Parker, R., & Hasbrouck, J. E. (2004). Effects of two tutoring programs on the English reading development of Spanish-English bilingual students. *The Elementary School Journal*, *104*(4), 289–305.

Denton, C. A., Tolar, T. D., Fletcher, J. M., Barth, A. E., Vaughn, S., & Francis, D. J. (2013). Effects of Tier 3 intervention for students with persistent reading difficulties and characteristics of inadequate responders. *Journal of Educational Psychology*, *105*(3), 633–648. https://doi.org/10.1037/a0032581

Diliberto, J. A., Beattie, J. R., Flowers, C. P., & Algozzine, R. F. (2008). Effects of teaching syllable skills instruction on reading achievement in struggling middle school readers. *Literacy Research and Instruction*, *48*(1), 14–27.

Edmonds, M. S., Vaughn, S., Wexler, J., Reutebuch, C., Cable, A., Tackett, K. K., & Schnakenberg, J. W. (2009). A synthesis of reading interventions and effects on reading comprehension outcomes for older struggling readers. *Review of Educational Research*, *79*(1), 262–300.

Ehri, L. C. (2005). Learning to read words: Theory, findings, and issues. *Scientific Studies of Reading*, *9*, 167–188.

Filderman, M. J., Toste, J. R., Didion, L. A., Peng, P., & Clemens, N. H. (2018). Data-based decision making in reading interventions: A synthesis and meta-analysis of the effects for struggling readers. *The Journal of Special Education*, *52*(3), 174–187.

Fountas, I. C., & Pinnell, G. S. (2009). *Leveled literacy intervention*. Heinemann.

Fuchs, D., & Fuchs, L. S. (2015). Rethinking service delivery for students with significant learning problems: Developing and implementing intensive instruction. *Remedial and Special Education*, *36*, 105–111. https://doi.org/10.1177/0741932514558337

Fuchs, D., Fuchs, L. S., Mathes, P. G., & Simmons, D. C. (1997). Peer-assisted learning strategies: Making classrooms more responsive to diversity. *American Educational Research Journal*, *34*, 174–206. https://doi.org/10.3102/00028312034001174

Fuchs, L. S., Fuchs, D., & Hamlett, C. L. (1989). Effects of instrumental use of curriculum-based measurement to enhance instructional programs. *Remedial and Special Education*, *10*(2), 43–52. https://doi.org/10.1177/074193258901000209

Fuchs, L. S., Fuchs, D., & Malone, A. (2017). The taxonomy of intervention intensity. *TEACHING Exceptional Children*, *50*, 35–43. https://doi.org/10.1177/0040059918758166

Gallagher, L., Means, B., & Padilla, C. (2008). *Teachers' use of student data systems to improve instruction: 2005 to 2007*. U.S. Department of Education.

Greenwood, C. R., Delquadri, J. C., & Hall, R. V. (1989). Longitudinal effects of classwide peer tutoring. *Journal of Educational Psychology*, *81*(3), 371–383. https://doi.org/10.1037/0022-0663.81.3.371

Hedges, L. V., & Hedberg, E. C. (2007). Intraclass correlation values for planning group-randomized trials in education. *Educational Evaluation and Policy Analysis*, *29*(1), 60–87. https://doi.org/10.3102/0162373707299706

Heggie, L., & Wade-Woolley, L. (2017). Reading longer words: Insights into multisyllabic word reading. *Perspectives of the ASHA Special Interest Groups*, *2*, 86–94. https://doi.org/10.1044/persp2.sig1.86

Hintze, J. M., Christ, T. J., & Methe, S. A. (2006). Curriculum-based assessment. *Psychology in the Schools*, *43*, 45–56. https://doi.org/10.1002/pits.20128

Houghton Mifflin. (2001). *Houghton Mifflin phonics/decoding screening test*. Author.

Hox, J. (2002). *Multilevel analysis: Techniques and applications*. Lawrence Erlbaum.

Lenz, B. K., & Hughes, C. A. (1990). A word identification strategy for adolescents with learning disabilities. *Journal of Learning Disabilities*, *23*, 149–158.

Jenkins, J., & Terjeson, K. J. (2011). Monitoring reading growth: Goal setting, measurement frequency, and methods of evaluation. *Learning Disabilities Research & Practice*, *26*(1), 28–35. https://doi.org/10.1111/j.1540-5826.2010.00322.x

Jung, P. G., McMaster, K. L., Kunkel, A. K., Shin, J., & Stecker, P. M. (2018). Effects of data-based individualization for students with intensive learning needs: A meta-analysis. *Learning Disabilities Research & Practice*, *33*, 144–155. https://doi.org/10.1111/ldrp.12172

Kearns, D. M., Steacy, L. M., Compton, D. L., Gilbert, J. K., Goodwin, A. P., Cho, E., Lindstrom, E. R., & Collins, A. A. (2016). Modeling polymorphemic word recognition: Exploring differences among children with early-emerging and late-emerging word reading difficulty. *Journal of Learning Disabilities*, *49*, 368–394. https://doi.org/10.1177/0022219414554229

Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher*, *49*(4), 241–253. https://doi.org/10.3102/0013189x20912798

Leach, J. M., Scarborough, H. S., & Rescorla, L. (2003). Late-emerging reading disabilities. *Journal of Educational Psychology*, *95*, 211–224. https://doi.org/10.1037/0022-0663.95.2.211

Luo, W., Cappaert, K. J., & Ning, L. (2015). Modelling partially cross-classified multilevel data. *British Journal of Mathematical and Statistical Psychology*, *68*(2), 342–362. https://doi.org/10.1111/bmsp.12050

Mathes, P., Denton, C., Fletcher, J., Anthony, J., Francis, D., & Schatschneider, C. (2005). The effects of theoretically different instruction and student characteristics on the skills of struggling readers. *Reading Research Quarterly*, *40*, 148–182. https://doi.org/10.1598/rrq.40.2.2

Nagy, W., Berninger, V. W., & Abbott, R. D. (2006). Contributions of morphology beyond phonology to literacy outcomes of upper elementary and middle-school students. *Journal of Educational Psychology*, *98*, 134–147.

National Center for Education Statistics. (2019). *2019 reading Grades 4 and 8 assessment report cards: Summary data tables for national and state average scores and achievement level results*. U.S. Department of Education.

National Center on Intensive Intervention. (2013). *Data-based individualization: A framework for intensive intervention*. U.S. Department of Education.

Penney, C. G. (2002). Teaching decoding skills to poor readers in high school. *Journal of Literacy Research*, *34*, 99–118.

Perfetti, C. A. (1985). *Reading ability*. Oxford University Press.

Perry, C., Ziegler, J. C., & Zorzi, M. (2010). Beyond single syllables: Large-scale modeling of reading aloud with the Connectionist Dual Process (CDP++) model. *Cognitive Psychology*, *61*, 106–151.

Roberts, C., & Roberts, S. A. (2005). Design and analysis of clinical trials with clustering effects due to treatment. *Clinical Trials*, *2*(2), 152–162. https://doi.org/10.1191/1740774505cn076oa

Ševa, N., Monaghan, P., & Arciuli, J. (2009). Stressing what is important: Orthographic cues and lexical stress assignment. *Journal of Neurolinguistics*, *22*, 237–249.

StataCorp. (2019). *Stata statistical software: Release 16* [Computer software]. StataCorp LLC.

Stecker, P. M., & Lembke, E. S. (2011). *Advanced applications of CBM in reading (K-6): Instructional decision-making strategies manual*. National Center on Student Progress Monitoring.

Swanson, E. A., Wexler, J., & Vaughn, S. (2009). Text reading and students with learning disabilities. In E. H. Hiebert (Ed.), *Reading more, reading better* (pp. 210–230). Guilford Press.

Torgesen, J. K., Wagner, R., & Rashotte, C. (2012). *Test of Word Reading Efficiency: (TOWRE-2)*. Pearson Clinical Assessment.

Toste, J. R., Capin, P., Vaughn, S., Roberts, G. J., & Kearns, D. M. (2017a). Multisyllabic word-reading instruction with and without motivational beliefs training for struggling readers in the upper elementary grades: A pilot investigation. *The Elementary School Journal*, *117*, 593–615. https://doi.org/10.1086/691684

Toste, J. R., Capin, P., Williams, K. J., Cho, E., & Vaughn, S. (2019). Replication of an experimental study investigating the efficacy of a multisyllabic word reading intervention with and without motivational beliefs training for struggling readers. *Journal of Learning Disabilities*, *52*, 45–58. https://doi.org/10.1177/0022219418775114

Toste, J. R., Williams, K. J., & Capin, P. (2017b). Reading big words: Instructional practices to promote multisyllabic word reading fluency. *Intervention in School and Clinic*, *52*(5), 270–278. https://doi.org/10.1177/1053451216676797

Vadasy, P. F., Sanders, E. A., & Peyton, J. A. (2006). Paraeducator-supplemented instruction in structural analysis with text reading practice for second and third graders at risk for reading problems. *Remedial and Special Education*, *27*, 365–378. https://doi.org/10.1177/07419325060270060601

VanDerHeyden, A. M., & Burns, M. K. (2018). Improving decision making in school psychology: Making a difference in the lives of students, not just a prediction about their lives. *School Psychology Review*, *47*, 385–395. https://doi.org/10.17105/spr-2018-0042.v47-4

Vaughn, S., Solís, M., Miciak, J., Taylor, W. P., & Fletcher, J. M. (2016). Effects from a randomized control trial comparing researcher and school-implemented treatments with fourth graders with significant reading difficulties. *Journal of Research on Educational Effectiveness*, *9*, 23–44. https://doi.org/10.1080/19345747.2015.1126386

Vaughn, S., & Wanzek, J. (2014). Intensive interventions in reading for students with reading disabilities: Meaningful impacts. *Learning Disabilities Research & Practice*, *29*(2), 46–53. https://doi.org/10.1111/ldrp.12031

Vaughn, S., Wanzek, J., Murray, C. S., & Roberts, G. (2012). *Intensive interventions for students struggling in reading and mathematics: A practice guide*. RMC Research Corporation.

Vaughn, S., Wexler, J., Roberts, G., Barth, A., Cirino, P., Romain, M., Francis, D., Fletcher, J. & Denton, C. (2011). Effects of individualized and standardized interventions on middle school students with reading disabilities. *Exceptional Children*, *77*, 391–407.

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., & Pearson, N. A. (2010). *Test of silent reading efficiency and comprehension*. Pro-Ed.

Wanzek, J., Vaughn, S., Scammacca, N. K., Metz, K., Murray, C. S., Roberts, G., & Danielson, L. (2013). Extensive reading interventions for students with reading difficulties after grade 3. *Review of Educational Research*, *83*(2), 163–195. https://doi.org/10.3102/0034654313477212

What Works Clearinghouse (2014). *Procedures and standards handbook* (Version 3.0).

White, T. G., Sowell, J., & Yanagihara, A. (1989). Teaching elementary students to use word-part clues. *The Reading Teacher*, *42*(4), 302–308.

Woodcock, R. W. (2011). *Woodcock Reading Mastery Tests—Third edition*. Pearson.

Yap, M. J., & Balota, D. A. (2009). Visual word recognition of multisyllabic words. *Journal of Memory and Language*, *60*(4), 502–529. https://doi.org/10.1016/j.jml.2009.02.001