Examining the Effects of Structured Word Inquiry on the Reading and Spelling Skills of Persistently Poor Grade 3 Readers

Abstract

Background: The purpose of this study was to examine the effects of Structured Word Inquiry (SWI; an orthography intervention with a focus on morphology and how morphology interrelates with phonology and etymology) and compare them to those of Simplicity intervention (a novel phonics intervention) on the reading and spelling skills of persistently poor Grade 3 readers.

Methods: Forty-eight English-speaking Canadian children (19 females, M_{age} = 8.73 years) with persistent reading difficulties were randomly assigned to one of three conditions: SWI, Simplicity, and Control. Interventions were delivered by trained graduate students over 10 weeks, 3 times a week, for 30 minutes. Children were tested three times (pre-test, post-test, and delayed post-test) on measures of phonological awareness, morphological awareness, reading (Word Reading, Morphological Relatedness, Word Attack), and spelling.

Results: Results of hierarchical linear modeling showed a significant main effect of condition for Morphological Relatedness at post-test favouring SWI and Simplicity over Control and a significant interaction between a latent variable of all secondary decoding measures and morphological awareness. Effect sizes (Cohen's *d*) showed medium to large effects for both interventions on primary outcomes of Word Reading and Morphological Relatedness.

Conclusions: These findings suggest that SWI and Simplicity can help persistently poor readers improve in some reading skills, but neither program is a panacea as a number of children continued to struggle after the intervention.

Keywords: intervention, morphological awareness, phonics, poor readers, structured word inquiry.

Implications for practice

a) What is already known about this topic

- Although English is a morphophonemic language, classroom instruction and 'pull out' intervention has mostly targeted phonology.
- Morphological instruction has the potential to be an effective method of intervention for children with reading difficulties.
- No studies have been conducted thus far to test the effectiveness of Structured Word Inquiry or Simplicity-based phonics intervention in persistently poor readers.

b) What this paper adds

- We offer the first empirical evidence on the effectiveness of Structured Word Inquiry in English and of Simplicity-based phonics intervention.
- We used Structured Word Inquiry and Simplicity-based phonics program as an intervention program with persistently poor readers so in effect we repurposed both programs.

c) Implications for theory, policy, or practice.

- Theory: If first attempts at a new form of instruction produce similar effects to what is typically treated as "best practice" (i.e., phonics), we can expect that an SWI intervention delivered over a longer period of time may produce greater long-term gains in persistently poor readers' word reading.
- Practice: One thing this study establishes is that even with minimal training of the SWI interventionists they were able to match the results of the phonics intervention that puts far less demands on teacher learning. This suggests that we should consider bringing SWI practice to wider use for persistently poor readers along with Simplicity-based content for remedial phonics programmes.

Examining the Effects of Structured Word Inquiry on the Reading and Spelling Skills of Persistently Poor Grade 3 Readers

A critical problem in reading intervention research is what interventions are effective with children who struggle with reading acquisition despite receiving high-quality Tier 1 and 2 instruction. These children, called *persistently poor readers*, have an increased risk for major academic difficulties and a cascading risk of socio-emotional, attentional and behavioural problems if their intervention needs are not met (e.g., Jordan & Dyer, 2017; Parhiala et al., 2015). In this study, we aimed (a) to examine the effects of a novel Tier 3 intervention that teaches the interrelations between phonology, morphology, and etymology (Structured Word Inquiry [SWI]; Bowers & Bowers, 2017) on the reading and spelling skills of persistently poor readers in Grade 3, and (b) to contrast the effects of SWI against those of a Simplicity intervention that involves teaching complex grapheme-phoneme correspondences (GPCs) selected through statistical optimality analysis of databases of children literature (Vousden, Ellefson, Solity, & Chater, 2011).

In the last two decades, both classroom reading instruction and 'pull-out' interventions in English have been heavily influenced by research on phonics, which emphasizes the teaching of connections between graphemes and phonemes. The adoption of phonics as the main instructional approach to reading drew on both theoretical and practical grounds. In terms of theory, it has been argued that English spellings are based on an alphabetic system in which the primary purpose of letters is to represent sounds (Byrne, 1998; Wyse & Goswami, 2008). By implication, all children should first be taught GPCs to be able to decode both familiar and unfamiliar words. Further, because the majority of children with dyslexia experience a phonological processing deficit (i.e., difficulty in identifying and manipulating

the sound structure of the language), it has been thought that reading intervention programs should target these skills in order to restore them, and phonics is doing exactly this. From a practical point of view, several intervention studies (including meta-analyses) have shown that phonics produces better results than alternative methods that do not target GPCs (e.g., Ehri, Nunes, Stahl, & Willows, 2001; Galuschka, Ise, Krick, & Schulte-Körne, 2014; Savage & Cloutier, 2018). For example, Galuschka et al. (2014) in their meta-analysis of 22 RCT reading interventions found that phonics was the only method that had a significant effect on word reading ($g' = .32$) and spelling ($g' = .34$).

However, both the theoretical and practical arguments in favour of phonics have recently been challenged (Bowers & Bowers, 2017, 2018). According to Bowers and Bowers (2018), claiming that English orthography is alphabetic or "close to alphabetic" is inaccurate. English is better described as a morphophonemic system in which spellings have evolved to represent both phonemes and meaning (through morphology and etymology). As Venezky (1967) explained "the present [English] orthography is not merely a letter-to-sound system riddled with imperfections, but, instead, a more complex and more regular relationship wherein phoneme and morpheme share leading roles" (p. 77). If we accept that the purpose of literacy instruction is to help children understand how their writing system works (e.g., Rayner, Foorman, Perfetti, Pesetsky, & Seidenberg, 2001), then children should be exposed not only to GPCs, but also to the interrelations between phonology, morphology, and etymology.

Some studies have further shown that not all children benefit from phonics instruction (Compton et al., 2014; Schaars, Segers, & Verhoeven, 2017) or that phonics/letter-sound instruction may not lead to significant effects unless it is combined with intensive training of

phonological awareness (Authors, 2019; Hatcher, Hulme, & Snowling, 2004). For example, in our intervention study, we provided 12-15 hours of explicit and systematic letter-sound instruction (guided by the Simplicity principle; Vousden et al., 2011) to a group of Grade 2 poor readers and found no significant main effects of the Simplicity intervention on word and pseudoword reading or on sentence comprehension. Only when we tested for possible interactions with phonological awareness, the effects of Simplicity intervention became significant for all aforementioned outcomes: the letter-sound instruction was effective only for children with stronger phonological awareness skills.

In view of these findings, some researchers turned their interest to interventions targeting other linguistic skills such as morphological awareness (e.g., Bar-Kochva & Hasselhorn, 2017; Lyster, Lervåg, & Hulme, 2016). In their meta-analysis, Bowers, Kirby, and Deacon (2010) reported significant effects of morphology instruction on literacy skills and further indicated that morphology interventions work better for struggling readers and children in early grades. However, Bowers et al. (2010) also indicated that very few of the identified studies were training children on the interrelations between phonology, morphology, and etymology.

An instructional approach that claims to do this is Structured Word Inquiry (SWI; see Bowers & Bowers, 2017; Bowers & Kirby, 2010). In SWI, children learn about the morphology and etymology of words together with the GPCs. Within this integrated context, students learn to generate and test hypotheses (hence the word ‘inquiry’ in SWI) about the spelling-meaning connection of words when there is a conflict (for example, Why is there a *w* in *two* but not in *too*? or Why does *action* use a *t* grapheme rather than the *sh*). This differentiates SWI from “conventional” morphological interventions that do not emphasize

inquiry on how the writing system works. In SWI, the inquiries include doing *word sums* (identifying morphemes in words and examining how they combine to make different words), building *word matrices* (matrices showing the base word with all its possible prefixes and suffixes), and learning about the etymology of words (e.g., by accessing online etymological dictionaries).

Research on SWI is still scarce. To our knowledge, only three studies have examined the effects of SWI or similar programs (Bowers & Kirby, 2010; Colenbrander et al., 2018; Devonshire, Morris, & Fluck, 2013). Bowers and Kirby (2010) examined the effect of SWI on vocabulary acquisition. They randomly assigned 81 Grade 4 and 5 children to either SWI (20 sessions at 50 minutes each) or a control condition that continued with regular classroom instruction. After the intervention, all children were tested on defining three kinds of words: (a) "Word Taught" (words that children may have been exposed to over the course of the intervention), (b) "Base Taught" (words that were never presented in lessons but that shared a base with a "Word Taught" word), and (c) "Affix Taught" (words that included trained affixes but with bases that had been carefully avoided). Results showed that the SWI group outperformed the control group in both "Word Taught" and "Base Taught" words. No significant differences between groups were obtained in the "Affix Taught" condition.

Devonshire et al. (2013) and Colenbrander et al. (2018) examined the effects of SWI on reading and spelling. Devonshire et al. (2013) compared an SWI-style intervention to a phonics intervention with 5 to 7-year-olds and reported statistically significant effects on standardized measures of word reading and spelling favouring the SWI-style intervention. However, Castles, Rastle, and Nation (2019) noted that the phonics condition in this study appeared to mix systematic phonics instruction with some at-home practice of rote learning of

whole words and encouragement to use context and picture cues. These are features that may have undermined the effectiveness of systematic phonics instruction. In Colenbrander et al. (2018), struggling readers and/or spellers (n=270) in Years 3 and 5 (ages 8 and 10) in the UK received either SWI or Motivated Reading (that involves reciprocal reading of a text and vocabulary instruction) over 24 weeks (three 20-minute sessions per week). Results showed no significant group differences, but there was a significant interaction: For children with lower pre-test scores in reading, Motivated Reading produced greater gains than SWI. In regard to spelling, there was an overall improvement in all types of words (taught and untaught real words and nonwords) for both interventions.

The Present Study

We aimed to examine the relative strengths of SWI and Simplicity principle based phonics intervention as plausible candidate interventions for persistently poor Grade 3 readers. We included a phonics intervention in our study (a) because some persistently poor readers still benefit from Tier 3 phonics interventions (see Waznek & Vaughn, 2010), and (b) because we adapted the phonics intervention to be more suitable for persistently poor readers by including phonological awareness activities. We expected a significant main effect of the interventions especially on the primary outcomes closer to the intervention content (i.e., SWI intervention would have a stronger effect on Morphological Relatedness and Simplicity intervention would have a stronger effect on Word Reading; see below for a description of these measures).

Because of our previous findings showing that the Simplicity intervention was more effective for children with stronger phonological awareness skills (Authors, 2019), we also explored here the effects of phonological awareness x Simplicity interaction on a wider set of

reading/spelling measures. Likewise, we explored the morphological awareness x SWI interaction as children with stronger morphological awareness skills may benefit more than children with weaker morphological awareness skills. Individual differences among young children in morphological awareness have already been reported in several studies (e.g., Manolitsis, Georgiou, Inoue, & Parrila, 2019). We expected that children with higher morphological awareness learn how to use morphemes with other sources of information in SWI to read better than children with lower morphological awareness.

The findings of the present study can add to the literature in four important ways. First, this is the first study on SWI with persistently poor readers. These children were recruited from public schools in XXXXXX where phonics is the medium of instruction and had received some kind of 'pull out' phonics intervention in the previous years (see below for more information). Second, there are few morphological awareness interventions that explicitly target the interrelations of phonology, morphology, and etymology as described in Bowers and Bowers (2018). Bowers et al. (2010) found that of the 22 studies in their meta-analysis only four explicitly targeted the fact that spellings of a base within a morphological family are consistent despite pronunciation changes (as in <sign> and <signal>). Third, we contrast for the first time the effects of SWI against the Simplicity program that was enhanced here by adding phonological awareness activities to the taught GPCs and by teaching children to blend the taught GPCs, thereby making it a phonics programme. Finally, neither of the previous SWI interventions included a delayed post-test. This is important in order to examine whether any gains made during the intervention are maintained over time.

Method

Participants

The participants of our study were 48 English-speaking Grade 3 children (19 females; M_{age} = 8.74 years; *SD* = .47) from Edmonton, Canada. They were recruited as follows: first, we invited children who participated in our Response to Intervention (RtI) study in Grades 1 and 2 (Authors et al., 2018, 2019), received a Tier 2 phonics intervention, but continued to struggle in reading (based on their WRAT Word Reading scores at the end of Grade 2). Second, we asked teachers from nine schools that were not part of our RtI project to nominate children from their classes who were poor readers and did not experience any intellectual, sensory, or behavioural difficulties. The schools are part of the same school division and use similar RtI practices as in our previous studies (Authors et al., 2018, 2019) and their at-risk children receive Tier 2 phonics intervention (e.g., Jolly Phonics) in Grades 1 and 2. Through this process, we identified 70 children who were then tested on WRAT-4 Word Reading (Wilkinson & Robertson, 2006) and on Matrix Reasoning and Vocabulary (Wechsler, 2011). Fifty-one children with a standard score lower than 90 in WRAT-4 Word Reading and with an average general intelligence were included and randomly assigned to one of three conditions: Simplicity, SWI, or Control with 17 children in each condition across 13 public schools. During the course of the intervention, one child from each condition withdrew from the study (they moved to a different school and could not be located), thus leaving our sample with 48 children (see Table 1 for sample characteristics). Parental and school consent was obtained prior to testing.

Materials

Parent measures. Parents filled out a questionnaire at the beginning of the study. The questionnaire included four questions: 1) Whether their child has learning disabilities (Yes, No, and if yes, then what kind of learning disability); 2) Whether their child has motor

difficulties; 3) What is mother's highest achieved educational level; and 4) What is the language spoken between parents and child at home.

Children measures. Children were assessed on measures of general intelligence (Vocabulary and Matrix Reasoning; Wechsler, 2011), phonological awareness (Elision; Wagner, Torgesen, Rashotte, & Pearson, 2013), morphological awareness (Word Analogies; Kirby et al., 2012), reading (WRAT-4 Word Reading [Wilkinson & Robertson, 2006], Morphological Relations [Mahony, Singson, & Mann, 2000], Word Attack [Schrank, Mather, & McGrew, 2014], and the Diagnostic Test of Word Reading Processes [Forum for Research in Literacy and Language, Institute of Education, 2012]), and spelling (WRAT-4 Spelling [Wilkinson & Robertson, 2006] and the Diagnostic Spelling Test [Kohnen, Colenbrander, Krajenbrink, & Nickels, 2015]). Of the reading and spelling tasks, WRAT-4 Word Reading and Morphological Relations were used as primary outcomes and the rest as secondary outcomes. A detailed description of the measures used in our study can be found in Supplementary Material.

Procedure

Test administration. There were three testing periods: pre-test (February of Grade 3), post-test (April of Grade 3), and delayed post-test (June of Grade 3). The measures of general intelligence, phonological awareness and morphological awareness were administered only at pre-test. The remaining measures were administered at all testing points.

Research assistants (RAs) conducted all assessments. They held or were studying for advanced degrees in Education or Psychology with most having teaching experience. The RAs participated in a 3-hour test administration training session delivered by one of the authors. The RAs scored all tests administered. The project coordinator checked the scoring

of the data, all data entry, and calculated derived scores. Additional training was provided in test administration and scoring by the project coordinator as required.

Interventions. One-on-one interventions took place in the winter semester of Grade 3 immediately after the pre-test. Around 60 to 70% of the time, the intervention was delivered during the Language Arts time in a quiet room at school for 30 minutes, 3 times a week. Children received an average of 12 hours of intervention over 10 school weeks while the Control group children received regular instruction. RAs were trained to run the intervention by two of the authors who are experienced in training SWI and Simplicity, respectively. Each training session lasted approximately 3 hours. In that meeting, the trainer gave an overview of the intervention goals and reviewed lesson plans one lesson at a time in detail and acted out scenarios that could arise. All RAs could, and did, contact the project coordinator and/or project leaders with questions at any point during the intervention. Observers also gave feedback directly following a lesson as part of the treatment integrity process (see below), if required.

The Simplicity intervention. The Simplicity intervention was researcher-designed (see Authors et al., 2019, for details) and involved five steps: 1) Practice phoneme blending and segmentation; 2) Introduction, definition, and spelling of a new "word of the day"; 3) Search for that word in authentic children's books selected to represent these words; 4) Shared reading of researcher-written texts that repeated the word of the day; and 5) Introduction of the "sound of the day" – a GPC within the word of the day articulated by the RA. Children said and wrote the grapheme and then identified the grapheme in texts that were often researcher-written and designed to incorporate a high density of taught GPCs. Finally, they practiced blending the new GPC with previously taught GPCs to read simple words and wrote

a sentence using the newly taught grapheme pattern in their notebook. The list of GPCs was adopted from Vousden et al. (2011). The inclusion of phonological awareness and blending of taught GPCs builds on what we did in our study with Grade 2 struggling readers in which the Simplicity intervention included only direct teaching of letter-sound correspondences.

The SWI intervention. The SWI intervention was developed by Bowers (2009). Like Simplicity, SWI included instruction of GPCs. However, in SWI, instruction of GPCs occurred with reference to the morphological context, and when appropriate, to etymology. In addition, while explicit, teaching of GPCs in SWI was not systematic in that there was no predetermined scope of GPCs that would be necessarily covered.

Four inquiry questions provided a framework to guide the spelling-meaning-pronunciation investigations in SWI: 1) What is the meaning of the word? 2) How is the word built (can any bases or affixes be identified with a word sum)? 3) What related words can you find (morphological relatives that share a base, or etymological relatives that share a historical root)? and 4) What graphemes function coherently? (representation of phonemes across morphological boundaries and influence of word origin on grapheme choice). During the course of the intervention, three main spelling conventions were targeted and explicitly taught: 1) Replacing the single, non-syllabic <e> before adding a vowel suffix; 2) Changing the <y> to <i> of a base word before adding a vowel suffix; and 3) Doubling a consonant of a base word before adding a suffix.

An integral component of SWI is the process of "writing-out-loud word sums" and the creation of a word matrix. For example, the interventionist would guide students to spell-out the base <knight> "kn - igh - t" and explicitly associate the <kn> digraph with the phoneme /n/, the trigraph <igh> with the phoneme /aɪ/, and the single-letter grapheme <t> with the

phoneme /t/. Within the framework of word sums, students were guided to announce prefixes (e.g., <de->, <re->) and suffixes (e.g., <-ed>, <-ment>, <-ure>, <-ing>).

Once the word sums were sufficiently practiced, students were guided to build a word matrix (see Appendix A). A matrix has a central base that can be combined with the other morphological elements (bases and affixes) to form complex words. Word sums show the morphological structure of individual words in a matrix. The vertical lines in the matrix correspond to the plus sign in the word sum. Horizontal lines cannot be crossed to form words. The matrix shows the full form of written morphemes in the cells, but the word sum can show suffixing changes (e.g., make/ + ing ➔ making; hop(p) + ing ➔ hopping; try/i + es ➔ tries). The matrix and word sums provide concrete representations of the underlying morphological structure of words and the surface realisations that we see in print. They provide "worked-out-examples" of complex information as recommended by cognitive load theory (Schnotz & Kürschner, 2007).

Treatment integrity. In order to assess treatment integrity (TI), both reading interventions were frequently observed. We created a TI rubric that reflected: 1) the specific Content of each intervention, 2) adherence to the specified Time Management, 3) the broader Teaching Quality, and 4) broader aspects of the one-on-one Learning Environment. Each component was assessed with a series of between 3 and 11 sub-components on a 3-point scale (0 = "not done", 1 = "partly done", 2 = "fully done"). Observations were audio recorded and independently assessed by the first and third author. Of these sessions, 100% of the intervention sessions were independently observed by two RAs to ensure ongoing inter-rater reliability. Analyses of inter-rater reliability on all scores showed 97% agreement in Simplicity, 95% in the SWI, and 98% in Control interventions. Mann-Whitney U tests for

each TI component by condition (Simplicity versus Control; SWI versus Control), adjusting for multiple contrasts were non-significant ($p > .10$ in all cases), confirming that both interventions were equally well implemented.

Results

Preliminary Data Analyses

All data were first screened for deviations from normality and no significant deviations from normality were detected. Little's MCAR test further showed that the data were missing (the total missing data in our study was 0.08%) completely at random, $x^2(34) = 40.38$, $p = .24$ *ns*. Inspection of standard scores in WRAT Word Reading also showed that 14 children from the two intervention conditions continued to have a standard score below 90 at post-test and 17 of them at delayed post-test.

Group Comparisons on Extraneous Variables

We performed a post-allocation checking to see if the three groups differed on any extraneous variables including pre-test attainment, general intelligence, home language, parent–reported developmental history, age, and gender. Results are reported in Table 1 and show no significant differences between groups in any of these variables.

Data Reduction for Secondary Measures

A total of eight subtests formed our battery of secondary measures (the three subtests of DTWRP, the three spelling subtests of DiST, WRAT Spelling, and Word Attack). Three exploratory factor analyses were performed for the pre-test, post-test, and delayed post-test scores, respectively. The total matrix sampling adequacy for factor analysis was sufficient in each cases (KMO = .78 pre-test; KMO = .84 post-test; KMO = .85 delayed post-test) and Bartlett's Test of Sphericity indicated the presence of a factor structure, (x^2 (28) = 202.93, $p <$

.001 pre-test; x^2 (28) = 276.40, p < .001 post-test; x^2 (28) = 249.72, p < .001 delayed post-test). Principal component analysis with a Varimax rotation yielded two factors with eigenvalues greater than one involving all items with factor loadings above .45, and together explained over 70% of the variance. Analysis with pre-test, post-test, and delayed post-test scores all yielded two factors with the same loading patterns. The two factors extracted at each phase of testing were labelled 'Reading and Spelling' and 'Decoding'. The 'Decoding' factor loaded most heavily on the pseudoword tasks, whereas the 'Reading and Spelling' factor loaded most strongly on real word reading and spelling tasks. Finally, the Anderson–Rubin method was used to save all latent factor *z*-scores. Anderson-Rubin factor scores are orthogonal and have a mean of 0 (*SD* = 1). Each pair of latent variables at pre-test, post-test, and delayed post-test was then modelled in secondary analyses in subsequent HLM.

Intervention Results

Data were analysed with HLM in order to account for the nested nature of data within schools (Hox, 2010; Raudenbush & Bryk, 2002). The final HLM models were built sequentially from preliminary analyses. Model 1 was an unconditional one-way ANOVA model with random effects. These unconditional models confirmed that one of the key assumptions of ordinary least squares (independence of scores) was violated as there was substantial school-level variance of up to 25% at post-test and delayed post-test on primary outcomes.

In the final two-level hierarchical models presented in Tables 3 and 4, we examined whether variance on post-test and delayed post-test outcome measures (after controlling for school-level mean pre-test classroom-level achievement variance at level 2), was explained by a main effect of Condition (SWI versus Simplicity versus Control). An ANCOVA model

was appropriate because controls for nested pre-test attainment improve the power of analyses even if the covariate is not statistically significant (Raudenbush et al., 2011). In addition, we ran analyses to explore whether any effects at post-test were moderated by (a) pre-test phonological awareness skills or (b) pre-test morphological awareness skills at leve1 1. A covariance model was also appropriate here as we explicitly sought to test Simplicity x pre-test Phonological Awareness skills and SWI x pre-test Morphological Awareness skills interactions in predicting growth in attainment, even after controlling for pre-test reading ability.

Equations 1 and 2 describe this final model at the student and school levels, for student *i* in school *j*, respectively for models including the interaction with phonological awareness skills.

1) Equation for Student Level 1 Model:

$$Y_{ij} = \beta_{0j} + \beta_{1j}\,(\textit{Phonological Awareness}) + \beta_{2j}\,(\textit{Intervention}) + \beta_{3j}$$

$$(\textit{Phonological Awareness} * \textit{Intervention}) + e_{ij}$$

2) Equations for School Level 2 Model:

$$\beta_{0j} = \gamma_{00} + {}_{01j}\,(\textit{Pre-test Attainment}) + r_{0j,}$$

For the parallel analyses exploring interaction effects involving Morphological Awareness skills the equations are depicted below.

3) Equation for Student Level 1 Model:

$$Y_{ij} = \beta_{0j} + \beta_{1j}\,(\textit{Morphological Awareness}) + \beta_{2j}\,(\textit{Intervention}) + \beta_{3j}$$

$$(\textit{Morphological Awareness} * \textit{Intervention}) + e_{ij}$$

4) Equations for School Level 2 Model:

$$\beta_{0j} = \gamma_{00} + {}_{01j}\,(\textit{Pre-test Attainment}) + r_{0j,}$$

In all analyses, predictor variables were grand mean-centered raw scores. Restricted Maximum Likelihood (REML) was employed to estimate variance components, with the aim of achieving unbiased errors of parameter estimates. Identity error matrices were specified for all HLM analyses.

The means and standard deviations are presented in Table 2. Inspection of these data suggests some advantages at post-test for the Simplicity and SWI conditions over the Control group on WRAT Word Reading and Morphological Relatedness compared to the pre-test scores. Results of the HLM analyses are reported in Table 3 where researcher-created dummy-coded Simplicity and SWI interventions are compared against the zero-coded Control condition.

The HLM results show a significant main effect of intervention for the Morphological Relatedness measure at post-test. Here, both SWI and Simplicity were advantaged over Controls when either morphological awareness or phonological awareness were included in the model ($p < .05$ in all cases). At delayed post-test the effect of condition approached significance when morphological awareness was included in the model. The effect reflected a larger mean estimate for SWI ($p = .051$) and Simplicity ($p = .073$) over Controls at delayed post-test. There were no significant intervention x phonological awareness or intervention x morphological awareness interaction effects for WRAT Word Reading or Morphological Relatedness at either post-test or delayed post-test.

The results of analyses with the latent factors representing the secondary outcomes are depicted in Table 4. These analyses showed that there were no significant main effects of intervention at post-test or delayed post-test. There was however a significant intervention x morphological awareness interaction effect for the ‘Decoding’ latent factor at post-test; both

Simplicity and SWI were advantaged over Controls across all decoding tasks when children had stronger morphological awareness skills at pre-test ($p < .05$ in both cases).

Effect Size Analysis

Effect sizes were calculated using the difference between the respective pre- and post-test means as the numerator and the pooled pre-test standard deviation as the denominator (Cohen, 1988). Cohen's *d* can be used to assess effect sizes within hierarchical models (see e.g., Snijders & Bosker 2012). Here, however, they were used outside of the hierarchical model to explore the practical significance of observed performance changes. Effect sizes for the latent secondary variable factor scores were not computed as they do not capture growth.

Inspection of Table 2 shows large effects for Simplicity at post–test and SWI at post–test and delayed post–test on the WRAT Word Reading. Small-to-medium effect sizes are evident for the Simplicity intervention at delayed post-test and for the Control condition at both post-tests. However, upper and lower 95% CIs at T1-T2 (.17 – 1.62) and T1-T3 (.33 – 1.81) were both positive only in the SWI condition. Effect sizes for Morphological Relatedness were negative or small in the Control condition at post-test and delayed post-test and ranged between small to medium and large for the SWI and Simplicity at post- and delayed post-test.

Discussion

The present study examined the effects of two alternative programs (SWI and Simplicity) on the word reading and spelling skills of persistently poor readers in Grade 3. We hypothesised that SWI and Simplicity interventions would improve persistently poor readers' performance with added benefits for each intervention observed when predicting the primary outcomes closer to the content of each intervention. In addition, we anticipated

significant interactions between children's phonological awareness skills and response to the Simplicity intervention, and between children's morphological awareness skills and response to the SWI intervention.

The results only partly confirmed our hypotheses. First, in contrast to our hypothesis, we found no significant effects of either intervention on WRAT Word Reading. Given that most words in this task are morphologically simple, the fact that SWI did not lead to a significant improvement is perhaps unsurprising. However, the Simplicity intervention did not improve WRAT Word Reading either. This implies that neither the inclusion of phonological awareness activities and GPC blending in the Simplicity intervention nor the increase in the frequency of delivering the intervention were sufficient to help this group of persistently poor readers who had already received phonics instruction (e.g., Jolly Phonics) in earlier grades. Second, in line with our hypothesis, both SWI and Simplicity produced better results in Morphological Relatedness in both post-test and delayed post-test compared to the control group. This is encouraging as it shows that Tier 3 interventions with very different foci can help children identify semantically related word pairs. Because Morphological Relatedness is a reading task, children in the Simplicity intervention may improve in it because they can read the words more accurately. In turn, children in the SWI intervention may improve because they understand the interaction with morphology and other features of language better.

The effect sizes in both WRAT Word Reading and Morphological Relatedness were relatively large for both intervention groups. According to What Works Clearinghouse (2014), effect sizes larger than .25 are considered to be substantively important, particularly for students at risk for reading difficulties. Effect sizes are important particularly in studies

with a ‘conservative’ design as ours (see Fletcher & Wagner, 2014, for a relevant discussion). Interestingly, effect sizes followed a different pattern in the intervention groups. Irrespective of the primary reading outcome, the effect sizes for SWI increased from post-test to delayed post-test. In contrast, the effect sizes for Simplicity decreased from post-test to delayed post-test. An inspection of the upper and lower 95% CIs in WRAT Word Reading also reveals that they were positive only in SWI and not in the other conditions including control. Thus, we are confident we have an effect of at least .17 (T2) and .33 (T3) here, which is not negligible in the intervention world.

The large effect sizes observed for SWI at delayed post-test (d = .71 for Morphological Relatedness and d = .98 for Word Reading) support the argument put forward by Bowers and Bowers (2017) that SWI takes advantage of the benefits predicted from cognitive load theory (i.e., developing well-represented mental representation of schema through deeper processing that comes with leveraging multiple memory routes; Schnotz & Kürschner, 2007). Relatedly, Kirby and Bowers’ (2017) binding agent model of morphology would predict that instruction targeting the interrelations of meaning (morphology and etymology), orthography and phonology results in higher quality and longer lasting lexical representations than what would be expected after instruction that targets fewer aspects (Simplicity lacks the direct semantic aspect included in SWI). These are interesting suggestions for mechanisms of impact that require further examination.

One hindrance to researchers suggesting the inclusion of morphology in instruction for struggling students is a fear that including morphology in literacy instruction would reduce time for studying GPCs and thus reduce learning opportunities for these already struggling students (e.g., Henbest & Apel, 2017). We do not need to show significant benefits of SWI

over Simplicity in this study to counter that fear. Showing equal effects (Simplicity did not outperform SWI in any outcomes) demonstrates that including morphology may not be as risky as initially thought. However, we need to acknowledge that all of our participants had already been exposed to two years of phonics instruction and it is possible that SWI builds on that foundation by giving children an alternative approach to remember and understand GPCs.

In contrast to our hypothesis, neither of the interaction terms (phonological awareness x Simplicity and morphological awareness x SWI) were significant in predicting the results in the two primary outcomes. SWI seemed to work for all as indicated by the significant main effect. In addition, unlike Authors (2019), we taught phonological awareness in the Simplicity condition. In Authors (2019), we explicitly tested the phonological awareness x Simplicity interaction by only teaching the GPCs and no phonological awareness. Teaching both in the Simplicity condition appears to help more children. Finally, the fact that the intervention was delivered on one-on-one basis may have allowed differentiation of instruction in both phonological awareness (in the Simplicity condition) and morphological awareness (in the SWI condition) that helped every child.

We did, however, find a significant interaction between the ‘Decoding’ factor and morphological awareness, indicating that higher morphological awareness predicted greater decoding factor scores. This is in line with recent arguments that decoding is built on a bed of strong language skills (Chang & Monaghan, 2019). It may also suggest that children who start with higher morphological awareness learn how to weave that knowledge with other sources in SWI to read.

Limitations and Future Studies

Some limitations of our study are worth noting. First, neither intervention program was a “magic bullet” as a number of children remained poor readers after the completion of the intervention. Second, there were no main effects on the “Reading and Spelling” factor indexing secondary outcomes for either SWI or Simplicity and none for secondary measures for SWI. This suggests that the effects of interventions did not generalise widely. Third, although there were no significant differences in vocabulary (see Table 1), the performance of children in SWI was somewhat lower than in the other two groups. This may have affected the results given that SWI integrates word meanings with spelling patterns. Finally, our intervention was relatively short (only 30 sessions). This may not be enough particularly for persistently poor readers. A future study should replicate our findings using longer and perhaps more frequently-delivered interventions.

Conclusion

Our findings add to a growing body of research with persistently poor readers (e.g., Denton et al., 2013; O’Connor, Fulmer, Harty, & Bell, 2005; van der Kleij, Segers, Groen, & Verhoeven, 2019) by suggesting that exposing children to multiple layers of their orthography (i.e., phonology, morphology, and etymology) can be beneficial (at least for some children). Clearly, we need larger studies to be confident about our results. However, statistically significant effects in small (conservative) well-conducted RCT designs are hard to obtain and we do achieve this. Significant effects in main analyses coincide with medium-to-large effect sizes for interventions in our analyses. The positive effect of the Simplicity intervention on aspects of reading in persistent poor readers in Grade 3 reported here replicates patterns reported elsewhere for Grade 1 and 2 ‘at-risk’ readers (Authors, 2019; Chen & Savage, 2014), and the positive effects of SWI replicate those reported in some earlier studies (Devonshire et

al., 2013), together increasing confidence in the results we report here. In future research we need to contrast SWI alone and Simplicity alone against an intervention combining the two to test the contention that using instruction like SWI that explicitly reflects the morphophonemic nature of English brings benefits compared to phonics instruction.

References

Authors (2018)

Authors (2019)

Bar-Kochva, I., & Hasselhorn, M. (2017). The training of morphological decomposition in word processing and its effect on literacy skills. *Frontiers in Psychology, 8*, 1–17. doi:10.3389.fpsychg. 2017.01583

Bowers, J. S., & Bowers, P. N. (2017). Beyond phonics: The case for teaching children the logic of the English spelling system. *Educational Psychologist, 52*, 124–141.

Bowers, J. S., & Bowers, P. N. (2018). Progress in reading instruction requires a better understanding of the English spelling system. *Current Directions in Psychological Science, 27*, 407–412. doi:10.1177/0963721418773749

Bowers, P. N. (2009). *Teaching how the written word works.* (available from www.wordworksinternational.com)

Bowers, P. N., & Kirby, J. R. (2010). Effects of morphological instruction on vocabulary acquisition. *Reading and Writing: An Interdisciplinary Journal, 23*, 515–537. doi:10.1007/ s11145-009-9172-z

Bowers, P. N., Kirby, J. R., & Deacon, S. H. (2010). The effects of morphological instruction on literacy skills: A systematic review of the literature. *Review of Educational Research, 80*, 144–179.

Byrne, B. (1999). *The foundations of literacy*. London, UK: Routledge.

Castles, A., Rastle, K., & Nation, K. (2018). Ending the reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest, 19*, 5–51. doi:10.1177/1529100618772271

Chen, V., & Savage, R. S. (2014). Evidence for a simplicity principle: Teaching common complex grapheme-phonemes improves reading and motivation in at-risk readers. *Journal of Research in Reading, 37*, 196–214. doi:10.1111/1467-9817.12022

Chang, Y.-N., & Monaghan, P. (2019). Quantity and diversity of preliteracy language exposure both affect literacy development: Evidence from a computational model of reading. *Scientific Studies of Reading, 23*, 235–253. doi:10.1080/10888438.2018.1529177

Colenbrander, D., Parsons, L., Murphy, S., Hon, Q., Bowers, J., & Davis, C. (2018, July). *Morphological intervention for children with reading and spelling difficulties*. Paper presented at the annual conference of the Society for the Scientific Studies of Reading, Brighton, UK.

Compton, D. L., Miller, A. C., Elleman, A. M., & Steacy, L. M. (2014). Have we forsaken reading theory in the name of 'quick fix' interventions for children with reading disability? *Scientific Studies of Reading, 18*, 55–73. doi:10.1080/10888438.2013.836200

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*, 633–648. doi: 10.1037/a0032581.

Devonshire, V., Morris, P., & Fluck, M. (2013). Spelling and reading development: the effect of teaching children multiple levels of representation in their orthography. *Learning and Instruction, 25*, 85–94. doi:10.1016/j.learninstruc.2012.11.007

Ehri, L. C., Nunes, S. R., Stahl, S. A., & Willows, D. M. (2001). Systematic phonics instruction helps students learn to read: Evidence from the National Reading Panel's meta-analysis. *Review of Educational Research, 71*, 393– 447.

Fletcher, J. M., & Wagner, R. K. (2014). Accumulating knowledge: When are reading intervention results meaningful? *Journal of Research on Educational Effectiveness, 7,* 294–299. doi:10.1080/19345747.2014.925310

Galuschka, K., Ise, E., Krick, K., & Schulte-Körne, G. (2014). Effectiveness of treatment approaches for children and adolescents with reading disabilities: A meta-analysis of randomized controlled trials. *Plos One, 9*(8), e105843. Retrieved from http://dx.doi.org/10.1371/journal.pone.0105843

Hatcher, P. J., Hulme, C., & Snowling, M. J. (2004). Explicit phoneme training combined with phonic reading instruction helps young children at risk of reading failure. *Journal of Child Psychology and Psychiatry, 45*, 338–358. doi:10.1111/j.1469-7610.2004.00225.x

Henbest, V. S., & Apel, K. (2017). Effective word reading instruction: what does the evidence tell us? *Communication Disorders Quarterly, 39,* 303–311. doi:10.1177/1525740116685183

Hox, J. J. (2010). *Multilevel analysis: Techniques and applications (2nd ed.)*. New York, NY: Routledge.

Jordan, J., & Dyer, K. (2017). Psychological well-being trajectories of individuals with dyslexia aged 3-11 years. *Dyslexia, 23,* 161–180. doi:10.1002/dys.1555

Kirby, R. J., & Bowers, P. N. (2017). Morphological instruction and literacy: Binding phonological, orthographic, and semantic features of words. In K. Cain, D. Compton, &

R. Parrila (Eds.), *Theories of reading development* (pp. 437–461). Amsterdam, The Netherlands: John Benjamins.

Kirby, J. R., Deacon, S. H., Bowers, P. N., Izenberg, L., Wade-Woolley, L & Parrila, R. (2012). Children's morphological awareness and reading ability. *Reading and Writing: An Interdisciplinary Journal. 25,* 389–410. doi:10.1007/s11145-010-9276-5

Kohnen, S., Colenbrander, D., Krajenbrink, T., & Nickels, L. (2015). Assessment of lexical and non-lexical spelling in students in grades 1–7. *Australian Journal of Learning Difficulties, 20,* 15–38. doi:10.1080/19404158.2015.1023209

Lyster, S.-A. H., Lervåg, A., & Hulme, C. (2016). Preschool morphological training produces long-term improvements in reading comprehension. *Reading and Writing: An Interdisciplinary Journal, 29,* 1269–1288. doi: 10.1007/s11145-016-9636-x

Mahony, D., Singson, M., & Mann, V. (2000). Reading ability and sensitivity to morphological relations. *Reading and Writing: An Interdisciplinary Journal, 12,* 191–218. doi:10.1023/A:1008136012492

Manolitsis, G., Georgiou, G., Inoue, T., & Parrila, R. (2019). Are morphological awareness and literacy skills reciprocally related? Evidence from a cross-linguistic study. *Journal of Educational Psychology, 111,* 1362–1381. https://doi.org/10.1037/edu0000354

O'Connor, R. E., Fulmer, D., Harty, K. R., & Bell, K. M. (2005). Layers of reading intervention in kindergarten through third grade: Changes in teaching and student outcomes. *Journal of Learning Disabilities, 38*, 440–455. doi:10.1177/002221940 50380050701

Parhiala, P., Torppa, M., Eklund, K., Aro, T., Poikkeus, A. M., Heikkilä, R., & Ahonen, T. (2015). Psychosocial functioning of children with and without dyslexia: A follow-up study from ages four to nine. *Dyslexia, 21,* 197–211. doi: 10.1002/dys.1486.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks, CA: Sage Publications.

Raudenbush, S. W., Spybrook, J., Congdon, R., Liu, X. F., Martinez, A., & Bloom, H. (2011). Optimal design software for multi-level and longitudinal research (Version 3.01) [Software]. Retrieved from www.wtgrantfoundation.org

Rayner, K., Foorman, B. R., Perfetti, C. A., Pesetsky, D., & Seidenberg, M. S. (2001). How psychological science informs the teaching of reading. *Psychological Science in the Public Interest, 2*, 31–74.

Savage, R. S., & Cloutier, E. (2017). Early reading interventions: The state of the practice and new directions in building causal theoretical models. In K. Cain, D. Compton, & R. Parrila (Eds.), *Theories of reading development* (pp. 409–436). Amsterdam, The Netherlands: John Benjamins.

Schaars, M. M. H., Segers, E., & Verhoeven, L. (2017). Word decoding development during phonics instruction in children at risk for dyslexia. *Dyslexia*, *23*, 141–160. doi:10.1002/dys.1556

Schnotz, W., & Kürschner, C. (2007). A reconsideration of cognitive load theory. *Educational Psychology Review, 19*, 496–508. doi: 10.1007/s10648-007-9053-4

Schrank, F. A., Mather, N., & McGrew, K. S. (2014). *Woodcock Johnson IV Tests of Achievement.* Rolling Meadows, IL: Riverside Publishing.

Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and*

advanced multilevel modelling (2nd ed.). London, UK: Sage Publishers.

van der Kleij, S. W., Segers, E., Groen, M. A., & Verhoeven, L. (2019). Post-treatment reading development in children with dyslexia. The challenge remains. *Annals of Dyslexia, 69*, 279–296. doi:10.1007/s11881-019-00186-6.

Venezky, R. (1967). English orthography: Its graphical structure and its relation to sound. *Reading Research Quarterly, 2*, 75–105.

Vousden, J. I., Ellefson, M. R., Solity, J., & Chater, N. (2011). Simplifying reading: Applying the simplicity principle to reading. *Cognitive Science, 35*, 34–78. doi:10.1111/j.1551-6709.2010.01134.x

Wagner, R.K., Torgesen, J.K., & Rashotte, C.A., & Pearson, N. A. (2013). *Comprehensive Test of Phonological Processing 2*. Austin, TX: Pro-Ed.

Wanzek, J., & Vaughn, S. (2010). Tier 3 interventions for students with significant reading problems. *Theory into Practice, 49*, 305–314. https://doi.org/10.1080/00405841.2010.510759

Wechsler, D. (2011). *Wechsler Abbreviated Scale of Intelligence, Second Edition (WASI-II)*. San Antonio, TX: NCS Pearson.

What Works Clearinghouse, Institute of Education Sciences, U.S. Department of Education (2014). *WWC procedures and standards handbook (Version 3.0)*. Retrieved from http://ies.ed.gov/ncee/wwc/documentsum.aspx?sidD19

Wilkinson, G. S., & Robertson, G. J. (2006). *Wide Range Achievement Test 4.* Wilmington, DE: Wide Range.

Wyse, D., & Goswami, U. (2008). Synthetic phonics and the teaching of reading. *British Educational Research Journal, 34*, 691–710. doi:10.1080/01411920802268912

Table 1

Characteristics of the Intervention Sample by Condition

Condition	SWI	Simplicity	Controls	*F* value and significance
Gender (*n* = female) [b]	5	7	7	.70 *ns*
Chronological Age yr month (range) [b]	8:80 (1:4)	8:10 (1:7)	8:70 (1:6)	.93 *ns*
Parent-reported motor problems [c]	14	9	9	.30 *ns*
Parent-reported LD [c]	44	39	50	.83 *ns*
Mother's education [d]	5	5	4.3	.67 *ns*
English as home language [c]	63	75	63	.38 *ns*
WRAT Word Reading [a]	81.60 (9.26)	81.19 (9.07)	82.19 (5.11)	.06 *ns*
WASI Vocabulary [b]	19.12 (7.20)	22.37 (4.35)	21.50 (5.88)	1.28 *ns*
WASI Matrix Reasoning [b]	10.00 (5.22)	11.00 (3.63)	11.68 (2.67)	.72 *ns*
Elision [b]	14.19 (8.32)	16.75 (7.07)	15.68 (4.84)	.58 *ns*
Word Analogies [b]	5.56 (3.69)	5.75 (4.74)	6.62 (3.15)	.33 *ns*

Note: SWI = Structured Word Inquiry; LD = Learning Disabilities; WRAT = Wide Range Achievement Test; WASI = Wechsler Abbreviated Scale of Intelligence. Values represented by (a) standard scores, (b) raw scores, (c) percentage, (d) median. Mother's education was measured on a 7-point scale (1 = finished elementary school, 2 = did not receive school graduate diploma, 3 = left school with graduation diploma, 4 = received technical training, 5 = finished college, 6 = university bachelor's degree, and 7 = completed graduate studies). ANOVA was used to test for group differences in continuous variables and chi square for group differences in frequencies.

Table 2

Means, Standard Deviations, and Effect Sizes with 95% Confidence Intervals for Pre-Test, Post-Test, and Delayed Post-Test Literacy Measures by Intervention Group

	SWI			Simplicity			Controls			SWI Effect Size		Simplicity Effect Size		Control Effect Size	
	T1	T2	T3	T1	T2	T3	T1	T2	T3	T1-T2	T1-T3	T1-T2	T1-T3	T1-T2	T1-T3
WR[a]	81.60 (9.26)	88.40 (9.45)	89.27 (10.89)	80.07 (9.18)	86.21 (9.26)	84.64 (11.44)	82.19 (5.11)	85.19 (4.33)	86.75 (6.44)	0.89 0.17 1.62	0.98 0.33 1.81	0.78 0.06 1.50	0.58 -0.06 1.37	0.38 -0.12 1.30	0.56 -0.01 1.38
MR[b]	14.53 (2.03)	15.20 (2.62)	16.20 (2.24)	13.93 (2.84)	14.36 (3.15)	15.21 (3.24)	14.38 (2.39)	13.13 (1.99)	14.88 (2.36)	0.28 -0.37 1.02	0.71 0.03 1.45	0.43 -0.26 1.14	0.36 -0.31 1.09	-0.59 -1.25 0.16	0.21 -0.49 0.90
RSF[c]	0.02 (1.07)	-0.18 (1.02)	-0.15 (1.12)	-0.14 (0.96)	.093 (1.15)	0.06 (1.10)	0.12 (1.02)	0.083 (0.86)	0.09 (0.75)						
DF[c]	0.24 (1.22)	.21 (1.11)	0.34 (1.17)	-0.02 (0.67)	-0.07 (0.91)	-0.22 (0.64)	-.22 (1.04)	-.14 (0.96)	-0.11 (1.05)						

Note: T1 = Pre-test; T2 = Post-test; T3 = Delayed Post-test; WR = Word Reading (WRAT-4); MR = Morphological Relatedness; RSF = Reading & Spelling factor; DF = Decoding factor. Values are represented by (a) standard scores, (b) raw scores, and (c) z-scores.

Table 3

HLM Results for Intervention Condition by Phonological/Morphological Ability on Primary Post-Test Attainment Measures

	Student-Level Model		School-Level Model	
Variable	Coefficient	SE	Coefficient	SE
DV = WRAT Word Reading post-test				
Intercept	28.32	(1.00)***		
Phonological Awareness	0.12	(0.18)		
Intervention				
Simplicity	1.59	(1.38)		
SWI	1.25	(1.41)		
Intervention x Phonological Awareness	0.12	(0.12)		
WRAT Word Reading pre-test			0.94	(0.22)***
Variance Components				
Child	14.52	(3.54)		
School			0.86	(2.01)
DV = WRAT Word Reading post-test				
Intercept	28.28	(1.05)***		
Morphological Awareness	0.03	(0.31)		
Intervention				
Simplicity	1.89	(1.43)		
SWI	1.05	(1.46)		
Intervention x Morphological Awareness	0.30	(0.24)		
WRAT Word Reading pre-test			0.95	(0.23)***
Variance Components				
Child	15.74	(3.89)		
School			0.90	(2.32)
DV = WRAT Word Reading delayed post-test				
Intercept	29.86	(1.24)***		
Phonological Awareness	1.86	(0.22)		
Intervention				
Simplicity	0.33	(1.64)		
SWI	1.07	(1.67)		
Intervention x Phonological Awareness	0.16	(0.14)		
WRAT Word Reading pre-test			0.87	(0.28)**
Variance Components				
Child	19.57	(4.96)		
School			2.76	(3.71)
DV = WRAT Word Reading delayed post-test				
Intercept	29.61	(1.33)***		
Morphological Awareness	0.10	(0.40)		
Intervention				
Simplicity	0.70	(1.79)		
SWI	0.83	(1.79)		
Intervention x Morphological Awareness	0.34	(0.29)		
WRAT Word Reading pre-test			0.90	(0.30)**
Variance Components				
Child	22.74	(5.91)		
School			2.91	(4.64)

DV = Morphological Relatedness post-test				
Intercept	12.99	(0.66)***		
Phonological Awareness	0.14	(0.11)		
Intervention				
Simplicity	1.82	(0.88)*		
SWI	2.22	(0.91)*		
Intervention x Phonological Awareness	0.01	0.08		
Morphological Relatedness pre-test			0.48	(0.28)
Variance Components				
Child	5.74	(1.41)		
School			0.74	(0.97)
DV = Morphological Relatedness post-test				
Intercept	12.97	(0.66)***		
Morphological Awareness	0.16	(0.19)		
Intervention				
Simplicity	2.15	(0.90)*		
SWI	2.10	(0.92)*		
Intervention x Morphological Awareness	0.09	(0.15)		
Morphological Relatedness pre-test			0.50	(0.27)
Variance Components				
Child	6.07	(1.49)		
School			0.45	(0.91)
DV = Morphological Relatedness delayed post-test				
Intercept	14.80	(0.70)***		
Phonological Awareness	0.23	(0.12)$^{+}$		
Intervention				
Simplicity	0.79	(0.95)		
SWI	1.46	(0.94)		
Intervention x Phonological Awareness	-0.22	(0.08)		
Morphological Relatedness pre-test			0.24	(0.30)
Variance Components				
Child	5.99	(1.70)		
School			1.08	(1.71)
DV = Morphological Relatedness delayed post-test				
Intercept	14.47	(0.75)***		
Morphological Awareness	0.46	(0.20)*		
Intervention				
Simplicity	1.74	(0.94)$^{+}$		
SWI	1.80	(0.89)$^{+}$		
Intervention x Morphological Awareness	-0.18	(0.14)		
Morphological Relatedness pre-test			0.43	(0.34)
Variance Components				
Child	5.02	(1.48)		
School			2.60	(2.42)

Note. DV = Dependent Variable; HLM = hierarchical linear modeling; WRAT = Wide Range Achievement Test; SWI = Structured Word Inquiry.

*** $p < .001$, ** $p < .01$, * $p < .05$, + $p < .10$

Table 4

HLM Results for Intervention Condition by Phonological/Morphological Ability on Secondary Post-Test Attainment Measures

	Student-Level Model		School-Level Model	
Variable	Coefficient	SE	Coefficient	SE
DV = Reading & Spelling factor post-test				
Intercept	0.08	(0.20)		
Phonological Awareness	0.02	(0.04)		
Intervention				
Simplicity	0.08	(0.28)		
SWI	0.26	(0.29)		
Intervention x Phonological Awareness	0.01	(0.02)		
Reading & Spelling factor pre-test			0.90	(0.17)***
Variance Components				
Child	0.62	(0.14)		
School			0	(0)
DV = Reading & Spelling factor post-test				
Intercept	0.06	(0.20)		
Morphological Awareness	0.09	(0.06)		
Intervention				
Simplicity	0.14	(0.06)		
SWI	0.37	(0.89)		
Intervention x Morphological Awareness	0.07	(0.05)		
Reading & Spelling factor pre-test			0.95	(0.17)***
Variance Components				
Child	0.62	(0.14)		
School			0	(0)
DV = Reading & Spelling factor delayed post-test				
Intercept	0.09	(0.21)		
Phonological Awareness	0.01	(0.04)		
Intervention				
Simplicity	0.04	(0.29)		
SWI	-0.36	(0.30)		
Intervention x Phonological Awareness	0.02	(0.03)		
Reading & Spelling factor pre-test			0.82	(0.18)***
Variance Components				
Child	0.66	(0.16)		
School			0.01	(0.09)
DV = Reading & Spelling factor delayed post-test				
Intercept	0.06	(0.21)		
Morphological Awareness	0.10	(0.06)		
Intervention				
Simplicity	0.09	(0.06)		
SWI	-0.37	(0.30)		
Intervention x Morphological Awareness	-0.05	(0.05)		
Reading & Spelling factor pre-test			0.86	(0.18)***
Variance Components				
Child	0.67	(0.15)		
School			0	(0)

DV = Decoding factor post-test				
Intercept	-0.18	(0.21)		
Phonological Awareness	0.08	(0.04)		
Intervention				
Simplicity	0.00	(0.29)		
SWI	0.33	(0.30)		
Intervention x Phonological Awareness	-0.01	(0.03)		
Decoding factor pre-test			0.33	(0.27)
Variance Components				
Child	0.64	(0.15)		
School			0.05	(0.09)
DV = Decoding factor post-test				
Intercept	-0.17	(0.21)		
Morphological Awareness	-0.11	(0.06)$^{+}$		
Intervention				
Simplicity	0.04	(0.29)		
SWI	0.32	(0.30)		
Intervention x Morphological Awareness	0.14	(0.30)**		
Decoding factor pre-test			0.67	(0.23)**
Variance Components				
Child	0.66	(0.16)		
School			0.02	(0.10)
DV = Decoding factor delayed post-test				
Intercept	-0.13	(0.21)		
Phonological Awareness	0.03	(0.04)		
Intervention				
Simplicity	-0.17	(0.29)		
SWI	0.48	(0.31)		
Intervention x Phonological Awareness	0.01	(0.03)		
Decoding factor pre-test			0.58	(0.26)*
Variance Components				
Child	0.69	(0.15)		
School			0	(0)
DV = Decoding factor delayed post-test				
Intercept	-0.13	(0.21)		
Morphological Awareness	-0.08	(0.07)		
Intervention				
Simplicity	-0.15	(0.30)		
SWI	0.46	(0.31)		
Intervention x Morphological Awareness	0.09	(0.05)		
Decoding factor pre-test			0.81	(0.23)**
Variance Components				
Child	0.72	(0.16)		
School			0	(0)

Note. DV = Dependent Variable; HLM = hierarchical linear modeling; SWI = Structured Word Inquiry.
*** $p < .001$, ** $p < .01$, * $p < .05$, + $p < .10$.

Appendix A

Word Sums for the <sign> matrix

sign + al → signal

as + sign + ment → assignment

re + de + sign → redesign

sign + ate/ + ure → signature

(slash shows that the "single, silent <e> is replaced by the vowel suffix <-ure> in <signature>

The word matrix of <sign>