Hammill Institute on Disabilities

Journal of Learning Disabilities
2019, Vol. 52(1) 45–58

DOI: 10.1177/0022219418775114
journaloflearningdisabilities.sagepub.com
SAGE

# Replication of an Experimental Study Investigating the Efficacy of a Multisyllabic Word Reading Intervention With and Without Motivational Beliefs Training for Struggling Readers

**Jessica R. Toste, PhD[1], Philip Capin, MA[1], Kelly J. Williams, PhD[1], Eunsoo Cho, PhD[2], and Sharon Vaughn, PhD[1]**

**Abstract**

This randomized control trial examined the efficacy of an intervention aimed at improving multisyllabic word reading (MWR) skills among fourth- and fifth-grade struggling readers ($n$ = 109, 48.6% male), as well as the relative effects of an embedded motivational beliefs training component. This study was a closely aligned replication of our earlier work. The intervention was replicated with a three-condition design: MWR only, MWR with a motivational beliefs component, and business-as-usual control. Students were tutored in small groups for 40 lessons (four 40-min lessons each week). When we combined performance of students in both MWR conditions, intervention students significantly outperformed controls on proximal measures of affix reading and MWR, as well as standardized measures of decoding, spelling, and text comprehension. Furthermore, there was a noted interaction between English learner status and treatment on spelling performance. There were no statistically significant main effects between the MWR groups on proximal or standardized measures of interest. Findings are discussed in terms of their relevance to MWR instruction for students with persistent reading difficulties and considerations for future research related to the malleability of motivation.

**Keywords**

word reading, multisyllabic words, motivation, intervention, replication

Students who lack proficiency in word reading by third grade are at increased risk for secondary school failure and school dropout (Brasseur-Hock, Hock, Kieffer, Biancarosa, & Deshler, 2011; D. J. Francis, Shaywitz, Stuebing, Shaywitz, & Fletcher, 1996). Although the importance of word reading with beginning readers is readily accepted, word reading instruction after second grade is not often provided (Vaughn et al., 2003). Instead, the primary instructional goal of language arts teachers in the upper elementary grades is to improve students' ability to gain knowledge from text. In recent years, the expectations for students to comprehend complex and varied texts have increased with the implementation of progressive standards (e.g., Common Core State Standards). But findings suggest that about one quarter of upper elementary students do not read words accurately or fluently at grade level and that their lack of automaticity influences their text comprehension (Cirino et al., 2013; Daane, 2005).

Based on the need for word reading instruction for upper elementary students with significant reading difficulties, it is essential that we better understand practices associated with improved reading outcomes for these students. Much of the research on reading interventions for older students addresses approaches that emphasize multicomponent reading interventions that include word reading and other components, such as vocabulary and reading comprehension, as a means of improving reading outcomes for students in upper elementary grades (Wanzek, Wexler, Vaughn, & Ciullo, 2010) and secondary grades (Torgesen et al., 2007). However, these multicomponent interventions do not allow us to determine the relative effects of various approaches to teaching word reading to these students. Thus, it is difficult to discern

[1]Department of Special Education and The Meadows Center for Preventing Educational Risk, The University of Texas at Austin, USA
[2]Department of Counseling, Educational Psychology, and Special Education, Michigan State University, East Lansing, USA

**Corresponding Author:**
Jessica Toste, Department of Special Education and The Meadows Center for Preventing Educational Risk, The University of Texas at Austin, 1 University Station, D5300, Austin, TX 78712, USA.
Email: jrtoste@austin.utexas.edu

which word reading practices are most beneficial for students with reading difficulties in the upper elementary grades. Few high-quality studies have examined the effects of word reading practices for students with reading difficulties in these grades (Wanzek et al., 2010).

## Multisyllabic Word Reading

The difference between good and poor decoders in the upper elementary grades is often evident when students read multisyllabic words (Just & Carpenter, 1987). Multisyllabic word reading (MWR) increases in importance as students move through the grades. The average number of syllables in the words that students read increases steadily (Kearns, 2015; Nagy & Anderson, 1984), and multisyllabic words often contain the meaning of content area texts, which can be discerned only if one can decode the word. Thus, it is essential to understand how we can improve advanced word reading skills for this age group.

Toste, Capin, Vaughn, Roberts, and Kearns (2017) designed an instructional approach aligned with research illustrating that word representations may be best acquired by building representations of meaningful linguistic units (i.e., morphemes) through repeated practice and experience with a large number of words. The researchers sought to reduce the cognitive demands associated with the memorization and application of numerous phonics rules by focusing on the application of only a few phonics rules within the context of the words in which they appear. The notion of reducing the cognitive demands associated with rule-based instruction is buttressed by evidence showing that students with reading difficulties have deficits in phonological memory, which makes it more difficult for them to simultaneously engage with multiple sources of information (Shankweiler, Crain, Brady, & Marcuso, 1992; Wagner & Torgesen, 1987).

## Motivational Processes

Studies investigating students from preschool through high school have reported that high levels of motivation are positively associated with reading (e.g., Guthrie & Wigfield, 2000; Logan, Medford, & Hughes, 2011; Morgan & Fuchs, 2007). There is likely bidirectionality between motivation and reading success in that development of reading skills can significantly influence motivational processes related to a student's sense of worth as a learner, active engagement within the classroom environment, and overall interest in one's school success. For example, students who report lower feelings of perceived competence are less likely to initiate tasks and more likely to show poor task persistence, resulting in lower levels of achievement (Zimmerman & Schunk, 2006).

To date, we know very little about how the relation between motivation and reading may be cultivated through intervention. In educational and psychological research, motivation remains an underspecified construct (e.g., Conradi, Jang, & McKenna, 2014). The nature of motivation and its associated processes have been examined from multiple perspectives—which has resulted in an extensive list of constructs and theoretical frameworks (e.g., Eccles & Wigfield, 2002; Pintrich, 2003; Schunk, Meece, & Pintrich, 2013). In education, these theories tend to be founded on various social-cognitive theories, which emphasize individuals' perceptions of themselves—achievement motives, perceived self-concept, self-efficacy, perceived value for and interest in an activity, goals, attributions of success and failure, and emotions (for review, see Schunk et al., 2013). In the absence of clear direction for how motivation may maximize students' gains within a reading intervention, Toste, Capin, et al. (2017) developed a training grounded in theory that was embedded within a reading intervention and designed to leverage motivational processes.

## Study Replication

A central mission of educational research is to identify evidence-based practices through rigorous scientific research (e.g., Coyne, Cook, & Therrien, 2016). As previously mentioned, Toste, Capin, et al. (2017) recently examined the effects of an MWR intervention for struggling readers in Grades 3 and 4. The results of the randomized controlled trial indicated that students receiving the intervention outperformed students in a control condition on a standardized measure of word reading fluency ($g$ = .73). This study provides support for further investigation of the effects of MWR instruction for nonproficient word readers in the upper elementary grades.

Toste, Capin, et al. (2017) also investigated the relative effects of an embedded motivational beliefs (MB) training component on motivation and reading outcomes with the use of a three-condition design: MWR only, MWR with MB training, and business-as-usual control. Findings of the three-condition design showed that students who received the embedded MB training outperformed the reading-only group on sentence-level reading comprehension and reported higher levels of attributions for success than students in the comparison condition. Taken together, these findings support a conceptual replication of the Toste, Capin, et al. (2017) study. Conceptual replications vary in one or more features from the original study (Schmidt, 2009), they support the accumulation of scientific knowledge (G. Francis, 2012), and can inform whether the effects of an intervention generalize when study components are altered (Brandt et al., 2014).

## Current Study

We sought to investigate four research questions. First, what are the effects of an MWR intervention as compared with a control on reading and motivation-related outcomes? We hypothesized that there would be significant positive effects in comparing the MWR intervention against the control on proximal and standardized measures of word reading. However, we did not expect for there to be significant effects on students' self-reported motivation. The success of interventions aimed at psychosocial processes (e.g., motivation) is dependent on a fluid self-reinforcing recursive process (Walton & Cohen, 2011; Yeager & Walton, 2011), and there was a relatively short period between pre- and postintervention.

We also had an interest in exploring potential student attributes that may moderate the impact of the intervention. Specifically, we were interested in whether English learners (ELs) might differentially respond to the MWR intervention. We asked: Does EL status influence the impact of the MWR intervention on measures of reading performance? We did not posit specific hypotheses related to this research question, as the MWR reading intervention was not developed with instructional practices to support ELs in mind. That said, we thought that EL status was an important student attribute to contribute, as the number of ELs in U.S. schools has been consistently increasing. Approximately 9.4% of students in K–12 public schools identify as ELs (Batalova, Fix, & Murray, 2007; McFarland et al., 2017), and ELs consistently perform lower than their native English-speaking peers on national assessments of reading (National Center for Education Statistics, 2016).

Next, to examine the potential value-added benefit of targeting motivational processes, we asked: What are the effects of the MWR intervention with an embedded MB training as compared with MWR only on reading and motivation-related outcomes? Given the findings reported by Toste, Capin, et al. (2017), we hypothesized that the students who received the MB training would have greater gains in reading. Again, we did not have reason to believe that we would find differences in reported motivation. Finally, to understand the value of study replication, we asked: To what extent do the effects of this replication study align with the findings of the original study? We consider recommendations outlined by Coyne et al. (2016) for the implementation and reporting of replication research in special education.

## Method

### Participants

Participants were recruited from three elementary schools located in a large district in the southeastern United States. Fourth- and fifth-grade students were selected to participate through a two-step screening process. First, on the basis of universal screening data, the school district nominated the lowest-performing students who were not already receiving intensive reading intervention. All nominated students ($n$ = 136) were then screened with the *Test of Word Reading Efficiency–Second Edition* (TOWRE-2). To be eligible for the study, students were required to display difficulties in word reading fluency, as evidenced by scoring ≤25th percentile on either subtest of the TOWRE-2. From this screening, 114 students were selected to participate in the study and were randomized to one of three conditions: MWR, MWR and MB training (MWR+MB), or business-as-usual control (BAU). Five students were pretested but not assessed during posttesting due to relocation. The amount of overall and differential attrition was not substantial according to the What Works Clearinghouse guidelines for attrition. Table 1 provides demographic data for the final sample of 109 students. There were no significant differences across three groups for grade, gender, ethnicity, and socioeconomic status as indicated by the free/reduced-price lunch status.

### Measures

All students were assessed at pre- and postintervention with measures of word reading, comprehension, spelling, and motivation. Approximately 20 hr of training on test administration and data collection were provided to research staff prior to pretesting. In a mock test administration session with a research coordinator, staff were required to demonstrate 98% reliability on each measure prior to working with students and collecting data. The same training and reliability procedures were used at posttesting.

*Word reading.* Word identification was measured with two untimed subtests of the *Woodcock-Johnson III* (WJIII; McGrew, Schrank, & Woodcock, 2007). Word Identification measures skill in reading words in isolation, and Word Attack assesses skill in using phonic and structural analysis to read nonsense words. Internal-consistency reliability exceeds .80 for each subtest. Word reading fluency was measured with two timed subtests of the TOWRE-2 (Torgesen, Wagner, & Rashotte, 2012): the Sight Word Efficiency subtest assesses the number of real printed words that can be accurately identified within 45 s, and the Phonetic Decoding Efficiency subtest measures the number of pronounceable printed nonwords that can be accurately decoded within 45 s. Test-retest reliability ranges from .83 to .96, and alternative-form reliability exceeds .90.

Students also completed a research-developed proximal measure of word reading: the *Big Word Reading Test*. This test includes two subtests wherein students read affixes in isolation and multisyllabic words from the intervention. Part I (Affix List) consisted of reading 48 affixes in isolation.

**Table 1.** Participant Demographic Information.

| | Control (*n* = 37) | | MWR (*n* = 34) | | MWR+MB (*n* = 38) | |
|---|---|---|---|---|---|---|
| Variables | *n* | % | *n* | % | *n* | % |
| Grade | | | | | | |
| Fourth | 20 | 54 | 17 | 50 | 22 | 58 |
| Fifth | 17 | 46 | 17 | 50 | 16 | 42 |
| Free/reduced-price lunch | | | | | | |
| No | 2 | 5 | 3 | 9 | 4 | 79 |
| Yes | 34 | 92 | 30 | 88 | 34 | 89 |
| Gender | | | | | | |
| Female | 19 | 51 | 15 | 44 | 20 | 53 |
| Male | 17 | 46 | 18 | 53 | 18 | 47 |
| Ethnicity | | | | | | |
| Hispanic | 32 | 86 | 27 | 79 | 34 | 89 |
| Black | 2 | 5 | 3 | 9 | 2 | 5 |
| White | 1 | 3 | 2 | 6 | 1 | 3 |
| Two or more | 1 | 3 | 1 | 3 | 1 | 3 |
| EL status[a] | | | | | | |
| No | 23 | 62 | 29 | 85 | 26 | 68 |
| Yes | 13 | 35 | 4 | 12 | 12 | 32 |

*Note.* The percentages do not sum to 100% for some variables due to missing data. EL = English learner; MB = motivational beliefs; MWR = multisyllabic word reading.
[a]EL status was reported by school district according to students' designiation as limited English proficiency (current or within the past 2 years).

Students received 1 point for each affix read correctly for a total summed score. In Part II (Big Word List), students read a list of 96 multisyllabic words and received a point for each word read fluently. For the current sample, internal reliability was calculated at pre- and posttest for the Affix List (.20 and .86, respectively) and the Big Word List (.94 and .93, respectively).

*Comprehension.* Comprehension was assessed through two measures. The WJIII Passage Comprehension subtest measures students' ability to identify a missing word from a passage (Schrank, McGrew, & Woodcock, 2001). This subtest has a median split-half reliability of .88. The second measure of comprehension was the Comprehension subtest of the *Gates-MacGinitie Reading Test–4* (GRMT-4; MacGinitie, MacGinitie, Maria, & Dreyer, 2000). This subtest has a Kuder-Richardson Formula 20 reliability coefficient of .93 for Grades 4 and 5.

*Spelling.* The WJIII Spelling subtest examines students' ability to spell words through dictation. This measure has split-half reliability of .90 (Schrank et al., 2001).

*Motivational processes.* As noted, we were cognizant of the difficulty with measuring motivation—and because we defined motivation broadly, we thought that it would be best to select a measure of motivational processes that was aligned with the training approach used in the MWR+MB intervention. The *Reading Self-Concept Scale* (RSCS; Chapman & Tunmer, 1995) is a 30-item measure that taps reading-related perceptions on three subscales: perceptions of competence, perceptions of difficulty, and attitudes toward reading. Scores for each item ranged from 1 (low self-concept) to 5 (high self-concept), and the three subscale scores were each calculated as the mean value of the 10 items. Internal reliability estimates have yielded high rates of reliability (.84–.87; Chapman & Tunmer, 1995). Internal reliability for the current study sample had similar estimates on each subtest at pre- and posttest: competence (.85 and .80), difficulty (.69 and .70), and attitudes (.89 and .87).

### Intervention Overview

Students were taught in small groups (three or four students), and lessons occurred four times each week for a total of 40 lessons. Students in both treatment groups received MWR instruction, and students in the MWR+MB tutoring group also received integrated MB training. All lesson activities were scripted to ensure standardization across tutoring groups. Each instructional lesson was 40 min. The MB training was embedded into the lesson scripts so that it did not add additional time to the MWR+MB treatment group's lessons.

*MWR instruction.* Lessons in both treatment conditions, MWR and MWR+MB, were taught with the same reading intervention curriculum. Each lesson consisted of seven instructional components: warm-up, affix bank, word play, beat the clock, write word, speedy read, and text reading. Due to space limitations, we provide brief descriptions of these components (for details, see Toste, Williams, & Capin, 2017).

*Warm-up.* Each lesson began with a 3-min warm-up activity where students practiced essential prerequisite skills for MWR. Students were taught target vowel patterns (i.e., short vowels, long vowels), vowel digraphs, r-controlled vowels, diphthongs, and variant correspondences) and then practiced reading the pattern in isolation and in nonsense words until mastery. Warm-up was discontinued after Lesson 33.

*Affix bank.* After reviewing prerequisite skills, tutors spent approximately 3 min explicitly teaching high-frequency affixes to the students. Three new prefixes were taught each day, selected from a list of the most commonly used prefixes and suffixes in Grades 3 to 9 (White, Sowell, & Yanagihara, 1989). Students then chorally read all previously learned affixes.

*Word play.* During word play, which took approximately 5 min, students focused on assembling or blending word parts, with lessons rotating through five word-building games that emphasized automaticity of the reading process that were used. Five "spotlight words," or base words, were used throughout the lesson (e.g., *judge, extend, thought, visible, strong*) to build real and pseudo (nonsense) words, ensuring multiple opportunities to work on the skills necessary for quick and accurate decoding of unknown words.

*Beat the clock.* Next, students practiced breaking or segmenting multisyllabic words into parts. Multisyllabic word lists were developed for students to read aloud, with each word appearing three times throughout the intervention. Students underlined the affixes in each word and then chorally read the affixes, while the tutor provided corrective feedback as needed. Then, the tutor and students read the whole list of words aloud. Following the practice readings of the affixes and words, students were given two timed opportunities to read the list of words, wherein they attempted to meet or beat their previous reading time.

*Write word.* During each lesson, students had the opportunity to work on encoding skills for 5 to 8 min. They practiced writing words with two or more syllables, using the targeted affixes in the lesson. Students were encouraged to write real words and nonsense words. The tutors provided corrective feedback on spelling of affixes, vowels, and other word parts.

*Speedy read.* In speedy read, students worked for 5 min on improving their reading accuracy and rate through timed word-list reading. In the first 20 lessons, students were given word lists with specific phonetic patterns, while later lessons included randomly generated lists of words from the previous lists. First, students read the lists aloud. Then, each student was given an opportunity to read for 30 s while the tutor tracked the accuracy of responses. The tutor provided corrective feedback by having students reread incorrectly pronounced words. After reading, students recorded the number of words read on a chart to help monitor their progress.

*Text reading.* The last lesson activity each day was text reading. For the first 20 lessons, students read sentences that were developed to have at least two multisyllabic words and spotlight words. Each sentence was read aloud two times with the choral read, echo read, or whisper read procedure. As of Lesson 21, students read connected texts. Tutors introduced and defined key words, had students read the passage aloud two times with the aforementioned variations, and then highlighted multisyllabic words and spotlight words. We used Level B QuickReads passages (Hiebert, 2012), slightly revised to include five to eight additional multisyllabic words (resulting in 20–50 additional total words per text).

*MB training.* In the MWR+MB condition, students received training to foster MB. As previously described, this training was grounded in motivation theory but did not subscribe to a limited motivation construct. This training utilized a cognitive-based approach (Toland & Boyle, 2008). Social-cognitive theories assume that thinking precedes feelings and behaviors and, as such, individuals are deeply influenced by inaccurate beliefs, cognitive distortions, and automatic thoughts (Beck, Hollon, Young, Bedrosian, & Budenz, 1985). Thus, the training that we designed targeted subconscious beliefs through positive statements, challenged negative beliefs, and replaced negative beliefs with positive thoughts (with the goal of those becoming students' automatic thoughts around reading). Each lesson began by asking students to think about their current readiness on a scale from 1 to 5. Throughout the 40 lessons, tutor modeling guided students in using positive self-talk and generated self-motivated statements to support their efforts while reading. Through use of scenarios and story vignettes, students were asked to identify the negative thoughts that a struggling reader in an upper elementary grade may be having and then help that student generate positive self-talk to support her or his learning. As students became comfortable with this process, we discussed real academic situations wherein they had experienced difficulty, the types of thoughts that they may have had during that situation, and how they could recognize and change negative thoughts when they arise in future.

Compared with the original study (Toste, Capin, et al., 2017), the current study expanded the MB training component by an additional 16 lessons. While the focus on positive self-talk continued throughout all lessons, we began to introduce explicit goal setting in Lesson 25. Tutors guided students through identifying their strengths and needs related to school activities. Then students selected strengths and needs specific to reading (e.g., "I know how to spell vowel teams," "I am able to recognize words without sounding them out," or "I can answer questions after reading to show that I understand"). On the basis of their identified needs, students developed reading goals, strengths and solutions to help them reach their goals, and positive thoughts that they would use when engaging in activities (likely challenging) around their goals.

*Tutors.* Eight tutors were hired and trained by the research team to deliver the intervention. Seven of the eight tutors were female, and all of the tutors had at least an undergraduate degree (with four having graduate degrees). Only one of the tutors was certified in elementary or special education; however, all tutors had at least 5 years' experience teaching or tutoring in private or public school settings.

*Implementation fidelity.* Tutors received a total of 8 hr of training (i.e., two 4-hr sessions). They were also provided professional development time to read scripted lessons and practice utilizing materials. The tutors completed a "mock lesson" with a research coordinator, with a requirement of achieving at least 90% adherence to protocols before tutoring began. During tutoring, each tutor was observed twice in person and four to six times through audio. Tutors who taught one condition had four fidelity checks in total, while tutors who taught both conditions had six checks. Implementation fidelity was scored on a 52-item checklist that detailed all intervention components and instructional routines. Project coordinators marked checklist items as *performed correctly, performed incorrectly*, or *not applicable*. Scores were calculated by dividing the number of items conducted correctly by the total number of items observed. The mean implementation adherence score across components and interventionists was 97.90% ($SD$ = 4.28). Tutors were also evaluated on three quality items (e.g., pacing, corrective feedback, behavior management). Each item was given a score between 1 (*ineffective*) and 3 (*highly effective*). The mean quality score across interventionists was 8.32 ($SD$ = 1.23).

## Business-as-Usual Control

Research staff met with the fourth- and fifth-grade teachers at participating schools to obtain information about the instruction received by the students in the control group, while students in the treatment conditions received reading intervention. Instruction during this time varied throughout the year, but teachers provided instruction that included small-group lessons and independent work time. Teachers reported that the majority of the time was spent on computer-based programming, guided reading, sustained silent reading, and preparation sessions for the State of Texas Assessments of Academic Readiness.

## Data Analyses

We investigated treatment effects using a series of regression models in Mplus 7.2 (Muthén & Muthén, 2014). First, we included two contrast codes and a covariate in the regression models. We created two contrast codes to compare both treatments with control and then to each other. The first contrast code (C1) estimated the difference between control and treatment (MWR and MWR+MB); the second contrast code (C2) estimated the difference between the treatment groups. We used a grand mean–centered pretest score for each outcome as a covariate in the models. In addition, because intraclass correlations at the classroom level ranged from .20 to .26 across the measures, the models accounted for classroom-level variance via the *type = complex* option with the maximum likelihood estimator with robust standard errors (MLR). Then, we used Benjamini-Hochberg (BH) corrections (Benjamini & Hochbert, 1995) to control for the heightened false discovery rate associated with multiple comparisons. Finally, effect sizes were calculated with the coefficients estimated from the model divided by the pooled within-group standard deviation across conditions.

Next, we created an interaction term by multiplying C1 with a binary EL variable. Pretest score, EL status, C1, C1 × EL, and C2 were simultaneously entered into a series of regression models to examine whether EL status moderated treatment effects. For the outcomes with a significant interaction, we report effect sizes separately for EL and non-EL students. Finally, to explore agreement of treatment effects between this replication and the original study, we compared statistical significance and the direction and magnitude of effects.

# Results

## Preliminary Analyses and Descriptive Statistics

Preliminary data analyses confirmed linear relations between pre- and posttest measures with one exception (Affix List [Part I of the *Big Word Reading Test*]), and data showed overall univariate and multivariate normality with two exceptions (Affix List and WJIII Spelling). We addressed these issues by using the MLR estimator. Preliminary analyses of the regression models indicated heteroscedastic residuals for two outcomes (Affix List and TOWRE-2 Phonetic Decoding Efficiency). For TOWRE-2

**Table 2.** Descriptive Statistics of Pre- and Posttest Measures.

| | Control (*n* = 37) | | | | MWR (*n* = 34) | | | | MWR+MB (*n* = 38) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Pretest | | Posttest | | Pretest | | Posttest | | Pretest | | Posttest | |
| | *n* | % | *n* | % | *n* | % | *n* | % | *n* | % | *n* | % |
| Reading variables | | | | | | | | | | | | |
| BW Affix List | | | | | | | | | | | | |
| Nonmastery | 34 | 92 | 32 | 89 | 32 | 94 | 0 | 0 | 30 | 79 | 0 | 0 |
| Mastery | 3 | 8 | 5 | 14 | 2 | 6 | 34 | 100 | 8 | 21 | 38 | 100 |
| | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| BW | | | | | | | | | | | | |
| Affix List | 38 | 3.79 | 39.50[a] | 3.31 | 38.68 | 4.73 | 46.91 | 1.00 | 38.89 | 5.16 | 47.13 | 0.96 |
| Big Word List | 62.81 | 14.46 | 66.89[a] | 13.06 | 67.00 | 16.43 | 80.29 | 12.09 | 67.63 | 15.25 | 79.92 | 9.72 |
| WJIII | | | | | | | | | | | | |
| WID | 94.78 | 5.43 | 93.42[a] | 5.90 | 96.44 | 8.57 | 97.09 | 9.85 | 99.08 | 9.41 | 99.53 | 8.88 |
| WAT | 94.62 | 7.21 | 92.97[a] | 6.16 | 95.09 | 6.91 | 96.50 | 7.48 | 98.63 | 9.41 | 98.34 | 7.83 |
| TOWRE-2 | | | | | | | | | | | | |
| SWE | 86.03 | 6.64 | 89.08[a] | 8.53 | 84.59 | 6.10 | 90.85 | 8.50 | 85.05 | 8.95 | 93.13 | 7.85 |
| PDE | 82.19 | 6.45 | 87.53[a] | 7.31 | 82.35 | 10.46 | 88.26 | 9.59 | 84.89 | 11.32 | 92.08 | 10.63 |
| Spelling | 95.41 | 9.37 | 94.50[a] | 10.82 | 94.91 | 11.10 | 96.38 | 11.92 | 96.84 | 8.91 | 99.00 | 9.56 |
| PC | 87.19 | 6.97 | 89.06[a] | 6.59 | 88.91 | 8.96 | 88.91 | 8.91 | 86.79 | 6.96 | 90.16 | 7.53 |
| GMRT-RC | 91.31 | 9.97a | 90.80[a] | 7.82 | 88.86 | 12.30 | 91.98 | 9.34 | 89.30 | 10.49 | 91.88 | 6.38 |
| Motivation variables | | | | | | | | | | | | |
| RSCS | | | | | | | | | | | | |
| Difficulty | 2.88 | 0.52 | 2.79 | 0.50 | 3.06 | 0.67 | 2.80 | 0.75 | 2.94 | 0.58 | 2.82 | 0.56 |
| Competence | 3.39 | 0.75 | 3.54 | 0.48 | 3.45 | 0.74 | 3.66 | 0.74 | 3.40 | 0.73 | 3.56 | 0.63 |
| Attitude | 3.82 | 0.90 | 3.81 | 0.74 | 3.78 | 0.89 | 3.52 | 0.94 | 3.66 | 0.84 | 3.44 | 0.87 |

*Note.* BW = *Big Word Reading Test*; GMRT = *Gates-MacGinitie Reading Test*; MB = motivational beliefs; MWR = multisyllabic word reading; PC = Passage Comprehension; PDE = Phonemic Decoding Efficiency; RC = Reading Comprehension; RSCS = *Reading Self-Concept Scale*; SWE = Sight Word Efficiency; TOWRE-2 = *Test of Word Reading Efficiency–2*; WAT = Word Attack; WID = Word Identification; WJIII = *Woodcock-Johnson III*.
[a]*n* = 36.

Phonetic Decoding Efficiency, eliminating two outliers resulted in homoscedasticity with the same pattern of results; thus, we report the results from the entire sample. For the Affix List, we considered 90% correct a criterion for mastery and used $\chi^2$ statistics to examine difference in percentage of students who reached mastery at posttest. Table 2 includes descriptive statistics for pre- and posttest measures. There were no pretreatment differences across three groups on the reading and motivation measures. For the Affix List, only 13 students showed mastery at pretest, and these numbers were not associated with the conditions, $\chi^2(2) = 4.71$, $p = .09$.

## *Effects of MWR Intervention*

Our proximal measure, *Big Word Reading Test*, had two parts that were scored and analyzed separately: Affix List and Big Word List. Due to the lack of linear relations and nonnormality of the Affix List, we report percentages of students who met the aforementioned mastery criterion. All students in both treatment conditions reached mastery, as opposed to only 14% of students in the control condition (see Table 2). As shown in Table 3, the effect of the treatment was statistically significant for the Big Word List ($b = 10.50$, $p < .01$, ES = .90).

We examined multiple measures of decoding and word recognition and found statistically significant treatment effects on WJIII Word Attack ($b = 3.09$, $p < .01$, ES = .43) and WJIII Word Identification ($b = 2.45$, $p < .05$, ES = .29). However, there were no statistically significant differences on measures of word reading fluency (i.e., TOWRE-2 subtests). After application of the readjusted alpha level after the BH correction (.014), the Big Word List and WJIII Word Attack subtests remained significant.

Furthermore, we found a statistically significant treatment effect on the *Gates-MacGinitie Reading Test* ($b = 2.04$, $p < .05$, ES = .26) but not on WJIII Passage Comprehension ($b = 0.05$, $p = .96$, ES = .01). The differences in *Gates-MacGinitie Reading Test* remained significant after the BH correction. A statistically significant treatment effect was also found for WJIII Spelling ($b = 2.65$, $p < .05$, ES = .25).

**Table 3.** Estimated Treatment Effects.

| Outcome variables | Coefficient | Robust *SE* | *t* | *p* | ES |
|---|---|---|---|---|---|
| Big Word List | | | | | |
| Pretest | 0.56 | 0.05 | 12.21 | .00 | |
| TX vs. C | 10.50 | 1.88 | 5.60 | .00 | 0.90 |
| TX2 vs. TX1 | −0.73 | 1.80 | −0.40 | .69 | −0.07 |
| Constant | 75.71 | 0.58 | 130.21 | .00 | |
| WJIII-WID | | | | | |
| Pretest | 0.84 | 0.07 | 13.07 | .00 | |
| TX vs. C | 2.45 | 1.24 | 1.97 | .05 | 0.29 |
| TX2 vs. TX1 | 0.21 | 1.19 | 0.18 | .86 | 0.02 |
| Constant | 96.72 | 0.57 | 168.56 | .00 | |
| WJIII-WAT | | | | | |
| Pretest | 0.64 | 0.06 | 11.49 | .00 | |
| TX vs. C | 3.09 | 1.03 | 2.99 | .00 | 0.43 |
| TX2 vs. TX1 | −0.41 | 1.14 | −0.36 | .72 | −0.05 |
| Constant | 95.98 | 0.50 | 193.58 | .00 | |
| TOWRE-2-SWE | | | | | |
| Pretest | 0.35 | 0.11 | 3.19 | .00 | |
| TX vs. C | 3.21 | 1.74 | 1.85 | .07 | 0.39 |
| TX2 vs. TX1 | 2.12 | 1.90 | 1.11 | .27 | 0.26 |
| Constant | 91.03 | 0.68 | 134.70 | .00 | |
| TOWRE-2-PDE | | | | | |
| Pretest | 0.68 | 0.10 | 6.62 | .00 | |
| TX vs. C | 1.59 | 1.35 | 1.18 | .24 | 0.17 |
| TX2 vs. TX1 | 2.10 | 2.39 | 0.88 | .38 | 0.21 |
| Constant | 89.32 | 0.67 | 133.51 | .00 | |
| WJIII-Spelling | | | | | |
| Pretest | 0.95 | 0.11 | 8.72 | .00 | |
| TX vs. C | 2.65 | 0.90 | 2.95 | .00 | 0.25 |
| TX2 vs. TX1 | 0.79 | 0.88 | 0.89 | .37 | 0.07 |
| Constant | 96.66 | 0.32 | 299.43 | .00 | |
| WJIII-PC | | | | | |
| Pretest | 0.63 | 0.10 | 6.56 | .00 | |
| TX vs. C | 0.05 | 0.87 | 0.05 | .96 | 0.01 |
| TX2 vs. TX1 | 2.59 | 1.64 | 1.58 | .12 | 0.32 |
| Constant | 89.35 | 0.57 | 158.17 | .00 | |
| GMRT-RC | | | | | |
| Pretest | 0.41 | 0.04 | 9.97 | .00 | |
| TX vs. C | 2.04 | 0.88 | 2.32 | .02 | 0.26 |
| TX2 vs. TX1 | −0.27 | 1.41 | −0.19 | .85 | −0.03 |
| Constant | 91.56 | 0.61 | 149.50 | .00 | |
| RSCS-Difficulty | | | | | |
| Pretest | 0.59 | 0.06 | 9.88 | .00 | |
| TX vs. C | −0.04 | 0.09 | −0.46 | .64 | −0.07 |
| TX2 vs. TX1 | 0.09 | 0.16 | 0.58 | .56 | 0.15 |
| Constant | 2.80 | 0.06 | 44.12 | .00 | |
| RSCS-Competence | | | | | |
| Pretest | 0.47 | 0.08 | 5.82 | .00 | |
| TX vs. C | 0.06 | 0.08 | 0.69 | .49 | 0.08 |
| TX2 vs. TX1 | −0.08 | 0.17 | −0.47 | .64 | −0.13 |
| Constant | 3.58 | 0.04 | 85.24 | .00 | |
| RSCS-Attitudes | | | | | |
| Pretest | 0.63 | 0.09 | 6.92 | .00 | |
| TX vs. C | −0.29 | 0.09 | −3.08 | .00 | −0.31 |
| TX2 vs. TX1 | −0.02 | 0.22 | −0.09 | .93 | −0.01 |
| Constant | 3.59 | 0.07 | 50.13 | .00 | |

*Note.* BW = *Big Word Reading Test*; C = control condition; GMRT = *Gates-MacGinitie Reading Test*; MB = motivational beliefs; MWR = multisyllabic word reading; PC = Passage Comprehension; PDE = Phonemic Decoding Efficiency; RC = Reading Comprehension; RSCS = *Reading Self-Concept Scale*; SWE = Sight Word Efficiency; TOWRE-2 = *Test of Word Reading Efficiency–2*; TX = treatment condition (MWR and MWR+MB); TX1 = MWR only; TX2 = MWR+MB; WAT = Word Attack; WID = Word Identification; WJIII = *Woodcock-Johnson III*.

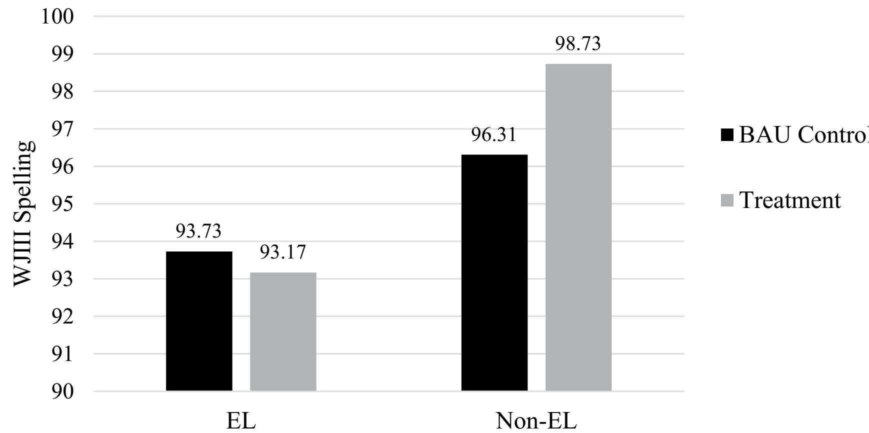


**Figure 1.** Posttest mean spelling scores for English learner (EL) and non-EL students. BAU = business as usual.

In examination of the three subtests of our motivation measure, regression analyses indicated that the treatment condition reported significantly lower scores on the RSCS Attitudes subscale as compared with those in the control condition ($\beta = -0.29$, $p < .01$, ES = −.31), which remained significant after the BH correction. No differences were found on the two additional RSCS subscales.

### *Interaction of EL Status*

The possible interaction between EL status and treatment in predicting reading outcomes was assessed. We found a significant interaction between EL status and treatment on WJIII Spelling ($b = -5.68$, $p < .05$), indicating that non-ELs benefited more than ELs from the MWR intervention (ES for non-ELs = .47, ES for ELs = −.11). See Figure 1 for mean posttest spelling scores. Treatment effects on other outcome measures did not differ according to EL status.

### *Comparison of MWR and MWR+MB Conditions*

We did not find any statistically significant differences between MWR and MWR+MB on any of the measured outcomes. Table 3 presents estimated treatment effects.

### *Replication Comparison*

When comparing effect sizes from the replication study with the original study, we found a similar pattern of effects on academic measures but disparate results on the motivation measure. Although there was only one statistically significant difference between treatment and comparison conditions on the WJIII Word Identification and Word Attack subtests, results showed that the effect sizes were identical for the Word Identification subtest (ES = .29) and comparable for the Word Attack subtest (ES = .30 in original study vs. ES = .43 in replication, $p < .05$). Results on the TOWRE-2 Phonemic Decoding Efficiency subtest in both studies were not statistically significant; however, both studies revealed positive effects sizes of similar magnitude (ES = .31 in original study vs. ES = .17 in replication). On the TOWRE-2 Sight Word Efficiency subtest, the original study found a statistically significant contrast between the treatment conditions and control, whereas the replication study did not. However, the effect sizes were moderate in both studies (ES = .73 in original study vs. ES = .39 in replication). Both studies also measured reading comprehension with sentence-level reading comprehension measures and reading motivation. Results showed that the MWR intervention treatment did not result in improved comprehension performance relative to the comparison conditions (ES = .01 for *Wide Range Achievement Test–4* [Wilkinson & Robertson, 2006] Sentence Comprehension subtest in original study vs. ES = .01 for WJIII Passage Comprehension subtest). In the original study, students who received the MWR+MB significantly and substantially outperformed those in the comparison condition (ES = .74) on reading motivation, as measured by *Reading Attribution Scale* (Berkeley, Mastropieri, & Scruggs, 2011) "success" items. Reading motivation results in the present study showed that students in the comparison condition scored better than those in the MWR+MB condition, indicating that the original study finding related to reading motivation was not reproduced in the replication study. See Figure 2 for comparisons of the overlapping measures, represented via forest plots.

## Discussion

We have described a 10-week (26-hr) experimental study of two versions of an MWR intervention (MWR-only and

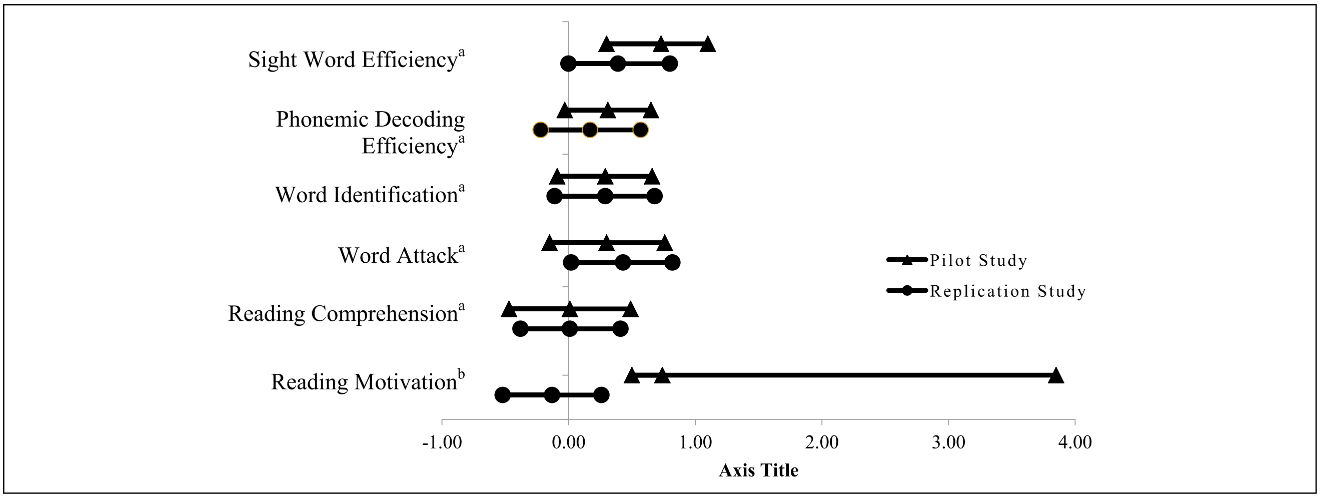


**Figure 2.** Examining agreement of effects from original and replication studies. Treatment groups were combined and contrasted with control. The multisyllabic word reading + motivational beliefs group is contrasted with the control group for the reading motivation measure. The Sight Word Efficiency and Phonemic Decoding Efficiency subtests were based on the *Test of Word Reading Efficiency–Second Edition*. The Word Identification and Word Attack subtests were taken from the *Woodcock-Johnson III*. Sentence-level reading comprehension was measured with *Wide Range Achievement Test–4* Sentence Comprehension subtest and *Woodcock-Johnson III* Passage Comprehension subtest during the pilot and replication studies, respectively. Reading motivation was assessed with the *Reading Attribution Scale* Attribution Success subscale and the *Reading Self-Concept Scale* Competence subscale during the pilot and replication studies, respectively.

MWR+MB) as compared with a business-as-usual control group. The study sought to (a) determine whether students in the MWR intervention outperformed controls, (b) examine the potential influence of EL status on effects of the MWR intervention, (c) evaluate the importance of an MB component by testing differences between the MWR conditions, and (d) examine the extent to which this replication study aligns with findings from the original study.

## *Effects of MWR Intervention*

When we combined performances of students in both MWR conditions, intervention students significantly outperformed controls on proximal measures of affix reading and MWR, as well as standardized measures of decoding, spelling, and text comprehension. Moreover, these effects were practically meaningful, with effect sizes ranging from to .25 to .90.

Although intervention students (MWR and MWR+MB) had higher performance than controls on the *Gates-MacGinitie Reading Test* as a measure of text comprehension at posttest, they were not superior to controls on the WJIII Passage Comprehension subtest. Our study findings suggest that improving upper elementary students' word-level reading skills may support their improved reading comprehension. However, as comprehension was not taught directly in this intervention, we note that this finding may be a consequence of the students accessing text more fluently and not a direct improvement to comprehension processes. Teaching reading comprehension can be more complex and challenging than teaching word-level skills (e.g., Kintsch & Kintsch, 2005; Perfetti, Landi, & Oakhill, 2005) and, as such, should be taught directly and systematically.

To tap motivational processes, we used a measure of reading self-concept (RSCS; Chapman & Tunmer, 1995) that was aligned with the MB training. While this measure was included for exploratory purposes, we found that students in the treatment conditions reported lower attitudes toward reading as compared with students in the control group. While this finding would seem contrary to expectations, there is some evidence that attitudes toward reading may decrease following intervention (Wanzek, Vaughn, Kim, & Cavanaugh, 2006). There is evidence to suggest that feedback can increase awareness of difficulties (Elbaum & Vaughn, 2001) and that children with and without disabilities often make more realistic evaluations of their performance once they receive positive feedback (e.g., Diener & Milich, 1997; Heath & Glen, 2005). In the current study, placement and participation in a reading intervention may have inadvertently triggered participants in the treatment groups to make more realistic evaluations of their reading.

## *English Learners*

We were also interested in exploring whether ELs might differentially respond to the MWR intervention. The intervention was not designed to specifically target the instructional needs of this group of students, but there is an increasing number of ELs in U.S. schools—and 27% of the current study sample was designated EL by the school district. ELs generally have more well-developed decoding skills and struggle in vocabulary and text comprehension (Geva & Massey-Garrison, 2013; Lesaux, Crosson, Kieffer, & Pierce, 2010); however, there is some evidence that ELs who are struggling readers may also have additional deficits. Lesaux and Kieffer (2010) conducted a study to examine the reading skills of sixth-grade struggling readers who were ELs and native English speakers and found that approximately 21.4% of their sample had deficits across multiple areas of reading, including word identification, decoding, and passage reading fluency.

Overall, the effects of the MWR intervention were similar across reading outcomes for ELs and non-ELs. Our findings did reveal a significant interaction between EL status and treatment on spelling performance. That is, for students who received the MWR intervention, non-ELs demonstrated greater gains in spelling than ELs. Past research has noted that native Spanish-speaking students tend to produce more spelling errors than native English-speaking peers (Fashola, Drum, Mayer, & Kang, 1996). Although reading and spelling are closely related processes (Graham, Harris, & Chorzempa, 2002; Noell, Connell, & Duhon, 2006; Santoro, Coyne, & Simmons, 2006), the actual process of spelling can be substantially more challenging (Westwood, 2008). It has also been noted that some students who have spelling deficits do not have comorbid word recognition deficits (Fletcher, Lyon, Fuchs, & Barnes, 2006), indicating that other skills and processes are involved when one is spelling words as opposed to reading words. ELs have been shown to often make spelling errors due to their knowledge of their first language (Figueredo, 2006), which may interfere with their ability to produce patterns from English orthography from memory. While this finding may support the need to further investigate the unique instructional needs for ELs who struggle with reading, we caution the interpretation of these findings as causal effects, as the ELs were not stratified across groups in this randomized controlled trial.

### *Value Added of MB Training*

In this study, we were interested in the contrast between MWR and MWR+MB to determine if the MB training component added value to the reading intervention. Although there were no statistically significant main effects between the MWR groups on proximal or standardized measures of interest, some of the effect sizes are moderate and could be considered meaningful for educational practice.

That said, the absence of reliable between-group differences prompted consideration about whether motivational training is indeed an efficacious approach to improving intervention outcomes. As previously described, research supports the notion that increased levels of motivation support students' reading performance (e.g., Guthrie & Wigfield, 2000; Logan et al., 2011; Morgan & Fuchs, 2007), but there is limited evidence in terms of how to elicit change in motivational processes in educational interventions. In this study, the MB training was grounded in motivation theory—but because motivation remains an underspecified construct, it is possible that the selected practices were inadequate. Similar to motivation research within the broader fields of education and psychology, research into reading motivation is fraught with terminology issues—and the present study is not exempt from these issues. Researchers have framed their investigations of motivational processes around diverse constructs, and although many overlap in their theoretical underpinnings, these diverse constructs focus on motivation through slightly different lenses (see Schutz & Pekrun, 2007). In a synthesis of motivation terminology in reading research, Conradi and colleagues (2014) proposed a framework that categorizes motivation within the constructs of goals (e.g., performance, mastery), beliefs (e.g., self-concept, self-efficacy, expectancy-value), and dispositions (e.g., attitude, interest). We argue that it is unlikely that all motivation theories hold equal relevance for students' reading performance (for review, see Toste, Didion, Peng, & McClelland, 2018); thus, future research must tease apart these constructs to better understand the relation between motivation and reading performance.

### *Replication Analysis*

We also examined the extent to which the effects of this replication study align with the findings of the original study (Toste, Capin, et al., 2017). Overall, we found that the studies' findings were aligned, although we note and discuss two observed discrepancies. First, when reading outcomes were compared between intervention students (MWR and MWR+MB) and controls, there was general agreement between the effects reported by both studies on standardized reading measures—as revealed by overlap in confidence intervals. The findings on measures of word reading fluency (TOWRE-2) were less aligned between the studies, with substantially higher effect sizes reported in the original study. The replication study was highly aligned with the original study. However, one key difference between these studies was the length of the interventions. In the original study, students completed 24 lessons of 40 min each for a total of 16 instructional hr, whereas the current study increased to 40 lessons of 40 min each for a total of 26.67 instructional hr. In a meta-analytic review of interventions for struggling readers in fourth through 12th grades, Scammacca, Roberts, Vaughn, and Stuebing (2015) reported that shorter interventions were associated with larger effect sizes. We caution that this finding should not be assumed to mean that shorter interventions are more beneficial to students, but we note that additional research is needed to monitor students' progress and determine how estimates of treatment effects change throughout the course of an intervention. It is possible that students in brief interventions experience an initial increase in performance but their underlying reading difficulties have not been ameliorated, as previous work has consistently found that students with significant reading problems make greater gains when provided extensive interventions (e.g., Vaughn, Solis, Miciak, Taylor, & Fletcher, 2016).

A second discrepancy arose when motivation outcomes were compared between students in the treatment conditions (MWR vs. MWR+MB) in the two studies. We did not observe agreement in findings between the original and replication studies. In the original study, the MWR+MB group outperformed the control group in ratings of attributions for success in reading. These findings were not replicated in the present study. Note that different motivation instruments were used in these studies: while each measures a facet of motivation related to reading performance, one may have inadvertently been more aligned with the intervention's MB training component. However, one of the other major differences between the interventions in the two studies was the structure of the MB training. In the present replication study, students received an additional 16 lessons of content. The cognitive-based approach continued throughout all lessons, but goal setting was also integrated in this replication study. It is possible that there was not sufficient instructional time to adequately address goal setting with students or that the changes to intrinsic motivational processes were not substantial enough to be measured through a self-report instrument.

### *Limitations and Future Research*

Finally, we note several study limitations. First, measurement issues plague the study of motivation. While the field relies almost exclusively on self-report measures, there is a question on the validity of these instruments with children and how well they align with specific motivation constructs of interest. In the current study, we used a measure of reading self-concept (beliefs and attitudes toward reading); however, motivation theory has demonstrated that changes in these processes are reliant on their interactions with recursive processes that are already present in students' environments (e.g., experiences of success and failure at school). As such, we would not presume that self-reported motivation would substantially change within a 10-week period.

The findings are also somewhat limited by sample size. While the study was sufficiently powered to examine between-group differences, we were unable to conduct more advanced analyses of potential mediators or moderators of treatment outcomes. Finally, there were differences in measures used to assess outcomes of interest between the original study (Toste, Capin, et al., 2017) and the present study, which may have influenced the replication analysis.

In summary, the absence of differences between the MWR-only and MWR+MB interventions raises a question about the value added of the MB component in its current form—and the training of motivational processes in general. How are studies of reading defining motivation? Are various motivation constructs differentially related to reading performance? Are specific reading domains (e.g., word reading vs. comprehension) differentially associated with motivation? How do student-level characteristics influence relations between reading and motivation? None of these questions, of course, suggests that motivation is an unfruitful direction for future research. Rather, more work is necessary to establish procedures to train complex cognitive processes and the measurement of these processes to promote efficacy of use.

## 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

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported in part by the Institute of Education Sciences, U.S. Department of Education, through Grant R305F100013 to The University of Texas at Austin as part of the Reading for Understanding Research Initiative. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

## ORCID iD

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

## References

Batalova, J., Fix, M., & Murray, J. (2007). *Measures of change: The demography and literacy of adolescent English learners—A report to the Carnegie Corporation of New York*. Washington, DC: Migration Policy Institute.

Beck, A. T., Hollon, S. D., Young, J. E., Bedrosian, R. C., & Budenz, D. (1985). Treatment of depression with cognitive therapy and amitriptyline. *Archives of General Psychiatry*, *42*(2), 142–148.

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*, *57*, 289–300.

Berkeley, S., Mastropieri, M. A., & Scruggs, T. E. (2011). Reading comprehension strategy instruction and attribution retraining for secondary students with learning and other mild disabilities. *Journal of Learning Disabilities*, *44*(1), 18–32.

Brandt, M. J., Ijzerman, H., Dijksterhuis, A., Farach, F. J., Geller, J., Giner-Sorolla, R., . . . Van't Veer, A. (2014). The replication recipe: What makes for a convincing replication? *Journal of Experimental Social Psychology*, *50*, 217–224.

Brasseur-Hock, I. F., Hock, M. F., Kieffer, M. J., Biancarosa, G., & Deshler, D. D. (2011). Adolescent struggling readers in urban schools: Results of a latent class analysis. *Learning and Individual Differences*, *21*(4), 438–452.

Chapman, J. W., & Tunmer, W. E. (1995). Development of young children's reading self-concepts: An examination of emerging subcomponents and their relationship with reading achievement. *Journal of Educational Psychology*, *87*, 154–167.

Cirino, P. T., Romain, M. A., Barth, A. E., Tolar, T. D., Fletcher, J. M., & Vaughn, S. (2013). Reading skill components and impairments in middle school struggling readers. *Reading and Writing*, *26*(7), 1059–1086.

Conradi, K., Jang, B. G., & McKenna, M. C. (2014). Motivation terminology in reading research: A conceptual review. *Educational Psychology Review*, *26*, 127–164.

Coyne, M. D., Cook, B. G., & Therrien, W. J. (2016). Recommendations for replication research in special education: A framework of systematic, conceptual replications. *Remedial and Special Education*, *37*(4), 244–253.

Daane, M. C. (2005). *The Nation's report card: Fourth-grade students reading aloud. NAEP 2002 special study of oral reading*. Washington, DC: National Center for Education Statistics.

Diener, M. B., & Milich, R. (1997). Effects of positive feedback on the social interactions of boys with attention deficit hyperactivity disorder: A test of the self-protective hypothesis. *Journal of Clinical Child Psychology*, *26*(3), 256–265.

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, *53*(1), 109–132.

Elbaum, B., & Vaughn, S. (2001). School-based interventions to enhance the self-concept of students with learning disabilities: A meta-analysis. *The Elementary School Journal*, *101*(3), 303–329.

Fashola, O. S., Drum, P. A., Mayer, R. E., & Kang, S. J. (1996). A cognitive theory of orthographic transitioning: Predictable errors in how Spanish-speaking children spell English words. *American Educational Research Journal*, *33*(4), 825–843.

Figueredo, L. (2006). Using the known to chart the unknown: A review of first-language influence on the development of English-as-a-second-language spelling skill. *Reading and Writing*, *19*(8), 873–905.

Fletcher, J. M., Lyon, G. R., Fuchs, L. S., & Barnes, M. A. (2006). *Learning disabilities: From identification to intervention*. New York, NY: Guilford Press.

Francis, D. J., Shaywitz, S. E., Stuebing, K. K., Shaywitz, B. A., & Fletcher, J. M. (1996). Developmental lag versus deficit models of reading disability: A longitudinal, individual growth curves analysis. *Journal of Educational Psychology*, *88*(1), 3–17.

Francis, G. (2012). The psychology of replication and replication in psychology. *Perspectives on Psychological Science*, *7*, 585–594.

Geva, E., & Massey-Garrison, A. (2013). A comparison of the language skills of ELLs and monolinguals who are poor decoders, poor comprehenders, or normal readers. *Journal of Learning Disabilities*, *46*(5), 387–401.

Graham, S., Harris, K. R., & Chorzempa, B. F. (2002). Contribution of spelling instruction to the spelling, writing, and reading of poor spellers. *Journal of Educational Psychology*, *94*(4), 669.

Guthrie, J. T., & Wigfield, A. (2000). Engagement and motivation in reading. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (pp. 403–422). Mahwah, NJ: Erlbaum.

Heath, N. L., & Glen, T. (2005). Positive illusory bias and the self-protective hypothesis in children with learning disabilities. *Journal of Clinical Child and Adolescent Psychology*, *34*(2), 272–281.

Hiebert, E. H. (2012). *QuickReads: A research-based fluency program*. Parsippany, NJ: Pearson.

Just, M. A., & Carpenter, P. A. (1987). *The psychology of reading and language comprehension*. Newton, MA: Allyn & Bacon.

Kearns, D. M. (2015). How elementary-age children read polysyllabic polymorphemic words. *Journal of Educational Psychology*, *107*(2), 364–390

Kintsch, W., & Kintsch, E. (2005). Comprehension. In S. G. Paris, S. A. Stahl, S. G. Paris, & S. A. Stahl (Eds.), *Children's reading comprehension and assessment* (pp. 71–92). Mahwah, NJ: Erlbaum.

Lesaux, N. K., Crosson, A. C., Kieffer, M. J., & Pierce, M. (2010). Uneven profiles: Language minority learners' word reading, vocabulary, and reading comprehension skills. *Journal of Applied Developmental Psychology*, *31*, 475–483.

Lesaux, N. K., & Kieffer, M. J. (2010). Exploring sources of reading comprehension difficulties among language minority learners and their classmates in early adolescence. *American Educational Research Journal*, *47*(3), 596–632.

Logan, S., Medford, E., & Hughes, N. (2011). The importance of intrinsic motivation for high and low ability readers' reading comprehension performance. *Learning and Individual Differences*, *21*, 124–128.

MacGinitie, W. H., MacGinitie, R. K., Maria, K., & Dreyer, L. G. (2000). *Gates-MacGinitie Reading Tests, Forms S and T* (4th ed.). Itasca, IL: Riverside.

McFarland, J., Hussar, B., de Brey, C., Snyder, T., Wang, X., Wilkinson-Flicker, S., . . . Hinz, S. (2017). *The condition of education 2017* (Publication No. 2017-144). Washington, DC: National Center for Education Statistics.

McGrew, K. S., Schrank, F. A., & Woodcock, R. W. (2007). *Technical Manual: Woodcock-Johnson III Normative Update*. Rolling Meadows, IL: Riverside.

Morgan, P. L., & Fuchs, D. (2007). Is there a bidirectional relationship between children's reading skills and reading motivation? *Exceptional Children*, *73*, 165–183.

Muthén, L. K., & Muthén, B. O. (2014). *Mplus* (Version 7.3). Los Angeles, CA: Authors.

Nagy, W. E., & Anderson, R. C. (1984). How many words are there in printed school English? *Reading Research Quarterly*, *19*(3), 304–330.

National Center for Education Statistics. (2016). *2015 reading assessment*. Retrieved from the NAEP Data Explorer: http://nces.ed.gov/nationsreportcard/naepdata

Noell, G. H., Connell, J. E., & Duhon, G. J. (2006). Spontaneous response generalization during whole word instruction: Reading to spell and spelling to read. *Journal of Behavioral Education*, *15*(3), 121–130.

Perfetti, C. A., Landi, N., & Oakhill, J. V. (2005). The acquisition of reading comprehension skill. In M. J. Snowling, M. S. Seidenberg, & C. Hume (Eds.), *The science of reading: Handbook of reading research* (pp. 227–247). Oxford, UK. Blackwell.

Pintrich, P. R. (2003). A motivational science perspective on the role of student motivation in learning and teaching contexts. *Journal of Educational Psychology*, *95*(4), 667.

Santoro, L. E., Coyne, M. D., & Simmons, D. C. (2006). The reading-spelling connection: Developing and evaluating a beginning spelling intervention for children at risk of reading disability. *Learning Disabilities Research & Practice*, *21*(2), 122–133.

Scammacca, N. K., Roberts, G., Vaughn, S., & Stuebing, K. K. (2015). A meta-analysis of interventions for struggling readers in Grades 4–12: 1980–2011. *Journal of Learning Disabilities*, *48*(4), 369–390.

Schmidt, S. (2009). Shall we really do it again? The powerful concept of replication is neglected in the social sciences. *Review of General Psychology*, *13*(2), 90.

Schrank, F. A., McGrew, K. S., & Woodcock, R. W. (2001). *Technical abstract* (Woodcock-Johnson III Assessment Service Bulletin No. 2). Itasca, IL: Riverside.

Schunk, D. H., Meece, J. L., & Pintrich, P. R. (2013). *Motivation in education: Theory, research, and applications* (4th ed.).Upper Saddle River,NJ:Pearson.

Schutz, P., & Pekrun, R. (Eds.). (2007). *Emotions in education*. San Diego, CA: Academic Press.

Shankweiler, D., Crain, S., Brady, S., & Macaruso, P. (1992). Identifying the cause of reading disability. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 275–305). Hillsdale, NJ: Erlbaum.

Toland, J., & Boyle, C. (2008). Applying cognitive behavioural methods to retrain children's attributions for success and failure in learning. *School Psychology International*, *29*(3), 286–302.

Torgesen, J. K., Houston, D. D., Rissman, L. M., Decker, S. M., Roberts, G., Vaughn, S., . . . Lesaux, N. (2007). *Academic literacy instruction for adolescents: A guidance document from the Center on Instruction*. Portsmouth, NH: RMC Research Corporation. Retrieved from http://files.eric.ed.gov/fulltext/ED521837.pdf

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (2012). *Test of Word Reading Efficiency* (2nd ed.). Austin, TX: Pro-ED.

Toste, J. R., Capin, P., Vaughn, S., Roberts, G. G., & Kearns, D. K. (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*(4), 593–615.

Toste, J. R., Didion, L., Peng, P., & McClelland, A. M. (2018).*Relations between reading performance and motivational processes: A meta-analytic review of the evidence*. Manuscript in final preparation.

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

Vaughn, S., Linan-Thompson, S., Kouzekanani, K., Bryant, D. P., Dickson, S., & Blozis, S. A. (2003). Reading instruction grouping for students with reading difficulties. *Remedial and Special Education*, *24*, 301–315.

Vaughn, S., Solis, 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*(1), 23–44.

Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin*, *101*(2), 192–212.

Walton, G. M., & Cohen, G. L. (2011). A brief social-belonging intervention improves academic and health outcomes of minority students. *Science*, *331*(6023), 1447–1451.

Wanzek, J., Vaughn, S., Kim, A., & Cavanaugh, C. L. (2006). The effects of reading interventions on social outcomes for elementary students with reading difficulties: A synthesis. *Reading and Writing Quarterly*, *22*, 121–138.

Wanzek, J., Wexler, J., Vaughn, S., & Ciullo, S. (2010). Reading interventions for struggling readers in the upper elementary grades: A synthesis of 20 years of research. *Reading and Writing*, *23*(8), 889–912.

Westwood, P. (2008). *What teachers need to know about spelling*. Camberwell, Australia: Australian Council for Education Research.

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

Wilkinson, G. S., & Robertson, G. J. (2006). *Wide Range Achievement Test—Fourth Edition*. Lutz, FL: Psychological Assessment Resources.

Yeager, D. S., & Walton, G. M. (2011). Social-psychological interventions in education: They're not magic. *Review of educational Research*, *81*(2), 267–301.

Zimmerman, B., & Schunk, D. H. (2006). Competence and control beliefs: Distinguishing the means and ends. In *Handbook of educational psychology* (pp. 349–367). Milton Park, England: Routledge.