Ann. of Dyslexia (2017) 67:219–258
DOI 10.1007/s11881-017-0140-z

The impact of multisensory instruction on learning letter names and sounds, word reading, and spelling

Nora W Schlesinger[1,2] **· Shelley Gray**[1]

Received: 7 August 2016 / Accepted: 30 January 2017 / Published online: 2 March 2017

Abstract The purpose of this study was to investigate whether the use of simultaneous multisensory structured language instruction promoted better letter name and sound production, word reading, and word spelling for second grade children with typical development (N = 6) or with dyslexia (N = 5) than structured language instruction alone. The use of non-English graphemes (letters) to represent two pretend languages was used to control for children's lexical knowledge. A multiple baseline, multiple probe across subjects single-case design, with an embedded alternating treatments design, was used to compare the efficacy of multisensory and structured language interventions. Both interventions provided explicit systematic phonics instruction; however, the multisensory intervention also utilized simultaneous engagement of at least two sensory modalities (visual, auditory, and kinesthetic/tactile). Participant's graphed data was visually analyzed, and individual Tau-U and weighted Tau-U effect sizes were calculated for the outcome variables of letter name production, letter sound production, word reading, and word spelling. The multisensory intervention did not provide an advantage over the structured intervention for participants with typical development or dyslexia. However, both interventions had an overall treatment effect for participants with typical development and dyslexia, although intervention effects varied by outcome variable.

Keywords Dyslexia · Multisensory · Orton-Gillingham · Structured language · Typical development

✉ Nora W Schlesinger
Nschlesi@kennesaw.edu

1 Department of Speech and Hearing Science, Arizona State University, Tempe, AZ, USA

2 Department of Elementary and Early Childhood Education, Kennesaw State University, 585 Cobb Avenue NW, MD 0121, Kennesaw, GA 30144-5591, USA

Introduction

Children with dyslexia have difficulty learning to decode. There is substantial evidence that the primary deficit in dyslexia is phonologically based (Liberman, Shankweiler, Fisher, & Carter, 1974; Wagner & Torgesen, 1987) and that this deficit negatively impacts children's ability to read and spell in English. This phonologically based deficit interferes with children's acquisition of letter knowledge (Ehri, 2014; Snowling & Hulme, 2012) and their ability to decipher the alphabetic code (Shaywitz, Morris, & Shaywitz, 2008), which leads to word decoding problems (e.g., Berninger, 2001; Ehri, 2014; Vellutino et al., 2004). For individuals with dyslexia, reading and spelling impairments are persistent (Shaywitz, 1998) and hinder the development of literacy skills (e.g., Berninger, Lee, Abbott, & Breznitz, 2013; Berninger et al., 2008).

Reading interventions that utilize systematic phonics have been shown to address the core phonological deficits found in dyslexia (e.g., Berninger & Amtmann, 2003; National Reading Panel [NRP], 2006). Further, phonics instruction has been shown to be effective for teaching word decoding and spelling to all young children (e.g., Adams, 1990; NRP, 2000; Snowling & Hulme, 2012). Systematic phonics instruction introduces phonics elements such as letter sound correspondence and spelling patterns in a planned, sequential manner (NRP, 2000).

A popular method of instruction for individuals with dyslexia is multisensory structured language programs (Clark & Uhry, 1995; Moats & Farrell, 2002; Ritchey & Goeke, 2006). These programs utilize direct and explicit instruction and include systematic phonics instruction. In addition to systematic phonics instruction, lesson activities incorporate the simultaneous engagement of at least two sensory modalities (visual, auditory, or kinesthetic/tactile) (Birsh, 2006; McIntyre & Pickering, 2001). These language programs are based on work by Samuel Orton, an early twentieth century neurologist, and his assistant Anna Gillingham (Hallahan & Mercer, 2007; Henry, 2006; McIntyre & Pickering, 2001). The term multisensory used in this study refers to the simultaneous engagement of sensory modalities and to programs that are Orton-Gillingham based or use Orton-Gillingham tenets.

Multisensory approaches to reading instruction have their basis in dual coding theory (Paivio, 1991; Sadoski & Paivio, 2001, 2013) that proposes there are two separate coding systems for the internal forms of mental representations used in memory. These include a verbal system for coding linguistic information and a nonverbal system for coding nonverbal mental images (Sadoski & Paivio, 2001). Based on this theory, teaching that engages a child's sensory modalities (e.g., visual, auditory, and tactile), as well as their linguistic system, may enhance learning. Experiments within the dual coding theory framework of multimodal instruction have been shown to enhance learning (Bell, 1991), and empirical results provide a theoretical explanation as to the possible pedagogical benefits of multisensory reading instruction (e.g., Block, Parris, & Whiteley, 2008; Mayer & Anderson, 1991).

Empirical evidence supports the structured systematic phonics element common to multisensory structured language instruction (Clark & Uhry, 1995, NRP 2000); however, scientific evidence is lacking that indicates the addition of multisensory input makes a significant difference (Clark & Uhry, 1995). Therefore, the body of research supporting multisensory structured language as efficacious for reading intervention is limited (Alexander & Slinger-Constant, 2004; Rose & Zirkel, 2007) and often inconclusive (Rose & Zirkel, 2007). Yet, parents often request it as a preferred form of reading intervention (Rose & Zirkel, 2007) and this has led to an increase in litigation requesting multisensory instruction under the Individuals with Disabilities Education Act (Bhat, Rapport, & Griffin, 2000; Rose & Zirkel, 2007).

Efficacy of multisensory programs for teaching decoding

Multisensory instruction versus non-explicit and nonsystematic phonics instruction Four studies have compared the efficacy of multisensory instruction and non-explicit and systematic phonics instruction in typical classrooms and a clinical setting for elementary-age children with typical development and with dyslexia. Results showed that multisensory instruction provided the best outcome for word decoding (Joshi, Dahlgren, Boulware-Gooden, 2002; Oakland, Black, Stanford, Nussbaum, & Balise, 1998; Uhry & Sheperd, 1993) and spelling (Post & Carreker, 2002; Uhry & Sheperd, 1993). One study used an experimental group design; the other three studies were quasi-experimental group designs. Of the four studies, only one specifically evaluated the impact of multisensory instruction on children with dyslexia.

In the experimental study by Uhry and Sheperd (1993), multisensory segmenting and spelling techniques were used. Eleven first grade children received multisensory spelling instruction and made significantly greater gains in decoding nonsense words, reading sight words, passage reading, and word spelling than their 11 counterparts receiving whole language. In the quasi-experimental studies, Joshi et al. (2002) found first grade children who received multisensory instruction (24) made greater end of the year gains in phonological awareness and decoding compared to the control group (32) who received nonsystematic phonics instruction. Post and Carreker (2002) compared explicit multisensory spelling instruction (70 children) to implicit spelling instruction (70 children) with second through fourth grade students. Students who received explicit multisensory instruction had fewer errors in consonant sound discrimination tasks and spelling generalization tasks. The one study of children with dyslexia took place in a clinical setting using a quasi-experimental design. The results of this study showed improved word reading (Oakland et al., 1998). In this longitudinal study, children received either multisensory instruction in a clinical setting (22) or reading instruction provided at their local school (26). The multisensory group, which initially performed lower, had significantly better word reading and polysyllabic nonsense word decoding after 2 years than students receiving instruction from the local school (Oakland et al., 1998).

Multisensory instruction versus other structured systematic instruction for children with dyslexia Three studies specifically compared the effects of multisensory instruction with other systematic phonics-based reading programs for children with dyslexia. Results suggest multisensory instruction was advantageous for phonological gains and word decoding (Campbell, Helf, & Cooke, 2008; Foorman et al., 1997; Torgesen et al., 2001). Using an experimental design, Torgesen et al. (2001) evaluated intensive remedial instruction for children with dyslexia, including children with comorbid dyslexia and ADHD and found multisensory instruction provided better outcomes for word attack and spelling. The results were not segregated for children with comorbid diagnoses. Participants were between 8 and 10 years of age and participated in either a multisensory program (26 children) or an embedded phonics program (24 children). The multisensory group showed significant gains on the rates of growth from pre to posttest in word attack, reading rate, and accuracy, but group differences were not present at the end of 2 years. In another study, an advantage for multisensory instruction was found in a three-treatment study of children with dyslexia using a quasi-experimental design (Foorman et al., 1997). The effects of the reading treatments were evaluated with second and third grade children: a synthetic multisensory phonics program

(28 children), a sight-word program (39 children), and an analytic phonics program based on rime patterns versus phonic rules (47 children). The multisensory group significantly outperformed the sight-word group, but not the analytic phonics group, in phonological gains.

The last study, a multiple baseline across subjects single-case design, showed that multisensory instruction input improved children with dyslexia's nonsense word and passage reading. Campbell et al. (2008) evaluated the effect of simultaneous multisensory input added to an evidence-based reading program for six second grade students with dyslexia. Students received lessons in *Early Reading Tutor* (Gibbs, Campbell, Helf, & Cooke, 2006) while also receiving instruction in *Open Court Reading* (Adams et al. 2000). During intervention sessions, a multisensory component was added to the Early Reading Tutor instruction. Results indicated vowel–consonant and consonant–vowel–consonant nonsense word decoding and second grade passage reading improved for all students when simultaneous multisensory components were added.

Limitations of current research

Of the studies that have evaluated multisensory instruction, there have been fundamental flaws limiting the generalization of results. This is due, in part, to the lack of well-controlled studies comparing multisensory instruction to an alternative remedial systematic approach (e.g., Uhry & Shepherd, 1993; Joshi et al., 2002; Post & Carreker; 2002). Only one study found, Campbell et al. (2008), specifically evaluated the impact of *simultaneous* multisensory input as a variable. In addition, lack of randomization of participants (e.g., Foorman et al., 1997; Joshi et al., 2002; Post & Carreker; 2002), unequal instructional time between interventions (e.g., Oakland et al., 1998), and level of instructor training or knowledge (e.g., Joshi et al., 2002; Oakland et al., 1998) may have inadvertently biased results in favor of multisensory instruction. Further, intervention fidelity was not always reported (e.g., Joshi et al., 2002).

Furthermore, to improve upon previous studies of multisensory instruction, this study met the What Works Clearinghouse criteria for *Meets Evidence Standards without Reservations* (What Works Clearinghouse [WWC], 2013). To meet this criteria, researchers (a) methodically manipulated the independent variable, (b) measured outcome variables over time by more than one assessor, (c) collected inter-assessor agreement on 20 % of the data points across phases for each condition (exceptions noted below) with inter-assessor agreements averaging at least 80 to 90 %, (d) included a minimum of three baseline conditions, (e) compared two alternating treatments with each other, and (f) collected at least five data points per phase and five alternating repetitions of the interventions. In addition, specific measurements for fidelity of implementation were reported. Due to logistics, inter-assessor agreement was collected for 14 % of baseline, 33 % of treatment, and 10 % of follow-up sessions for participants with typical development for both interventions. For participants with dyslexia, inter-assessor agreement was collected for 10 % of baseline, 43 % of treatment, and 30 % of follow-up sessions.

Purpose, hypotheses, and research questions

The purpose of this study was to determine whether simultaneous multisensory input, in addition to structured language instruction, would promote better letter name, letter sound production, word decoding, and encoding in young children with typical development and dyslexia than structured language instruction alone. To provide control of the independent variable, simultaneous multisensory input, participants were taught two created alphabets. Participants were taught non-English grapheme names and their associated English phonemes (sounds) for two alphabets, one each for structured language and multisensory lessons (explained below). The structured language alphabet was known as *Saraf* and multisensory alphabet was known as *Rasaf*. For the multisensory lessons, but not the structured lessons, participants utilized simultaneous multisensory input during learning and lesson activities. For example, participants utilized mirrors to see how their mouths looked and felt when saying letter names and sounds, manipulated 3D representations of letters, etc.

Letter name, letter sound, word reading, and word spelling were selected as dependent variables because knowledge of letter names is an important fundamental literacy skill that promotes awareness of letter sounds (NELP, 2008; Treiman, et al., 1998). In turn, grapheme–phoneme correspondence allows children to shift from reliance on visual cues to phonetic processing (Ehri & Wilce, 1985). Utilization of grapheme–phoneme correspondences allows for the formation of orthographic mapping for sight-word reading and spelling from memory (Ehri, 2014). Collectively, these dependent variables give a range of early literacy skills to assess the impact of multisensory instruction.

The inclusion of children with typical development was advantageous on two levels. First, it provided an evaluation of whether multisensory instruction was effective for teaching foundational literacy skills for children with typical development. Second, it allowed for an evaluation of whether lesson factors, such as number of letters or words taught per session, were efficacious. Third, it helped establish that the level of learning for participants with dyslexia was not due to the interventions taught but rather reflective of their disability.

Based on a visual inspection of slope, level, immediacy of effects, and the Tau-U nonoverlap index of effect, the specific research questions and hypotheses were as follows:

1. Will the multisensory intervention be more effective than the structured language intervention for learning letter names, letter sounds, word reading, and spelling for participants with typical development? We hypothesized that there would be a multisensory intervention advantage for learning among all participants with typical development.
2. Will the multisensory intervention be more effective than the structured language intervention for learning letter names, letter sounds, word reading, and spelling for participants with dyslexia? We hypothesized that there would be a multisensory intervention advantage for learning among participants with dyslexia.
3. Will the multisensory intervention be more effective than the structured language intervention for maintenance of letter names, letter sounds, word reading, and spelling for participants with typical development? We hypothesized that there would be a multisensory intervention advantage for maintenance among all participants with typical development.
4. Will the multisensory intervention be more effective than the structured language intervention for maintenance of letter names, letter sounds, word reading, and spelling for participants with dyslexia? We hypothesized that there would be a multisensory intervention advantage for maintenance among all participants with dyslexia.

Method

Participant recruitment and selection

Children were recruited from public, private, and charter schools and organizations in the Phoenix metropolitan area. Flyers were distributed in the community through schools, tutoring centers, and through Facebook. Several schools also sent home parent consent packets. The parents of 30 children consented to participation per university institutional review board requirements for human subjects and each child assented to participation. Participants received an incentive of up to 50 dollars for participating. Of the 30 children, 19 were excluded for one or more of the following reasons: Spanish was the predominate language, standardized language scores were too low, unintelligible speech, diagnosis of ADHD, or the parents withdrew prior to data collection because of schedule restrictions or illness. Eleven second graders met inclusionary criteria and participated in the study—six with typical development (four girls) and five with dyslexia (two girls). Children ranged in age from 7 years and 8 months to 8 years and 8 months. One consented participant with typical development began the study but was not able to complete the study. The partial data for this participant is included in the analysis and displayed as indicated in the "Results" section. The participant sample was predominately White (7) but also included one African American and three children identifying as more than one race. Eleven children were non-Hispanic and one was Hispanic.

To qualify for inclusion, all children were required to pass a bilateral hearing screening at 20 dB HL at 500, 1 K, 2 K, and 4 K Hz (ASHA, 1997), a near vision acuity screening (20/32) in both eyes with glasses, and to be monolingual English speakers with no history of a neurologically based disorder other than dyslexia per parent report. Participants were required to demonstrate nonverbal intelligence within the average range as indicated by a standard score of 75 or higher (70 + 1 SEM) on the nonverbal portion of the Kaufman Assessment Battery for Children-Second Edition (KABC-2; Kaufman & Kaufman, 2004) to assess nonverbal intelligence and were required to demonstrate adequate language performance and to ensure participants did not have language impairments as specified by a standard score of 88 or higher on the Core Language of the Clinical Evaluations of Language Fundamentals, Fourth Edition (CELF-4; Semel, Wiig, & Secord, 2003). In addition, participants were required to demonstrate intelligible speech, with a score higher than the 31st percentile on the Goldman-Fristoe Test of Articulation-Second Edition (GFTA-2; Goldman & Fristoe, 2000).

Participants with typical development in this study were required to show adequate word reading with a standard score of 96 or higher on the Test of Word Reading Efficiency (TOWRE-2; Torgesen, Wagner, & Rashotte, 2012). Participants with dyslexia were required to score 88 or lower on the TOWRE-2 (see Table 1 for test scores). For descriptive purposes, each participant also completed the following subtests from the Readiness cluster of the Woodcock Reading Mastery Test-Third Edition (WRMT-III; Woodcock, 2011): Letter Identification, Phonological Awareness, and Rapid Automatic Number and Letter Naming.

Study design

This study used a multiple baseline multiple probe single-case design with alternating treatments structured language and multisensory intervention. Two groups of children completed treatment—one with typical development and one with dyslexia. The dyslexia group began 1 week before the typical development group. The independent variable was the type of

Table 1 Participant description information including age, mother's years of education, mean standard scores, and standard deviations on norm-referenced tests

Measure	Typical development Mean (SD) (*n* = 6)	Dyslexia Mean (SD) (*n* = 5)
Age in months	99 (3.9)	94 (6.6)
Mother's Ed	15 (3.5)	15 (1.1)
Attention	20* (3.8)	30* (6.7)
CELF-4	114.5** (8.6)	96.6** (9.6)
KABC-2[a]	112.8 (19)	112 (11.50)
GFTA-2[b]	51 (6.2)	40 (20)
WRMT-III LID[c]	100 (0)	100 (0)
WRMT-III PA	111* (12.7)	91* (16.30)
WRMT-III RAN	100* (6.7)	92* (3.10)
TOWRE-2	108** (6.0)	78.4** (3.6)

Attention Attention Questionnaire, *CELF-4* Clinical Evaluations of Language Fundamentals, Fourth Edition (Semel, Wiig, & Secord, 2003), *KABC-2* Nonverbal Scale of the Kaufman Assessment Battery for Children-Second Edition (Kaufman & Kaufman, 2004), *GFTA-2* Goldman-Fristoe Test of Articulation-Second Edition (Goldman & Fristoe, 2000), *WRMT-III* Woodcock Reading Mastery Test-Third Edition (Woodcock, 2011), *WRMT-III LID* Letter Identification Subtest, *WRMT-III PA* Phonological Awareness Subtest, *WRMT-III RAN* Rapid Automatic Naming Subtest, *TOWRE-2* Total Word Reading Efficiency of the Test of Word Reading Efficiency-Second Edition (Torgesen et al. 2012)

*$p \leq 0.05$, **$p \leq 0.01$

[a] KABC-2 scores for one participant with typical development were unavailable, per parents, and from observations, cognition was not a concern

[b] Less than 31 % allowed if treatment phonemes were articulated consistently

[c] Standard scores for subtest were unavailable for age of participants; scores reflect percent correct

treatment. This study consisted of three phases, baseline, intervention, and follow-up, which are explained below. Performance on the multisensory letter name production was used as the mastery criterion variable to determine when participants moved from the baseline to intervention phase of the study; however, letter sound production, word reading, and spelling variables were also assessed.

Baseline phase All participants within a group entered baseline simultaneously, with three baseline data points taken over the same week for each participant. One randomly selected participant within each group then completed two additional baseline data points. After the fifth baseline probe, the randomly selected participant proceeded to the treatment phase if a stable baseline pattern was evident for letter name production in the multisensory treatment condition. Baseline stability was determined by data analysts who evaluated data separately and were not privy to the study's purpose. A stable baseline pattern was required to demonstrate (a) a consistent level, (b) little variability (e.g., consistent data range), and (c) lack of a positive trend using a minimum of three consecutive data points (Barlow & Hersen, 1984; WWC, 2013). All participants, other than the first participant from each group, received six baseline probes to ensure stable baselines. Once the first randomly selected participant in a group entered the treatment phase, the next randomly selected participant completed three more baseline probes prior to moving into the treatment phase. Per WWC pilot standards for multiple baseline multiple probes, at least one baseline data point from the second randomly selected participant was taken within the same session time frame, 1 week, in which the preceding participant first received intervention (WWC, 2013).

Treatment phase Participants could not enter the treatment phase until the participant ahead of them in the treatment phase demonstrated stable letter name production in the multisensory treatment based on the following criteria: (a) data mean was above that of baseline mean using a minimum of three data points and (b) there were at least three consecutive data points trending in the same direction (Barlow & Hersen, 1984). This introduction format continued for all remaining participants from both groups. Once they entered the treatment phase, all participants completed six treatment sessions during which the structured and multisensory interventions were presented in each session in randomized order to control for sequencing effects (Barlow & Hayes, 1979). To progress forward and learn additional letters in the structured language or the multisensory intervention, participants had to meet the mastery criteria. The mastery criteria required participants to correctly name the newly taught letters in an intervention session two times in a row during assessments (explained below). Therefore, a participant could meet the mastery criteria for one intervention and move on but not meet mastery and repeat lessons in the other.

Follow-up phase There were two assessment sessions in the follow-up phase. The first was conducted 1 week after intervention ended and the second 2 weeks after intervention ended.

Intervention overview

During the study, children attended baseline and follow-up sessions that were 30 min and intervention sessions that were approximately 1 h in length, one to three times per week. Sessions were completed over a 6- to 7-week period. Participants were taught two created alphabets using non-English grapheme names and their associated English phonemes (sounds). Two intervention treatments were delivered in random order within each treatment session—structured language and multisensory. The interventions were adapted from Orton-Gillingham-based programs and followed a systematic sequential structured language approach. Structured language activities did not include simultaneous sensory engagement, but the multisensory intervention used simultaneous engagement of at least two of the three sensory modalities (visual, auditory, and kinesthetic/tactile) during each lesson.

The interventionists in this study implemented both the structured language and multisensory treatments. Other than the first author, all interventionists were speech-language pathology assistants. Interventionists received training for teaching procedures for each treatment.

Intervention materials

Graphemes and grapheme names A total of 18 non-English graphemes were used. The graphemes were symbols drawn from ancient alphabets, letter forms developed by Gibson (Gibson, Pick, & Osser, 1962), and Aurebesh letters from *Star Wars* © & ™ Lucasfilm Ltd. The names for graphemes were randomly assigned from the set of phonemes (see below) that was in turned randomly assigned to either the structured or

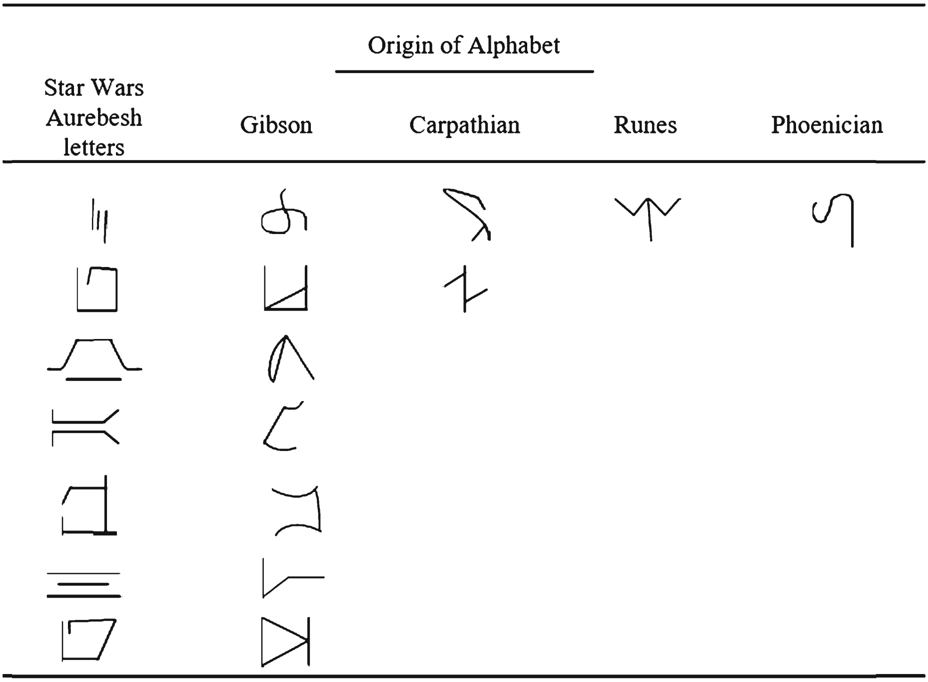

Fig. 1 Courtesy of Lucasfilm Ltd. permission granted for the use of the Aurebesh Letters, *Star Wars* © and ™ Lucasfilm Ltd. Aurebesh letters from, "Star Wars Miniature Battles Imperial Entanglements," by S. Crane, 1996. Gibson letter forms from "A Developmental Study of the Discrimination of Letter-Like Forms," by Gibson, Gibson, Pick, and Osser, 1962, *Journal of Comparative and Physiological Psychology, 55*, pp. 897–906. Carpathian letters adapted from "Heritage of Scribes: The Relation of Rovas Scripts to Eurasian Writing Systems," by Hosszu, 2012. Rune letter from "The Old English Rune for S," by Nicholson, 1982, *The Journal of English and Germanic Philology, 81*, pp. 313–319. Phoenician letter from "The Languages of the World Ancient and Modern," by Wemyss, 1950

multisensory intervention. Graphemes (Fig. 1) and their names were unique for each participant.

Phonemes Eighteen English phonemes (Fig. 2) were used and divided into two sets of seven consonants and two lax vowels per set. Research shows a grapheme's sound is easier to learn if the grapheme name contains the phoneme (Treiman, et al., 1998). Therefore, one third of grapheme names contained the phoneme at the beginning (e.g., /n/ beginning the grapheme name /nɛ/), one third at the end (e.g., /p/ ending the grapheme name / p/), and one third did not contain the phoneme (e.g., /k/ for a grapheme named / z/). For each child, the sets were randomly assigned to one intervention or the other and within each set the grapheme–phoneme pairings were randomized. The only exceptions were graphemes representing /b/and /d/ phonemes, which were assigned mirror image forms so that they were visually similar as they are in English.

Words Children were asked to decode and spell 12 words in each intervention. They included vowel–consonant (vc), consonant–vowel (cv), consonant–vowel–consonant (cvc), and vowel–consonant–vowel (vcv) constructions.

Fig. 2 Classification of speech sounds from Zemlin, W. (1998). *Speech and hearing science anatomy and physiology, fourth edition.* Boston: Allyn and Bacon

Place of articulation Manner of articulation Voiced or unvoiced Phoneme	Place of articulation Manner of articulation Voiced or unvoiced Phoneme
Bilabial Stop Unvoiced /p/	Bilabial Stop Voiced /b/
Labiodental Fricative Voiced /v/	Labiodental Fricative Unvoiced /f/
Alveolar Stop Voiced /d/	Alveolar Stop Unvoiced /t/
Velar Stop Voiced /g/	Velar Stop Unvoiced /k/
Alveolar Fricative Unvoiced /s/	Alveolar Fricative Voiced /z/
Bilabial Nasal (stop) /m/	Alveolar Nasal (stop) /n/
Alveolar Liquids /l/	Palatal Liquids /r/
Front high vowel sound /I/	Front medium vowel sound / ɛ /
Front low vowel sound /æ/	Back low vowel sound / ɑ/

Teaching cards Four sets of color-coded cards (Fig. 3) were used for the interventions. White grapheme cards and word cards were used for teaching and assessing letter names and words. Green phoneme cards and spelling cards were used for teaching and assessing phonemes and word spelling.

Spelling matrix A 6-in. bingo-like spelling matrix with three columns and three rows was used for spelling activities and assessments. Graphemes were placed on the matrix and participants placed selected graphemes onto a line below the matrix during spelling practice and assessment activities. For the structured language intervention, each grapheme was written on a white, 2-in. square piece of cardstock. For the multisensory intervention, 3D plastic graphemes (Fig. 4) were placed in each of the squares. The plastic 3D shapes were approximately 1 to 1 ¼ in. high by 1-in. wide and ¼ in. thick. A supplemental appendix of graphemes, phonemes, teaching cards, and the spelling matrix is available online.

Teaching procedures

Interventions took place at the university lab, a local library or center, or the participant's home. All sessions were audio recorded and implemented by trained interventionists. Children received instruction from at least two different interventionists who provided no fewer than two intervention sessions. Interventionists were naive to the research hypotheses (Barlow & Hersen, 1984) and completed an implementation checklist each session.

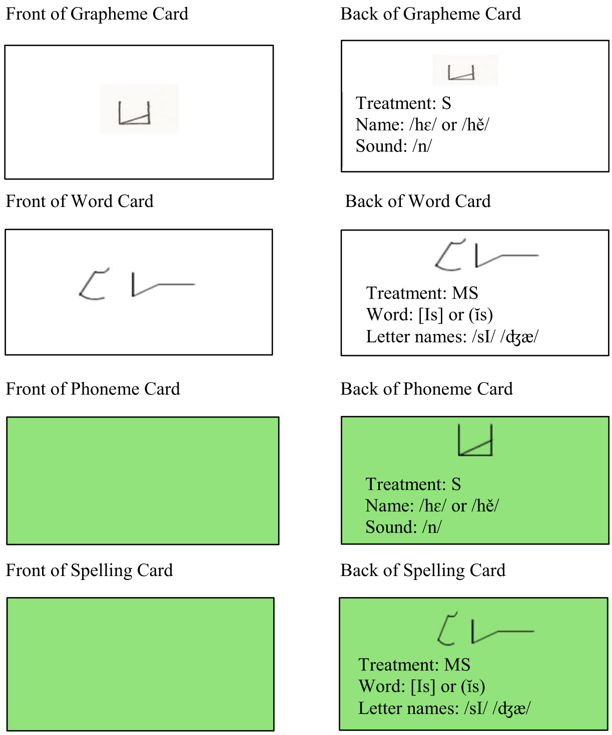

Fig. 3 *S* = structured language intervention; MS = multisensory intervention. Cards were 4 × 6 in. for the structure language intervention and 3 × 5 in. for the multisensory intervention to make them easy to differentiate for interventionists. Graphemes were printed in *black ink* (approximately 1 × 1 in.). Words were coded phonetically and with common phonics symbols for interventionist use. Cards were based on the *Initial Reading Deck* and *Instant Spelling Deck* from Alphabetic Phonics (Cox, 1992)

Within each research session, children completed two teaching sessions with assessments given after each teaching session. During teaching, interventionists highlighted the scripted text as it was read. The order of intervention types was randomized with the caveat that no

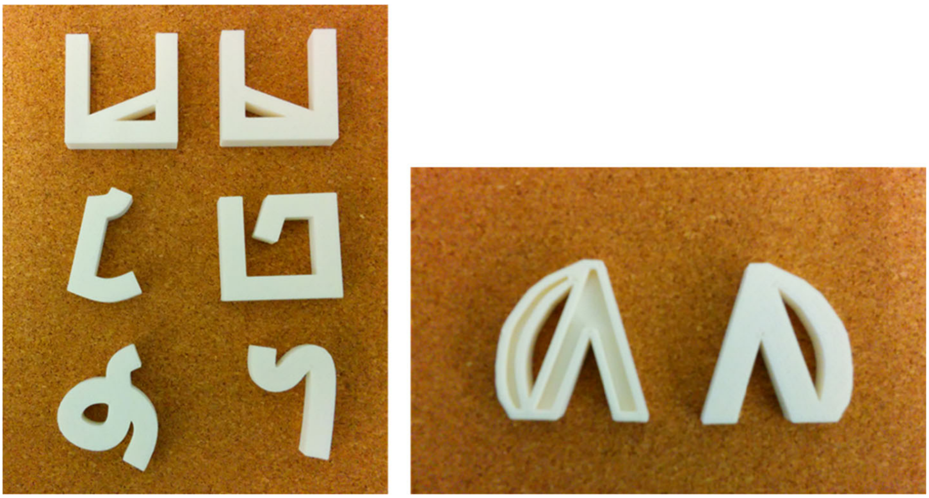

Fig. 4 Example of 3D plastic graphemes used in the multisensory intervention. The *left* picture depicts the front of the graphemes; the top two graphemes are mirror images. The picture on the *right* depicts the back and front of the graphemes. *3D* = three-dimensional

intervention may be presented in the same order more than two consecutive times. During the first intervention, session children were informed they would be learning two pretend languages (Hulme, Monk, & Ives, 1987). The structured language was called *Saraf* and multisensory was called *Rasaf*. Each teaching session followed the same lesson schedule; however, activities varied between interventions.

The first lesson began with *new learning*. In this activity, two graphemes were taught. Using a grapheme card, the first grapheme was presented and the letter name and sound were taught, followed by the second new grapheme. For the multisensory lessons, but not the structured lessons, participants utilized mirrors to see how their mouths looked and felt when saying letter names and sounds. Participants were also taught how to write the letter. In structured language lessons, participants traced over the letter twice, after tracing it the first time the participant named the letter and after the second trace participants gave the letter's sound. In contrast, during multisensory lessons, participants were guided to skywrite the letter using gross motor movements while simultaneously looking at the letter and saying the letter's name. Then, participants traced over the letter while simultaneously saying the letter's sound. In the second activity, *word reading*, participants practiced reading words. In the structured lessons, participants sounded out each phoneme then read the word. During multisensory lessons, participants looked at the word and simultaneously tapped and sounded out each phoneme by sequentially tapping their index, middle, and ring fingers to their thumb. Participants then looked at the word and read it by scooping their finger under it while simultaneously blending the phonemes. *Spelling practice* was then introduced which consisted of two activities. In the first activity, *sound dictation*, participants repeated a given letter sound then selected the correct grapheme from the spelling matrix. For structured language, participants selected the grapheme tile, then named it. In contrast, for multisensory, 3D plastic letters were selected then traced over by participants while simultaneously saying the letter's name. For the second activity, *word spelling*, participants were required to repeat the word given then selected the word's graphemes from the matrix and placed them on the line from left to right. During structured language lessons, participants sounded out each phoneme, next they said the letter names, then selected the letter tiles, after which they named them, and lastly read the word. During multisensory, participants simultaneously tapped and sounded out each phoneme, next simultaneously tapped and named each letter, then selected 3D letters while simultaneously naming them, then scooped under the word and read it.

The same lesson schedule was followed for all subsequent lessons, except three review activities were presented before new learning. In order of presentation, the activities were *alphabet review*, *grapheme practice*, and *sound dictation*. In alphabet review, all previously learned letter names or sounds were reviewed through various activities. Structured lesson activities used grapheme cards to elicit letter names or sounds. Activities included turning cards print side up, selecting cards from the interventionist's hand, or touching a card print side up. Multisensory lessons incorporated 3D letters to prompt letter names or sounds and included tossing then turning 3D letters face side up, holding and feeling each letter with eyes closed, or tracing over each letter while simultaneously responding. The second review activity, grapheme practice, was the same for both interventions. In this activity, grapheme cards were presented one at a time and the participant recited the letter name and sound. In the last review activity, sound dictation, previously learned letter sounds were presented using phoneme cards. In the structured lesson, participants wrote the letter for each sound on the table top with their index finger, then named the letter. Multisensory lessons required the participant to look in their mirror while repeating the letter sound, then wrote the letter with their index finger while simultaneously naming it. Different materials were used for writing the letter on and included a small carpet square, tray of sand, or a wipe-off board. After these three review activities, new learning was introduced. Following this, assessments were administered (see below).

Treatment integrity To ensure treatment integrity, direct observation of 20 % of intervention lessons was performed by another interventionist or trained observer who also completed an implementation checklist (Fiske, 2008; Schoenwald, et al., 2011). The two intervention checklists were evaluated for adherence using point-by-point agreement. The average percent of agreement (Kershener, et al., 2014) indicated structured language interventions were implemented for participants with typical development with 99 % (range 94–100 %) fidelity and the multisensory with 98 % (range 96–100 %) fidelity. The structured language fidelity for participants with dyslexia was 96 % (range 83–99), and the multisensory fidelity was 96 % (87–99).

Assessments Assessments were administered during each phase of the study. Within all phases, the order of assessment presentation and the items within each assessment measure were randomly determined. During the treatment phase interventions, assessments were conducted immediately after the teaching session for each intervention. At follow-up, assessment measures were given at minimum 1 and 2 weeks after the participant's last teaching session.

For each assessment, participants were asked to produce letter names and letter sounds, to read words, and to spell words. A score of zero was given for incorrect responses and one point for correct responses. To assess letter name and letter sounds, participants were shown nine grapheme cards, one at a time, and asked to give the letter name or sound. Nine points were possible for letter names and nine for letter sounds. Responses were written phonetically by the interventionist. For word reading, 12 word cards were presented, one at a time. Children's responses were recorded phonetically by the interventionist, with 12 points possible. To assess word spelling, participants used the spelling matrix to select letter tiles or 3D letters for structured language and multisensory, respectively. Interventionist marked participant's grapheme selections in numerical order on spelling boxes containing all available graphemes to the right of each spelling word. Word spelling had 12 points possible.

Reliability Twenty percent of sessions in the baseline, treatment, and follow-up phases were attended by a trained observer. Both the interventionist and trained observer recorded and scored participants' responses. The recordings by the interventionist and observer were evaluated for point-by-point agreement. Point-by-point inter-rater agreement was calculated between interventionists' and observers' records. Inter-observer agreement was calculated by dividing the number of agreements by the number of agreements plus disagreements and multiplying by 100 (Caro, Roper, Young, & Dank, 1979). The average agreement for participants with typical development in the structured language intervention was 99 % (range 88–100%) and for multisensory was 98 % (range 83–100 %). The average agreement for participants with dyslexia in the structured language intervention was 99 % (range 87–100 %) and for multisensory was 98 % (range 78–100 %).

Analytic approach

Each participant's data was evaluated using visual analysis to compare the effects of structured language and multisensory interventions on letter name, letter sound, word reading, and spelling. *Tau-U* and *weighted Tau-U* effect sizes were also calculated for each dependent variable to permit comparison of the effects for both interventions.

Visual inspection Within- and between-phase data patterns were examined for each dependent variable in each condition to address the research questions. The visual analyses included the projected data, which refers to the hypothetical continuation of a data pattern from the previous phase, and the observed data within each phase evaluated based on data features for (a) level (mean), (b) trend (slope of the best-fitting line), and (c) variability (range of the data about the best-fitting line). In addition, data patterns across phases were examined for (a) immediacy of effect (visible distinction between the data features of the last three baseline data points and the first three treatment data points), (b) the proportion of data points overlapping between baseline and treatment with low overlap suggesting larger treatment effects (Horner, Swaminathan, & Smolkowski, 2012), and (c) inter-case (across participants) replication of data patterns. In this study, three inter-case replications indicated an experimental effect (Horner, et al., 2005; WWC, 2013).

To address research questions 1 and 2 regarding whether the multisensory intervention would be more effective than the structured language intervention for learning letter names, letter sounds, word reading, and spelling for participants with typical development and dyslexia, the slope, level, immediacy of effects, and Tau-U effect sizes of each participant's data from baseline to intervention were subjected to visual inspection. For research questions 3 and 4, regarding whether the multisensory intervention would be more effective than the structured language intervention for maintenance of letter names, letter sounds, word reading, and spelling for participants with typical development and dyslexia, the data in each participant's follow-up phase was visually inspected.

Tau-U effect size

Tau-U effect size, a nonoverlap index of effect, was used to determine whether individual participants had a statistically significant treatment effect across the dependent variables for each intervention. Tau-U is distribution free, utilizes trend and level, and can control for baseline trends (Brossart, Vannest, Davis, & Patience, 2014). Rather than measurements of central tendency, Tau-U takes into account the individual values of all data points in pairwise comparisons across phases. Conceptually, it is the percentage of data showing improvement from the baseline and treatment phases for each participant (Parker et al., 2011a). The *individual* Tau-U contrast was calculated for each dependent variable for each participant.

A *weighted* Tau-U effect also was calculated for both interventions for each of the dependent variables. Weighted Tau-U is all of the individual participants' phase contrasts between baseline and treatment, for a specific dependent variable and intervention, combined to reflect the overall effect for each intervention resulting in a combined weighted average for each dependent variable (Vannest, Parker, & Gonen, 2011). The tentative benchmarks for the individual and weighted Tau-U range are in percentages from 0 to 100, with a weak to small effect size indicated by results of 65 % or less, moderate to high effect sizes range from 66 to

92 %, and large effect sizes of 93 % or greater. Both the individual Tau-U and the weighted Tau-U effect sizes were calculated using the Single Case Research™ web-based calculator (Parker & Vannest, 2009). The statistical power for Tau-U ranges from 91 to 95 % (Parker, Vannest, Davis, & Sauber, 2011b; Vannest et al., 2011).

Results

Results were analyzed using visual analysis, individual Tau-U, and weighted Tau-U effect sizes. A summary of observed means, range, overlap, and the individual Tau-U effect sizes for the dependent variables for both interventions for participants with typical development is presented in Tables 2 and 3, and a summary for participants with dyslexia is presented in Tables 4 and 5. Baseline data were stable for all participants for both interventions, and observed treatment levels were above projected levels based on baseline scores. In addition, visual analyses of data features indicated both interventions had a positive effect as indicated by slope, level, linear trends, and immediacy of effect. Exceptions are presented below.

Letter name production

Participants with typical development Graphs of visual analyses are presented in Fig. 5. The positive response data patterns and the inter-case treatment replications indicated a positive experimental effect for both interventions, which is supported by the individual Tau-U effect sizes. Thus overall, both interventions had a positive treatment effect. In addition,

Table 2 Letter name and sound production outcome variables for children with typical reading across interventions and phases

	Structured language intervention					Multisensory intervention				
DV	BL *M* (range)	TX *M* (range)	FU *M* (range)	Overlap	Tau-U ES	BL *M* (range)	TX *M* (range)	FU *M* (range)	Overlap	Tau-U ES
Letter name production										
1[a]	0 (0)	1.3 (0–3)	1.5(1–2)	17	83*	0 (0)	2.5 (2–4)	1 (0–2)	0	100**
2	0 (0)	1.3 (0–3)	–	50	50	0.3 (0–1)	2 (0–4)	_	25	42
3	0 (0)	3.7 (0–7)	6 (5–7)	17	83*	0 (0)	3.7 (0–7)	6 (5–7)	17	83**
4	0 (0)	2.5 (0–5)	3.5 (3–4)	33	67	0 (0)	2 (0–4)	2.5 (2–3)	33	67
5	0 (0)	5 (2–9)	8 (NR)	0	100**	0 (0)	5.7 (2–9)	8 (NR)	0	100**
6	0 (0)	4.7 (1–8)	8 (7–9)	0	100**	0 (0)	5.2 (2–8)	7 (NR)	0	100**
Letter sound production										
1[a]	0.6 (0–2)	2.5 (0–5)	1.5 (0–3)	50	53	0.8 (0–1)	2.2 (0–5)	2 (0–4)	50	13
2	0 (0)	2.3 (0–4)	_	17	75	0.2 (0–1)	3.3 (2–4)	–	0	100**
3	0 (0)	6 (2–9)	9 (NR)	0	100**	1 (NR)	5.7 (3–8)	8 (7–9)	0	100**
4	0.7 (0–1)	4.2 (0–8)	7 (6–8)	17	56	0.2 (0–1)	1.5 (0–3)	1.5 (0–3)	67	64
5	0.8 (0–2)	6.2 (3–9)	9 (NR)	0	92*	0.2 (0–1)	6 (2–9)	8.5 (8–9)	0	86
6	0 (0)	5.5 (2–9)	8 (NR)	0	100*	1 (0–1)	6.2 (3–9)	9 (NR)	0	100**

Table outcomes reflect measures specifically created for this study. Overlap is expressed in percent data overlap between baseline and treatment

DV dependent variables, *BL* baseline mean, *TX* treatment mean, *FU* follow-up mean, *Tau-U ES* individual Tau-U effect size expressed in percentage

$*p \leq 0.05$, $**p \leq 0.01$

[a] Participant 1 inadvertently received six less review activities for the multisensory intervention

Table 3 Word reading and spelling outcome variables for children with typical reading across interventions and phases

DV	Structured language intervention BL *M* (range)	TX *M* (range)	FU *M* (range)	Overlap	Tau-U ES	Multisensory intervention BL *M* (range)	TX *M* (range)	FU *M* (Range)	Overlap	Tau-U ES
Word reading										
1[a]	0 (0)	2.3 (0–7)	1 (NR)	17	83*	0.4 (0–1)	3.2 (0–6)	2.5 (2–3)	50	57
2	0 (0)	3.8 (2–5)	_	0	100**	0 (0)	3.5 (2–6)	_	0	25
3	0 (0)	8 (1–12)	12 (NR)	0	100**	0 (0)	6.7 (2–12)	10 (NR)	0	100**
4	0 (0)	6.3 (3–10)	11 (10–12)	0	100**	0 (0)	3.7 (1–9)	7.5 (6–9)	0	100**
5	0.2 (0–1)	9 (4–12)	12 (NR)	0	86**	0 (0)	8 (2–12)	11.5 (11–12)	0	100**
6	0 (0)	6.5 (3–12)	11 (11–12)	0	100**	0.3 (0–1)	6.7 (3–10)	12 (NR)	0	78*
Word spelling										
1[a]	0 (0)	2 (0–9)	2 (NR)	50	50	0 (0)	2.5 (0–6)	3.5 (2–5)	50	50
2	0 (0)	4 (2–6)	_	0	100**	0 (0)	6.5 (4–9)	–	0	100**
3	0.3 (0–1)	8.5 (2–12)	11.5 (11–12)	0	78*	0 (0)	7.3 (2–12)	11.5 (11–12)	0	100**
4	0 (0)	6.2 (1–10)	10 (8–12)	0	100**	0.5 (0–1)	5.2 (2–10)	5.5 (5–6)	0	81*
5	0 (0)	8.2 (5–12)	11.5 (11–12)	0	100**	0 (0)	8.2 (5–12)	11.5 (11–12)	0	100**
6	0 (0)	6.2 (0–12)	12 (NR)	17	83*	0 (0)	6.2 (0–12)	12 (NR)	17	83*

Table outcomes reflect measures specifically created for this study. Overlap is expressed in percent data overlap between baseline and treatment

DV dependent variables, *BL* baseline mean, *TX* treatment mean, *FU* follow-up mean, *Tau-U ES* individual Tau-U effect size expressed in percentage

[a] Participant 1 inadvertently received six less review activities for the multisensory intervention

*$p \leq 0.05$, **$p \leq 0.01$

Table 4 Letter name and sound production outcome variables for children with dyslexia across interventions and phases

	Structured language intervention					Multisensory intervention				
DV	BL *M* (range)	TX *M* (range)	FU *M* (range)	Overlap	Tau-U ES	BL *M* (range)	TX *M* (range)	FU *M* (range)	Overlap	Tau-U ES
Letter name production										
7	0 (0)	1.8 (0–4)	3.5 (3–4)	33	67	0 (0)	1.8 (0–4)	1 (0–2)	33	67
8	0 (0)	2.5 (0–5)	4 (NR)	17	83*	0 (0)	2.2 (0–4)	1 (NR)	17	83*
9	0 (0)	0 (NR)	0 (NR)	100	0	0 (0)	0 (NR)	0 (NR)	100	0
10	0 (0)	1 (0–2)	0 (NR)	17	83*	0 (0)	0.3 (0–2)	0 (NR)	83	17
11	0 (0)	0.5 (0–2)	1 (0–2)	67	33	0 (0)	0.5 (0–1)	1 (0–2)	50	50
Letter sound production										
7	0.6 (0–1)	3 (0–5)	3 (2–4)	17	67	0.2 (0–1)	3 (0–6)	2.5 (0–5)	17	67
8	0 (0)	3.7 (2–5)	3.5 (3–4)	0	100**	0.2 (0–1)	2.8 (0–6)	2.5 (2–3)	17	78*
9	0.2 (0–1)	0.3 (0–1)	0.5 (0–1)	100	8	0.8 (0–2)	1.7 (0–4)	2 (1–3)	17	40
10	0 (0)	0.7 (0–2)	0 (NR)	50	50	0 (0)	0.7 (0–1)	0 (NR)	33	67
11	0 (0)	1.3 (0–3)	2 (0–4)	50	50	0.2 (0–1)	1.2 (0–3)	0 (NR)	50	42

Table outcomes reflect measures specifically created for this study. Overlap is expressed in percent data overlap between baseline and treatment

DV dependent variables, *BL* baseline mean, *TX* treatment mean, *FU* follow-up mean, *Tau-U ES* Tau-U effect size expressed in percentage

$^{*}p \leq 0.05$, $^{**}p \leq 0.01$

the weighted Tau-U for letter names showed statistically significant, moderate effect sizes for both interventions (structured language *Tau-U* = 82, multisensory *Tau-U* = 83). However, multisensory did not appear to have an overall greater effect than structured language.

Table 5 Word reading and spelling outcome variables for children with dyslexia across interventions and phases

	Structured language intervention					Multisensory intervention				
DV	BL *M* (range)	TX *M* (range)	FU *M* (range)	Overlap	Tau-U ES	BL *M* (range)	TX *M* (range)	FU *M* (range)	Overlap	Tau-U ES
Word reading										
7	0 (0)	3.2 (1–5)	4 (NR)	0	100**	0 (0)	3.5 (1–7)	8 (NR)	0	100**
8	0 (0)	2 (0–5)	3.5 (3–4)	50	50	0 (0)	2.3 (0–4)	3 (2–4)	17	83*
9	0 (0)	0 (NR)	0 (NR)	100	0	0 (0)	0 (NR)	0 (NR)	100	0
10	0 (0)	0.3 (0–1)	0.5 (0–1)	67	33	0 (0)	0 (NR)	0 (NR)	100	0
11	0 (0)	0.8 (0–3)	1 (0–2)	50	50	0 (0)	0.8 (0–2)	1.5 (1–2)	33	50
Word spelling										
7	0 (0)	2.8 (1–4)	4 (NR)	0	100**	0 (0)	4.8 (1–9)	8 (NR)	0	100**
8	0.2 (0–1)	3.5 (0–7)	6 (4–8)	33	78*	0 (0)	1.5 (0–4)	1 (0–2)	33	67
9	0 (0)	0.3 (0–1)	0 (NR)	67	33	0 (0)	0.8 (0–1)	0 (NR)	17	83*
10	0 (0)	0.7 (0–2)	0 (NR)	50	50	0.2 (0–1)	0 (NR)	0 (NR)	100	−31
11	0 (0)	2.2 (0–4)	3.5 (3–4)	17	83*	0 (0)	1.5 (0–3)	2 (NR)	17	83*

Table outcomes reflect measures specifically created for this study. Overlap is expressed in percent data overlap between baseline and treatment

DV dependent variables, *BL* baseline mean, *TX* treatment mean, *FU* follow-up mean, *Tau-U ES* Tau-U effect size expressed in percentage

$^{*}p \leq 0.05$, $^{**}p \leq 0.01$

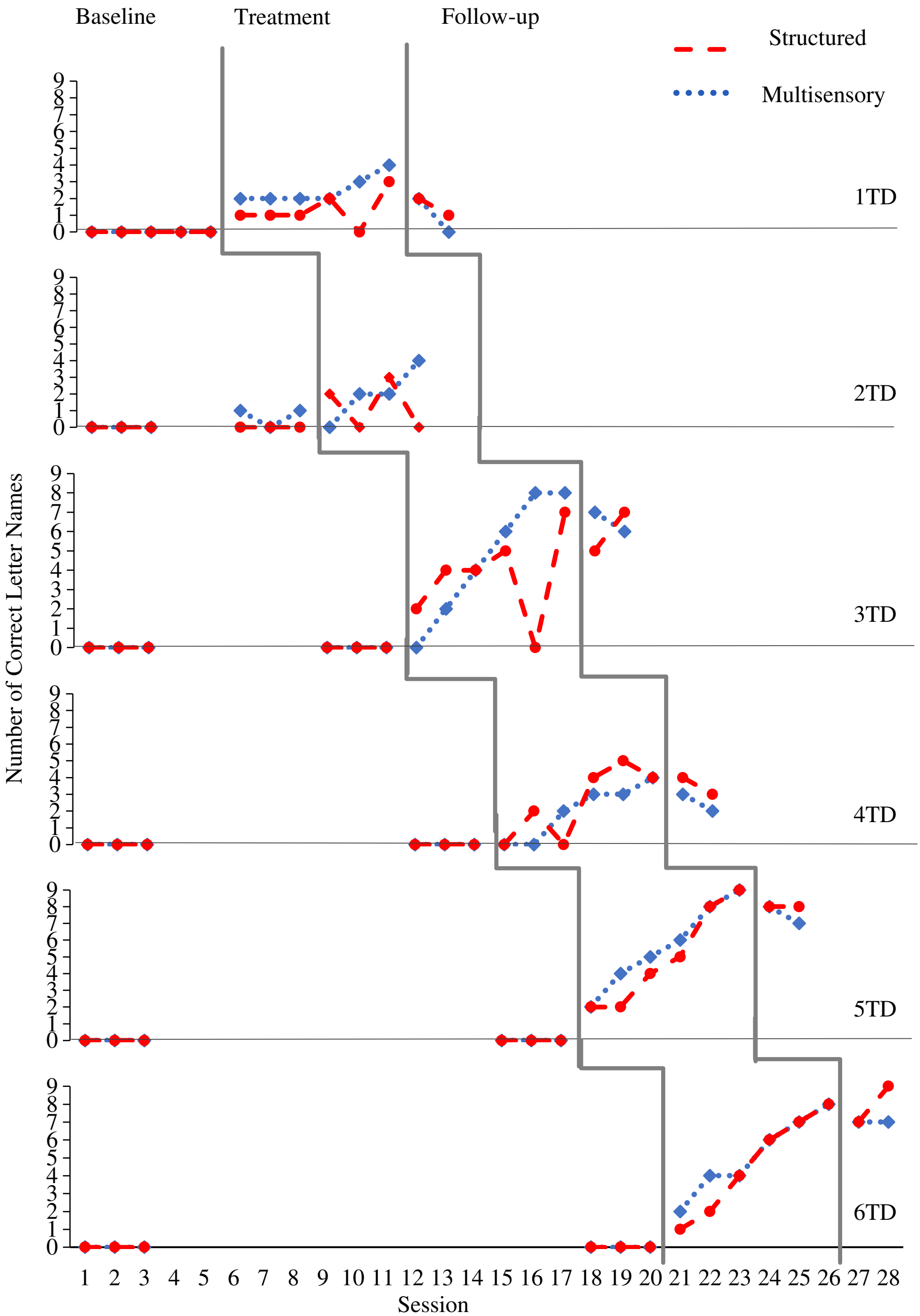

Fig. 5 Accuracy of the number of letter names produced for each participant with typical development across three phases. *TD* = typical development

Follow-up data showed that all participants demonstrated maintenance for letter names in each treatment condition, with somewhat higher maintenance in structured language intervention for all participants except three with typical development. This participant showed a higher maintenance in multisensory intervention.

Participants with dyslexia Graphs of visual analyses are presented in Fig. 6. Participant 9 remained at baseline levels across all three phases of the study. Visual analyses of data features indicated structured language intervention had a small overall effect, but multisensory did not have an effect as demonstrated by slope, level, linear trends, or immediacy of effect. For all participants, neither intervention showed a clear advantage, which is supported by individual Tau-U effect sizes. However, for participant 10, the structured language intervention appeared more effective. The three inter-case replications for participants 7, 8, and 10 indicated structured language intervention showed an experimental effect, but the lack of three inter-case replications indicated there was not an experimental treatment effect for multisensory intervention. The inter-case replications for structured language suggested a structured language advantage over multisensory intervention. However, the less conservative weighted Tau-U indicated statistically significant, small effect sizes for both interventions (structured language $Tau\text{-}U$ = 53, multisensory $Tau\text{-}U$ = 43).

Follow-up data showed participants 7, 8, and 11 demonstrated maintenance for letter names in each intervention, with higher maintenance for structured language compared to multisensory for participants 7 and 8. Participant 11 had higher maintenance in multisensory intervention. Data for participant 10 did not demonstrate maintenance in either intervention.

Letter sound production

Participants with typical development Graphs of visual analyses are presented in Fig. 7. Neither treatment appeared to be more effective except for participant 2 who showed a multisensory advantage and participant 4 who showed a structured advantage. The positive response data patterns and inter-case treatment replications indicated a positive experimental effect for both interventions, which is supported by individual Tau-U effect sizes. This positive effects of both interventions are supported by the weighted Tau-U, which indicated statistically significant, moderate effect sizes for both interventions (structured language $Tau\text{-}U$ = 80, multisensory $Tau\text{-}U$ = 77).

Follow-up data showed all participants demonstrated maintenance for letter sounds in each treatment condition. Participants 3, 4, and 5 had higher maintenance in the structured language intervention. Participants 1 and 6 had higher maintenance in the multisensory intervention.

Participants with dyslexia Graphs of visual analyses are presented in Fig. 8. Visual analyses of data features indicated the multisensory intervention had a small overall effect as demonstrated by slope, level, linear trends, and immediacy of effect, but the structured language did not have an effect. There were only two structured inter-case replications, participants 7 and 8, which indicated structured language did not have an experimental effect. For multisensory, participants 7, 8, and 9 demonstrated inter-case replications. Only one participant, 8, had individual statistically significant Tau-U effect sizes; for this participant, both structured language and multisensory were

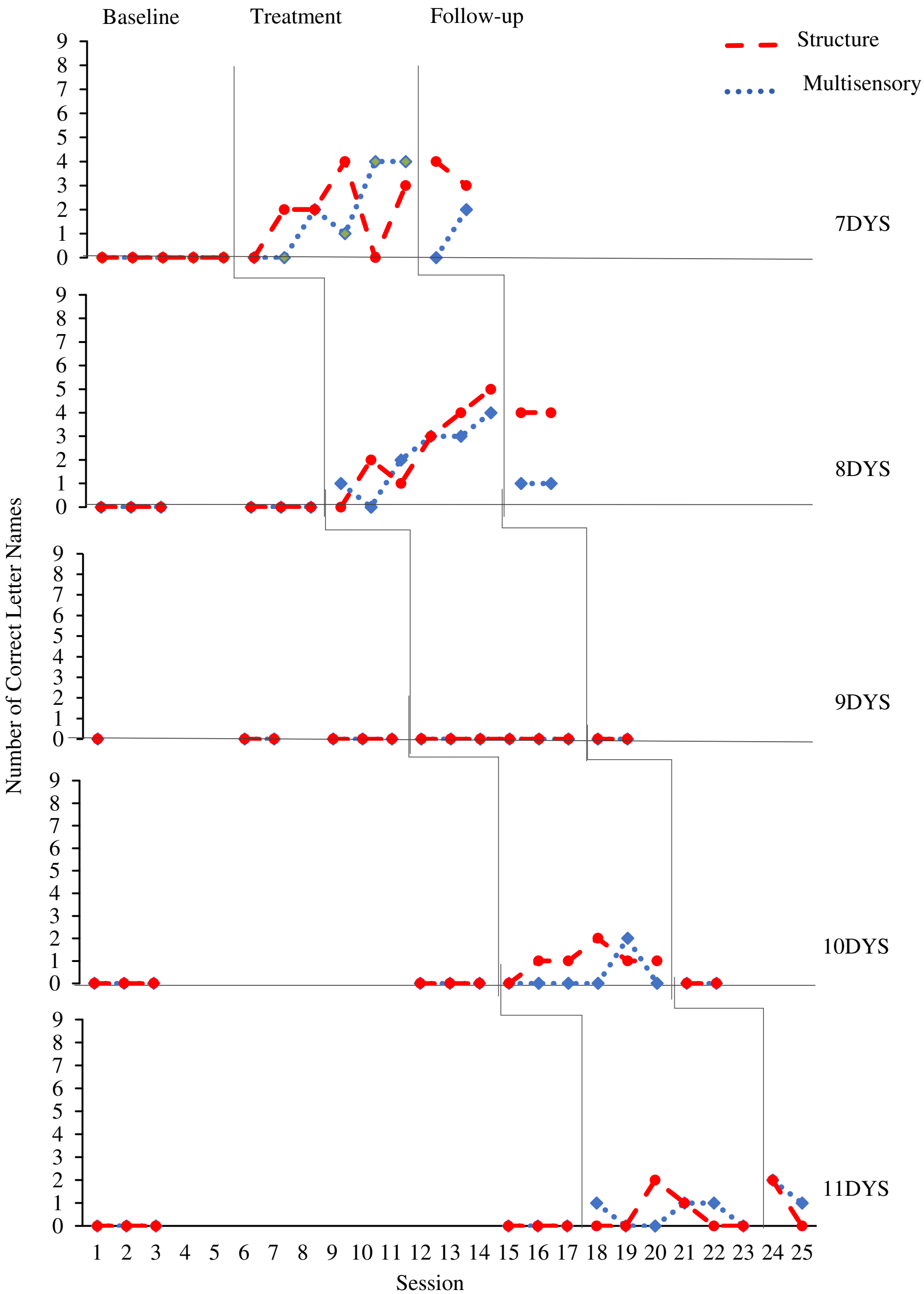

Fig. 6 Accuracy of the number of letter names produced for each participant with dyslexia across three phases. *DYS* = dyslexia

significant. The response patterns and inter-case replications indicated an experimental effect for multisensory but not for structured language interventions. However, the less

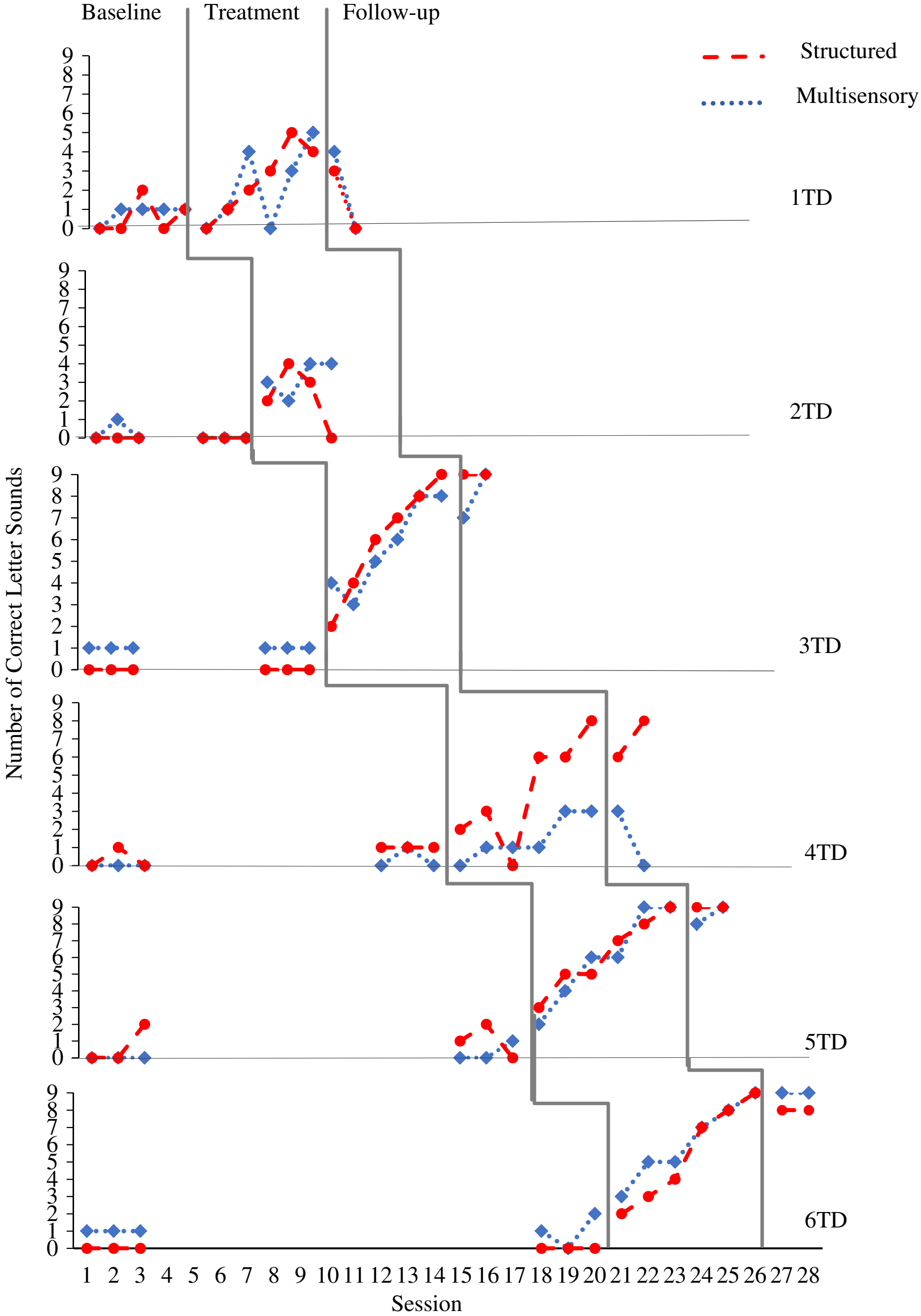

Fig. 7 Accuracy of the number of letter sounds produced for each participant with typical development across three phases. *TD* = typical development

conservative weighted Tau-U for letter sounds showed small, statistically significant effect sizes for both interventions (structured *Tau-U* = 55, multisensory *Tau-U* = 58).

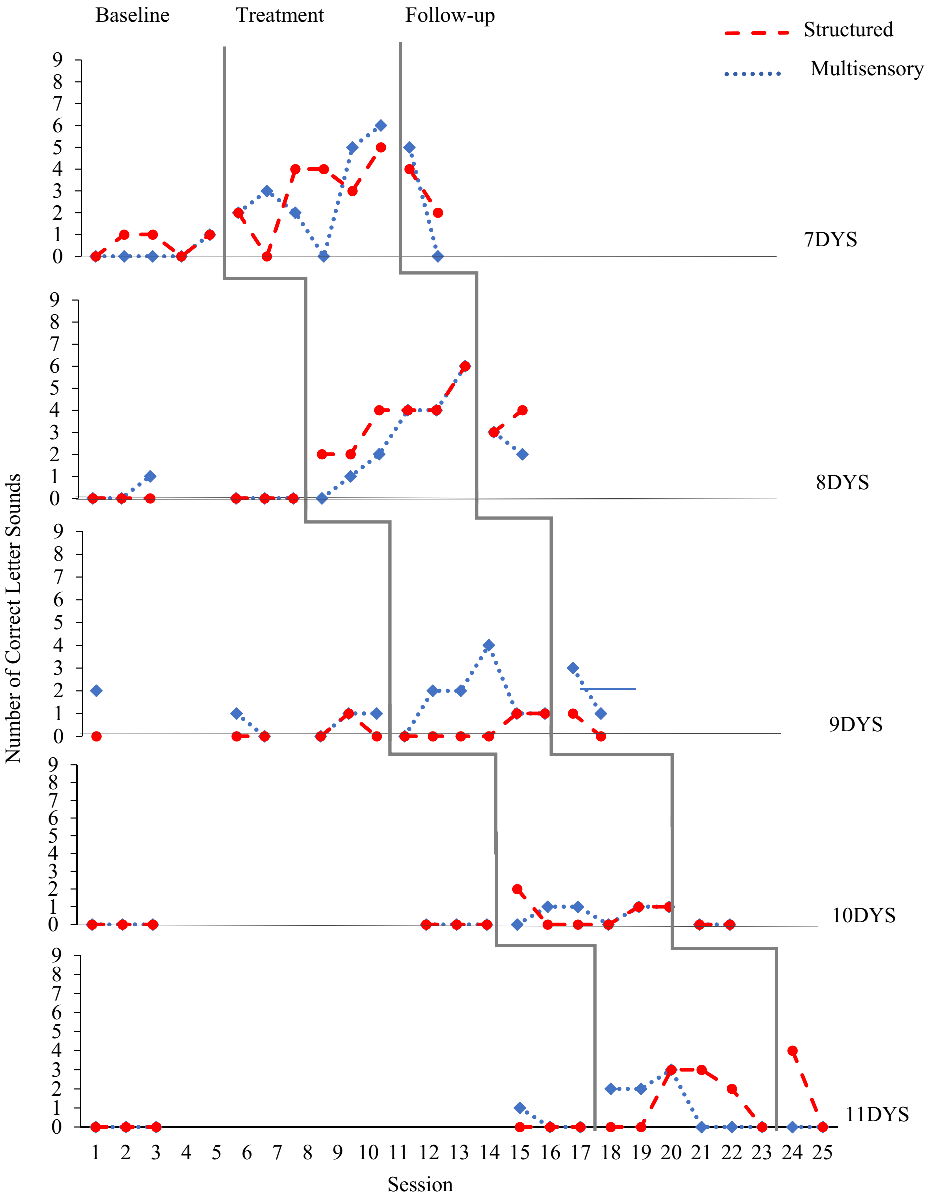

Fig. 8 Accuracy of the number of letter sounds produced for each participant with dyslexia across three phases. *DYS* = dyslexia

Follow-up data showed participants 7, 8, 9, and 11 demonstrated maintenance of letter sounds. Participants 7, 8, and 11 had higher maintenance in structured language than multisensory. Participant 11 did not demonstrate maintenance in multisensory intervention but did in structured language. Participant 9 had higher maintenance in multisensory than structured intervention.

Words read correctly

Participants with typical development Graphs of visual analyses are presented in Fig. 9. Positive response data patterns and inter-case treatment replications indicated a positive experimental effect for both interventions, which was supported by individual Tau-U effect sizes. However, visual analyses indicated a structured language intervention advantage. This was supported by the weighted Tau-U for words read correctly which indicated there was a statistically significant, large effect size for structured intervention (*Tau-U* = 95), and a significant, moderate effect size for multisensory (*Tau-U* = 78).

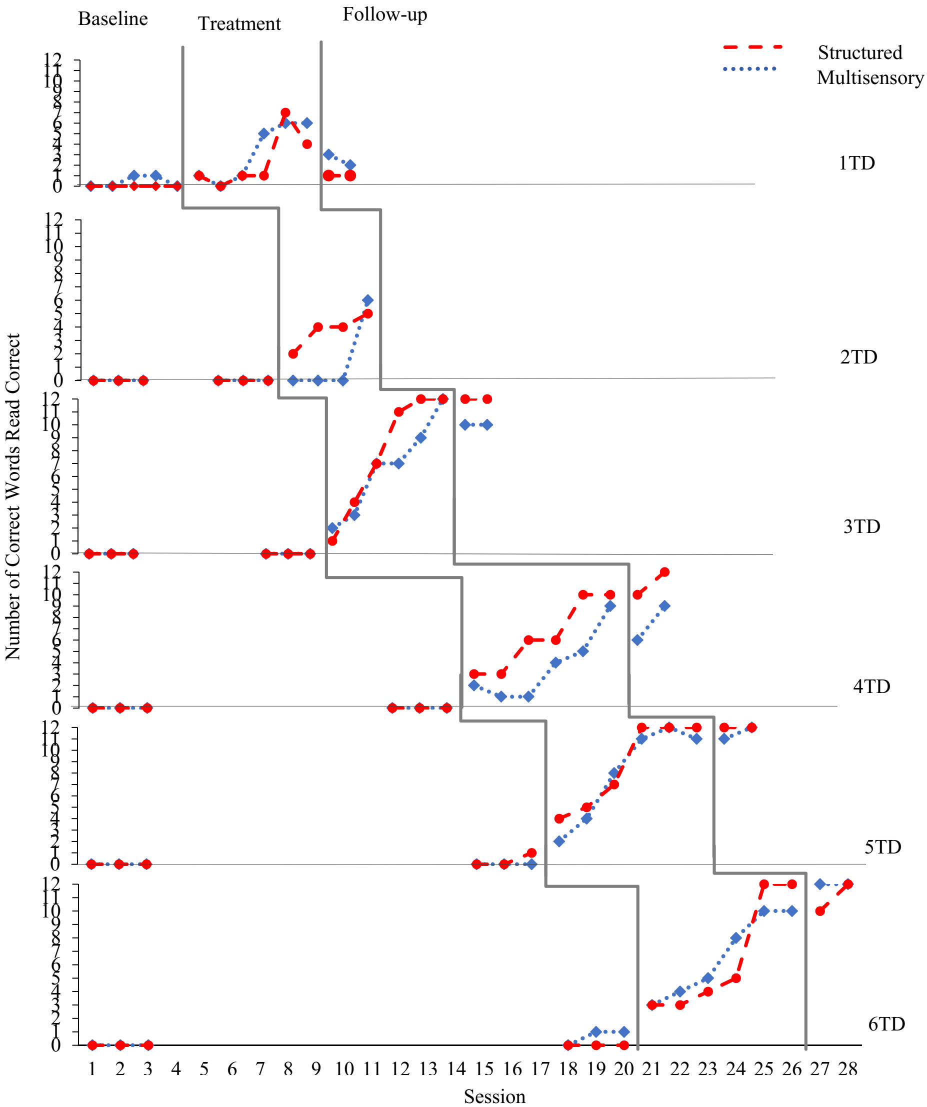

Fig. 9 Accuracy of words read correct for each participant with typical development across three phases. *TD* = typical development

Follow-up data showed all participants demonstrated maintenance for words read correct in each treatment condition, with somewhat higher maintenance in structured language intervention for participants 3, 4, and 5. Participants 1 and 6 had higher maintenance in multisensory intervention; follow-up data was not available for participant 2.

Participants with dyslexia Graphs of visual analyses are presented in Fig. 10. Observed treatment levels were greater than projected levels at baseline for all participants except for participants 9 and 10. Participant 9 remained at baseline levels for both interventions throughout the three phases, and participant 10 remained at baseline levels for multisensory throughout

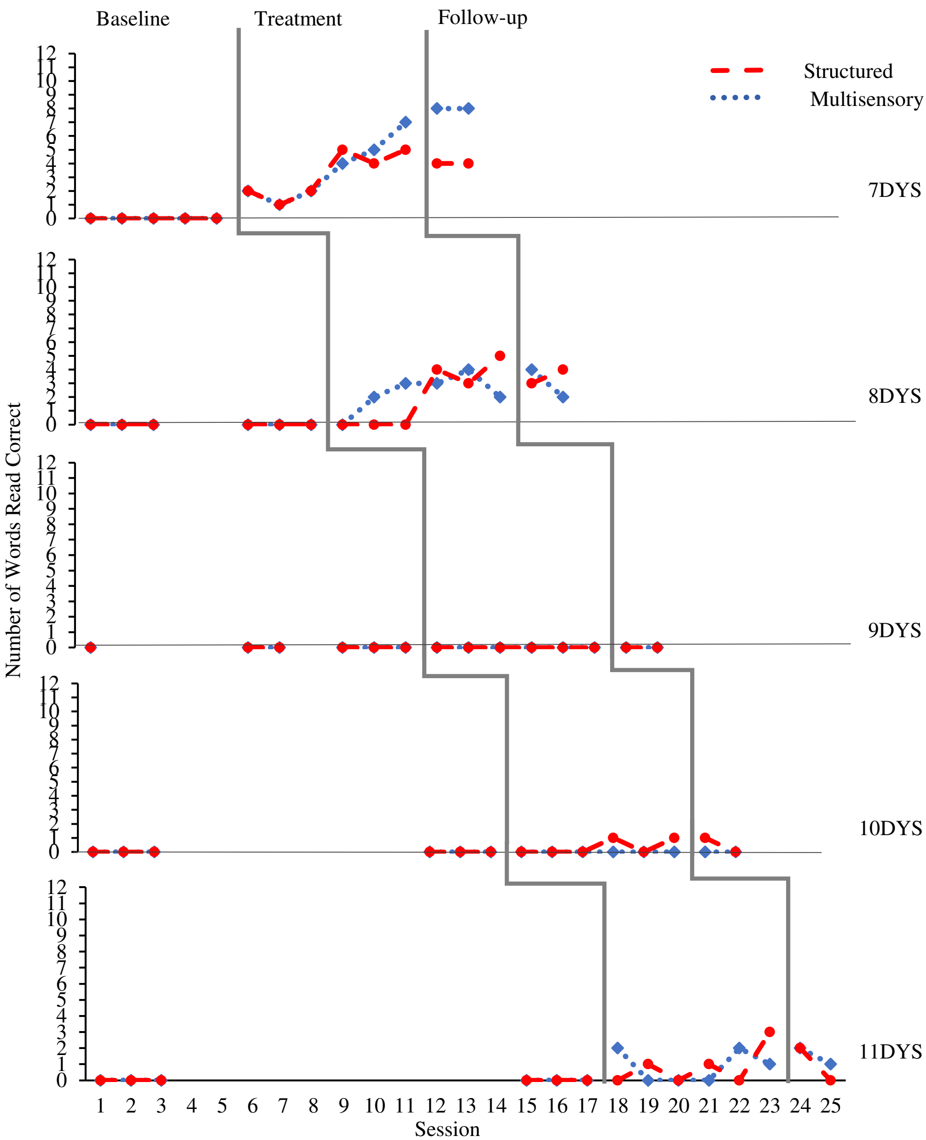

Fig. 10 Accuracy of words read correct for each participant with dyslexia across three phases. *DYS* = dyslexia

the three phases. For the remaining participants, neither structured language nor multisensory appeared effective, which was supported by individual Tau-U effect sizes. Only participants 7 and 8 had inter-case replications. Thus, per visual inspection, neither intervention appeared to have had an experimental effect for words read correctly since three inter-case replications were not evident. However, the weighted Tau-U indicated significant, small effect sizes for both interventions (structured language and multisensory *Tau-U* = 46).

Follow-up data showed participants 7, 8, 10, and 11 demonstrated structured language maintenance for words read correctly. Participants 7, 8, and 11 demonstrated maintenance in multisensory intervention. Participants 8 and 10 demonstrated higher structured language maintenance, and participants 7 and 11 had higher multisensory maintenance. Participant 9 remained at baseline for each intervention.

Words spelled correctly

Participants with typical development Graphs of visual analyses are presented in Fig. 11. For participant 4, multisensory intervention had lower scores and more variable data patterns, which indicated structured intervention was more effective than multisensory. For the remaining participants, neither intervention showed a clear advantage. The positive response data patterns and inter-case treatment replications indicated a positive treatment effect for both interventions, which was supported by individual Tau-U effect sizes. Based on visual analyses, both interventions had a positive overall experimental effect. The weighted Tau-U for words spelled correctly indicated statistically significant, moderate effect sizes for both interventions (structured language *Tau-U* = 85, multisensory *Tau-U* = 86).

Follow-up data showed all participants demonstrated maintenance for words spelled correctly in each treatment condition. For participant 1, multisensory had a higher follow-up level, and for participant 4, structured language had a higher follow-up level. For the remaining participants, structured language and multisensory follow-up levels were similar.

Participants with dyslexia Graphs of visual analyses are presented in Fig. 12. For both interventions, the observed treatment levels were above projected levels based on baseline scores, except for participant 10 who did not respond to multisensory treatment and remained at baseline levels. Visual analyses of data features indicated both interventions had a small effect as demonstrated by slope, level, linear trends, and immediacy of effect. There were three structured language inter-case treatment replications for participants 7, 8, and 11 compared to two multisensory inter-case replications for participants 7 and 8. Therefore, structured language had an experimental effect, but the multisensory intervention did not. The individual Tau-U effect sizes supported the structured language experimental effect for participants 7, 8, and 11. Based on visual analyses and inter-case replications, the structured language intervention showed an advantage over multisensory. The weighted Tau-U for words spelled correctly showed a statistically significant, moderate effect size for structured (*Tau-U* = 69) and a significant, small effect size for multisensory (*Tau-U* = 60).

Follow-up data showed that participants 7, 8, and 11 demonstrated maintenance for words spelled correctly in each treatment condition. Structured language had higher maintenance for participants 8 and 11, and 7 had higher maintenance for multisensory. Participants 9 and 10 did not demonstrate maintenance for either intervention.

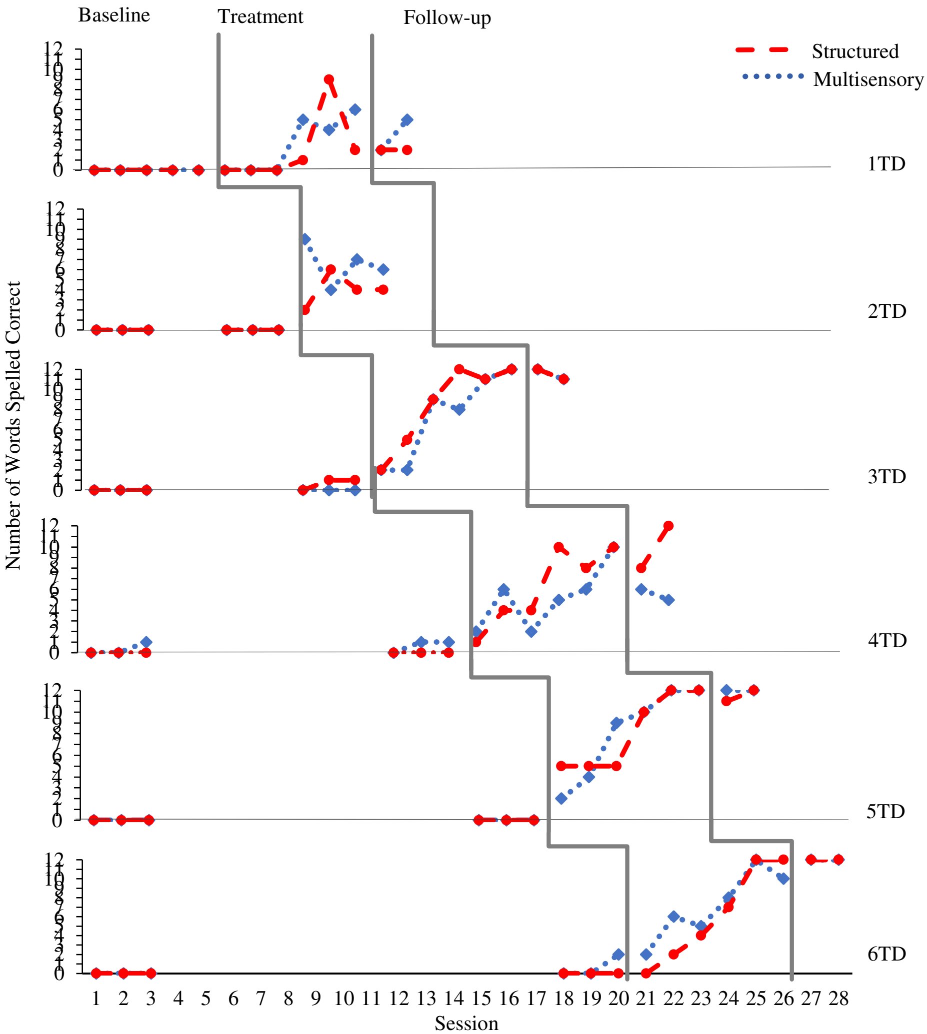

Fig. 11 Accuracy of words spelled correct for each participant with typical development across three phases. *TD* = typical development

Discussion

This study investigated the efficacy of two reading interventions for teaching letter name, sound production, word reading, and word spelling for 11 second grade students, six with typical development and five with dyslexia. The first intervention was a structured language program with restricted use of simultaneous multisensory input; the second intervention was a multisensory structured language program that included simultaneous multisensory input. Due

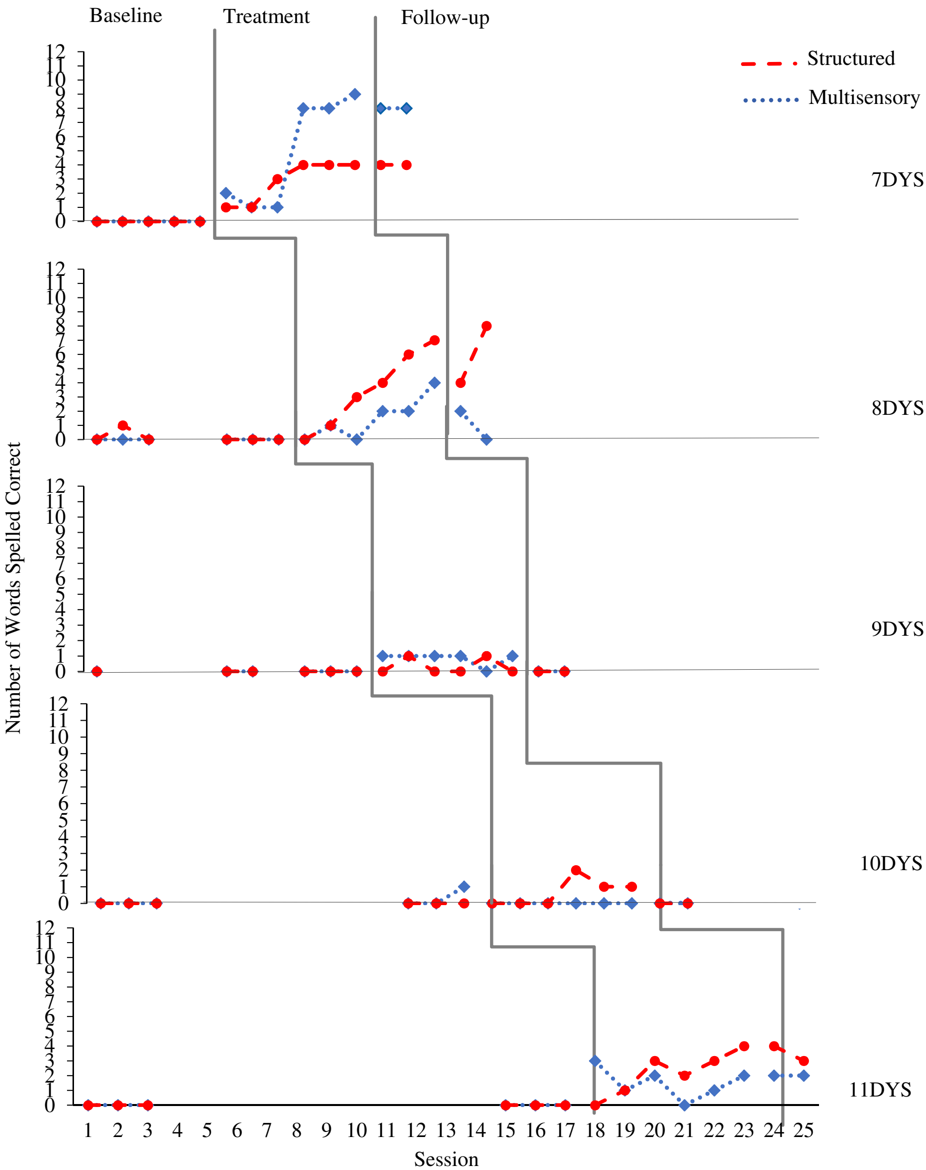

Fig. 12 Accuracy of words spelled correct for each participant with dyslexia across three phases. *DYS* = dyslexia

to the lack of experimental control in previous multisensory intervention studies (Ritchey & Goeke, 2006; Rose & Zirkel, 2007), this empirical study was designed to add to the limited scientific evidence testing the efficacy of simultaneous multisensory instruction for improving reading skills.

Based on research that has shown simultaneous multisensory instruction to be effective for teaching foundational reading skills to children with typical development (Joshi et al., 2002; Uhry & Shepherd, 1993) and dyslexia (Campbell et al., 2008; Hulme, 1981; Torgesen et al., 2001), the hypotheses were that multisensory intervention would provide an advantage over structured instruction for the learning and maintenance of letter names, letter sounds, word reading, and spelling. Hypotheses were in line with the principles of dual coding theory and supported by research that multimodal instruction has been shown to provide more opportunity to build representational connections for both visual and verbal information and enhanced referential connections between logogens and imagens (Mayer & Anderson, 1991).

Visual analysis and Tau-U effect sizes indicated that both structured language and multisensory instruction had a positive treatment effect for participants. However, there did not appear to be a clear overall advantage for one type of instruction. Likewise, visual analyses indicated that maintenance was apparent for both interventions, but there did not appear to be an overall advantage for either intervention.

Effectiveness of multisensory compared to structured language intervention for letter names and sounds, words read, and words spelled for participants with typical development

Letter names and letter sounds for participants with typical development Visual analyses and Tau-U effect sizes indicated both structured language and multisensory instructions had an overall treatment effect for letter names and sounds. Though there appeared to be a multisensory intervention advantage in some cases for letter name, the multisensory intervention was not more effective for all participants than structured language. Therefore, the hypothesis that multisensory would be more efficacious than the structured language for all participants with typical development for letter names and sounds per visual inspection of slope, level, immediacy of effects, and Tau-U effect sizes was not supported.

The results of this study were consistent with findings in the extant literature that explicit, structured, and systematic instruction is effective for teaching basic literacy skills (e.g., Adams, 1990; Ehri, 2014; NRP, 2006). However, it did not appear that the addition of simultaneous multisensory input provided an overall advantage. During auditory learning of words for objects, such as in learning letter names in this study, phonological representations are cognitively processed (Gupta & Tisdale, 2009). For participants with typical development, with intact phonological processing abilities, it appeared they learned grapheme names because they were able to effectively build, for each grapheme in each intervention, a phonological representation, a semantic representation, as well as links between the representations (Gupta & Tisdale, 2009).

Although participants with typical development did well learning letter names, during structured language letter name assessments, error responses of participants 1, 2, 3, and 4 frequently were due to responding with the correct letter sound instead of the letter name. These errors changed the level and trend in structured language for participants' letter names. This type of incorrect response did not result in changes in the data trajectory for multisensory intervention letter name. In a series of experiments with similar aged children, Hulme (1987) found simultaneous tracing and naming of letter-like forms resulted in improved visual recognition as well as significantly more correct letter form names. Results were interpreted to imply that simultaneous multisensory tracing and naming improved the recognition phase of paired-associate learning. Perhaps the simultaneous multisensory tracing activities resulted in better visual-verbal paired-associate learning for the multisensory letter name, and therefore, participants did not incorrectly give multisensory letter sounds during assessments as in structured language.

Furthermore, as indicated by research, children utilize phonological skills to learn grapheme sounds. Sounds are learned easiest if the name contains the phoneme represented by the grapheme (Treiman, et al., 1998). Participants with typical development frequently appeared to be vocalizing sounds before responding to assessment prompts. Because of intact phonological processing, it is possible that participants were able to use sub-vocalization, to rehearse letter sounds (Baddeley, Gathercole, & Papagno, 1998). In addition, because a majority of the grapheme names contained the grapheme's sound, either at the beginning or the end of the letter name, interventions further promoted letter sound learning.

The lower multisensory data features for letter sounds for participant 4 were likely due to the grapheme to phoneme randomization process; a grapheme with a similar shape to the English uppercase letter *A* was paired with a lax vowel for this participant in the multisensory intervention. Participant 4 associated the incorrect vowel sound, /æ/, to the grapheme during baseline assessments. It appeared that participant 4 was unable to inhibit interference for the /æ/ sound during the treatment phase, despite being taught the correct sound. This lack of interference control influenced multisensory assessment responses for letter name until the third teaching session, when the /æ/ sound was taught in structured language. However, letter sounds and words read continued to be affected throughout the treatment phase. As a result, the multisensory data had lower levels and trends in letter sound and words read for participant 4, because multiple words contained the grapheme that was incorrectly encoded as /æ/.

Words read and words spelled correct for typical development Per visual analysis and individual Tau-U effect sizes, structured language showed an advantage for words read. However, upon closer inspection of the data, the advantage appeared to be related to data patterns for two participants, 2 and 4. Participant 2 only attended four teaching sessions, and participant 4's lower multisensory data patterns were due to the interference of the /æ/ sound in word reading. This calls into question the structured language advantage for words read. Visual analyses and Tau-U effect sizes for words spelled indicated both structured language and multisensory instructions had an overall treatment effect for all participants. Therefore, the hypothesis that multisensory would be more efficacious than the structured language for words read and spelled for all participants with typical development per visual inspection of slope, level, immediacy of effects, and Tau-U effect sizes was not supported.

In this paper, orthographic knowledge includes mental representations of written words stored in memory and how speech is represented in writing, including the alphabetic principle (Apel, 2011). To learn words, participants utilized phonological awareness and newly developed orthographic knowledge for both interventions to effectively read words. Participants demonstrated phonological awareness in their ability to decode words by segmenting then blending sounds during reading activities and assessments. Orthographic knowledge was evident by participants' ability to correctly read newly taught words and utilize analogy to read unknown words. For example, participants were able to use a previously taught rime /ɛz/ and correctly insert an untaught phoneme for a word containing the rime (e.g., the /k/ in /kɛz/). For participants with typical development, their proficiency in phonological awareness, phonological recoding, and phonetic recoding allowed them to learn to read words fluently (Wagner & Torgesen, 1987) in both interventions.

Much of the same underlying knowledge used for reading is also used in spelling (Moats, 2006). Orthographic processing is the ability to acquire, store, and use orthographic knowledge (Apel, 2011). For both interventions, participants' well-developed phonological awareness and orthographic processing were able to be employed to spell words. During session activities, participants with typical development segmented spelling words into phonemes correctly selected graphemes, then checked spelling accuracy by decoding the word. Furthermore, via a process of elimination, participants appeared to use orthographic processing to help spell words. For example, participants would spell a three phoneme word containing two letters in which letter sound correspondence had been taught and correctly guess about another sound in which the letter sound had not been directly taught. Participants were able to map phonological and orthographic connections between words and develop orthographic word forms (Berninger et al., 2006a).

Effectiveness of multisensory compared to structured language intervention for maintenance of letter names and sounds, words read, and words spelled for participants with typical development

Letter names and letter sounds for typical development For participants with typical development, it was hypothesized there would be a multisensory intervention advantage for maintenance of letter names and sounds. Both interventions proved to be effective for the retention of letter names and sounds; however, the multisensory intervention did not provide an overall advantage. Therefore, the hypothesis was not supported. All participants demonstrated maintenance for letter names and sounds for both interventions, although follow-up levels were not available for participant 2. However, for participant 4, the incorrect /æ/ sound continued to be given during multisensory assessments and as a consequence multisensory letter sound level was lower than structured level. Participants with typical development, because of their ability to efficiently learn letter names, were able to meet the mastery criteria (correctly name new letters twice in assessments) and learn new letter names in subsequent sessions. All material previously taught was reviewed at the beginning of each teaching session. The repeated exposure and practice of previously taught information likely strengthened participants' mental and semantic representations and links for letter names in both interventions. The results were well-retained letter names.

Research has shown knowledge of letter names promotes awareness of letter sounds (NELP, 2008; Treiman, et al., 1998). Participants overall were successful at maintaining letter sounds, in part due to their ability to successfully utilize existing phonological skills and strategies for learning letter sound correspondence, such as letter name and phonemic awareness, which they appeared to apply equally well to both interventions. Because of intact phonological processing, it is probable that participants utilized articulatory rehearsal to recover auditory memory (Baddeley, 2000; Berninger et al., 2006b). This would explain the consistent maintenance results found across interventions for letter sound.

Words read and words spelled correct for typical development For participants with typical development, it was hypothesized there would be a multisensory intervention advantage for maintenance of words read and spelled. Both interventions proved to be effective for maintenance of word reading and spelling; however, the multisensory intervention did not provide an overall advantage. Therefore, the hypothesis was not supported. The available follow-up data showed all participants demonstrated maintenance for words read and spelled correct for both interventions. For participants with typical development, word reading was fluent and accurate for both interventions. Repeated word reading practice during teaching sessions gave participants the opportunity to develop strong phonological and orthographic connections for words. Participants' initial and repeated practice using decoding and analogy word reading strategies lead to maintenance, as demonstrated by their fluent word reading (Ehri & McCormick, 1998) at follow-up.

For both interventions, participants utilized their phonological awareness and orthographic processing abilities to spell word forms. This allowed participants to map phonological and orthographic connections between words and allowed participants to not only develop but maintain orthographic word forms (Berninger et al., 2006a).

Summary for participants with typical development

Research supports use of structured, systematic instruction for teaching basic reading skills (e.g., Adams, 1990; Ehri, 2014; NRP, 2006). For participants with typical development, the explicit and systematic instruction common to both interventions provided a possible explanation for the positive overall intervention effects found across dependent variables and interventions. The results of this study reinforce extant literature that evidence-based reading instruction should incorporate structured, systematic instruction (NRP, 2006). Results extend the literature by demonstrating structured language, and multisensory interventions were efficacious for teaching foundational literacy skills. Lack of overall multisensory advantage suggested overall positive effects for both interventions were likely not due to the simultaneous multisensory input but to the embedded Orton-Gillingham structured language components common to both reading interventions, such as phonemic spelling.

Effectiveness of multisensory compared to structured language intervention for letter names and sounds, words read, and words spelled for participants with dyslexia

Letter names and letter sounds for participants with dyslexia Visual analyses and Tau-U effect sizes indicated structured language appeared to be more advantageous than multisensory instruction for letter names. The visual analyses for letter sound indicated multisensory intervention was more effective based on three inter-case replications. However, the weighted Tau-U effect sizes indicated both interventions were similarly advantageous. Therefore, the hypothesis that multisensory would be more efficacious than structured language for letter name and sound for all participants with dyslexia per visual inspection of slope, level, immediacy of effects, and Tau-U effect sizes was not supported.

It is unclear why structured language was more advantageous for letter name. Although interventions were based on best practices, participants had considerable difficulty learning letter names for both interventions. Learning letter names for the graphemes required participants to cognitively process an internal phonological representation of the word (Gupta & Tisdale, 2009). In order to learn the letter name, participants had to store both a phonological and a semantic representation of the word. In addition, participants had to develop strong phonological–semantic links, to produce the letter name during assessments (Gray, 2005).

Participants with dyslexia, who by definition have difficulty with phonological processing, appeared to have difficulty encoding phonological information and developing phonological and semantic representations and links (Gray, Pittman, & Weinhold, 2014; Gupta & Tisdale, 2009). This was evident by the difficulty all participants had meeting mastery criteria. Participants frequently required three or more re-introductions of a letter name before meeting the mastery criteria. Participants with dyslexia needed repetitive practice to allow them to develop sufficient phonological and semantic representations and links in order to recall the novel letter names during assessments (Gray, 2005).

Furthermore, for letter name, participants 9, 10, and 11 did not respond in the same manner as participants 7 and 8 to either intervention. It appears these three participants were *treatment resisters,* which meant they did not respond or were slower to respond to interventions (Alexander & Slinger-Constant, 2004; Berninger et al., 2000; Shaywitz et al., 2008), based on their lower levels and trends for letter name and sound and words read and spelled. Empirical evidence suggests treatment resisters require differentiated instruction that meet their individual needs (Berninger et al., 2000; Alexander & Slinger-Constant, 2004). The multiple re-introduction of letter names in order to meet the mastery criteria suggested the need for differentiated instruction. Perhaps the presentation of only one intervention per session or the introduction of only one letter per intervention may have improved data response patterns for these three participants.

For letter sound, deficiencies associated with dyslexia have been known to affect the ability of individuals to develop the alphabetic principle (Shaywitz et al., 2008). Research has indicated that children do not memorize letter sound correspondences as rote pairs; rather they evaluate and utilize their knowledge of the letter's name to develop an understanding of the letter's sound. Therefore, both letter knowledge and phonological awareness are utilized in learning letter sounds (Treiman, et al., 1998). Because of poor phonological awareness and difficulty establishing letter knowledge, participants exhibited a difficult time making necessary links between a letter's name and sound. Due to their difficulty encoding phonological information, participants were not able to effectively utilize the phonological loop as a resource for learning and later recovering stimuli via auditory rehearsal for letter sound (Baddeley, 2000; Berninger et al., 2006b).

Words read and words spelled correct for dyslexia Visual analysis and individual Tau-U effect sizes for words read indicated that both interventions were effective for two participants only; therefore, due to the lack of three interclass replications, neither intervention appeared to be effective. Visual analyses and Tau-U effect sizes for words spelled indicated structured language instruction had an overall treatment effect for participants, but multisensory intervention did not. Thus, the hypothesis that multisensory would be more effective than structured language for words read and spelled for all participants with dyslexia per visual inspection of slope, level, immediacy of effects, and Tau-U effect sizes was not supported.

Only participants 7 and 8 appeared to learn in both interventions for word reading. The remaining participants, 9, 10, and 11, did not learn well in either intervention, per their levels and trends for words read. These participants appeared to struggle with letter sound decoding as evident by their difficulty segmenting words during teaching sessions and their inability to demonstrate one to one letter sound correspondence for two and three phoneme words. Participants also required frequent redirection to maintain focus.

Participant 10 presented additional concerns. Participant 10 had trouble accurately repeating spelling words two to three phonemes long. An inspection of participant 10's inclusionary phonemic decoding assessment revealed it was in the first percentile, indicating a possible issue with phonological recoding. Additional examination of descriptive assessments indicated phonological awareness was in the fifth percentile and rapid automatic was naming in the 25th percentile. Poor phonological awareness and rapid automatic naming may indicate a possible double deficit. The double deficit hypothesis suggests children who have deficits in both phonological awareness and rapid automatic naming are the most severely impaired readers (Norton & Wolf, 2012). A double deficit profile would provide a hypothesis for participant 10's lack of response to either intervention.

Participants 9, 10, and 11 difficulties with phonological awareness, poor attention, and the possible double deficit for participant 10 are indicative of the heterogeneous nature of dyslexia (Tobia & Marzocchi, 2014; Ramus, 2004). Alexander and Slinger-Constant (2004) suggest reading requires attention to sensory input to map representations, such as phonological and orthographic representations, to neural substrates (Alexander & Slinger-Constant, 2004). It appears the lack of attention to incoming stimuli for participants 9, 10, and 11 made it difficult for them to map phonological and orthographic representations for word reading. The lack of or poorly developed representations resulted in deficit input from phonological and orthographic components and affected the holding and manipulation of information for processing in working memory (Alexander & Slinger-Constant, 2004). This further may help explain the response patterns for these three participants in word reading.

It appeared participants with dyslexia had difficulty mapping phonological and orthographic relationships and therefore were not able build orthographic word forms for correct word spelling (Berninger et al., 2006a). For participants with dyslexia, their poorly developed phonological awareness made it difficult to acquire, store, and use orthographic knowledge. For example, participants would select three or more graphemes for a two phoneme spelling word or select a grapheme they had been taught but place it the incorrect position. In addition, participants with dyslexia often would not attempt to spell a three phoneme word containing an untaught phoneme, even though they had demonstrated correct spelling of a word containing the two taught phonemes.

Multisensory effectiveness compared to structured language for maintenance of letter names and sounds, words read, and spelled for participants with dyslexia

Letter name and letter sounds for dyslexia For participants with dyslexia, it was hypothesized there would be a multisensory intervention advantage for maintenance of letter names and sounds. Only three participants demonstrated maintenance for letter names, and only four participants had maintenance of letter sounds. Therefore, the hypothesis was not supported. Participants 7, 8, and 11 had follow-up levels higher letter name than treatment phase levels in structured language. The remaining participants did not demonstrate maintenance of information. Due to the inability of participants with dyslexia to meet the mastery criteria, their exposure to new letter names was limited. However, in spite of repeated exposures and practice on a limited number of letter names, some participants with dyslexia had difficulty maintaining letter names. It appeared, for both interventions, the phonological and semantic representations and links for letter names were tenuously established and subject to decay without the repeated practice of treatment sessions. Despite the implementation of evidence-based practices, the instruction was not differentiated based on individual needs, which further may explain the poor retention for participants with dyslexia for letter names.

Three participants, 7, 8, and 11, demonstrated higher structured language maintenance for letter sound. Poor maintenance by participants with dyslexia for letter sounds was in part due to their difficulty establishing secure phonological and semantic representations and links for letter names. Their poorly developed letter name representations made linking a sound to the letter difficult, which resulted in poor retrieval and production. It required multiple repetitions for participants with dyslexia before information about the letter sound was able to activate a short-term trace in the phonological store that was able to influence long-term retention (Baddeley, 2000; Baddeley et al., 1998).

Words read and words spelled correct for dyslexia For participants with dyslexia, it was hypothesized there would be a multisensory intervention advantage for maintenance of words read and spelled. Three participants, 7, 8, and 11, demonstrated maintenance for words read in both interventions. Participant 10 only demonstrated maintenance for the structured language. However, follow-up levels for participants 10 and 11 were quite low due to lack of maintenance at follow-up two. Only participants 7, 8, and 11 appeared to retain information of words spelled, with follow-up levels higher than treatment levels for structured language. Only participant 7 had follow-up levels higher than treatment levels in multisensory. Therefore, multisensory intervention did not provide an overall advantage for either words read or spelled. Therefore, the hypothesis was not supported. The low follow-up levels for both interventions by participants with dyslexia may reflect their poor phonological awareness and orthographic knowledge (Lyon, Shaywitz, & Shaywitz, 2003; Richards, et al., 2006). For participants with dyslexia, few words were read fluently in either intervention, despite repeated exposures to a limited set of reading words. Participants were unable to develop strong phonological and orthographic connections to words. Also, participants were not able to efficiently store phonological representations (Gupta & Tisdale, 2009) and therefore were not able to maintain words during follow-up.

Maintenance of spelling words was only evident for three participants despite repeated practice with limited spelling words and one to one correspondence for each grapheme with an English phoneme. Participants with dyslexia were not able to utilize orthographic knowledge (e.g., NRP, 2000; Snowling & Hulme, 2011) to correctly spell words containing two to three phonemes and therefore were not able to maintain the phonological and orthographic connections they had developed.

The short duration of the study may not have provided enough instructional time for either intervention to effectively address the participants' learning differences for word reading and spelling. Research suggests for individuals with dyslexia to demonstrate overall gains in reading requires time-intensive intervention (Torgesen, et al., 2001). Research has shown intensive reading remediation, 100 min per day for 8 weeks, in small groups with explicit systematic instruction improves reading outcomes and promotes maintenance (Gabrieli, 2009). Neuroimaging studies of individuals with dyslexia have shown normalization of activity in left temporoparietal and frontal regions with intensive differentiated remediation (Gabrieli, 2009).

Summary for participants with dyslexia

Participants with dyslexia demonstrated varied data patterns for both interventions. For participants with dyslexia, the multisensory instruction did not provide an overall advantage. It appeared neither intervention adequately overcame participants' poor orthographic knowledge, phonological awareness, or phonological recoding to help them learn the targeted literacy skills. Though both interventions were evidence based, participants 9, 10, and 11 did not appear to respond in the same manner as participants 7 and 8. The response patterns for participants 9, 10, and 11 highlighted the concern for treatment resisters (Alexander & Slinger-Constant, 2004; Berninger et al., 2000; Shaywitz et al., 2008). For these participants, the lack of differentiated instruction, which has been shown to be important in current research (Berninger et al., 2000; Alexander & Slinger-Constant, 2004), may have been a factor in poor learning. Furthermore, the individual profile characteristics for participants 9, 10, and 11 showed the multifactorial nature of dyslexia (Ramus, 2004; Tobia & Marzocchi, 2014).

Educational implications and future research

This study evaluated the impact of simultaneous multisensory input on developing basic reading skills within an Orton-Gillingham based structured language frame work. It is one of the first studies to evaluate simultaneous multisensory input in a well-controlled study. This scientific study supported structured language instruction, within Orton-Gillingham based programs, as efficacious in promoting basic decoding and encoding skills for children with typical development and dyslexia. However, this study did not show that simultaneous multisensory input improved learning over structured language intervention alone. In fact, results suggested that other components within the Orton-Gillingham framework, such as phonemic spelling, or reciprocal teaching of reading and spelling, may play critical roles in the effectiveness of structured language programs. Research has shown children benefit from explicit and systematic structured language instruction (e.g., Adams, 1990; Ehri, 2014; NRP, 2006). For individuals with dyslexia, it is especially critical to provide literacy instruction based on sound empirical evidence (e.g., Berninger et al., 2013; Berninger et al., 2000; MacArthur & Graham, 1987; Torgesen, et al., 2001). It is important to empirically study other elements of Orton-Gillingham multisensory structured language programs, for example the diagnostic teaching to mastery, embedded phonological awareness activities, and reciprocal teaching of reading and spelling, to determine which components promote learning.

The divergent characteristics profiled by inclusionary and descriptive assessments and session observations for participants with dyslexia support research demonstrating dyslexia is a multifactorial deficit, e.g., presenting multiple deficits (e.g., Berninger, 2008; Berninger et al., 2013; Norton & Wolf, 2012; Shaywitz et al., 2008; Tobia & Marzocchi, 2014), and provided insight as to the impact individual profiles have on learning basic literacy skills. There is a worthwhile opportunity to extend this research by using diagnostic and prescriptive protocol to determine how best to meet the needs of treatment resister. Evaluating the impact of differentiated instruction in longitudinal studies for individuals with dyslexia would develop a clearer picture of how to best support treatment resisters across the continuum.

Strengths and limitations

Strengths This empirical study was highly controlled, ensuring scientifically valid results. First, non-English graphemes were used to better target simultaneous multisensory input as a variable and provide control for participants' prior letter and lexical knowledge. We elected to use this control as it allowed participants to be on more equal footing since all participants lacked prior exposure to the created Saraf and Rasaf alphabets. Second, outside of the first author, interventionists received the same amount of training. Third, percentages for fidelity of implementation and reliability were high and directly reported. Fourth, the use of visual analysts naïve to the study's purpose helped control for type I errors (Ferron & Jones, 2006). Fifth, the design followed WWC (2013) established criteria. Lastly, the more stringent experimental effect of three inter-case replications was used versus the minimally accepted two inter-case replications (Horner et al., 2012).

Limitations In both interventions, participants were introduced to reading and spelling practices not commonly utilized in classrooms (Berninger et al., 2000), for example spelling isolated phonemes. Dual coding research has shown the lack of experience with a task can increase demands on the central executive and in turn hinder working memory (Constantinidou, Danos, Nelson, & Baker, 2011). This was likely compounded by participants' lack of experience with the created alphabets. Practice with strategies prior to beginning the study could have helped control for new learning techniques and strategies and their undue influence on outcomes. Furthermore, poor response by participants with dyslexia to either intervention may have been due to the short duration of the treatment phase. A more pronounced advantage for one intervention over the other may have become more evident had the study included a longer treatment phase. Research suggests intensive intervention is needed for individuals with dyslexia to demonstrate reading gains. Intensity of intervention includes explicit and systematic instruction in small groups and increases in instructional time (Gabrieli, 2009; Torgesen, et al., 2001).

Conclusions

This well-controlled study using two created alphabets provided important missing information regarding simultaneous multisensory input as efficacious reading intervention. Using a single-case design allowed each participant to act as their own control in the learning of the two created alphabets, Saraf and Rasaf. Furthermore, the use of the created alphabets helped regulate the influence of extraneous variables on the independent variable of simultaneous multisensory instruction since participants' exposure to the alphabets were restricted to treatment sessions.

Results supported structured language instruction within an Orton-Gillingham based program as effective in promoting basic literacy skills. However, simultaneous multisensory input did not provide a treatment effect above and beyond the structured language effect. This is an important finding, as significant time is spent training instructors on the implementation of simultaneous multisensory aspects. In addition, considerable teaching time is utilized implementing simultaneous multisensory activities during lesson instruction. If multisensory features are not critical, instructor training and teaching of lessons might be easier and more efficient to implement. Other components inherent to structured language may have directly impacted treatment effects. This study supported extant literature that explicit systematic language instruction is important for developing foundational decoding and encoding skills for both children with typical development and re. Importantly, the multiple deficit nature of dyslexia was amplified in this study. The critical need for individuals with dyslexia, especially treatment resisters, to be provided with differentiated instruction that is diagnostic, prescriptive, and empirically based was accentuated.

Acknowledgements We greatly appreciate Lucasfilm Ltd. for the approval to use the Aurebesh letters from *Star Wars* © & ™ Lucasfilm Ltd. Many thanks are due to the participants and their families, the reading centers and schools who supported this study, and the visual analysts, research assistants, and volunteers whose support and hard work helped complete the study. We are grateful to the interventionists who went above and beyond to ensure a quality study.

References

Adams, M. (1990). *Beginning to read: Thinking and learning about print*. Urbana-Champaign, IL: University of Illinois at Urbana-Champaign.

Adams, M. J., Bereiter, C., Carruthers, I., Case, R., Hirshberg, J., McKeough, A., et al. (2000). *Open court reading*. Columbus, OH: SRA/McGraw-Hill.

Alexander, A. W., & Slinger-Constant, A.-M. (2004). Current status of treatments for dyslexia: Critical review. *Journal of Child Neurology, 19*, 744–758.

Apel, K. (2011). What is orthographic knowledge? *Language, Speech, and Hearing Services in Schools, 42*, 592–603.

ASHA Panel on Audiologic Assessment. (1997). Guidelines for audiologic screening. Retrieved from American Speech-Language-Hearing Association: http://www.asha.org/docs/html/GL1997-00199.html

Baddelely, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105*, 158–173.

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Science, 4*, 417–423.

Barlow, D. H., & Hayes, S. C. (1979). Alternating treatment design: One strategy for comparing the effects of two treatments in a single subject. *Journal of Applied Behavior Analysis, 12*, 199–210.

Barlow, D. H., & Hersen, M. (1984). *Single-case experimental designs strategies for studying behavior change*. Elmsford, NY: Pergamon Press Inc..

Bell, N. (1991). Gestalt imagery: A critical factor in language comprehension. *Annals of Dyslexia, 41*, 246–260.

Berninger, V. W., Abbott, R. D., Jones, J., Wolf, B., Gould, L., Anderson-Youngstrom, M., & Apel, K. (2006a). Early development of language by hand: Composing, reading, listening, and speaking connections; three letter-writing modes; and fast mapping in spelling. *Developmental Neuropsychology, 29*, 61–92.

Berninger, V., Abbott, R., Thomson, J., Wagner, R., Swanson, L., Wijsman, E., & Raskind, W. (2006b). Modeling phonological core deficits within a working memory architecture in children and adults with developmental dyslexia. *Scientific Studies of Reading, 10*, 165–198.

Berninger, V. W. (2001). Understanding the 'Lexia' in dyslexia: A multidisciplinary team approach to learning disabilities. *Annals of Dyslexia, 51*, 23–48.

Berninger, V., & Amtmann, D. (2003). Preventing written expression disabilities through early and continuing assessment and intervention for handwriting and/or spelling problems: Research into practice. In L. Swanson, K. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (pp. 345–363). New York: The Guildford Press.

Berninger, V., Lee, Y.-L., Abbott, R., & Breznitz, Z. (2013). Teaching children with dyslexia to spell in a reading-writer's workshop. *Annals of Dyslexia, 63*, 1–24.

Berninger, V., Nielsen, K., Abbott, R., Wijsman, E., & Raskind, W. (2008). Writing problems in developmental dyslexia: Under-recognized and under-treated. *Journal of School Psychology, 46*, 1–21.

Berninger, V., Vaughan, K., Abbott, R., & Brooks, A. (2000). Language-based spelling instruction: Teaching children to make multiple connections between spoken and written words. *Learning Disability Quarterly, 23*, 117–135.

Bhat, P., Rapport, M. K., & Griffin, C. (2000). A legal perspective on the use of specific reading methods for students with learning disabilities. *Learning Disability Quarterly, 23*, 283–297.

Birsh, J. (2006). What is multisensory structured language? *Perspectives, Fall*, 15–20.

Block, C. C., Parris, S. R., & Whiteley, C. S. (2008). CPMs: Akinesthetic comprehension strategy. *The Reading Teacher, 6*, 460–470.

Brossart, D., Vannest, K. J., Davis, J. L., & Patience, M. A. (2014). Incorporating nonoverlap indices with visual analysis for quantifying intervention effectiveness in single-case experimental designs. *Neuropsychological Rehabilitation, 24*, 464–491.

Campbell, M. L., Helf, S., & Cooke, N. L. (2008). Effects of adding multisensory components to a supplemental reading program on the decoding skills of treatment resisters. *Education and Treatment of Children, 31*, 267–295.

Caro, T. M., Roper, R., Young, M., & Dank, G. R. (1979). Inter-observer reliability. *Behaviour, 69*, 303–315.

Clark, D. B., & Uhry, J. K. (1995). *Dyslexia theory and practice of remedial instruction* (second ed.). Baltimore: York Press.

Constantinidou, F., Danos, M., Nelson, D., & Baker, S. (2011). Effects of modality presentation on working memory in school-age children: Evidence for the pictorial superiority hypothesis. *Child Neuropsychology, 17*, 173–196.
Cox, A. (1992). *Foundations for literacy structures and techniques for multisensory teaching of basic written english language skills*. Cambridge: Educators Publishing Service, Inc.
Crane, S. (1996). *Star wars miniature battles imperial entanglements*. West End Games.
Ehri, L., & McCormick, S. (1998). Phases of word learning: Implications for instruction with delayed and disabled readers. *Reading and Writing Quarterly, 14*, 135–163.
Ehri, L. C., & Wilce, L. S. (1985). Movement into reading: Is the first stage of printed word learning visual or phonetic? *Reading Research Quarterly, 20*, 163–179.
Ehri, L. C. (2014). Orthographic mapping in the acquistion of sight word reading, spelling memory, and vocabulary learning. *Scientific Studies of Reading, 18*, 5–21.
Fiske, K. E. (2008). Treatment integrity of school-based behavior analytic interventions: A review of research. *Behavior Analysis in Practice, 1*, 19–25.
Ferron, J., & Jones, P. K. (2006). Tests for the visual analysis of response-guided multiple-baseline data. *The Journal of Experimental Education, 75*, 66–81.
Foorman, B. R., Francis, D. J., Winikates, D., Mehta, P., Schatschneider, C., & Fletcher, J. M. (1997). Early interventions for children with reading disabilities. *Scientific Studies of Reading, 1*, 255–276.
Gabrieli, J. (2009). Dyslexia: A new synergy between education and cognitive neuroscience. *Science, 325*, 280–283.
Gibbs, S., Campbell, M., Helf, S., & Cooke, N. (2006). *Early reading tutor*. Columbus, OH: SRA/McGraw-Hill.
Gibson, E., Gibson, J., Pick, A., & Osser, H. (1962). A developmental study of the discrimination of letter-like forms. *Journal of Comparative and Physiological Psychology, 55*, 897–906.
Goldman, R., & Fristoe, M. (2000). *Goldman-Fristoe Test of Articulation 2*. Circle Pines, MN: Pearson.
Gray, S. (2005). Word learning by preschoolers with specific language impairment: Effect of phonological or semantic cues. *Journal of Speech, Language, and Hearing Research, 48*, 1452–1467.
Gray, S., Pittman, A., & Weinhold, J. (2014). Effect of phonotactic probability and neighborhood density of word-learning configuration by preschoolers with typical development and specific language impairment. *Journal of Speech, Language, and Hearing Research, 57*, 1011–1025.
Gupta, P., & Tisdale, J. (2009). Word learning, phonological short-term memory, phonotactic probability and long-term memory: Towards an integrated framework. *Philosophical Transactions of the Royal Society, 364*, 3755–3771.
Hallahan, D. P., & Mercer, C. D. (2007, November 19). Learning disabilities: Historical perspectives. Retrieved from National Research Center on Learning Disabilities: http://www.nrcld.org/resources/ldsummit/hallahan2.html
Henry, M., & Hook, P. (2006). Multisensory instruction: Then and now. *Perspectives, 32*, 9–11.
Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Council for Exceptional Children, 71*, 165–179.
Horner, R., Swaminathan, K., Sugai, G., & Smolkowski, K. (2012). Considerations for the systematic analysis and use of single-case research. *Education and Treatment of Children, 35*, 269–290.
Hosszu, G. (2012). *Heritage of scribes: The relation of Rovas scripts to Eurasian writing systems*. Budapest: Rovas Foundation, Hungary.
Hulme, C. (1981). The effects of manual tracing on memory in normal and retarded readers: Some implications for multi-sensory teaching. *Psychological Research, 43*, 179–191.
Hulme, C., Monk, S., & Ives, S. (1987). Some experimental studies of multi-sensory teaching: The effects of manual tracing on children's paired-associate learning. *British Journal of Developmental Psychology, 5*, 299–307.
Joshi, M. R., Dahlgren, M., & Boulware-Gooden, R. (2002). Teaching reading in an inner city school through a multisensory teaching approach. *Annals of Dyslexia, 52*, 229–242.
Kaufman, A. S., & Kaufman, N. l. (2004). *Kaufman assessment battery for children* (Second ed.). Circle Pines, MN: AGS.
Kershener, S., Flynn, S., Prince, M., Potter, S. C., Craft, L., & Alton, F. (2014). Using data to improve fidelity when implementing evidence-based programs. *Journal of Adolescent Health, 54*, S29–S36.
Liberman, I. Y., Shankweiler, D., Fisher, F. W., & Carter, B. (1974). Explicit syllable and phoneme segmentation in the young child. *Journal of Experimental Child Psychology, 18*, 201–212.
Lyon, R. G., Shaywitz, S. E., & Shaywitz, B. A. (2003). A definition of dyslexia. *Annals of Dyslexia, 53*, 1–14.
MacArthur, C., & Graham, S. (1987). Learning disabled students' composing under three methods of text production: Handwriting, word processing, and dictaiton. *The Journal of Special Education, 21*, 22–42.
Mayer, R. E., & Anderson, R. B. (1991). Animations need narrations: An experimental test of a dual-coding hypothesis. *Journal of Educational Psychology, 83*, 484–490.
McIntyre, C., & Pickering, J. (2001). *Clinical studies of multisensory structured language education*. Dallas: International Multisensory Structured Language Education Council.

Moats, L. (2006). How spelling supports reading. *American Educator, 29*, 12–43.
Moats, L., & Farrell, M. (2002). Multisensory instruction. In J. Birsh (Ed.), *Multisensory teaching of basic language skills* (pp. 1–18). Baltimore: Brookes Publishing Co..
National Institute for Literacy. (2008). *Developing early literacy: Report of the National Early Literacy Panel*. Jessup: Eds Pubs.
National Reading Panel. (2000). Teaching children to read: An evidence based assessment of scientific research literature on reading and its implications for reading instruction. Retrieved from National Institute of Child Health and Human Development: http://www.nichd.nih.gov/publications/nrp/report.htm
National Reading Panel. (2006). *Retrieved from Eunice Kennedy Shriver National Institute of Child Health and Human Development* .www.nichd.nih.gov/publications/nrp/pages/smallbook.aspx
Nicholson , P. (1982). The old English rune for s. *The Journal of English and Germanic Philology, 81*, 313–319.
Norton, E., & Wolf, M. (2012). Rapid automatized naming (RAN) and reading fluency: Implication for understanding and treatment of reading disabilities. *Annual Review of Psychology, 63*, 427–452.
Oakland, T., Black, J. L., Stanford, G., Nussbaum, N., & Balise, R. R. (1998). An evalation of the dyslexia training program: A multisensory method for promoting reading in students with reading disabilities. *Journal of Learning Disabilities, 31*, 140–147.
Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology, 45*, 255–287.
Parker, R., & Vannest, K. (2009). An improved effect size for single-case research: Nonoverlap of all pairs. *Behavior Therapy, 40*, 357–367.
Parker, R. I., Vannest, K. J., & Davis, J. L. (2011a). Effect size in single-case research: A review of nine nonoverlap techniques. *Behavior Modifications, 35*, 303–322.
Parker, R. I., Vannest, K. J., Davis, J. L., & Sauber, S. B. (2011b). Combining nonoverlap and trend for single-case research: Tau-U. *Behavior Therapy, 42*, 284–299.
Post, Y., & Carreker, S. (2002). Orthographic similarity and phonological transparency in spelling. *Reading and Writing: An Interdisciplinary Journal, 15*, 317–340.
Ramus, F. (2004). Neurobiology of dyslexia: A reinterpretation of the data. *Trends in Neuroscience, 27*, 720–726.
Richards, T. L., Aylward, E. H., Field, K. M., Grimme, A. C., Raskind, W., Richards, A., et al. (2006). Converging evidence for triple word form theory in children with dyslexia. *Developmental Neuropsychology, 30*, 547–589.
Ritchey, K. D., & Goeke, J. L. (2006). Orton-Gillingham and Orton-Gillingham-based reading instruction: A review of the literature. *Journal of Special Education, 40*, 171–183.
Rose, T. E., & Zirkel, P. (2007). Orton-Gillingham methodology for students with reading disabilities: 30 years of case law. *The Journal of Special Education, 41*, 171–185.
Sadoski, M., & Paivio, A. (2001). *Imagery and text*. Mahwah, NJ: Lawrence Erlbaum Associates.
Sadoski, M., & Paivio, A. (2013). A dual coding theoretical model of reading. In D. Alvermann, N. Unrau, & R. Ruddell (Eds.), *Theoretical models and processes of reading* (sixth ed., pp. 886–922). Newark, DE: International Reading Association.
Schoenwald, S. K., Garland, A. F., Chapman, J. E., Frazier, S. L., Sheidow, A. J., & Southam-Gerow, M. A. (2011). Toward the effective and efficient measurement of implementation fidelity. *Administration and Policy in Mental Health Services Research, 38*, 32–43.
Semel, E., Wiig, E. H., & Secord, W. A. (2003). *Clinical evaluation of language fundamentals* (Fourth ed.). Toronto: The Psychological Corporation.
Shaywitz, S. (1998). Dyslexia. *The New England Journal of Medicine, 338*, 307–312.
Shaywitz, S. E., Morris, R., & Shaywitz, B. A. (2008). The education of dyslexic children from childhood to young adulthood. *Annual Reveiw of Psychology, 59*, 451–475.
Snowling, M. J., & Hulme, C. (2011). Evidence-based interventions for reading and language difficulties: Creating a virtuous circle. *British Journal of Educational Psychology, 81*, 1–23.
Snowling, M. J., & Hulme, C. (2012). Children's reading impairments: From theory to practice. *Japanese Psychological Research, 55*, 186–202.
Tobia, V., & Marzocchi, G. M. (2014). Cognitive profiles of Italian children with developmental dyslexia. *Reading Research Quarterly, 49*, 437–452.
Torgesen, J. K., Alexander, A. W., Wagner, R. K., Roshotte, C. A., Voeller, K., & Conway, T. (2001). Intensive remedial instruction for children with severe reading disabilities: Immediate and long-term instructional approaches. *Journal of Learning Disabilites, 34*, 33–58.
Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (2012). *Test of word reading efficiency* (second ed.). Austin: Pro-Ed.
Treiman, R., Tincoff, R., Rodriquez, K., Mouzaki, A., & Francis, D. (1998). The foundations of literacy: Learning the sounds of letters. *Child Development, 69*, 1524–1540.

Uhry, J., & Shepherd, M. (1993). Segmentation/spelling instruction as part of a first-grade reading program: Effects on several measures of reading. *Reading Research Quarterly, 28*, 218–233.

Vannest, K.J., Parker, R.I., & Gonen, O. (2011). Single Case Research: web based calculators for SCR analysis. (Version 1.0) [Web-based application]. College Station, TX: Texas A&M University. Retrieved Friday 19th February 2016. Available from singlecaseresearch.org.

Vellutino, F., Fletcher, J., Snowling, M., & Scanlon, D. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology and Psychiatry, 45*, 2–40.

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

Wemyss, S. (1950). *The languages of the world ancient and modern*. Philadelphia: Stanley Wemyss.

What Works Clearinghouse. (2013). *What works clearinghouse reviewer training in single-case design standards: Reference and practice materials*(WWC Publication No. 40214.G32). Washington, DC: What Works Clearinghouse.

Woodcock, R. (2011). *Woodcock reading mastery tests* (Third ed.). Toronto: Pearson Education.

Zemlin, W. (1998). *Speech and hearing science anatomy and physiology, fourth edition*. Boston: Allyn and Bacon.