THE IMPACT OF EMBEDDING BEHAVIORAL SUPPORTS INTO READING INSTRUCTION FOR UPPER ELEMENTARY STUDENTS WITH READING DIFFICULTIES AND INATTENTION

ABSTRACT

The present multiprobe single-case design study investigated the impact of integrating evidence-based behavioral supports into a small-group reading intervention on student engagement and disruptive behavior for six upper elementary students with co-occurring reading difficulties and inattention. Visual analysis suggested a functional relation for all six students on engagement and five students on disruptive behavior. Furthermore, there was a statistically significant mean effect of the intervention on engagement (Tau-U = 0.76) and disruptive behavior (Tau-U = 0.81). At the student level, the intervention led to a statistically significant improvement in engagement and decline in disruptive behavior for three and five students, respectively. Study outcomes suggested that evidence-based behavior supports, integrated into an evidence-based commercially available reading curriculum during small-group instruction, are both feasible and can lead to improved student behavior. Limitations and implications for research and practice are discussed.

Garrett J. Roberts
Senaida Mehmedovic
UNIVERSITY OF DENVER

Brooke Cote
COLORADO SPRINGS SCHOOL DISTRICT 11

Jade Wexler
UNIVERSITY OF MARYLAND

Phil Strain
UNIVERSITY OF DENVER

Published online February 7, 2023

Elementary School Journal, volume 123, number 3, March 2023.
 https://doi.org/10.1086/723397

READING difficulties and attention-deficit hyperactivity disorder (ADHD) co-occur as much as 40% of the time (Willcutt & Pennington, 2000). Students with co-occurring reading difficulties and ADHD are likely to have more severe reading deficits than students with only reading difficulties and have more inattention and social impairments than students with only ADHD (Friedman et al., 2020; Mayes & Calhoun, 2007; Willcutt, 2012). For students with the most commonly displayed subtype of ADHD, inattention (e.g., lack focus, engage in off-task behavior; Willcutt, 2012; Willcutt et al., 2012), numerous research studies have shown that inattentive behaviors are negatively associated with reading comprehension outcomes and that these students are particularly vulnerable to not adequately responding to evidence-based reading instruction (e.g., Lin et al., 2013; Macdonald et al., 2020; Roberts et al., 2021). Due to the critical role of attention to yield adequate growth during reading instruction, interventions that support both reading and attention continue to be an area needing further investigation (Lemons et al., 2016; Macdonald et al., 2020; National Center on Intensive Intervention, 2013; Roberts et al., 2015, 2020).

Interventions for Students with Co-occurring Reading Difficulties and Inattention

Currently, it remains unclear how to provide students with co-occurring reading difficulties and inattention the supports they need to make adequate progress, as reading interventions alone do not always mitigate years of compounded reading failure and problem behaviors (Denton et al., 2020; Macdonald et al., 2020; Roberts et al., 2021). This is due, in part, to the fact that only limited reading intervention research has been conducted with students with co-occurring reading difficulties and inattention (Roberts et al., 2020; Stewart & Austin, 2020). Therefore, it is not surprising that there is a need for research that focuses on the integration of academic and behavioral interventions (e.g., Kuchle & Riley-Tillman, 2019; Lemons et al., 2016; Roberts et al., 2015, 2020).

Upper Elementary Reading Interventions

Despite the fact that research for students with co-occurring reading difficulties and inattention (or ADHD) is limited (Roberts et al., 2020; Stewart & Austin, 2020), there is an abundance of research supporting reading interventions for students with reading difficulties (without co-occurring inattention). According to the simple view of reading (SVR; Gough & Tunmer, 1986), reading comprehension is the product of decoding and linguistic comprehension. As students transition in the upper elementary grades from "learning to read" to "reading to learn," there is an increased demand on linguistic comprehension (e.g., vocabulary, background knowledge, inference making) to comprehend increasingly complex text (Chall, 1996). Thus, this shift in the upper elementary grades also represents a change in instruction from more malleable word and sentence reading tasks to more complex and difficult to remediate reading comprehension tasks (Chall, 1996; Compton et al., 2008).

For students in grades 3–5, it is important for reading instruction to address both decoding and linguistic comprehension skills. Therefore, it is not surprising that

multicomponent interventions, which include both decoding and comprehension skills (National Reading Panel, 2000), have been found to be effective for upper elementary students (e.g., Donegan & Wanzek, 2021; Scammacca et al., 2015). In addition, there is widespread agreement that the most efficacious method to deliver multicomponent interventions is through supplemental explicit and systematic reading instruction with an evidence-based reading curriculum delivered in a small-group setting (e.g., National Center on Intensive Intervention, 2013; Vaughn et al., 2010).

Interventions to Support Student Engagement

The more engaged students are, the more likely they will make progress in reading (Guthrie et al., 2007; Kim et al., 2017). Whereas many children can be easily engaged, other children, such as those with or at risk for inattention, require systematic supports to promote engagement. Based on numerous reviews (Collier-Meek et al., 2019; DuPaul et al., 2011; Harrison et al., 2019; Sayeski & Brown, 2011; Simonsen et al., 2015), antecedent- (i.e., a manipulation of events before the behavior) and consequence-based (i.e., a manipulation of events that occur after the behavior) strategies have been found to be highly effective at improving student engagement and reducing challenging behaviors in classroom settings.

Highly effective antecedent-based strategies that should be considered for all classroom management systems include establishing and explicitly teaching group expectations, providing students opportunities to practice the expectations, and frequently reviewing expectations (Myers et al., 2017; Simonsen et al., 2015). When children understand the reading lesson expectations, they also develop a clearer understanding of the expectations for their participation. This leads to an increase in child engagement and independence, and a decrease in challenging behavior (Myers et al., 2017).

When students are reinforced for meeting expectations (i.e., consequence-based strategy), they are more likely to adhere to classroom expectations. Highly effective methods to reinforce appropriate behaviors include using differential reinforcement (i.e., appropriate behaviors are reinforced with inappropriate behaviors ignored), behavior-specific praise (e.g., providing positive descriptive feedback statements to children who are engaged and meeting expectations), and token economies (e.g., points are later exchanged for a reward; DuPaul et al., 2011; Harrison et al., 2019; Simonsen et al., 2015). When token economies are in place, rewards can be earned based on an individual student's behavior (i.e., independent group contingent), a single student's behavior from a group (i.e., dependent group contingent), or all the students' behavior in the group (i.e., interdependent group contingencies; Simonsen et al., 2008).

Based on literature reviews by Roberts and colleagues (2020) and Stewart and Austin (2020), a limited number of group and single-case design (SCD) studies have delivered a reading intervention to upper elementary students with co-occurring reading difficulties and ADHD (or inattention-only). Across the upper elementary studies identified by Roberts and colleagues (2020) and Stewart and Austin (2020), findings point to reading instruction aligning to SVR (i.e., decoding, comprehension) leading to improvements in reading outcomes. However, neither review identified a study that tested the impact of an embedded behavioral support into a small-group reading setting for elementary students with co-occurring reading difficulties and ADHD (or inattention-only).

Even though the integration of behavioral supports into reading instruction for upper elementary students with co-occurring reading difficulties and ADHD (or inattention-only) is limited, several studies have integrated behavior support strategies into the general education setting (e.g., Kamps et al., 2015; Sutherland et al., 2020; Wills et al., 2018) and, for students with emotional and behavioral disorders (EBD), into small-group reading settings (see McKenna et al., 2017). For example, in the general education setting, Class-Wide Function-related Intervention Teams (CW-FIT; Kamps et al., 2015; Wills et al., 2009, 2018) teach and review class expectations, implement token economies (using points) with group contingencies, and deliver rewards for meeting predetermined point goals. CW-FIT has been shown to have positive effects on behavior and has been replicated in large-scale, high-quality randomized controlled trials (RCTs; e.g., Wills et al., 2018).

In addition, McKenna and colleagues (2017) conducted a review of SCD reading interventions for students with EBD and identified two reading intervention with behavior support studies that (*a*) embedded behavior supports into a small-group reading instruction for elementary students and (*b*) measured a behavior outcome. Both studies (i.e., Barton-Arwood et al., 2005; Lane et al., 2001) tested the impact of a combined reading intervention with behavior support. Lane and colleagues (2001) implemented an interdependent group contingency token economy embedded into reading instruction for grade 1 students. Barton-Arwood and colleagues (2005) implemented an independent group contingency token economy embedded into reading instruction with grade 3 students. Both found an overall improvement in reading and behavior outcomes, with reading and behavior outcome variability present at the student level.

Study Purpose and Research Questions

Students who display co-occurring reading difficulties and inattention are vulnerable to inadequate response to evidence-based reading instruction (Denton et al., 2020; Macdonald et al., 2020; Roberts et al., 2021). This has led researchers to call for reading intervention research on how best to integrate supports that mitigate behaviors that interfere with access to instruction (e.g., low engagement, disruptive behavior; Burns et al., 2012; Roberts et al., 2021). To date, questions on the extent to which such a combined reading and behavior support intervention would lead to improved reading and behavior outcomes have been inadequately addressed. Even though such an approach holds promise to improve reading and behavior, ultimately, meeting the needs of these students with reading difficulties and inattention requires a better understanding of the extent to which a combined evidence-based reading and behavior intervention is able to improve student behavior, relative to an evidence-based reading intervention without behavior support (e.g., Cook et al., 2012; Roberts et al., 2021).

Therefore, to address the reading and inattention needs of these students concurrently, we aimed to integrate a set of behavioral supports into evidence-based reading instruction for upper elementary students with co-occurring reading difficulties and inattention. The behavior supports (more fully discussed in the Method section) were designed to be integrated into an evidence-based reading curriculum to improve student engagement through antecedent- and consequence-based strategies. To

identify an evidence-based reading curriculum, we reviewed the *What Works Clearinghouse* and recently published research for upper elementary students with reading difficulties. Following this review, we chose to integrate the behavior supports into the evidence-based curriculum Voyager Passport (Voyager Sopris Learning, 2008), as RCT studies have shown Voyager Passport to be an efficacious intervention on reading comprehension outcomes (Wanzek et al., 2016, 2017). In addition, Voyager Passport is aligned to SVR and addresses the essential components of reading instruction—decoding, background knowledge, vocabulary, and comprehension skills—through explicit and systematic instruction (National Reading Panel, 2000). This curriculum is more thoroughly described in the baseline phase section. Through integrating antecedent- and consequence-based behavior supports into Voyager Passport, we sought to answer the following research question: What are the effects of integrating behavior supports into a reading intervention on student engagement and disruptive behavior relative to a reading intervention without behavior supports for upper elementary with co-occurring reading difficulties and inattention?

Method

Setting

This study was conducted at two urban elementary schools in the Rocky Mountain region of the United States. Schools were chosen based on previously developed partnerships with the first author. The first elementary school had approximately 600 students, with 56% White, non-Hispanic; 22% Hispanic; 11% Black, non-Hispanic; and 11% either multiple race, Asian, Native American, or Native Alaskan. Of these students, 12% were English-language learners, 36% qualified for free or reduced lunch, and 15% received special education services. The second elementary school had approximately 430 students, with 14% White, non-Hispanic; 50% Hispanic; 28% Black, non-Hispanic; and 8% either multiple race, Asian, Native American, or Native Alaskan. In addition, 47% of the students were English-language learners, 77% qualified for free or reduced lunch, and 11% received special education services.

Screening Procedures

To identify upper elementary student participants with reading difficulties and inattention and in consultations with school principals, grade 4 teachers at School 1 nominated 10 students and grade 5 teachers at School 2 nominated 5 students. A double-gating procedure was used to verify the presence of reading difficulties and inattentive behavior of nominated students. This procedure required a standard score equal to or less than 85 on a standardized reading fluency and comprehension measure (i.e., Test of Silent Reading Efficiency and Comprehension [TOSREC]; Wagner et al., 2010) and a greater than or equal to 115 standard score on a standardized teacher survey behavior measure (i.e., Behavior Assessment Scale for Children, Third Edition, Teacher Rating Scale [BASC-3 TRS]; Reynolds & Kamphaus, 2015) on the inattention subscale (i.e., high inattention). Each measure is further described in the Measures section of this article. To support the school's scheduling of student interventions, the intervention was delivered to all teacher-nominated students; however, for the purpose of this study,

we only present data for students displaying reading and inattentive behaviors. Pseudonyms were used for all students in this study.

Participants

At School 1, six nominated grade 4 students met the inclusion criteria. Four nominated students did not meet the inclusion criteria due to not meeting the reading criterion ($n = 1$), the attention problems criterion ($n = 1$), or both the reading and attention criteria ($n = 2$). The 10 students were divided equally into two groups of 5 for the intervention based on the Voyager Passport word reading placement test (Voyager Sopris Learning, 2008). At School 1, Groups 1 and 2 had two and four students, respectively, qualify for the study. In Group 1, Nia was present for the entire baseline and three intervention sessions prior to withdrawing from the study due to moving schools. At School 2, Group 3, one nominated student moved to a different school prior to the baseline phase. The remaining four nominated students had reading difficulties, but only one student also had inattentive behaviors. Therefore, this group only had one student qualify for the study. Table 1 presents the demographic information for the participating students.

Interventionists

Three female graduate students served as interventionists. The interventionist for Group 1 was a doctoral student in curriculum and instruction and was previously a high school English language arts (ELA) teacher. The interventionist for Group 2 was in a licensed early childhood special education program. The final interventionist for Group 3 was a doctoral student in curriculum and instruction, who previously served as an elementary classroom teacher and reading coach.

Experimental Design

The study used a concurrent multiprobe design to determine the impact of the reading with behavior supports intervention condition on engagement and disruptive behavior. Based on the *What Works Clearinghouse SCD Guidelines* (Institute of Education Sciences [IES], 2020), the aim was for (*a*) each student and phase to have five probes per phase, (*b*) probes for the first three consecutive initial baseline and intervention sessions, and (*c*) probes for the last three consecutive sessions just prior to the intervention. In addition, the aim was for each case not receiving an intervention to

Table 1. Student Demographics

Student	Grade	Gender	LEP Status	Race/Ethnicity	SPED Status
Aliah	4	Female	No	Black	No
Nia	4	Female	No	Black	No
Sofia	4	Female	No	Black	No
Kiyana	4	Female	No	Black	No
Isabella	4	Female	No	Hispanic	No
Emma	4	Female	No	White	No
Mateo	5	Male	Yes	Hispanic	No

Note.—LEP = limited English proficiency; SPED = special education.

have a probe point where a different case first receives the intervention. At times, these aims were not met due to across-school schedule variations (e.g., Teacher Grading Day at School 2 but not at School 1) and student absences. In these instances, the next possible session included a probe.

Analysis

The analysis was based on comparing direct measures of engagement and disruptive behavior in the intervention with the baseline phase. Student training sessions were not included in the analysis. Outcomes from this study were based on visual analysis (IES, 2017) and supplemented with the Tau-U (Parker et al., 2011) effect size. To determine a functional relation with visual analysis, the *What Works Clearinghouse Standards Handbook* (IES, 2017) recommends using within- and across-phase characteristics. Within-phase characteristics include the level (i.e., mean score), trend line (i.e., slope of the best-fitting straight line), and variability of data around the trend line (i.e., range or standard deviation). Across-phase characteristics included immediacy of effect after a phase change (i.e., difference in level between last three data points in the baseline phase to the first three data points in the intervention phase) and the extent to which data overlapped across phases. For each student, the level, trend, and variability were assessed to identify patterns within each phase and compared across phases. Immediacy of effect and overlap will then supplement within phase data characteristics to establish if a functional relation is present.

The Tau-U (Parker et al., 2011) and baseline-corrected Tau effect sizes (Tarlow, 2017) were used to measure data overlapped across phases (i.e., baseline, intervention). Tau-U (Parker et al., 2011) is a nonparametric technique that compares each data point in the baseline phase with each data point in the intervention phase and adjusts for trends in data (Vannest & Ninci, 2015). This effect size was chosen for three reasons: (*a*) it is strong performing relative to other SCD effect sizes (Brossart et al., 2014; Parker et al., 2011; Tarlow, 2017; Vannest & Ninci, 2015), (*b*) it is well suited for small data sets (Parker et al., 2011; Vannest & Ninci, 2015), and (*c*) it is widely used in systematic reviews with SCD research (e.g., Harrison et al., 2019; Stewart & Austin, 2020). Regarding this last point, the fact that Tau-U is commonly used in meta-analyses allows for a mean effect size benchmark to be obtained from a highly relevant meta-analysis. In the case of the present study, Harrison and colleagues (2019) conducted a meta-analysis on classroom-based SCD interventions for students with ADHD and found the mean behavior intervention effect size to be 0.67. Therefore, for the purposes of this study, intervention effect sizes were compared with the 0.67 effect size benchmark, as the 0.67 effect size is more contextually relevant than the commonly used Tau-U effect size benchmarks of small (0.20 or less), moderate (0.21–0.59), and large (0.60 or greater; Vannest and Ninci, 2015).

To calculate within case and weighted overall average Tau-U effect sizes, raw data were entered into the singlecaseresearch.org calculator (Vannest et al., 2016). For within-case Tau-U effect sizes, observation data (i.e., engagement, disruptive behavior) were entered by student and phase and compared across phases for each student. For each case, a rank-order correlation between outcomes and time within each baseline phase was also calculated to identify monotonic trends (i.e., upward or downward linear, curvilinear, or mixed pattern trends) in baseline data. For cases with statistically

significant ($p < .05$) monotonic trends in baseline data, we report a baseline corrected Tau effect size (Tarlow, 2017; as compared with a non-baseline corrected Tau effect size) via a Theil–Sen regression (Sen, 1968; Theil, 1950). The study weighted average included Tau-U effect sizes or, in cases with statistically significant baseline rank-order correlations, a baseline corrected Tau effect size. For the presented data, positive effect sizes represent an increase in engagement and a decrease in disruptive behavior.

Measures

Table 2 presents the student pretest and posttest reading and behavior scores.

Pretest and posttest reading measures. The reading measures included word reading, reading fluency, and reading comprehension. Prior to the delivery of the reading measures, assessment team members participated in a 1-hour training led by the first author or a research team member with expertise in their respective assessment. At the conclusion of the training, all assessment team members were able to deliver the measure with 100% reliability with the training lead. Due to posttesting time constraints, the Gates-MacGinitie Reading Test (GMRT) reading comprehension subtest (MacGinitie et al., 2000) was delivered at pretest only. All other reading measures were delivered at pretest and posttest.

Word reading. The Test of Word Reading Efficiency, Second Edition, Sight Word Efficiency (TOWRE-2 SWE; Torgesen et al., 2012) is a 45-second, individually administered measure of word reading. This measure has students read a list of increasingly difficult words. The test-retest reliability has a coefficient of 0.87.

Reading fluency. The easyCBM passage reading fluency measure is a 1-minute timed reading passage assessment with benchmark and progress monitoring forms (Alonzo et al., 2006). The test-retest reliabilities are 0.86–0.96.

Reading comprehension. Two reading comprehension measures were delivered. The TOSREC (Wagner et al., 2010) is a group-administered test of reading fluency and comprehension. Students are given 3 minutes to read and verify the accuracy (circling "yes" or "no") of as many sentences as possible. The alternate-form reliability coefficients exceed 0.85 across all grades and forms. The GMRT (MacGinitie et al., 2000) is a timed, group-administered assessment measuring reading comprehension and targeted inference making, summarization, literal understanding, and vocabulary. The test-retest reliability ranges from 0.88 to 0.92.

Pretest behavior measure. The BASC-3 TRS (Reynolds & Kamphaus, 2015) contains 156 questions and takes approximately 10–20 minutes to complete. The median Cronbach's alpha for the BASC-3 TRS clinical and adaptive scales for children ages 8–11 is 0.90, and the reliability coefficients range from 0.86 to 0.94. This measure was completed by each student's ELA teacher at pretest only, as we did not hypothesize that an intervention of this duration would lead to generalized improved student behavior in the general education setting.

Direct measures of engagement and disruptive behavior. Engagement and disruptive behavior were coded from video-recorded sessions. Similar to recent SCD research during academic instruction (e.g., Harris et al., 2005; Wills et al., 2018), the operational definition of engagement and disruptive behavior were as follows. Engagement included (*a*) having eyes oriented toward a given assignment or the teacher during instruction, directions, or on-topic comments or questions; (*b*) working on

Table 2. Pretest and Posttest Standard Scores[a]

	TOSREC		TOWRE-2 SWE		easyCBM PRF[b,c]		GMRT NCE	BASC-3 ADHD Probability	BASC-3 Attention Problems	BASC-3 Externalizing	BASC-3 Internalizing
Student	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest	Pretest	Pretest	Pretest	Pretest	Pretest
Aliah	76	80	91	101	106	144	70	117	126	108	90
Nia	85	N/A	91	N/A	90	N/A	74	127	129	135	120
Sofia	81	93	85	86	107	119	81	121	129	117	87
Kiyana	81	77	89	88	97	101	76	144	142	136	123
Isabella	69	89	72	76	79	106	70	120	126	120	96
Emma	58	91	70	67	82	113	77	109	120	91	120
Mateo	61	64	71	72	78	101	78	115	120	105	106
Mean (*SD*)[b]	71.00 (9.98)	82.33 (10.98)	79.67 (9.71)	81.67 (12.44)	91.50 (13.49)	114.00 (16.30)	75.14 (4.10)	121.86 (11.23)	127.43 (7.44)	116.00 (16.27)	106.00 (15.26)

Note.—TOSREC = Test of Silent Reading Efficiency and Comprehension (Wagner et al., 2010); TOWRE-2 SWE = Test of Word Reading Efficiency, Second Edition, Sight Word Efficiency (Torgesen et al., 2012); PRF = passage reading fluency; GMRT = Gates-MacGinitie Reading Test reading comprehension subtest (MacGinitie et al., 2000); NCE = normal curve equivalent; BASC-3 = Behavior Assessment Scale for Children, Third Edition (Reynolds & Kamphaus, 2015); ADHD = attention-deficit hyperactivity disorder; N/A = not available (moved schools).

[a] All scores are reported in standard scores unless otherwise noted.

[b] easyCBM is reported in correct words per minute.

[c] Nia is not included in the mean or *SD* due to withdrawal from study.

an assigned task; (*c*) using the materials appropriately (e.g., writing on a paper with a pencil, opening a binder, having the book opened to the correct page); and (*d*) interacting with teachers or peers about academic topics relevant to completing assignments. Disruptive behavior included (*a*) arguing or name-calling with peers or adults, (*b*) gestures or movements with the potential to disrupt other students' learning, or (*c*) any other verbal or gestural action that interrupts the learning activity. All sessions were video recorded. These recorded sessions were used to code behaviors with a 15-second momentary time sampling recording system. Coders were blind to the study design and the phase of each session being coded. On the coding sheet, 15-second intervals were scored with a 1 or 0 if the student met the given behavior criteria or did not, respectively. To calculate the percentage of engagement, the sum of the intervals with engagement was divided by the total number of intervals and multiplied by 100.

Social validity. The student social validity survey used a 7-point Likert rating scale ranging from *strongly disagree* (score = 1) to *strongly agree* (score = 7) to measure student views on the perceived impact and the acceptability of the intervention. The student survey measured responses to the following statements: (*a*) "improving my reading skills is important to me," (*b*) "I enjoy coming to this reading group," (*c*) "this reading group is helping me improve my reading skills," and (*d*) "I will be able to use the strategies I am learning in this reading group when I read on my own." The interventionist social validity survey used the same 7-point Likert rating scale to measure the interventionist views on the perceived impact, the acceptability, and the feasibility of the intervention. The interventionist survey measured responses to the following statements: (*a*) "the intervention phase was easy to implement," (*b*) "the intervention training prepared me to fully implement the intervention," (*c*) "the intervention was effective at improving reading comprehension," and (*d*) "the intervention was effective at improving student behavior." Both surveys were completed at the conclusion of the study.

Procedures

Intervention Overview

The intervention had two phases: a reading-only baseline phase and reading with behavior support intervention phase. There were three intervention student training sessions prior to the implementation of the intervention. In the following sections, the professional development, the baseline and intervention phases, and the student training sessions are fully described. During all phases, weekly meetings to review fidelity to the respective phase were conducted by a member of the research team. All intervention groups were conducted in the period following lunch and recess in a non–general education classroom setting.

Professional Development

All interventionists participated in 4- and 2-hour professional development sessions prior to the baseline and intervention phases, respectively. The baseline phase professional development included information on how to deliver each component

of the Voyager Passport (Voyager Sopris Learning, 2008). The intervention phase professional development included information on the four behavior support components (more thoroughly described in the Intervention Phase section): (*a*) teaching and reviewing group expectations, (*b*) using behavior-specific praise and precorrections (i.e., a reminder of an expected behavior before the behavior should occur), (*c*) implementing a token economy (i.e., awarding points contingent on appropriate behavior), and (*d*) providing point goals with a reward for obtaining the point goal. Both professional development sessions included modeling, guided practice, and independent demonstration of mastery by each interventionist by delivering one lesson to a member of the intervention team with 90% or greater fidelity.

Baseline Phase

The baseline phase utilized Voyager Passport. This curriculum is divided into 12 adventures, with 10 lessons per adventure. Lessons are designed to be 30 minutes and delivered in a small-group setting. Each lesson is semiscripted and designed to be sequential to build decoding skills, background knowledge, and comprehension strategies (e.g., cause and effect). Lessons also integrate active learning strategies such as discussing questions, utilizing graphic organizers, and increasing students' opportunities to engage in the lesson through oral and written practice.

Lessons 1–4 and 6–9 of each adventure have two parts: word study and connected text. During word study, there is a 2-minute warm-up (e.g., vowel combinations, sight words) and advanced word study activities (e.g., prefixes, suffixes, compound words, synonyms). During connected text, students preview the text, receive vocabulary instruction, engage in repeated readings, use comprehension strategies, and are provided multiple opportunities to check for understanding. The connected text activities, vocabulary, and comprehension strategies were explicitly taught and practiced. In the curriculum, students also discuss questions, utilize graphic organizers, and make connections both orally and through written practice. Every adventure has checkpoints at lessons 5 and 10 to progress monitor reading fluency and comprehension (see Wanzek et al., 2016, for a more detailed description of Voyager Passport).

Student Training Session Prior to the Intervention Phase

Prior to the intervention phase, there were three training sessions to teach the group expectations and introduce students to the token economy. Prior to teaching each expectation, the interventionist completed the following steps. First, upon student entry to the group, the interventionist started a 3-minute timer. Every time that the timer sounded during the lesson, the instructor delivered a point (as part of the token economy) and provided students with behavior-specific praise for appropriate behavior. This was followed by the timer being reset for 3 minutes and restarted. Second, the instructor reminded the students of the group expectations. Third, the interventionist and students agreed on a point goal for the lesson. The agreed-upon point goal was for all students in the group, although rewards were earned as an independent group contingency (i.e., each student independently earned a reward for meeting the point goal). After the point goal was established, the interventionists used 10 minutes to teach and review the group expectations. To teach each group expectation (i.e., raise your hand

to get the teacher's attention, follow directions the first time, ignore peers' inappropriate behaviors), the interventionist completed the following steps: (*a*) posted the new group expectation in a visible location, (*b*) reviewed previously taught expectations, (*c*) discussed the importance of the expectation (e.g., to help us learn), and (*d*) provided opportunities to practice the expectation. Following each of the three 10-minute training sessions, students completed the reading lesson with the reading intervention in place. When the reading lesson ended, points were tallied. Finally, the students who met their point goal would engage in a 3-minute game with the instructor (e.g., Uno, Go Fish) until the session ended. Students not meeting their goal continued with independent reading work. Interventionists were encouraged to use behavior-specific praise and precorrections throughout the lesson.

Intervention Phase

The intervention phase session began each lesson with the 3-minute timer starting to signal that points were able to be earned at 3-minute intervals (i.e., when the timer sounded) for appropriate behavior. Next, the interventionist reviewed the expectations with the students for 2 minutes by each student stating the definition of one expectation and modeling the appropriate behavior of that expectation. After the expectations were reviewed, a group point goal was established, and the reading lesson began. Throughout the lesson, points, paired with behavior-specific praise, were delivered to individual students for appropriate behavior when the timer sounded every 3 minutes. At the conclusion of every lesson, students who met their point goal engaged in the 3-minute game with the instructor (e.g., Uno). Students not meeting their point goal were instructed to read independently.

Procedural Fidelity

Using direct observation methods (Lane et al., 2004) from video-recorded sessions, the procedural fidelity protocol (Ledford & Gast, 2018) measured the alignment of the teacher practices to the components of the behavior intervention. The behavior support procedural fidelity checklist included the following components: (*a*) class expectations were posted and reviewed, (*b*) points were in sight of students, (*c*) point goals were discussed and posted, (*d*) timer was used with 3-minute intervals, (*e*) points were delivered when the timer sounded, and (*f*) points were calculated and desired activity was provided (when applicable) at the end of the session. Components were scored as 0 (not present), 1 (implemented with partial fidelity; e.g., points were delivered at some [but not all] of the 3-minute intervals, points goals were posted and not discussed), or 2 (implemented with full fidelity). Procedural fidelity was conducted on two randomly selected sessions per instructor per phase and one of the three training sessions (i.e., 33% of the training sessions). Procedural fidelity was collected for an average of 26% (SD = 2%, range: 25%–29%) and 28% (SD = 5%, range: 25%–33%) of the sessions per instructor in the baseline and intervention phases, respectively. For all instructors, the baseline phase procedural fidelity was 0% and 100% in all training phase sessions. The procedural fidelity in the intervention phase, across all instructors, averaged 93% (SD = 8%, range: 83%–100%).

Interobserver Agreement

Graduate students were trained to reliably code engagement, disruptive behavior, and procedural fidelity. An engagement and disruptive behavior coding team (blind to study purpose and phases) and a procedural fidelity coding team participated in separate 1-hour trainings. Both trainings defined relevant variables, identified examples and nonexamples, reviewed filling out relevant forms, and independently reached an interobserver agreement (IOA) of 85% or above with the lead IOA trainer. Interval-by-interval comparisons were used to calculate IOA by summing the number of intervals with agreements, dividing the sum by the total number of intervals (i.e., agreements plus disagreements), and converting the result to a percentage. All engagement and procedural fidelity IOA data were collected on one session per phase. All coding discrepancies were discussed between coders following each IOA coding, although the presented IOA data are based on the original coding.

Following the *What Works Clearinghouse SCD Guidelines* (IES, 2020) guidelines, a minimum of 20% of the sessions were coded per phase. For both engagement and disruptive behavior, two sessions were coded per phase per student. The average percentage of sessions coded for all students averaged 34% ($SD = 15\%$, range: 25%–67%) and 38% ($SD = 20\%$, range: 22%–67%) for the baseline and intervention phases, respectively. For the training phase, IOA was conducted for one session per student. The training phase IOA across all students was 88% ($SD = 4\%$, range: 83%–94%) and 88% ($SD = 2\%$, range: 84%–90%) for engagement and disruptive behavior, respectively. Tables 3 and 4 present the IOA per student per baseline and intervention phase for engagement and disruptive behavior, respectively. Procedural fidelity IOA was collected on one session per phase per instructor (i.e., 50% of baseline and intervention sessions and 100% of training sessions). Across all interventionists, the procedural fidelity IOA for the baseline was 100%. The training and intervention phase procedural fidelity IOA averaged 95% ($SD = 8\%$, range: 83%–100%) and 95% ($SD = 8\%$, range: 83%–100%), respectively.

Results

Table 2 presents the student pretest and posttest reading and behavior scores. Tables 3 and 4 presents the means, standard deviations, IOA, Tau-U, and statistical significance for engagement and disruptive behavior, respectively. The study weighted Tau-U averages for engagement and disruptive behavior equaled 0.76 ($p < .0001$) and 0.81 ($p < .0001$), respectively. These weighted Tau-U averages do not include Nia, as she withdrew from the study. Figures 1 and 2 visually display the percentage of intervals of student engagement and disruptive behavior, respectively. The following sections provide the visual analysis and effect sizes for engagement followed by disruptive behavior. At the conclusion of the Results section, the social validity data are presented.

Engagement

Group 1 had two students, Aliah and Nia. For Aliah, the baseline condition had a slight upward trend with high variability. The intervention phase had a higher level

Table 3. Engagement Data

Student	Phase	*M (SD)*	Range	IOA *M (SD)*	Tau-U	*p* Value
Aliah	Baseline	69 (11.2)	46–79	88 (2.8)		
	Intervention	88 (4.5)	83–93	96 (3.5)	1.00	.001
Nia[a]	Baseline	55 (10.4)	47–67	86 (.0)		
	Intervention	66 (21.7)	43–86	87 (4.9)	.33	.51
Sofia	Baseline	53 (22.7)	26–81	83 (3.5)		
	Intervention	68 (2.6)	66–71	85 (5.7)	.94[b]	.03
Kiyana	Baseline	48 (15.3)	24–71	80 (7.1)		
	Intervention	58 (12.3)	47–78	85 (4.9)	.38	.21
Isabella	Baseline	64 (18.6)	40–82	82 (9.9)		
	Intervention	80 (8.7)	69–94	84 (3.5)	.52	.11
Emma	Baseline	53 (11.9)	36–69	82 (9.9)		
	Intervention	74 (13.5)	53–91	86 (4.2)	1.14[b]	.0001
Mateo	Baseline	78 (8.1)	66–89	87 (2.8)		
	Intervention	87 (4.4)	85–92	90 (3.5)	.57	.09
				Mean effect	.76	<.0001

Note.—IOA = interobserver agreement.

[a] Participant withdrew from study; data are not included in study effect size or *p* value.

[b] This is the baseline corrected Tau.

and low variability. Aliah had a statistically significant Tau-U effect size of 1.00 ($p < .001$). For Nia, it was difficult to draw conclusions from the visual analysis due to frequent absences and an early withdrawal from the study. Nia had a Tau-U effect size of 0.33 ($p > .05$).

Group 2 had four students, Sofia, Kiyana, Isabella, and Emma. All students had a downward trend in the baseline phase. During the intervention phase, all students demonstrated a higher level and less variability than their baseline phase. Sophia and Emma had statistically significant baseline-corrected Tau effect sizes of 0.94

Table 4. Disruptive Behavior Data

Student	Phase	*M (SD)*	Range	IOA *M (SD)*	Tau-U	*p* Value
Aliah	Baseline	38 (7.4)	28–44	87 (.0)		
	Intervention	22 (8.5)	8–33	85 (.7)	.91	.003
Nia[a]	Baseline	48 (30)	14–71	91 (2.8)		
	Intervention	28 (11.8)	15–38	84 (7.1)	.33	.51
Sofia	Baseline	31 (26.7)	14–84	90 (5.7)		
	Intervention	23 (12.3)	14–30	79 (2.8)	.22	.61
Kiyana	Baseline	54 (13.0)	38–73	89 (3.5)		
	Intervention	29 (18.1)	4–59	84 (1.4)	.78	.009
Isabella	Baseline	35 (15.4)	20–53	83 (3.5)		
	Intervention	8 (4.7)	3–17	88 (5.7)	1.00	.003
Emma	Baseline	38 (21.8)	31–75	88 (7.1)		
	Intervention	12 (5.8)	3–21	90 (.7)	.92	.002
Mateo	Baseline	42 (11.8)	26–58	88 (1.4)		
	Intervention	19 (10.1)	12–32	88 (4.2)	.85	.008
				Mean	.81	<.0001

Note.—IOA = interobserver agreement.

[a] Participant withdrew from study; data are not included in study effect size or *p* value.

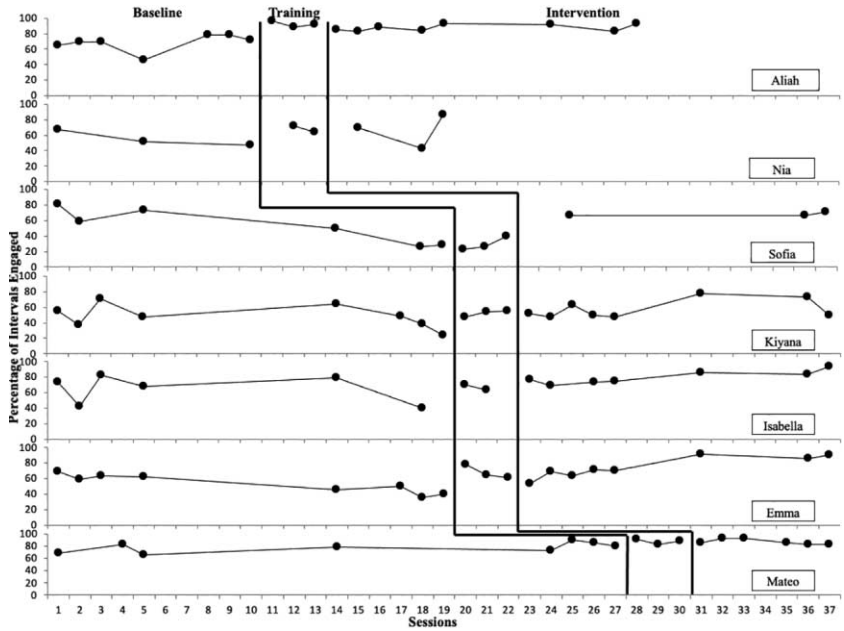

Figure 1. Percentage of intervals with engagement.

($p < .05$) and 1.14 ($p < .001$), respectively. Kiyana and Isabella had Tau-U effect sizes of 0.38 ($p > .05$) and 0.52 ($p > .05$), respectively.

Group 3 had one student, Mateo. For Mateo, the baseline condition had a slight downward trend with variability. The intervention phase had a higher level and less variability. Mateo had a Tau-U effect size of 0.57 ($p > .05$).

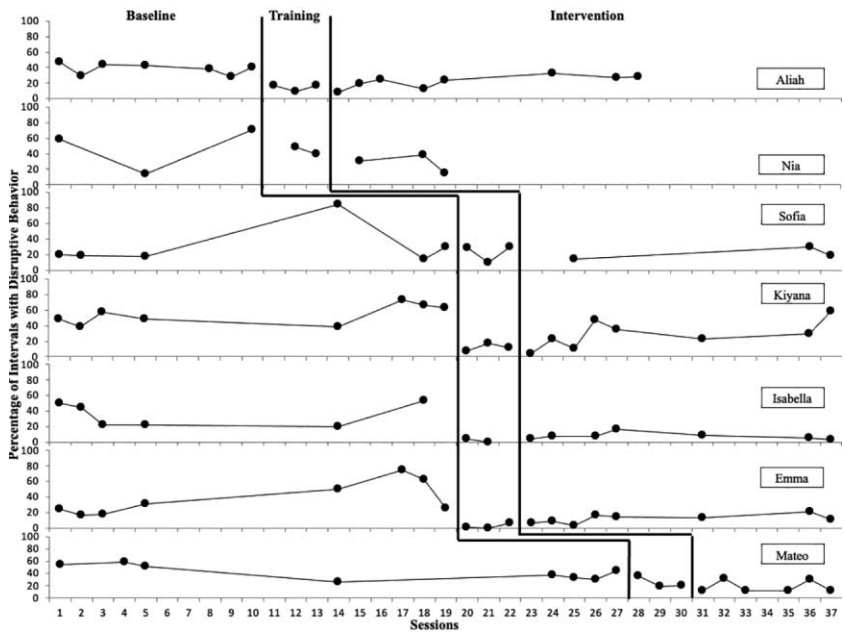

Figure 2. Percentage of intervals with disruptive behavior each session by student.

Disruptive Behavior

In Group 1, Aliah had a baseline condition with a slight downward trend and variability. The intervention phase had lower levels of disruptive behavior, although the variability was similar to the baseline phase. Again, for Nia it was difficult to draw conclusions from the visual analysis due to frequent absences and an early withdrawal from the study. Aliah had a Tau-U effect size of 0.91 ($p < .01$). Nia's Tau-U effect size was 0.33 ($p > .05$).

In Group 2, all students had an upward or flat trend in baseline with high variability. Following the introduction of the intervention phase, all students had a lower level of disruptive behavior, with Sofia, Isabella, and Emma also having reduced variability. Kiyana, Isabella, and Emma had a statistically significant Tau-U effect size of 0.78 ($p < .01$), 1.00 ($p < .01$), and 0.92 ($p < .001$), respectively. Sofia had a Tau-U effect size of 0.22 ($p > .05$).

In Group 3, Mateo had a baseline condition with a slight downward trend and some variability. The intervention phase had a flat trend, with a lower level and less variability. Mateo had a statistically significant Tau-U effect size of 0.85 ($p < .01$).

Social Validity

As previously stated, the social validity measures used a 7-point Likert-type scale. Across all items and students, student social validity averaged 6.71 ($SD = 0.55$, range: 5–7), with all items having an average (across students) of 6.67 or greater. Across all items and interventionists, interventionist social validity averaged 5.73 ($SD = 1.16$, range: 4–7). Two items for the interventionists had an average of less than 6.0: "the intervention training prepared me to fully implement the intervention" averaged 5.33 ($SD = 1.53$, range: 4–7) and "the intervention was effective at improving reading comprehension" averaged 4.67 ($SD = 1.15$, range: 4–6). All other items averaged greater than or equal to 6.0.

Discussion

Given that a student's ability to remain attentive during reading instruction is positively associated with reading comprehension outcomes, it is not surprising that students with co-occurring reading difficulties and inattention are particularly vulnerable to inadequate reading growth following evidence-based reading instruction. Therefore, to support the behavior of students with co-occurring reading difficulties and inattention, this study investigated the impact of integrating evidence-based behavioral supports into a small-group evidence-based reading intervention. The impact of the behavior support intervention was evaluated with visual analysis and supported by Tau-U effect sizes and tests of statistically significant differences between the baseline and intervention phase. Using visual analysis, all six students demonstrated a functional relation on engagement and five students demonstrated a functional relation on disruptive behavior. For disruptive behavior, only Kiyana did not demonstrate a clear functional relation due to an increasing trend in disruptive behavior in the intervention phase. These findings were also supported by statistically significant study Tau-U mean effect sizes of 0.76 and 0.81, for engagement and

disruptive behavior, respectively. Descriptively, this finding suggests that relative to other classroom-based SCD interventions for students with ADHD (Harrison et al., 2019), adding this study's behavior supports to a reading intervention led to an above-average effect size on student behavior. Overall, this study found that for most students, the intervention led to an increase in student engagement and a decrease in disruptive behavior. Such findings align with previous research suggesting that embedding antecedent- and consequence-based strategies into academic settings can lead to improvements in student behavior (Kamps et al., 2015; Sutherland et al., 2020; Wills et al., 2009, 2018).

At the student level, there was variability that warrants additional explanation. Kiyana, Isabella, and Mateo had positive, yet non–statistically significant engagement effect sizes below 0.67. For Kiyana in particular, the lack of clear improvement from the baseline to intervention phase suggests a need for additional support (e.g., social skills, self-monitoring) in addition to what was provided in the intervention. Isabella had an upward trend in the intervention phase, therefore more sessions may have yielded a larger Tau-U effect size. Finally, for Mateo, the lack of statistical significance was most likely due to a consistently higher level of engagement in his baseline phase, suggesting that the added structure of the small-group explicit reading instruction provided in the baseline may have been a sufficient level of engagement support. Only Sofia had a non–statistically significant disruptive behavior effect size that was most likely due to a consistently low level of disruptive behavior in her baseline phase and a limited number of data points in the intervention phase. Furthermore, the visual analysis of Kiyana's disruptive behavior again suggests the need for additional support beyond this intervention.

Overall, these findings demonstrated that integrating antecedent- and consequence-based behavior supports into reading instruction can improve student engagement beyond what would be expected with an evidence-based reading curriculum alone. Such findings are particularly important as obtaining meaningful reading gains for upper elementary students with reading difficulties requires considering many factors on how to best intensify reading instruction (Miciak et al., 2017; Vaughn et al., 2019). For many students with reading difficulties, a contributor to their inadequate response to instruction may be low attention (e.g., Macdonald et al., 2020). Therefore, identifying mechanisms to improve student engagement during reading instruction has the potential to play a role in improving student response to reading instruction (e.g., Roberts et al., 2015, 2021).

Procedural Fidelity and Social Validity

Procedural fidelity was 100% in all training phase sessions and averaged 93% (SD = 8%, range: 83%–100%) during the intervention phase across all instructors. When low procedural fidelity occurred, it was due to an interventionist having instances of not immediately resetting the timer to signal the 3-minute interval. Overall, a mean of 93% procedural fidelity represents high adherence to the behavior support intervention, suggesting that the intervention can be implemented with fidelity following minimal training.

In terms of student social validity, scores were overall high (i.e., mean score of 6.71 on a 7-point Likert-type scale). For the interventionist social validity, the interventionists

only slightly agreed that the intervention improved reading comprehension, which is not unexpected given the brevity of the intervention and the challenges associated in remediating reading comprehension outcomes for upper elementary students, even after yearlong interventions (Miciak et al., 2017; Vaughn et al., 2019). The interventionists also felt the behavior intervention was effective at improving student behavior. Despite high levels of fidelity, the interventionists felt that more training was needed, perhaps with more opportunities for practice prior to implementing the intervention.

Limitations

This study has several limitations. The first limitation was that the behavior screening measure was completed by a general education ELA teacher. It may have been more beneficial to have a teacher who works with these students in a small-group reading setting complete the survey to identify students with inattentive behaviors in a setting more closely resembling the baseline phase. The second limitation was that the study did not evaluate the impact of the intervention on reading outcomes. Ultimately, the aim of this research is to improve student engagement to allow students greater access to instruction, and thus improve reading outcomes. Descriptively, we found that students improved their word reading, reading fluency, and reading comprehension from pretest to posttest. However, we cannot attribute any reading gains to the intervention. Third, there is still much debate on the use of standardized effect sizes in SCD research, and even though Tau-U is commonly used, caution is warranted when using this metric. Finally, we acknowledge that the social validity measures should have included questions on the extent to which students felt the intervention phase helped them improve their behavior.

Future Research and Implications for Practice

This study provides evidence to support the development of future studies. First, future studies could consider measuring reading in addition to the behavior outcomes. By including reading progress monitoring data, it could be possible to test the impact of the intervention on reading outcomes as well. Furthermore, to better understand the impact of a combined reading with behavior support intervention, future studies could include RCTs of a reading with behavior support, a reading-only, and a comparison condition. Through this design, it would allow researchers to test the impact of the integrating behavior supports into reading instruction on distal reading outcomes, as well as behavior.

In considering future intervention development, we would note that students with co-occurring reading difficulties and inattention often have other co-occurring behavioral difficulties (e.g., Willcutt et al., 2012). According to the BASC-3, the students in this study had mean ADHD probability, externalizing behavior, and internalizing behavior scores that were all above-average levels (see Table 2). Based on these findings, when researchers develop standardized behavior supports for small-group academic instruction, they may need to consider addressing inattention, hyperactivity, and externalizing and internalizing behaviors, concurrently.

Finally, even though outcomes varied by student, this study demonstrated that it is feasible and can be effective to integrate group expectations, precorrections, praise,

and a token economy with point goals and rewards into evidence-based reading instruction. For students who do not adequately respond to the behavior supports outlined in this study, individualized supports may be necessary. Furthermore, because the behavior supports were not designed specifically for the Voyager Passport reading curriculum, they hold promise to be integrated with other explicit and systematic evidence-based reading curriculums, and potentially other academic curriculums (e.g., mathematics). More research is needed to demonstrate that integrating the behavior supports outlined in this study can lead to improved reading outcomes and for whom. Yet, for educators wanting to improve student engagement and reduce disruptive behavior during small-group reading instruction, the behavior support strategies outlined in this study hold promise as a feasible and effective method that can be implemented with minimal training.

Note

The research reported here was supported by the Institute of Education Sciences, US Department of Education, through grant R324B190010 to the University of Denver. The opinions expressed are those of the authors and do not represent views of the Institute or the US Department of Education. The authors declare that they have no conflict of interest. Garrett J. Roberts is an associate professor at the University of Denver; Senaida Mehmedovic is a doctoral student at the University of Denver; Brooke Cote is a teaching and learning coach at Colorado Springs School District 11; Jade Wexler is an associate professor of special education at the University of Maryland; Phil Strain is a James C. Kennedy Endowed Chair and Professor at the University of Denver. Correspondence may be sent to Garrett J. Roberts at garrett.roberts@du.edu.

References

Alonzo, J., Tindal, G., Ulmer, K., & Glasgow, A. (2006). *EasyCBM online progress monitoring assessment system*. Center for Educational Assessment Accountability.

Barton-Arwood, S. M., Wehby, J. H., & Falk, K. B. (2005). Reading instruction for elementary-age students with emotional and behavioral disorders: Academic and behavioral outcomes. *Exceptional Children*, **72**(1), 7–27. https://doi.org/10.1177/001440290507200101

Brossart, D. F., 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**(3–4), 464–491. https://doi.org/10.1080/09602011.2013.868361

Burns, M. K., Riley-Tillman, T. C., & VanDerHeyden, A. M. (2012). RTI applications: Academic and behavioral interventions (Vol. 1). Guilford Press.

Chall, J. S. (1996). American reading achievement: Should we worry? *Research in the Teaching of English*, **30**(3), 303–310. https://www.jstor.org/stable/40171366

Collier-Meek, M. A., Johnson, A. H., Sanetti, L. H., & Minami, T. (2019). Identifying critical components of classroom management implementation. *School Psychology Review*, **48**(4), 348–361. https://doi.org/10.17105/SPR-2018-0026.V48-4

Compton, D. L., Fuchs, D., Fuchs, L. S., Elleman, A. M., & Gilbert, J. K. (2008). Tracking children who fly below the radar: Latent transition modeling of students with late-emerging reading disability. *Learning and Individual Differences*, **18**(3), 329–337. https://doi.org/10.1016/j.lindif.2008.04.003

Cook, C. R., Dart, E., Collins, T., Restori, A., Daikos, C., & Delport, J. (2012). Preliminary study of the confined, collateral, and combined effects of reading and behavior interventions: Evidence

for a transactional relationship. *Behavioral Disorders*, **38**(1), 38–56. https://doi.org/10.1177/019874291203800104

Denton, C. A., Tamm, L., Schatschneider, C., & Epstein, J. N. (2020). The effects of ADHD treatment and reading intervention on the fluency and comprehension of children with ADHD and word reading difficulties: A randomized clinical trial. *Scientific Studies of Reading*, **24**(1), 72–89. https://doi.org/10.1080/10888438.2019.1640704

Donegan, R. E., & Wanzek, J. (2021). Effects of reading interventions implemented for upper elementary struggling readers: A look at recent research. *Reading and Writing: An Interdisciplinary Journal*. Advance online publication. https://doi.org/10.1007/s11145-021-10123-y

DuPaul, G. J., Weyandt, L. L., & Janusis, G. M. (2011). ADHD in the classroom: Effective intervention strategies. *Theory Into Practice*, **50**(1), 35–42. https://doi.org/10.1080/00405841.2011.534935

Friedman, L. M., McBurnett, K., Dvorsky, M. R., Hinshaw, S. P., & Pfiffner, L. J. (2020). Learning disorder confers setting-specific treatment resistance for children with ADHD, predominantly inattentive presentation. *Journal of Clinical Child and Adolescent Psychology*, **49**(6), 854–867. https://doi.org/10.1080/15374416.2019.1644647

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, **7**(1), 6–10. https://doi.org/10.1177/074193258600700104

Guthrie, J. T., McRae, A., & Klauda, S. L. (2007). Contributions of concept-oriented reading instruction to knowledge about interventions for motivations in reading. *Educational Psychologist*, **42**(4), 237–250. https://doi.org/10.1080/00461520701621087

Harris, K. R., Friedlander, B. D., Saddler, B., Frizzelle, R., & Graham, S. (2005). Self-monitoring of attention versus self-monitoring of academic performance: Effects among students with ADHD in the general education classroom. *Journal of Special Education*, **39**(3), 145–157. https://doi.org/10.1177/00224669050390030201

Harrison, J. R., Soares, D. A., Rudzinski, S., & Johnson, R. (2019). Attention deficit hyperactivity disorders and classroom-based interventions: Evidence-based status, effectiveness, and moderators of effects in single-case design research. *Review of Educational Research*, **89**(4), 569–611. https://doi.org/10.3102/0034654319857038

Institute of Education Sciences. (2017). *What works clearinghouse standards handbook* (Version 4.0). https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_standards_handbook_v4.pdf

Institute of Education Sciences. (2020). *What works clearinghouse standards handbook* (Version 4.1). https://ies.ed.gov/ncee/wwc/Docs/referenceresources/WWC-Standards-Handbook-v4-1-508.pdf

Kamps, D., Wills, H., Dawson-Bannister, H., Heitzman-Powell, L., Kottwitz, E., Hansen, B., & Fleming, K. (2015). Class-wide function-related intervention teams "CW-FIT" efficacy trial outcomes. *Journal of Positive Behavior Interventions*, **17**(3), 134–145. https://doi.org/10.1177/1098300714565244

Kim, J. S., Hemphill, L., Troyer, M., Thomson, J. M., Jones, S. M., LaRusso, M. D., & Donovan, S. (2017). Engaging struggling adolescent readers to improve reading skills. *Reading Research Quarterly*, **52**(3), 357–382. https://doi.org/10.1002/rrq.171

Kuchle, L. B., & Riley-Tillman, T. C. (2019). Integrating behavior and academics in intervention planning. In R. Z. Edmonds, A. G. Gandhi, & L. Danielson (Eds.), *Essentials of intensive intervention* (pp. 51–68). Guilford.

Lane, K. L., Bocian, K. M., MacMillan, D. L., & Gresham, F. M. (2004). Treatment integrity: An essential—But often forgotten—Component of school-based interventions. *Preventing School Failure: Alternative Education for Children and Youth*, **48**(3), 36–43. https://doi.org/10.3200/PSFL.48.3.36-43

Lane, K. L., O'Shaughnessy, T. E., Lambros, K. M., Gresham, F. M., & Beebe-Frankenberger, M. E. (2001). The efficacy of phonological awareness training with first-grade students who have behavior problems and reading difficulties. *Journal of Emotional and Behavioral Disorders*, **9**(4), 219–231. https://doi.org/10.1177/106342660100900402

Ledford, J. R., & Gast, D. L. (2018). *Single case research methodology: Applications in special education and behavioral sciences*. Routledge. https://doi.org/10.4324/9781315150666

Lemons, C. J., Al Otaiba, S., Conway, S. J., & Mellado De La Cruz, V. (2016). Improving professional development to enhance reading outcomes for students in special education. *New Directions for Child and Adolescent Development*, **154**, 87–104.

Lin, Y. C., Morgan, P. L., Farkas, G., Hillemeier, M., & Cook, M. (2013). Reading, mathematics, and behavioral difficulties interrelate: Evidence from a cross-lagged panel design and population-based sample of US upper elementary students. *Behavioral Disorders*, **38**(4), 212–227. https://doi.org/10.1177/019874291303800404

Macdonald, K. T., Barnes, M. A., Miciak, J., Roberts, G., Halverson, K. K., Vaughn, S., & Cirino, P. T. (2020). Sustained attention and behavioral ratings of attention in struggling readers. *Scientific Studies of Reading*. Advance online publication. https://doi.org/10.1080/10888438.2020.1826950

MacGinitie, W. H., MacGinitie, R. K., Maria, K., Dreyer, L. G., & Hughes, K. E. (2000). *Gates-MacGinitie reading tests* (4th ed.). Riverside.

Mayes, S. D., & Calhoun, S. L. (2007). Learning, attention, writing, and processing speed in typical children and children with ADHD, autism, anxiety, depression, and oppositional-defiant disorder. *Child Neuropsychology*, **13**(6), 469–493. https://doi.org/10.1080/09297040601112773

McKenna, J. W., Kim, M. K., Shin, M., & Pfannenstiel, K. (2017). An evaluation of single-case reading intervention study quality for students with and at risk for emotional and behavioral disorders. *Behavior Modification*, **41**(6), 868–906. https://doi.org/10.1177/0145445517701896

Miciak, J., Roberts, G. J., Taylor, W. P., Solis, M., Vaughn, S., & Fletcher, J. M. (2017). The effects of a two-year, intensive reading intervention implemented with late elementary struggling readers. *Learning Disabilities Research and Practice*, **33**, 24–36. https://doi.org/10.1111/ldrp.12159

Myers, D., Freeman, J., Simonsen, B., & Sugai, G. (2017). Classroom management with exceptional learners. *TEACHING Exceptional Children*, **49**(4), 223–230. https://doi.org/10.1177/0040059916685064

National Center on Intensive Intervention. (2013). *Implementing intensive intervention: Lessons learned from the field*. US Department of Education, Office of Special Education Programs. https://files.eric.ed.gov/fulltext/ED575660.pdf

National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups*. National Institute of Child Health and Human Development. https://www.nichd.nih.gov/sites/default/files/publications/pubs/nrp/Documents/report.pdf

Parker, R. I., Vannest, K. J., Davis, J. L., & Sauber, S. B. (2011). Combining nonoverlap and trend for single-case research: Tau-U. *Behavior Therapy*, **42**(2), 284–299. https://doi.org/10.1016/j.beth.2010.08.006

Reynolds, C. R., & Kamphaus, R. W. (2015). *Behavior assessment system for children* (3rd ed.). Pearson.

Roberts, G., Rane, S., Fall, A.-M., Denton, C. A., Fletcher, J. M., & Vaughn, S. (2015). The impact of intensive reading intervention on level of attention in middle school students. *Journal of Clinical Child and Adolescent Psychology*, **44**(6), 942–953. https://doi.org/10.1080/15374416.2014.913251

Roberts, G. J., Cho, E., Garwood, J. D., Goble, G. H., Robertson, T., & Hodges, A. (2020). Reading interventions for students with reading and behavioral difficulties: A meta-analysis and evaluation of co-occurring difficulties. *Educational Psychology Review*, **32**(1), 17–47. https://doi.org/10.1007/s10648-019-09485-1

Roberts, G. J., Vaughn, S., Roberts, G., & Miciak, J. (2021). Problem behaviors and response to reading intervention for upper elementary students with reading difficulties. *Remedial and Special Education*, **42**(3), 169–181. https://doi.org/10.1177/0741932519865263

Sayeski, K. L., & Brown, M. R. (2011). Developing a classroom management plan using a tiered approach. *TEACHING Exceptional Children*, **44**(1), 8–17. https://doi.org/10.1177/004005991104400102

Scammacca, N. K., Roberts, G., Vaughn, S., & Stuebing, K. K. (2015). A meta-analysis of interventions for struggling readers in grades 4–12: 1980–2011. *Journal of Learning Disabilities*, **48**(4), 369–390. https://doi.org/10.1177/0022219413504995

Sen, P. K. (1968). Estimates of the regression coefficient based on Kendall's tau. *Journal of the American Statistical Association*, **63**(324), 1379–1389. https://doi.org/10.2307/2285891

Simonsen, B., Fairbanks, S., Briesch, A., Myers, D., & Sugai, G. (2008). Evidence-based practices in classroom management: Considerations for research to practice. *Education and Treatment of Children*, **31**(3), 351–380. https://doi.org/10.1353/etc.0.0007

Simonsen, B., Freeman, J., Goodman, S., Mitchell, B., Swain-Bradway, J., Flannery, B., & Putman, B. (2015). *Supporting and responding to behavior: Evidence-based classroom strategies for teachers*. US Office of Special Education Programs. https://osepideasthatwork.org/sites/default/files/ClassroomPBIS_508.pdf

Stewart, A. A., & Austin, C. R. (2020). Reading interventions for students with or at risk of attention deficit hyperactivity disorder: A systematic review. *Remedial and Special Education*, **41**(6), 352–367. https://doi.org/10.1177/0741932519849660

Sutherland, K. S., Conroy, M. A., McLeod, B. D., Granger, K., Broda, M., & Kunemund, R. (2020). Preliminary study of the effects of BEST in CLASS–Elementary on outcomes of elementary students with problem behavior. *Journal of Positive Behavior Interventions*, **22**(4), 220–233. https://doi.org/10.1177/1098300719900318

Tarlow, K. R. (2017). An improved rank correlation effect size statistic for single-case designs: Baseline corrected Tau. *Behavior Modification*, **41**(4), 427–467. https://doi.org/10.1177/0145445516676750

Theil, H. (1950). A rank-invariant method of linear and polynomial regression analysis, I, II, and III. *Proceedings of the Royal Netherlands Academy of Sciences*, **53**, 386–392, 521–525, 1397–1412.

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (2012). *Test of word reading efficiency*. Pro-Ed.

Vannest, K. J., & Ninci, J. (2015). Evaluating intervention effects in single-case research designs. *Journal of Counseling and Development*, **93**(4), 403–411. https://doi.org/10.1002/jcad.12038

Vannest, K. J., Parker, R. I., Gonen, O., & Adiguzel, T. (2016). *Single case research: Web based calculators for SCR analysis* (Version 2.0) [Web-based application]. Texas A&M University. http://www.singlecaseresearch.org/

Vaughn, S., Denton, C. A., & Fletcher, J. M. (2010). Why intensive interventions are necessary for students with severe reading difficulties. *Psychology in the Schools*, **47**(5), 432–444. https://doi.org/10.1002/pits.20481

Vaughn, S., Roberts, G. J., Miciak, J., Taylor, P. M, & Fletcher, J. M. (2019). Efficacy of a word- and text-based intervention for students with significant reading difficulties. *Journal of Learning Disabilities*, **52**(1), 31–44. https://doi.org/10.1177/0022219418775113

Voyager Sopris Learning. (2008). *Passport*. Voyager Sopris Learning.

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

Wanzek, J., Petscher, Y., Al Otaiba, S., Kent, S. C., Schatschneider, C., Haynes, M., Rivas, B. K., & Jones, F. G. (2016). Examining the average and local effects of a standardized treatment for fourth graders with reading difficulties. *Journal of Research on Educational Effectiveness*, **9**(Suppl. 1), 45–66. https://doi.org/10.1080/19345747.2015.1116032

Wanzek, J., Petscher, Y., Al Otaiba, S., Rivas, B. K., Jones, F. G., Kent, S. C., Schatschneider, C., & Mehta, P. (2017). Effects of a year-long supplemental reading intervention for students with reading difficulties in fourth grade. *Journal of Educational Psychology*, **109**(8), 1103–1119. https://doi.org/10.1037/edu0000184

Willcutt, E. G. (2012). The prevalence of DSM-IV attention-deficit/hyperactivity disorder: A meta-analytic review. *Neurotherapeutics*, **9**(3), 490–499. https://doi.org/10.1007/s13311-012-0135-8

Willcutt, E. G., Nigg, J. T., Pennington, B. F., Solanto, M. V., Rohde, L. A., Tannock, R., Loo, S. K., Carlson, C. L., McBurnett, K., & Lahey, B. B. (2012). Validity of DSM-IV attention deficit/hyperactivity disorder symptom dimensions and subtypes. *Journal of Abnormal Psychology*, **121**(4), 991–1010. https://doi.org/10.1037/a0027347

Willcutt E. G., & Pennington, B. F. (2000). Psychiatric comorbidity in children and adolescents with reading disability. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, **41**(8), 1039–1048. https://doi.org/10.1111/1469-7610.00691

Wills, H., Wehby, J., Caldarella, P., Kamps, D., & Swinburne Romine, R. (2018). Classroom management that works: A replication trial of the CW-FIT program. *Exceptional Children*, **84**(4), 437–456. https://doi.org/10.1177/0014402918771321

Wills, H. P., Kamps, D., Hansen, B., Conklin, C., Bellinger, S., Neaderhiser, J., & Nsubuga, B. (2009). The classwide function-based intervention team program. *Preventing School Failure: Alternative Education for Children and Youth*, **54**(3), 164–171. https://doi.org/10.1080/10459880903496230