



HHS Public Access

Author manuscript

Behav Modif. Author manuscript; available in PMC 2015 April 01.

Published in final edited form as:

Behav Modif. 2015 January ; 39(1): 8–42. doi:10.1177/0145445514561318.

Reading Interventions With Behavioral and Social Skill Outcomes: A Synthesis of Research

Garrett J. Roberts¹, Michael Solis², Stephen Ciullo³, John W. McKenna⁴, and Sharon Vaughn¹

¹The University of Texas at Austin, USA

²University of Virginia, Charlottesville, USA

³Texas State University, San Marcos, USA

⁴St. Johns University, Queens, NY, USA

Abstract

Research findings have suggested that reading deficits and problem behaviors are positively related. This synthesis investigated how reading interventions impact behavioral/social skill outcomes by reviewing studies that included (a) a reading intervention without behavioral/social skill components, (b) behavioral/social skill dependent variables, and (c) students in Grades K-12. Fifteen articles were evaluated by the type of reading intervention, associations between positive reading effects and behavioral/social skill outcomes, and The What Works Clearinghouse (WWC) determinants of study ratings. Findings suggested that reading interventions tended to have positive reading outcomes, while behavioral/social skill outcomes were small or negative. Research did not suggest an association between improved reading and behavioral performance, regardless of the WWC study determinants rating. Implications include reading instruction may not be sufficient to improve behavioral and social skill outcomes. Additional research is warranted to investigate the long-term impact of reading on behavioral and social skill outcomes.

Keywords

reading difficulties; learning disability; behavior difficulties; behavior disorder; behavior; social skills; reading intervention; at risk students; synthesis; all levels

Many students who display persistent problem behaviors also struggle academically (Kauffman & Landrum, 2009). Academic and social behaviors (e.g., reading, math, social

© The Author(s) 2014

Reprints and permissions: sagepub.com/journalsPermissions.nav

Corresponding Author: Garrett J. Roberts, The Meadows Center for Preventing Educational Risk, College of Education, The University of Texas at Austin, SZB 228, 1912 Speedway D4900 Austin, TX 78712-1284, USA. garrett.roberts@utexas.edu.

Authors' Note

The content is solely the responsibility of the authors and does not necessarily represent the official views of the Eunice Kennedy Shriver National Institute of Child Health and Human Development or the National Institutes of Health.

Declaration of Conflicting Interests

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

skills, problem behaviors) are related and reciprocal (Malecki & Elliot, 2002), with a positive relationship between social behavior and reading (J. W. Adams, Snowling, Hennessy, & Kind, 1999). A meta-analysis by Kavale and Forness (1996) reported that 75% of students with Learning Disabilities (LD) could be differentiated from nonlearning disabled peers by social competence measures alone.

Studies have also suggested that early behavior problems are predictive of later reading difficulties and early reading deficits can predict future behavior problems. This was demonstrated with higher levels of behavior problems in 6-year-olds predicting later reading deficits in 17-year-olds (Breslau et al., 2009). In addition, two other studies demonstrated that first-grade reading difficulties predicted third-grade behavior problems (Miles & Stipek, 2006; Morgan, Farkas, Tufis, & Sperling, 2008). Moreover, when behavior problems persist and students are unable to develop prosocial behavior by approximately third grade, children are likely to display antisocial behavior throughout their lives (Bullis & Walker, 1994; Kazdin, 1987).

Problem behaviors are also negatively associated with the likelihood that a student will respond to intensive reading interventions (Hagan-Burke et al., 2011). Measures of attention and problem behaviors have been found to be significantly and positively related to students' nonresponsiveness to intensive reading interventions (Al Otaiba & Fuchs, 2002). Furthermore, in a study by Torgesen et al. (1999), problem behavior was identified as one of the most reliable predictors of student outcomes (or lack of positive outcomes) in response to intervention. The authors noted that even in one-to-one teaching situations, problem behavior made it difficult for students to benefit from intervention.

Problem Behaviors and Reading Research

The relationship between problem behaviors and academic performance manifests in various forms including (a) externalizing behavior (Nelson, Benner, Lane, & Smith, 2004), (b) internalizing behavior (Harris, Oakes, Lane, & Rutherford, 2009), and (c) deficits in social skills (Kavale & Forness, 1996; Malecki & Elliot, 2002). See Table 1 for operational definitions of these behaviors. Reading achievement tends to be defined in two stages referred to as (a) learning to read in Grades K-3 and (b) reading to learn or understand in Grades 4 to 12 (Chall, 1996; Torgesen et al., 2007). Related literature on problem behaviors and reading interventions will be summarized according to these two categories.

Grades K-3

One syntheses and one meta-analysis investigated preschool to third-grade students who failed to respond to reading interventions. In a synthesis by Al Otaiba and Fuchs (2002), they investigated the characteristics of students who were unresponsive to early literacy intervention. Nine of 23 studies in this synthesis reported attention and behavior problems as a factor. Seven of these 9 studies reported a relationship between unresponsiveness to an early reading intervention and attention or behavior problems. In a similar meta-analysis, Nelson, Benner, and Gonzalez (2003) reported problem behavior as having the third largest magnitude in predicting treatment effectiveness (or nonresponsiveness) in literacy interventions with rapid naming and phonological awareness having larger magnitudes.

Grades 4 to 12

Students with reading difficulties in Grades 4 to 12 often demonstrate declining engagement and motivation to read, possibly affecting growth in adolescent reading proficiency in two ways (Torgesen et al., 2007). First, students with lower motivation spend less time reading than students with higher motivation. Second, students who are less motivated to read are less engaged during reading, thus negatively influencing comprehension (Torgesen et al., 2007). Furthermore, Torgesen and colleagues (2007) identified motivation and engagement as areas of skill and knowledge that need to continually improve in Grades 4 to 12, stating that "... motivation and engagement in reading and completing reading-based assignments must be part of any comprehensive plan for improving levels of academic literacy in adolescents" (p. 10).

In recent meta-analyses pertaining to 4th- to 12th-grade reading interventions (Edmonds et al., 2009; Flynn, Zheng, & Swanson, 2012; Scammacca, Roberts, Vaughn, & Stuebing, 2013; Wanzek et al., 2013), there were not a sufficient number of studies included in any of these systematic reviews to examine the moderating effects of problem behavior. Edmonds et al. (2009) acknowledged this by suggesting that future research should investigate social, affective, engagement, and motivational variables to better understand their role in student reading comprehension outcomes.

Theoretical Explanations: Behavior and Reading

There are four hypothesized models (Hinshaw, 1992; Spira & Fischel, 2005) that address the causality between problem behaviors and reading outcomes: (a) Reading difficulties lead to problem behaviors, as student's behavior is escape maintained to avoid difficult tasks; (b) problem behaviors result in reading deficits, as students are not academically engaged and do not access the academic content; (c) a transactional relationships exists, where both problem behaviors and reading underachievement simultaneously influence each other; and (d) a "common cause" such as inattentiveness can lead to reading and behavior problems. This synthesis will further investigate the hypothesis that reading difficulties lead to problem behaviors by examining the evidence from reading intervention studies that include behavioral outcomes (Hinshaw, 1992; Spira & Fischel, 2005). If the hypothesis is supported, improving a student's reading, through a reading intervention, may lead to a decrease in problem behaviors, as behaviors are escape maintained by avoiding difficult tasks (Morgan et al., 2008).

Reading Interventions and Problem Behaviors

Two syntheses investigated the impact of reading interventions on social outcomes (Nelson, Lane, Benner, & Kim, 2011; Wanzek, Vaughn, Kim, & Cavanaugh, 2006). Wanzek et al. (2006) located 27 intervention studies that investigated students with learning or reading disabilities in Grades K-6 and reported the following conclusions: (a) The majority of the studies had small positive effects on social outcomes, (b) reading interventions and social outcomes were positively associated, and (c) further research was needed. Nelson and colleagues (2011) investigated the effects of reading instruction on the social adjustment for students who have or are at risk of having reading and/or behavior problems. The inclusion criteria for this synthesis required the studies to (a) be a randomized controlled trial (RCT) or

quasi-experimental, (b) include enough quantitative information to have a calculable effect size, and (c) have dependent measures in literacy and social adjustment. With this inclusion criterion, they located four studies in Grades K-6, and found that reading interventions consistently improved reading skills ($g = 0.49$) but did not improve the social adjustment of students ($g = -0.14$).

Rationale and Purpose

The rationale for this investigation is to expand findings from previous related syntheses (Nelson et al., 2011; Wanzek et al., 2006) based on research suggesting that there is a (a) high co-occurrence rate of behavior problems and reading problems; (b) need to improve reading outcomes, especially for students who fail to respond to adequate and intensive interventions; and (c) need to further investigate the theoretical explanations that posit the idea that reading difficulties lead to problem behaviors by exploring the relationship between reading and behavioral outcomes in the context of reading interventions aimed to improve reading skills.

While this synthesis shares common features with Wanzek et al. (2006) and Nelson et al. (2011), our purpose differs from Wanzek et al. (2006) in that we expand their inclusion criteria to include all students in Grades K-12 and isolate the reading component by including only reading interventions without behavioral or social skill independent variables (e.g., token economy). In addition, we present a more comprehensive review than Nelson et al. (2011) by expanding the inclusion criteria in the following areas: (a) participant selection; (b) type of reading instruction provided is not restricted to phonological awareness, phonics, fluency, comprehension, and vocabulary; (c) research design (i.e., single-case research designs); (d) broadening the dependent measure requirement; and (e) not requiring a calculable effect size.

This synthesis will evaluate reading interventions with behavioral/social skill outcomes to answer the following research questions:

Research Question 1: What are the effects of reading interventions on behavioral/social skill outcomes for students in Grades K-12 and do these effects differ when disaggregated by the quality of the study as determined by the What Works Clearinghouse (WWC) determinants of study rating (Institute of Education Sciences [IES], 2014)?

Research Question 2: How are positive reading effects from reading interventions related to behavioral/social skill outcomes for students in Grades K-12?

Method

Criteria for Inclusion

Studies were included if they met the following criteria: (a) included reading intervention as an independent variable, (b) did not include a behavioral or social skill component (e.g., token economy, social skills training) as an independent variable, (c) included a behavioral or social skills outcome (see Table 1), (d) included students in Grades K-12, (e) occurred during regular school hours in the United States, (f) published in English, (g) published in a peer-reviewed

journal from 1975 to 2013, and (h) was a group design (single-group designs were excluded) or single-case design.

Literature Search Procedures

First, computer-assisted searches for relevant literature were conducted for articles published between January 1975 and September 2013, using Educational Resources Information Center (ERIC), Education Full Text, PsycINFO, and Academic Search Complete with the following search terms: *reading, behavior, English language learner, culturally and linguistically diverse, peer mediated, peer tutoring, and social adjustment* in various combinations. This search resulted in 2,783 articles that were reviewed. Second, a hand search was conducted from 2011 to 2013 of the following journals that published the articles that met the criteria in the first step: *Annals of Dyslexia, Behavior Disorders, Education and Treatment of Children, Exceptional Children, Journal of Applied Behavior Analysis, Journal of Emotional and Behavioral Disorders, Journal of Learning Disabilities, Learning Disability Quarterly, Preventing School Failure, Psychological Reports, Psychology in the Schools, and Remedial and Special Education*. Third, reference sections of 11 relevant literature reviews were searched (Al Otaiba & Fuchs, 2002; Bruhn, Lane, & Hirsch, 2014; Cook et al., 2008; Joseph & Eveleigh, 2011; Mooney, Ryan, Uhing, Reid, & Epstein, 2005; Ryan, Reid, & Epstein, 2004; Simonsen, Fairbanks, Briesch, Myers, & Sugai, 2008; Spencer, 2006; Vannest, Temple-Harvey, & Mason, 2009; Wanzek et al., 2006; Wanzek, Wexler, Vaughn, & Ciullo, 2010). Finally, the reference sections of all the studies meeting the inclusion criteria were examined. Following the initial computer search, no additional studies were identified that met the inclusion criteria. The search procedures resulted in 15 studies from 1981 to 2009.

Coding Procedures

A code sheet was developed based on previous intervention synthesis and included (a) participant information, (b) design, (c) treatment and design group characteristics, (d) clarity of causal inference, (e) precision of outcome, and (f) effect size measurements/descriptive findings when applicable (Edmonds et al., 2009). All articles were double coded after interrater agreement was established at or above 90% for all coders on two articles (i.e., single-case design and group design). Interrater agreement was calculated by dividing the number of agreements by the sum of the number of agreements and disagreements. In addition, all tables were double coded for accuracy using the criteria described in the following section.

Data Analysis

The effect sizes for the group design studies were compared descriptively due to the low number of studies with calculable effect sizes ($n = 3$; Lane et al., 2007; Nelson, Stage, Epstein, & Pierce, 2005; O'Shaughnessy & Swanson, 2000). For studies where effect sizes could be recalculated, Hedges's g (Cooper & Hedges, 1994) was used to recalculate the effect sizes by taking the posttest divided by the pooled standard deviation.

Data analysis was descriptively provided for the single-case design studies using the visual analysis procedures recommended by Horner, Swaminathan, Sugai, and Smolkowski (2012). Visual analysis was selected for data analysis as effect size estimations for single-case designs currently have no widely agreed upon method (Horner et al., 2012) and many researchers prefer

visual analysis for basing their inferences (Kratochwill et al., 2010). Using Horner and colleagues (2012) proposed procedures for visual analysis, we investigated five variables including level, slope variability, immediacy of effect, and/or degree of overlap.

For all single-case designs (e.g., alternating treatment, multiple baseline), three variables that evaluated outcomes within and across adjacent phases were used including (a) level (i.e., measuring central tendency within phases), (b) trend or slope of data points, and (c) variability or deviation in scores from the trend line. For designs that included a baseline phase (e.g., withdrawal, multiple baseline) changes between two adjacent phases were evaluated by (a) immediacy of effect (i.e., how quickly data patterns change following a manipulation of the independent variable) and (b) degree of overlap, which is calculated by determining the percentage of intervention data points not having a larger positive outcome than the baseline data point with the largest positive outcome (i.e., proportion of data points in Phase 2 that overlap with Phase 1; lower number indicates greater effect). This synthesis evaluated these variables with the results available in Table 4.

Study Design Evaluations

To assess the quality of each study, The WWC procedure and standards handbook, version 3.0 (IES, 2014) determinants of study ratings was used. The WWC uses a multistep evaluation process to classify studies into (a) meets WWC standards without reservations, (b) meets WWC standards with reservations, and (c) does not meet WWC standards. The criteria set forth by the WWC vary by study design (i.e., group design, single-case design) and are described in this synthesis.

Group design evaluations—The determinants of study ratings for group designs include (a) study design (i.e., RCT or quasi-experimental), (b) sample attrition with liberal boundary (i.e., low, high), and (c) baseline equivalence (reported differences greater than 0.25 pooled standard deviations on any baseline characteristic do not meet baseline equivalence criteria).

The WWC flowchart of the multistep evaluation process (IES, 2014), only allows an RCT design without differential attrition to receive the highest study determinant, which is meets WWC standards without reservations. Group studies receive the study determinant rating of meets WWC standards with reservations if they were either an RCT with high differential attrition and had a baseline equivalence or a quasi-experimental study with a baseline equivalence. RCTs that had a high differential attrition and no baseline equivalence or quasi-experimental studies that had no baseline equivalence were classified as does not meet WWC standards.

Single-case design evaluations—Based on the WWC (IES, 2014) study rating determinants, for a single-case design study to meet the criteria for meets WWC standards without reservations, (a) they must have a systematically manipulated independent variable, (b) interrater agreement must be systematically measured in at least 20% of data points in each condition (if percentage of data points is not reported per condition, studies still meet requirement, but this absence of information must be documented), (c) a minimum interrater agreement threshold must be met (percentage agreement is .80–.90, Cohen's kappa is .60), and (d) attempts to demonstrate the effect of the independent variable on the

dependent variable over time and with data points in each phase. The final step in determining the study rating varied by design (e.g., alternating treatment, multiple baseline) and included the evaluation of the total number of phases/conditions and total number of data points per phase/condition.

Results

This synthesis reports the results of 15 reading intervention studies and their reading and behavioral/social skill outcomes. First, a summary of studies is provided, followed by interventions disaggregated and evaluated by the type of intervention that was conducted (e.g., phonics, multicomponent with comprehension). Next, studies that report both reading and behavior/social skill outcomes are investigated to report on the associations between reading and behavioral/social skill outcomes. This is followed by reporting on the visual analysis evaluation based on the proposed criteria of Horner and colleagues (2012) for the eight single-case design studies. We conclude with evaluating each study to determine if it meets the WWC standards without reservations, with reservations, or does not meet WWC standards.

Summary of Studies

Research designs and outcome measures of the 15 studies meeting the inclusion criteria included (a) 4 RCT studies, with 3 reporting both behavior and reading outcomes (Gest & Gest, 2005; Lane, Fletcher, Carter, Dejud, & DeLorenzo, 2007; Nelson et al., 2005) and 1 reporting behavior outcomes with no reading outcomes (Strayhorn & Bickel, 2002); (b) 2 quasi-experimental studies, with 1 reporting both behavior and reading outcomes (Scruggs & Osguthorpe, 1986) and 1 reporting behavior outcomes with no reading outcomes (Feldman, 1981); (c) 8 single-case studies, with 6 reporting both behavior and reading outcomes (Kamps, Barbeta, Leonard, & Delquadri, 1994; Lane, Little, Redding-Rhodes, Phillips, & Welsh, 2007; Lane et al., 2002; Lingo, Slaton, & Jolivet, 2006; Scott & Shearer-Lingo, 2002; Wehby, Falk, Barton-Arwood, Lane, & Cooley, 2003) and 2 reporting behavior outcomes with no reading outcomes (Beck, Burns, & Lau, 2009; Burke, Hagan-Burke, & Sugai, 2003); and (d) 1 study that used an RCT and single-case design (for oral reading fluency [ORF]) that included both reading and behavior measures (O'Shaughnessy & Swanson, 2000). The total number of students was 304 with 68% males and 32% females (2 studies did not report gender). Thirteen studies included students ($n = 294$) in Grades K-5, and 2 studies included students ($n = 10$) in Grades 6 to 7.

A total of 14 studies included students who were either at risk of or had been identified as having a reading disability and 13 studies included students who were either at risk or had been identified as having either a behavioral, attention, or social skills deficit. Of the 12 studies that included the number of treatment sessions, the range was 2 to 137. The primary implementers of treatments were teachers ($n = 6$), researchers ($n = 5$), and paraprofessionals ($n = 3$). One study did not report who implemented the treatment. The reading interventions included 9 studies with a commercially available curriculum (e.g., Corrective Reading; Engelmann et al., 1999), and 6 studies used researcher-developed curricula. Fidelity was reported on treatments in 9 studies.

Effect of Reading Intervention on Behavior and Social Skill Outcomes

Out of the 15 studies meeting the inclusion criteria, 3 studies included a phonics intervention, 6 studies included a phonemic awareness (PA) and phonics intervention, 1 study included a phonics and fluency intervention, 3 studies included a multicomponent with comprehension intervention, and 2 studies implemented interventions categorized as “other” (i.e., based on Functional Behavior Assessment [FBA], Sentence Approach; Jansky, 1981). For an overview of the studies see Table 2 (group design) and Table 3 (single-case design).

Phonics—All three phonics-based interventions utilized a single-case design. In two studies, Beck et al. (2009) pretaught letter-sounds and words to students with Emotional Behavior Disorder (EBD) or who were at risk of behavior difficulty and Lane, Little, et al. (2007) implemented Peer Assisted Learning Strategies (PALS; Fuchs, Fuchs, Mathes, & Simmons, 1997), focusing on segmenting and blending words, decoding words, story reading, and partner reading with students at risk of reading and behavior difficulty. One study implemented a phonics and writing intervention (Lane et al., 2002), where phonics chapter books were used to teach PA, connect sound symbols, high frequency words, chapter reading, dictation, and writing. Two of these studies reported reading and behavioral outcomes (Lane, Little, et al., 2007; Lane et al., 2002) and one study only reported behavioral outcomes (Beck et al., 2009).

Across these three studies, two found student performance improved for all students in (a) on-task behavior (Beck et al., 2009) and (b) total disruptive behavior (TDB; Lane et al., 2002), while in the third study (Lane, Little, et al., 2007) student performance in academic engaged time (AET) had high variability, no clear trends across phases, and a total of 88% of the data points overlapping (i.e., 88% of the AET data points during the intervention were lower than the highest AET data point during baseline). For studies that included reading outcomes (Lane, Little, et al., 2007; Lane et al., 2002), the reading outcomes (i.e., ORF, nonsense word fluency [NWF]) tended to have higher levels and improved trends during the reading intervention phase as compared with the baseline phase.

PA—Six multicomponent studies targeted PA. Four of these studies were group designs (Lane, Fletcher, et al., 2007; Nelson et al., 2005; O’Shaughnessy & Swanson, 2000; Scruggs & Osguthorpe, 1986) and two studies were single-case design (Lingo et al., 2006; Wehby et al., 2003). In the four studies that implemented a group design, two studies used PA and phonics; one study used PA, phonics, and working memory; and one study used PA and spelling. One of the two studies that used a PA and phonics intervention, was a two-treatment and comparison group design with one treatment condition (Phonological Awareness Training Group [PAT]) focused on instruction sound blending, sound segmenting, and letter-sound correspondences, and the second treatment condition (Word Analogy Training Group [WAT]) focused on instruction in rhyming, decoding strategies, whole words, written language, and high frequency spelling (O’Shaughnessy & Swanson, 2000). The other study (Scruggs & Osguthorpe, 1986) used peer tutoring and the curriculum Beginning Reading I and II (Harrison, 1982). This study only included students who had met a criterion for a LD and a behavior disorder.

In the PA, phonics, and working memory study, the curriculum Stepping Stones to Literacy (Nelson, Cooper, & Gonzalez, 2004) was used to deliver a scripted program including PA, phonics, and serial processing to kindergarten students who were at risk of reading and behavior

difficulties (Nelson et al., 2005). Finally, in an intervention study that was to be delivered to first-grade students who were at risk of reading and behavior difficulties, students received Phonological Awareness Training for Reading (PATR; Torgesen & Bryant, 1994a) to teach rhyming, blending, segmenting, and spelling (Lane, Fletcher, et al., 2007).

Out of the four studies using a group design, there were no behavioral measures with significant posttest group differences. Three of these studies reported enough information to recalculate an effect size (Lane, Fletcher, et al., 2007; Nelson et al., 2005; O'Shaughnessy & Swanson, 2000) and the other study (Scruggs & Osguthorpe, 1986) did not report enough information to recalculate an effect size (see Table 2). In the three studies that reported effect sizes for nonstatistically significant results, the range was $g = 0.39$ (positive effect associated with negative outcomes, negative social interaction; Lane, Fletcher, et al., 2007) to $g = 0.63$ (intrapersonal skills; Nelson et al., 2005).

For the reading measures, three of the studies found at least one significant interaction effect between time and treatment (Lane, Fletcher, et al., 2007; Nelson et al., 2005; O'Shaughnessy & Swanson, 2000) favoring the treatment condition, while one study did not find any significant differences between groups. On measures where groups significantly differed at posttest, effect sizes ranged from $g = 0.40$ on the Woodcock Reading Mastery Test–Revised, Word Attack (R. Woodcock, 1987) to $g = 2.07$ on a trained content measure of phonemic deletion (Stanovich, Cunningham, & Cramer, 1984) both of which were in the O'Shaughnessy and Swanson (2000) study (see Table 2).

Two PA, phonics, and word-reading studies were implemented using single-case designs (Lingo et al., 2006; Wehby et al., 2003). In the Lingo et al. (2006) study, Corrective Reading (Engelmann et al., 1999) was delivered to seven students in the sixth to seventh grade who all had reading and behavior objectives on their Individualized Education Plan (IEP). In the Wehby et al. (2003) study, they used Open Court (M. J. Adams et al., 2000) and PALS (Fuchs et al., 1997) with eight students in the second to fourth grade in a self-contained classroom for students with EBD.

For behavioral measures, in Lingo et al. (2006) the overlapping data points for appropriate and inappropriate behaviors were not calculated due to some participants having only one data point during baseline, while Wehby et al. (2003) had a pooled overlap of data points across students and behavioral measures (i.e., time attending, total inappropriate behavior) of 186/192 (97%). For Wehby et al. (2003), the overlapping data point percentage suggests that 3% of the data points during treatment outperformed the data point with the greatest positive effect during baseline. These results also varied at the student level when comparing the intervention and baseline conditions on level, trend, and variability. For reading outcomes, Lingo et al. (2006) found (a) all students had an increase in level and trend on instructional- and grade-level text ORF measures; (b) overlapping data points, pooled across students, on grade-level ORF measures were 4/34 (12%) and instructional-level ORF measures were 27/88 (31%); and (c) six out of seven students had pretest to posttest gains on the Woodcock Reading Mastery Test, Normative Update (WRMT-R NU; R. W. Woodcock, 1998b). The Wehby and colleagues (2003) study had a pooled overlap of data points across students of (a) 29/59 (49%) on NWF, (b) 19/52 (37%) on blending sounds, (c) 45/59 (83%) on sound naming, and (d) 21/59 (36%)

on sight words. In addition, results from all the reading measures varied at the student level when comparing the intervention and baseline conditions on level, trend, and variability, with the exception that sound blending had a higher level for all students during the intervention condition.

Phonics and fluency—One study implemented a phonics and fluency intervention (Scott & Shearer-Lingo, 2002) with two phonics-based treatment conditions: (a) Teach Your Child To Read in 100 Easy Lessons (Engelmann, Haddox, & Bruner, 1986), which included 10- to 15-min lessons focusing on letter-sound correspondences, and (b) Great Leaps (Campbell & Mercer, 1994), which included 10-min lessons on sounds, phrases, and a story. The second treatment placed a larger emphasis on fluency. Visual analysis was limited due to the frequency of data points for on-task behavior (pooled across three students there was a total of six data points during baseline, four data points during Intervention 1, and seven data points during Treatment 2) and ORF measures (pooled across three students there was a total of 10 data points during baseline, seven data points during Intervention 1, and 39 data points during Treatment 2). With four data points for on-task behavior and seven data points for ORF across three students, analysis of the effectiveness of the first treatment is unavailable. The second treatment condition showed a positive trend in on-task behavior for all three students and ORF had an increase in level and trend with a pooled overlap of data points across students of 1/39 (3%).

Multicomponent with comprehension—Three studies implemented a multicomponent with comprehension intervention. In a group design intervention that included PA, phonics, and listening comprehension, three 30-min sessions per week focused on beginning and ending sounds, knowledge of letter-sound correspondences, phonics, word studies, and deriving meaning from text (Gest & Gest, 2005). Results were reported descriptively, with the majority of the students in the treatment group showing greater gains from the first week to the last week on behavioral (i.e., on-task behavior) and reading (i.e., letter identification, letter-sound knowledge, word identification, ORF) measures compared with the control condition. No comprehension measures were administered.

In a two-treatment group design study with no comparison condition (Strayhorn & Bickel, 2002), a PA, phonics/word reading, and comprehension intervention was implemented. The amount of time students received the intervention varied in each of the two treatment conditions (102 hr of intervention vs. 19 hr of intervention). The behavior measure for this study included the Attention Deficit/Hyperactivity Disorder (ADHD) Symptom Checklist 4 (Gadow & Sprafkin, 1997), which had items for ADHD and Oppositional Defiant Disorder (ODD). This study did not find a significant difference between conditions and no reading measures were reported.

The final multicomponent with reading comprehension intervention study (Kamps et al., 1994) implemented a single-case design study for students with autism. This study had 25- to 30-min lessons using a classwide peer-tutoring model with the following components: (a) students read a passage, (b) feedback from peers included comprehension questions, (c) error correction, and (d) scores were read to the teacher and scores are posted in the classroom (Kamps et al., 1994). The authors reported that all students improved in the duration and mean

length of time of their social interactions. Two out of three students improved on ORF and comprehension measures, and one out of three students improved on the number of errors while reading a passage. For the measures where visual analysis was available, social interaction levels were higher during the intervention condition than the baseline condition and the pooled overlap of data points across students was 35/86 (41%), while reading comprehension had a higher level during treatment than baseline, there was an immediacy of effect, and the pooled overlap of data points across students was 51/81 (63%).

Other interventions—Two studies were categorized as “other” due to unique qualities of each of these reading interventions. The first study (Burke et al., 2003) used a single-case design and a FBA to investigate the hypothesis that a single student’s problem behavior was maintained through escaping a difficult task. Based on the FBA, the student was pretaught the vocabulary words prior to the beginning of the lesson. The results ranged from a 53% to a 61% increase in on-task behaviors. No reading measures were given.

The second study was quasi-experimental (Feldman, 1981), using the “Sentence Approach,” which builds on a student’s understanding of language in larger units and the understanding of language in context (Jansky, 1981). Significant group differences were found by the researcher on posttest measures of on-task behavior and students ignoring teacher prompts favoring the treatment condition. However, not enough information was provided to calculate an effect size or *p* value. No reading measures were administered.

Reading and Behavioral Outcome Associations

To evaluate the association between reading and behavioral outcomes, studies with both reading and behavioral outcomes were descriptively evaluated comparing outcomes both within and across studies. This synthesis included 11 studies with reading and behavioral outcomes, including five studies that were group design and six studies that were single-case design.

Group design—Five studies used a group design and reported reading and behavioral outcomes. Three of these studies (Lane, Fletcher, et al., 2007; Nelson et al., 2005; O’Shaughnessy & Swanson, 2000) reported statistical group differences at posttest on reading measures with effect sizes ranging from $g = 0.40$ to $g = 2.07$. One additional study descriptively reported group differences at posttest (Gest & Gest, 2005), and one quasi-experimental study reported nonsignificant group differences at posttest (Scruggs & Osguthorpe, 1986). Of the studies that reported statistical group differences on reading measures at posttest, none of these studies found statistical group differences on behavioral measures at posttest (see Table 2).

In a reading intervention that reported outcomes descriptively, Gest and Gest (2005) focused on beginning and ending sounds, knowledge of letter-sound correspondences, phonics, word studies, and deriving meaning from text. They reported improved reading outcomes for the treatment condition at posttest. They also reported on-task performance by disaggregating data based on the students with the lowest and highest pretest reading scores (i.e., letter-sounds, word reading). The students with the lowest pretest reading scores all improved from the beginning to the end of the intervention in time on-task (total percentage of time on-task per student ranged from an improvement of 6% to 11%), while the comparison condition showed

little to no change in their time on-task (total percentage of time on-task per student ranged from a decrease of 10% to an increase of 1%). For the students with the highest pretest reading scores, the authors reported (scores not provided), four of the six students in the treatment condition improved in on-task behavior, while three of the four students in the comparison condition decreased in their time on-task.

In a quasi-experimental study (Scruggs & Osguthorpe, 1986) using a peer-tutoring intervention, researchers did not report statistically significant group differences at posttest on criterion tests of percentage of words read correctly or the Woodcock Johnson Psycho Educational Battery (word attack, sight word reading, reading comprehension; R. Woodcock, 1978). This intervention consisted of using modified versions of Beginning Reading I and II (Harrison, 1982). There were also no statistically significant group differences on the behavior measures (Attitude Toward School; Marascuilo & Levin, 1968), yet they reported pretest to posttest statistically significant differences for the treatment condition.

Overall, interventions that utilized group designs provide evidence of improved reading outcomes for students receiving the aforementioned treatments. However, these results indicate that various dependent variables pertaining to engagement and behavior were not associated with statistically significant differences.

Single-case design evaluations—Reading and behavioral outcomes were reported in six studies. Furthermore, all six interventions included at least one reading outcome (e.g., fluency, reading errors, comprehension) where an increase in level was observed for all students (Lane, Little, et al., 2007; Lingo et al., 2006; Scott & Shearer-Lingo, 2002; Wehby et al., 2003), with two of these studies showing positive gains in behavioral measures (Kamps et al., 1994; Lane et al., 2002). In the Kamps et al. (1994) study, a classwide peer tutoring, multicomponent with reading comprehension intervention was implemented. Higher levels of reading comprehension (researcher-designed measure) and social interactions during the treatment condition as compared with baseline or withdrawal phases were reported. The pooled overlapping data points across students for the (a) comprehension measure was 51/81 (63%) and (b) social interaction measure was 35/86 (41%). In the other study with positive gains in reading and behavioral measures, Lane et al. (2002) found lower levels of TDB and higher levels on Dynamic Indicators of Basic Early Literacy Skills (DIBELS) (Kaminski & Good, 1996) NWF when comparing baseline and intervention. Pooled overlapping data points across students were as follows: (a) TDB was 4/18 (22%) and (b) DIBELS NWF was 1/18 (6%). The other four studies (Lane, Little, et al., 2007; Lingo et al., 2006; Scott & Shearer-Lingo, 2002; Wehby et al., 2003) observed behavioral outcomes that included no clear differences in conditions on level, trend, variability, and immediacy of effect, while also having high levels of overlapping data points (ranging from 90% to 97% pooled across students per measure) in AET (Lane, Little, et al., 2007) and time attending and total inappropriate behavior (Wehby et al., 2003). In addition, two studies had a lack of sufficient data points (<1 data point in any condition) to visually analyze behavioral measures (Lingo et al., 2006; Scott & Shearer-Lingo, 2002).

Visual Analysis of Single-Case Design

Visual analysis, based on the proposed criteria by Horner and colleagues (2012), was used to evaluate the eight single-case design studies and 19 outcome measures on level, trend, variability, immediacy of effect, and overlap. The 19 outcome measures included eight behavioral measures from six studies (Beck et al., 2009; Burke et al., 2003; Kamps et al., 1994; Lane, Little, et al., 2007; Lane et al., 2002; Wehby et al., 2003) and 11 reading measures from six studies (Kamps et al., 1994; Lane, Little, et al., 2007; Lane et al., 2002; Lingo et al., 2006; Scott & Shearer-Lingo, 2002; Wehby et al., 2003). Outcomes from four studies (Kamps et al., 1994; Lane et al., 2002; Lingo et al., 2006; Scott & Shearer-Lingo, 2002) were excluded from visual analysis based on (a) data not reported visually and/or (b) not enough data points available for analysis (i.e., < 2 points in a condition; see Table 4).

Visual analysis of level on the behavioral measures included (a) four measures with positive findings in the intervention phase (Beck et al., 2009; Burke et al., 2003; Kamps et al., 1994; Lane et al., 2002), (b) three measures with results that varied across students (Lane, Little, et al., 2007; Wehby et al., 2003), and (c) no measures that favored the baseline phase. Visual analysis of level on the reading measures included (a) six measures with positive findings in the intervention phase (Kamps et al., 1994; Lane, Little, et al., 2007; Lane et al., 2002; Lingo et al., 2006; Scott & Shearer-Lingo, 2002; Wehby et al., 2003), (b) four measures with results that varied across students (Lane, Little, et al., 2007; Wehby et al., 2003), (c) one measure with no change across phases (Lingo et al., 2006), and (d) no measures that favored the baseline phase.

No change in trend between phases or conditions was observed across all students on a single measure in any study reviewed. Variability that favored the intervention phase for all students on behavioral measures (i.e., task engagement, TDB) was observed in two interventions (Burke et al., 2003; Lane et al., 2002). Visual analysis of variability in data points across all the other studies, phases, and conditions were either equal or varied at the student level. Immediacy of effect analysis was available on 16 measures. These data suggest an immediacy of effect was present across all students in one reading measure (i.e., reading comprehension; Kamps et al., 1994) and no behavioral measures.

In determining the overlap of data points between baseline and treatment, 16 outcomes measures were available for analysis. The pooled overlap was 516/929 (56%) for 5 behavior and 11 reading measures, suggesting that across all measures of reading and behavior 56% of the data points during the intervention phase had a greater positive outcome than the greatest positive outcome during the baseline phase.

The pooled overlap of data points (low overlap suggesting larger effect size) for behavioral measures was 271/341 (79%). The behavioral measures included (a) time attending/AET with a pooled overlap of data points equaling 136/145 (96%), (b) disruptive or inappropriate behaviors with a pooled overlap of data points equaling 100/116 (86%), and (c) duration of social interaction with a pooled overlap of data points at 35/86 (41%).

For the reading measures, the pooled overlap of data points was 245/588 (42%). Only one reading measure investigated reading comprehension with a pooled overlap of data points

equaling 51/81 (63%) and one measure investigated sight word reading with a pooled overlap of data points at 45/59 (76%). The remaining reading measures investigated ORF, NWF, or sound naming, which had a pooled overlap of data points at 149/448 (33%).

Overall, findings across these single-case studies included outcomes that varied across level, trend, variability, and immediacy of effect. However, for overlapping data points, reading effects were larger than behavioral measures. The pooled overlap of data points for reading measures was 42% as compared with 79% for behavior measures. In addition, the behavioral data points accounted for 53% of the total overlapping data points, yet only 37% of the total data points, while the reading data points accounted for 47% of the overlapping data points and 63% of the total data points.

Evaluation of Research Designs

The WWC determinants of study ratings were used to categorize studies as (a) meets WWC standards without reservations, (b) meets WWC standards with reservations, and (c) does not meet WWC standards. Of the seven group design studies, four studies were classified as meets WWC standards without reservations (Lane, Fletcher, et al., 2007; Nelson et al., 2005; O'Shaughnessy & Swanson, 2000; Strayhorn & Bickel, 2002), and three studies were classified as does not meet WWC standards (Feldman, 1981; Gest & Gest, 2005; Scruggs & Osguthorpe, 1986; see Table 5). For the single-case design studies, two studies met the criteria for meets WWC standards without reservations (Kamps et al., 1994; Lane, Little, et al., 2007), three studies did not meet WWC standards (Beck et al., 2009; Burke et al., 2003; Scott & Shearer-Lingo, 2002), one study did not include enough information for evaluation (Lane et al., 2002; see Table 6), and two studies' classification varied depending on which measure was evaluated (Lingo et al., 2006; Wehby et al., 2003).

In summary, while reading interventions tended to improve outcomes for participants on reading measures, they were not associated with improvements in behavioral measures. This lack of association between improvements in reading measures and behavioral measures does not change even if study findings are disaggregated based on studies that met the WWC with or without reservations and studies that did not meet the WWC standards.

Discussion

We reviewed the extant research on the effects of reading interventions on reading, behavioral, and social skill outcomes. Specifically, we isolated the independent variable to reading-only interventions for studies that investigated dependent measures of reading and behavior and social skills outcomes. We extended the findings of previously conducted syntheses (Nelson et al., 2011; Wanzek et al., 2006) by (a) including all students in Grades K-12, (b) broadening the criteria for type of literacy instruction, (c) including and systematically evaluating single-case design studies, (d) broadening the dependent measure requirement, and (e) not requiring a calculable effect size. The findings from group design studies reconfirmed the findings from the previous syntheses, suggesting that outcomes for students participating in reading interventions have positive effect sizes for reading outcomes and small or negative effect sizes for behavior and social skill outcomes. The findings of single-case design studies utilizing Horner and colleagues (2012) criteria for visual analysis indicated little evidence in support of

reading interventions positively affecting behavior or social skill outcomes with the exception of two studies (see Table 4; Kamps et al., 1994; Lane et al., 2002). This finding was contrary to findings from Wanzek et al. (2006), which, based on single-case and single-group studies of reading interventions, indicated the possibility of improving behavior and social skill outcomes.

The findings from this synthesis indicated that reading interventions tended to improve outcomes for participants on reading measures but were not associated with improvements in behavioral or social skill measures. The reading interventions represented in this synthesis included phonics, PA and phonics, phonics and fluency, multicomponent comprehension, and other approaches (e.g., FBA). Regardless of intervention type, improvements in reading outcomes were consistent. The majority of studies did not report improvements in behavior or social skill outcomes. In addition, when comparing the associations between reading-only interventions and outcomes of reading and behavior or social skills, the findings consistently revealed improvements in reading with no improvements in behavior or social skill outcomes with the exception of two studies (see Table 4) (Kamps et al., 1994; Lane et al., 2002).

When considering the WWC determinants of study ratings, four of the group design studies were classified as meets WWC standards without reservations (Lane, Fletcher, et al., 2007; Nelson et al., 2005; O'Shaughnessy & Swanson, 2000; Strayhorn & Bickel, 2002) and three were classified as does not meet WWC standards (Feldman, 1981; Gest & Gest, 2005; Scruggs & Osguthorpe, 1986). For the single-case design studies, two studies met the criteria for meets WWC standards without reservations (Kamps et al., 1994; Lane, Little, et al., 2007), three studies did not meet WWC standards (Beck et al., 2009; Burke et al., 2003; Scott & Shearer-Lingo, 2002), two studies had ratings that varied by measure used (Lingo et al., 2006; Wehby et al., 2003), and one study's visual analysis was unavailable, therefore no study rating was determined (Lane et al., 2002). However, regardless of quality classification, the finding of improvements on reading measures with no associated improvements for behavior or social skills remained consistent.

In summary, findings from this synthesis suggest that reading instruction alone may not be sufficient to improve behavior or social skill outcomes. Studies included in this study provide insufficient evidence of the hypothesized model that reading problems cause behavior problems (Hinshaw, 1992; Spira & Fischel, 2005) as the many methods of analysis showed that improving reading did not tend to also improve behavior or social skill outcomes.

Limitations

This synthesis has limitations similar to Wanzek et al. (2006) and Nelson et al. (2011) in that there was, overall, (a) a lack of total number of studies, (b) a lack of high quality of studies based on the WWC study determinants (IES, 2014), and (c) methodological limitations (e.g., insufficient information provided to recalculate effect sizes). In addition, it is possible that some studies did not report the use of a behavioral independent variable in their methods such as the use of a token economy with some or all participants. The lack of studies warrants caution when interpreting the results from this synthesis. Methodological issues in comparing across single-case design studies arise as there is no agreed upon effect size calculation for these studies (Horner et al., 2012), and with a limited number of group design studies that included

enough information for a calculable effect size ($n = 3$; Lane et al., 2007; Nelson et al., 2005; O'Shaughnessy & Swanson, 2000), comparing findings across these studies is problematic.

Overall, more research needs to be done in exploring the relationship between reading deficits and problem behavior to address these limitations.

Implications for Future Research and Practice

The findings from this synthesis suggest that further research is needed to adequately answer the research questions raised in this synthesis. Considering that (a) problem behaviors are one of the most reliable predictors of negative outcomes in response to intervention (Torgesen et al., 1999), (b) there is a negative relationship between students who respond to intensive reading interventions and problem behavior (Hagan-Burke et al., 2011), and (c) a limited number of studies in this synthesis were classified as meets WWC standards without reservations, new studies with robust research designs are needed to explore the relationship between reading and behavioral outcomes. This relationship should also be further explored in Grades 6 to 12 as only two studies (Lingo et al., 2006; Scott & Shearer-Lingo, 2002) investigated the effects of reading interventions on behavior or social skill outcomes in older students. We also suggest examining the long-term effects of improved reading outcomes on social-behavioral functioning as the effects of improved reading may require more time to influence social-behavioral outcomes.

Finally, future research could explore differences in behavior and social skill outcomes based on the type of measures used (i.e., direct, indirect) and the behavior measured (i.e., externalizing, internalizing, social skills). For type of measure used, direct measures (e.g., observations) were used in all single-case design studies and two of the group design studies (Feldman, 1981; Lane, Fletcher, et al., 2007) and measured either externalizing behavior or social skills. Indirect measures (e.g., teacher survey) were used in the remaining group design studies with only two studies measuring internalizing behaviors (Nelson et al., 2005; Scruggs & Osguthorpe, 1986). Differences in outcomes between single-case design and group designs could be based on measure type and/or the behavior measured as direct measures of social behavior have low to moderate correlations with indirect measures (Cost & Simpson, 2004) and externalizing and internalizing behaviors each have unique behavior patterns (Hinshaw, 1992)

Acknowledgments

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by Grant P50 HD052117-07 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development.

Biographies

Garrett J. Roberts is a project coordinator and doctoral student at The Meadows Center for Preventing Educational Risk at The University of Texas at Austin. Currently, he coordinates large-scale research projects. His research interests include academic and behavioral interventions, as well as effective instructional practices to support all learners.

Michael Solis is an assistant professor of special education in the Curry School of Education at the University of Virginia. He is experienced in running experimental and single-case design studies. His line of research focuses on vocabulary and reading comprehension interventions for students with reading difficulties in Grades 4 to 12.

Stephen Ciullo is an assistant professor at Texas State University in the curriculum and instruction department. His research focuses on improving content-area literacy outcomes for students in Grades 4 to 8 with learning disabilities and students in high-needs schools.

John W. McKenna is an assistant professor at St. John's University. His primary research interests are positive behavior supports and effective instructional strategies for students with Emotional Behavior Disorder (EBD) and students at risk. He is also interested in responsible inclusion and wraparound services.

Sharon Vaughn, H. E. Hartfelder/Southland Corp. Regents chair in human development, is the executive director of The Meadows Center for Preventing Educational Risk, an organized research unit at The University of Texas at Austin. She is the recipient of numerous awards and has authored more than 250 research articles.

References

References marked with an asterisk indicate studies included in the meta-analysis.

- Adams JW, Snowling MJ, Hennessy SM, Kind P. Problems of behavior, reading and arithmetic: Assessments of comorbidity using the Strengths and Difficulties Questionnaire. *British Journal of Educational Psychology*. 1999; 69:571–585. [PubMed: 10665170]
- Adams, MJ.; Bereiter, C.; Brown, A.; Campione, J.; Carruthers, I.; Case, R.; Treadway, GH, Jr. Open court reading. Columbus, OH: SRA/McGraw-Hill; 2000.
- Al Otaiba S, Fuchs D. Characteristics of children who are unresponsive to early literacy intervention: A review of the literature. *Remedial and Special Education*. 2002; 23:300–316.
- *. Beck M, Burns MK, Lau M. The effect of preteaching reading skills on the on-task behavior of children identified with behavioral disorders. *Behavioral Disorders*. 2009; 34(2):91–99.
- Breslau J, Miller E, Breslau N, Bohnert K, Lucia V, Schweitzer J. The impact of early behavior disturbances on academic achievement in high school. *Pediatrics*. 2009; 123:1472–1476. [PubMed: 19482756]
- Bruhn AL, Lane KL, Hirsch SE. A review of tier 2 interventions conducted within multitiered models of behavioral prevention. *Journal of Emotional and Behavioral Disorders*. 2014; 22(3):171–189.
- Bullis, M.; Walker, HM. Comprehensive school-based systems for troubled youth. Eugene: Center on Human Development, University of Oregon; 1994.
- *. Burke MD, Hagan-Burke S, Sugai G. The efficacy of function-based interventions for students with learning disabilities who exhibit escape-maintained problem behaviors: Preliminary results from a single-case experiment. *Learning Disability Quarterly*. 2003; 26:15–25.
- Campbell, SC.; Mercer, C. Great leaps reading. Micanopy, FL: Diarmuid; 1994.
- Chall, JS. Stages of reading development. 2. Fort Worth, TX: Harcourt Brace; 1996.
- Cook CR, Gresham FM, Kern L, Barreras RB, Thornton S, Crews SD. Social skills training for secondary students with emotional and/or behavioral disorders: A review and analysis of the meta-analytic literature. *Journal of Emotional and Behavioral Disorders*. 2008; 16:131–144.
- Cooper, HM.; Hedges, LV., editors. The handbook of research synthesis. New York, NY: Russell Sage; 1994.
- Cost HC, Simpson RG. A behavioral assessment investigation: Direct observations and behavioral ratings. *Research in Schools*. 2004; 11(2):51–60.

- Edmonds MS, Vaughn S, Wexler J, Reutebuch C, Cable A, Tackett KK, Schnakenberg JW. A synthesis of reading interventions and effects on reading comprehension outcomes for older struggling readers. *Review of Educational Research*. 2009; 79:262–300. [PubMed: 20072704]
- Engelmann, S.; Haddox, P.; Bruner, E. *Teach your child to read in 100 easy lessons*. New York, NY: Simon & Schuster; 1986.
- Engelmann, S.; Meyer, L.; Carnine, L.; Becker, W.; Eisele, J.; Johnson, G. *Corrective reading program*. Columbus, OH: SRA/McGraw-Hill; 1999.
- Epstein, MH.; Sharma, JM. *Behavioral and emotional rating scale: A strength-based approach to assessment*. Austin, TX: PRO-ED; 1998.
- *. Feldman RS. Classroom observations of first-graders and teachers in an experimental reading program. *Annals of Dyslexia*. 1981; 31:165–174.
- Flynn LJ, Zheng X, Swanson HL. Instructing struggling older readers: A selective meta-analysis of intervention research. *Learning Disabilities Research & Practice*. 2012; 27:21–32.
- Fuchs D, Fuchs LS, Mathes PG, Simmons DC. Peer-assisted learning strategies: Making classrooms more responsive to diversity. *American Educational Research Journal*. 1997; 34:174–206.
- Gadow, KD.; Sprafkin, J. *ADHD Symptom Checklist–4 manual*. Stony Brook, NY: Checkmate Plus; 1997.
- *. Gest SD, Gest JM. Reading tutoring for students at academic and behavioral risk: Effects on time on-task in the classroom. *Education & Treatment of Children*. 2005; 28:25–47.
- Gresham FM, Cook CR, Crews SD, Kern L. Social skills training for children and youth with emotional and behavioral disorders: Validity considerations and future directions. *Behavioral Disorders*. 2004; 30(1):32–46.
- Gresham, FM.; Elliot, SN. *Social skills rating system*. Circle Pines, MN: American Guidance Service; 1990.
- Gresham, FM.; Kern, L. Internalizing behavior problems in children and adolescents. In: Rutherford, RB.; Quinn, MM.; Mathur, SR., editors. *Handbook of research in emotional and behavioral disorders*. New York, NY: Guilford Press; 2004. p. 262-281.
- Gresham FM, Sugai G, Horner RH. Interpreting outcomes of social skills training for students with high-incidence disabilities. *Exceptional Children*. 2001; 67:331–344.
- Hagan-Burke S, Kwok OM, Zou Y, Johnson C, Simmons D, Coyne MD. An examination of problem behaviors and reading outcomes in kindergarten students. *The Journal of Special Education*. 2011; 45:131–148.
- Harris PJ, Oakes WP, Lane KL, Rutherford RB Jr. Improving the early literacy skills of students at risk for internalizing or externalizing behaviors with limited reading skills. *Behavioral Disorders*. 2009; 34(2):72–90.
- Harrison, GV. *Beginning reading: A professional guide for the lay tutor*. Blanding: METRA, Utah Navajo Industries; 1982.
- Hinshaw SP. Externalizing behavior problems and academic underachievement in childhood and adolescence: Causal relationships and underlying mechanisms. *Psychological Bulletin*. 1992; 111:127–155. [PubMed: 1539086]
- Horner R, Swaminathan H, Sugai G, Smolkowski K. Considerations for the systematic analysis and use of single-case research. *Education & Treatment of Children*. 2012; 35:269–290.
- Institute of Education Sciences. *What Works Clearinghouse study review standards*. 2014. Retrieved from http://ies.ed.gov/ncee/wwc/pdf/reference_resources/wwc_procedures_v3_0_standards_handbook.pdf
- Jansky JJ. The clinician in the classroom: A first-grade intervention study. *Bulletin of the Orton Society*. 1981; 31:145–164.
- Joseph LM, Eveleigh EL. A review of the effects of self-monitoring on reading performance of students with disabilities. *The Journal of Special Education*. 2011; 45:43–53.
- Kaminski RA, Good RH III. Toward a technology for assessing basic early literacy skills. *School Psychology Review*. 1996; 25:215–227.

- *. Kamps DM, Barbetta PM, Leonard BR, Delquadri J. Classwide peer tutoring: An integration strategy to improve reading skills and promote peer interactions among students with autism and general education peers. *Journal of Applied Behavior Analysis*. 1994; 27:49–61. [PubMed: 8188563]
- Kauffman JM, Landrum TJ. Politics, civil rights, and disproportional identification of students with emotional and behavioral disorders. *Exceptionality*. 2009; 17:177–188.
- Kavale KA, Forness SR. Social skill deficits and learning disabilities: A meta-analysis. *Journal of Learning Disabilities*. 1996; 29:226–237. [PubMed: 8732884]
- Kazdin AE. Treatment of antisocial behavior in children: Current status and future directions. *Psychological Bulletin*. 1987; 102:187–203. [PubMed: 3310058]
- Kratochwill, TR.; Hitchcock, J.; Horner, RH.; Levin, JR.; Odom, SL.; Rindskopf, DM.; Shadish, WR. Single-case designs technical documentation: What Works Clearinghouse. 2010. Retrieved from http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf
- *. Lane KL, Fletcher T, Carter EW, Dejud C, DeLorenzo J. Paraprofessional-led phonological awareness training with youngsters at risk for reading and behavioral concerns. *Remedial and Special Education*. 2007; 28:266–276.
- *. Lane KL, Little MA, Redding-Rhodes J, Phillips A, Welsh MT. Outcomes of a teacher-led reading intervention for elementary students at risk for behavioral disorders. *Exceptional Children*. 2007; 74:47–70.
- *. Lane KL, Wehby JH, Menzies HM, Gregg RM, Doukas GL, Munton SM. Early literacy instruction for first-grade students at-risk for antisocial behavior. *Education & Treatment of Children*. 2002; 25:438–458.
- *. Lingo AS, Slaton DB, Jolivet K. Effects of corrective reading on the reading abilities and classroom behaviors of middle school students with reading deficits and challenging behavior. *Behavioral Disorders*. 2006; 31:265–283.
- Malecki CK, Elliot SN. Children's social behaviors as predictors of academic achievement: A longitudinal analysis. *School Psychology Quarterly*. 2002; 17:1–23.
- Marascuilo, L.; Levin, JR. Evaluation report for the Berkeley Unified School District. Berkeley: University of California; 1968.
- Markwardt, FC. Peabody Individual Achievement Test–Revised. Circle Pines, MN: American Guidance Service; 1989.
- Miles SB, Stipek D. Contemporaneous and longitudinal associations between social behavior and literacy achievement in a sample of low-income elementary school children. *Child Development*. 2006; 77:103–117. [PubMed: 16460528]
- Mooney P, Ryan JB, Uhing BM, Reid R, Epstein MH. A review of self-management interventions targeting academic outcomes for students with emotional and behavioral disorders. *Journal of Behavioral Education*. 2005; 14:203–221.
- Morgan PL, Farkas G, Tufis PA, Sperling RA. Are reading and behavior problems risk factors for each other? *Journal of Learning Disabilities*. 2008; 41:417–436. [PubMed: 18768774]
- Nelson JR, Benner GJ, Gonzalez J. Learner characteristics that influence the treatment effectiveness of early literacy interventions: A meta-analytic review. *Learning Disabilities Research & Practice*. 2003; 18:255–267.
- Nelson JR, Benner GJ, Lane K, Smith BW. Academic achievement of K-12 students with emotional and behavioral disorders. *Exceptional Children*. 2004; 71:59–73.
- Nelson, JR.; Cooper, P.; Gonzalez, J. Stepping stones to literacy. Longmont, CO: Sopris West; 2004.
- Nelson JR, Lane KL, Benner GJ, Kim O. A best evidence synthesis of literacy instruction on the social adjustment of students with or at-risk for behavior disorders. *Education & Treatment of Children*. 2011; 34(1):141–162.
- *. Nelson JR, Stage SA, Epstein MH, Pierce CD. Effects of a prereading intervention on the literacy and social skills of children. *Exceptional Children*. 2005; 72:29–45.
- *. O'Shaughnessy TE, Swanson HL. A comparison of two reading interventions for children with reading disabilities. *Journal of Learning Disabilities*. 2000; 33:257–277. [PubMed: 15505964]
- Ryan JB, Reid R, Epstein MH. Peer-mediated intervention studies on academic achievement for students with EBD: A review. *Remedial and Special Education*. 2004; 25:330–341.

- Scammacca NK, Roberts G, Vaughn S, Stuebing KK. A meta-analysis of interventions for struggling readers in Grades 4–12: 1980–2011. *Journal of Learning Disabilities*. 2013 Advance online publication. 10.1177/0022219413504995
- *. Scott TM, Shearer-Lingo A. The effects of reading fluency instruction on the academic and behavioral success of middle school students in a self-contained EBD classroom. *Preventing School Failure: Alternative Education for Children and Youth*. 2002; 46:167–173.
- *. Scruggs TE, Osguthorpe RT. Tutoring interventions within special education settings: A comparison of cross-age and peer tutoring. *Psychology in Schools*. 1986; 23:149–193.
- Simonsen B, Fairbanks S, Briesch A, Myers D, Sugai G. Evidence-based practices in classroom management: Considerations for research to practice. *Education & Treatment of Children*. 2008; 31:351–380.
- Spencer VG. Peer tutoring and students with emotional or behavioral disorders: A review of the literature. *Behavioral Disorders*. 2006; 31:204–222.
- Spira EG, Fischel JE. The impact of preschool inattention, hyperactivity, and impulsivity on social and academic development: A review. *Journal of Child Psychology and Psychiatry*. 2005; 46:755–773. [PubMed: 15972069]
- Stanovich KE, Cunningham AE, Cramer B. Assessing phonological awareness differences in early reading acquisition: Issues of task comparability. *Journal of Experimental Child Psychology*. 1984; 38:175–190.
- *. Strayhorn JM Jr, Bickel DD. Reduction in children's symptoms of attention deficit hyperactivity disorder and oppositional defiant disorder during individual tutoring as compared with classroom instruction. *Psychological Reports*. 2002; 91:69–80. [PubMed: 12353805]
- Torgesen, JK.; Bryant, BR. Phonological awareness training for reading. Austin, TX: PRO-ED; 1994a.
- Torgesen, JK.; Bryant, BR. Test of phonological awareness. Austin, TX: PRO-ED; 1994b.
- Torgesen, JK.; Houston, DD.; Rissman, LM.; Decker, SM.; Roberts, G.; Vaughn, S.; Lesaux, N. Academic literacy instruction for adolescents: A guidance document from the Center on Instruction. Portsmouth, NH: Center on Instruction, RMC Research Corporation; 2007. Retrieved from <http://files.eric.ed.gov/fulltext/ED521837.pdf>
- Torgesen JK, Wagner RK, Rashotte CA, Rose E, Lindamood P, Conway T, Garvan C. Preventing reading failure in young children with phonological processing disabilities: Group and individual responses to instruction. *Journal of Educational Psychology*. 1999; 91:579–593.
- Vannest KJ, Temple-Harvey KK, Mason BA. Adequate yearly progress for students with emotional and behavioral disorders through research-based practices. *Preventing School Failure: Alternative Education for Children and Youth*. 2009; 53:73–84.
- Wagner, RK.; Torgesen, JK.; Rashotte, GA. Comprehensive Test of Phonological Processing (CTOPP). Longmont, CO: Sopris West; 1999.
- Wanzek J, Vaughn S, Kim AH, Cavanaugh CL. The effects of reading interventions on social outcomes for elementary students with reading difficulties: A synthesis. *Reading & Writing Quarterly*. 2006; 22:121–138.
- Wanzek J, Vaughn S, Scammacca NK, Metz K, Murray CS, Roberts G, Danielson L. Extensive reading interventions for students with reading difficulties after grade 3. *Review of Educational Research*. 2013; 83:163–195.
- Wanzek J, Wexler J, Vaughn S, Ciullo S. Reading interventions for struggling readers in the upper elementary grades: A synthesis of 20 years of research. *Reading and Writing*. 2010; 23:889–912. [PubMed: 21072128]
- *. Wehby JH, Falk KB, Barton-Arwood S, Lane KL, Cooley C. The impact of comprehensive reading instruction on the academic and social behavior of students with emotional and behavioral disorders. *Journal of Emotional and Behavioral Disorders*. 2003; 11:225–238.
- Woodcock, R. Woodcock-Johnson Psychoeducational Battery. Boston, MA: Teaching Resources; 1978.
- Woodcock, R. Woodcock Reading Mastery Tests–Revised. Circle Pines, MN: American Guidance Service; 1987.
- Woodcock, RW. Woodcock Reading Mastery Tests–Revised. Itasca, IL: Riverside; 1998a.

Woodcock, RW. Woodcock Reading Mastery Test–Revised Normative Update. Circle Pines, MN: American Guidance Service; 1998b.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1

Behavioral and Social Skills Definition.

Behavioral	Externalizing	Behaviors that are directed outwardly at the social environment (e.g., opposition, defiance, aggression; Gresham & Kern, 2004)
	Internalizing	Behaviors that are directed inwardly at the individual exhibiting the behavior (e.g., withdrawal, depression, anxiety; Gresham & Kern, 2004)
Social skills	Social skills	The behavior(s) that are performed after being taught and learned (Gresham, Sugai, & Horner, 2001). These behaviors include three domains: (a) social interaction, (b) prosocial behavior, and (c) social-cognitive skills (Cook et al., 2008; Gresham, Cook, Crews, & Kern, 2004)

Table 2

Overview of Group Design Studies.

Citation	Design	<i>n</i> / <i>M</i> age in years/ Grade	Risk type	Description of reading intervention	Behavioral measures	Behavioral findings	Reading measures	Reading findings
Gest and Gest (2005)	RCT	<i>N</i> = 17 Age = NR Grade 2	At risk of reading and behavior difficulty	PA, phonics, and LC	On-task behavior	Descriptive results provided. Students with low pretest scores: All four treatment students improved, while the control group showed almost no change. Students with high pretest scores: Four out of six treatment students improved, while three out of four control group students decreased.	Basic Reading Inventory: Letter ID letter-sound knowledge	Descriptive results provided. Students in the tutored group tended to outperform students in the control group
Lane, Fletcher, Carter, Dejud, and DeLorenzo (2007)	RCT	<i>N</i> = 24 Age = 6.5 and 6.7 Grade 2	Poor reading skills and at risk of behavior difficulty	PA and spelling	NSI ^b TDB	<i>g</i> = 0.39 (<i>ns</i>) <i>g</i> = 0.34 (<i>ns</i>)	TOPA DIBELS NWF	<i>g</i> = 0.75* <i>ns</i>
Nelson, Stage, Epstein, and Pierce (2005)	RCT	<i>N</i> = 63 Age = 5.1 and 5.2 Grade K	At risk of reading and behavior difficulty	PA, phonics, and working memory	BERS: School functioning Interpersonal Intrapersonal	<i>g</i> = 0.56 (<i>ns</i>) <i>g</i> = 0.00 (<i>ns</i>) <i>g</i> = 0.63 (<i>ns</i>)	CTOPP-PA WRMT-R WI WRMT-R WA DIBELS LNF CTOPP-RN	<i>g</i> = 0.55* <i>g</i> = 0.99* <i>g</i> = 0.94* <i>g</i> = 0.88* <i>ns</i>
Strayhorn and Bickel (2002)	RCT	<i>N</i> = 22 Age = 7.5 Grade K-4	Reading difficulty/ ADHD	PA, phonics, word reading, and RC	Symptom checklist for ODD and ADHD	Two-treatment groups (high frequency & low frequency tutoring) with no control group. Researchers reported the high frequency group outperformed the low frequency group with no statistical significance	NA	NA
Feldman (1981)	Quasi-experimental	<i>N</i> = 40 Age = NR Grade 1	Struggling reader	Other (the sentence approach)	On-task behavior	<i>g</i> = NA* ; Not enough information to calculate <i>p</i> value or calculate an effect size <i>g</i> = NA* ; Not enough information to	NA	NA

Citation	Design	n/M age in years/ Grade	Risk type	Description of reading intervention	Behavioral measures	Behavioral findings	Reading measures	Reading findings
Scruggs and Osguthorpe (1986)	Quasi- experimental	N = 67 ^a Age = NR Grades 1 to 5	LD/EBD	PA and phonics	Attitude toward school	calculate <i>p</i> value or calculate an effect size g = NA (ns)	Criterion Test WJ-Psycho Educational Battery (WA, WR, & RC)	ns ns
O'Shaughnessy and Swanson (2000)	RCT with SCD for ORF	N = 45 Age = 7.8 Grade 2	Reading disability	PA and phonics	SSRS; Academic competence	T1 vs. C; g = 0.26 (ns) T2 vs. C; g = 0.36 (ns)	PAT word list WAT word list TOPA	T1 vs. C; g = 1.51*; T2 vs. C; g = 1.28*; T1 vs. C; g = 1.03*; T2 vs. C; g = 1.79*; T1 vs. C; g = 2.00*; T2 vs. C; g = 1.16*; T1 vs. C; g = 2.07*; T2 vs. C; g = 1.50*
					Problem behavior Social skills	T1 vs. C; g = -0.36 (ns) T2 vs. C; g = -0.46 (ns) T1 vs. C; g = 0.22 (ns) T2 vs. C; g = 0.27 (ns)	Phonemic deletion WRMT-R WI WRMT-R WA WRMT-R PC PIAT-R spelling Rhyming words sentence span ORF (SCD)	T1 vs. C; g = ns; T2 vs. C; g = ns T1 vs. C; g = 0.40*; T2 vs. C; g = 0.42* T1 vs. C; g = ns; T2 vs. C; g = 0.63 T1 vs. C; g = ns; T2 vs. C; g = 0.63 T1 vs. C; g = 1.16*; T2 vs. C; g = 1.06*; T1 vs. C; g = 1.40*; T2 vs. C; g = 0.83* T1 vs. C; g = ns; T2 vs. C; g = ns

Note. RCT = Randomized Controlled Trial; NR = Not Reported; PA = Phonological Awareness; LC = Listening Comprehension; NSI = Negative Social Interaction; TOPA = Test of Phonological Awareness (Torgesen & Bryant, 1994b); NWF = nonsense word fluency; NS = nonsignificant group differences at posttest; BERS = The Behavioral and Emotional Rating Scale (Epstein & Sharma, 1998); CTOPP-PA = Comprehensive Test of Phonological Processing Phonological Awareness Composite (Wagner et al., 1999); WRMT-R = Woodcock Reading Mastery Test-Revised (Woodcock, 1998a); WI = Word Identification; WA = Word Attack; PC = Passage Comprehension; LNF = Letter Naming Fluency; CTOPP-RN = Comprehensive Test of Phonological Processing Rapid Naming Composite; RC = Reading Comprehension; ODD = Oppositional Defiance Disorder; ADHD = Attention Deficit/Hyperactivity Disorder; Exp. = Experimental; LD = Learning Disabilities; EBD = Emotional Behavior Disorder; WR = Word Reading; ORF = oral reading fluency; SSRS = Social Skills Rating System (Gresham & Elliot, 1990); PAT = Phonological Awareness Training (i.e., T1); WAT = Word Analogy Training Group (i.e., T2); PIAT-R = Peabody Individual Achievement Test-Revised (Markwardt, 1989); TDB = total disruptive behavior; WJ = Woodcock-Johnson; SCD = single-case design.

^a Attrition not reported.

^b observation occurred on playground.

* $p < .05$ between groups at posttest.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3

Overview Single-Case Design Studies.

Citation	Design	n/Age (in years)/Grade	Risk type	Description of reading intervention	Behavioral measures	Behavioral findings	Reading measures	Reading findings
Beck, Burns, and Lau (2009)	AT	N = 2 Age = 6 and 9 Grades K and 3	EBD, At risk of behavior difficulty	Phonics	On-task behavior	The intervention condition outperformed the baseline condition.	NA	NA
Burke, Hagan-Burke, and Sugai (2003)	AT	N = 1 Age = 7 Grade 3	LD, reading and behavior difficulty	Other (vocabulary intervention based on FBA)	Task engagement	The treatment condition outperformed the baseline condition.	NA	NA
Kamps, Barbeta, Leonard, and Delquadri (1994)	MBL with ABAB	N = 3 Age = 8, 8, and 9 Grade 1/2 grade split class, 2 and 3	High functioning autism	Multicomponent with RC	Social interaction	The intervention condition outperformed the baseline condition.	ORF, reading errors, RC	Results were mixed across all measures
Lane, Little, Redding-Rhodes, Phillips, and Welsh (2007)	MBL across classrooms	N = 7 Age (M) = 7 Grade 1	At risk of reading and behavior difficulty	Phonics	Academic engaged time	During the intervention condition, four students showed decrease in variability but no change in level or slope. Experimental control was not established for three students.	DIBELS NWF and ORF	During the intervention condition, five out of seven students increased in NWF and all seven students increased in ORF
Lane et al. (2002)	MBL across groups	N = 5 Age (M) = 6.9 Grade 1	At risk of reading and behavior difficulty	Phonics and writing	TDB, NSI	TDB: All students had positive effects. NSI: All students were low during baseline and four out of five students had positive effects during the intervention.	DIBELS NWF and ORF	NWF: All students made growth. ORF: Results were mixed across students
Lingo, Slaton, and	Multiprobe	N = 7 Age (range)	LD, EBD, OHI, reading and behavior difficulty	PA, phonics, word reading	Appropriate behavior inappropriate behavior	Results were mixed across all measures.	ORF from Corrective Reading, Reading	The intervention condition outperformed the baseline condition

Citation	Design	n/Age (in years)/Grade	Risk type	Description of reading intervention	Behavioral measures	Behavioral findings	Reading measures	Reading findings
Jolivette (2006)		N = 11 Age = 6-7					Errors, and WRMT-R NU	on ORF and WRMT-R (NU)
Scott and Shearer-Lingo (2002)	MBL across students	N = 3 Age = NR Grade 7	EBD, reading difficulty	Phonics and fluency	On-task behavior	Limited behavioral analysis due to lack of data points Intervention 1: Results were mixed across students. Intervention 2: The intervention condition outperformed the baseline condition.	ORF	The intervention condition outperformed the baseline condition
Wehby, Falk, Barton-Arwood, Lane, and Cooley (2003)	MBL across groups	N = 8 Age (M) = 7.8 Grade 4	EBD, LD, ID, OHI, SLI	PA, phonics, word reading	Time attending, total inappropriate behavior	Results were mixed across students and measures.	NWF, blending, sound naming, sight words, segmentation probes, WRMT-R (Letter ID, WI, WA), CTOPP (PA, PM)	Results were mixed across all measures

Note. AT = Alternating Treatment; EBD = Emotional Behavior Disorder; LD = Learning Disabilities; FBA = functional behavior assessment; MBL = Multiple Baseline; ABAB = Reversal/Withdrawal; RC = Reading Comprehension; NR = Not Reported; ORF = oral reading fluency; NWF = nonsense word fluency; TDB = total disruptive behavior; NSI = Negative Social Interaction; OHI = Other Health Impairment; NU = Normative Update; WRMT-R = Woodcock Reading Mastery Test-Revised (Woodcock, 1998a); ID = Intellectual Disability; SLI = Speech Language Impairment; WI = Word Identification; WA = Word Attack; CTOPP = Comprehensive Test of Phonological Processing; PA = Phonological Awareness; PM = Phonological Memory.

Table 4

Visual Analysis of Single-Case Design Studies.

Study	DV	Level	Trend	Variability	Immediacy of effect	Overlap
Beck, Bums, and Lau (2009)	On-task behavior	Higher in intervention than baseline	Neutral in 3/4 of phases; 1 descending trend in baseline	Low across phases	NA	NA
Burke, Hagan-Burke, and Sugai (2003)	Task engagement	Higher in intervention than baseline	Neutral in all phases	Greater for baseline	NA	NA
Kamps, Barbetta, Leonard, and Delquadri (1994) ^a	Social interaction	Higher in intervention than baseline	Mixed across phases	High across phases	Mixed	14/40; 20/31; 1/15
Lane, Little, Redding-Rhodes, Phillips, and Welsh (2007)	Reading comp	Higher in intervention than baseline	Mixed across phases	High across phases	Yes, between baseline and intervention	21/35; 14/30; 16/16
	Academic engaged time	Mixed across phases and participants	Neutral across phases	Mixed across phases	No	6/7; 7/7; 7/7; 5/6; 5/7; 6/7
	DIBELS NWF	Higher in intervention than baseline	Baseline: Neutral or descending intervention; Neutral or increasing	Low across phases	No	6/7; 1/7; 1/7; 2/7; 6/6; 2/7; 0/7
	DIBELS ORF	Mixed; Slightly higher for intervention	Neutral across all phases	Mixed across phases	No	4/7; 0/7; 2/7; 4/7; 6/6; 2/7; 6/7
Lane et al. (2002) ^a	Total disruptive behavior	Lower in intervention than baseline	Baseline: Mixed across participants intervention; Neutral across participants	Higher during baseline than intervention	Mixed across participants	2/9; 4/9
	DIBELS NWF	Higher in intervention than baseline	Increasing in all phases	Low across phases	No	1/9; 0/9
Lingo, Slaton, and Jolivette (2006) ^{a,c}	ORF (grade level)	Slightly higher in intervention than baseline	Baseline: Neutral or descending intervention; Neutral or increasing	Low across phases	No	0/7; 0/5; 0/3; 0/3; 3/7; 1/6; 0/7
	ORF (instructional level)	Higher in intervention than baseline	Baseline: Neutral or descending intervention; Neutral or increasing	Low across phases	Mixed	10/21; 0/11; 0/6; 1/5; 16/19; 0/13; 0/13
	Reading error (grade and instructional level)	Low levels	Neutral across all phases	Low across all phases	No	Unable to determine overlap
Scott and Shearer-Lingo (2002) ^b	ORF from great leaps	Higher in intervention than baseline	Baseline: Neutral Intervention: Neutral or increasing	Low across all phases	NA (insufficient data points)	0/15; 1/19; 0/5

Study	DV	Level	Trend	Variability	Immediacy of effect	Overlap
Webby, Falk, Barton-Arwood, Lane, and Cooley (2003)	Time attending	Mixed across all phases	Baseline: Neutral intervention: Neutral	High across all phases	No	16/16; 12/12; 15/15; 14/14; 8/11; 9/9; 10/10; 10/10
	Total Inappropriate Behavior	Mixed across all phases	Baseline: Neutral or increasing intervention: Neutral or increasing	High across all phases	No	11/11; 16/16; 15/15; 15/15; 9/9; 11/11; 7/11; 10/10
	Nonsense word fluency	Mixed across all phases	Baseline: Neutral or increasing intervention: Neutral or increasing	Mixed across all phases	Mixed across all phases	9/9; 2/9; 8/9; 1/8; 3/6; 1/6; 5/6; 5/6
	Blending sounds	Higher in intervention than baseline	Baseline: Neutral or increasing intervention: Neutral or increasing	Mixed across all phases	Mixed across all phases	0/9; 3/9; 2/9; 1/8; 2/4; 3/4; 4/4; 4/4
	Sound naming	Mixed & slightly higher in intervention than baseline	Baseline: Increasing intervention: Neutral or increasing	Low across all phases	Low	3/9; 2/9; 3/8; 1/9; 3/6; 6/6; 2/6; 1/6
	Sight words	Mixed across all phases	Mixed across all phases	Mixed across all phases	No	8/9; 2/9; 6/9; 6/8; 6/6; 6/6; 5/6; 6/6

Note. DV = Dependent Variable; mixed results varied by student; NWF = nonsense word fluency; ORF = oral reading fluency.

^aVisual analysis not available for one or more dependent variables.

^binsufficient data points for visual analysis on one or more dependent variables.

Table 5

Study Determinants and Ratings (Group Design).

Study	Study design	Sample Attrition (low, high)	Baseline equivalence	Evaluation result
Gest and Gest (2005)	RCT	Total: 7/24 (29%); treatment: 2/12 (17%); comparison: 5/12 (41%); difference: 24%; sample attrition: High	No	DNMS
Lane, Fletcher, Carter, Dejud, and DeLorenzo (2007)	RCT	Total: 1/25 (4%); treatment: 0%; comparison: 1/12 (8%); difference: 8%; sample attrition: Low	Not required for determining study rating	MS
Nelson, Stage, Epstein, and Pierce (2005)	RCT	Total: 21/84 (25%); treatment: 17/64 (27%); comparison: 4/20 (20%); difference: 7%; sample attrition: Low	Not required for determining study rating	MS
O'Shaughnessy and Swanson (2000)	RCT	No attrition; sample attrition: Low	Not required for determining study rating	MS
Strayhorn and Bickel (2002)	RCT	Total: 3/27 (11%); treatment: 1/10 (10%); comparison: 2/17 (12%); difference: 2%; sample attrition: Low	Not required for determining study rating	MS
Feldman (1981)	Quasi	Not required for determining study rating	No	DNMS
Scruggs and Osguthorpe (1986)	Quasi	Not required for determining study rating	No	DNMS

Note. RCT = Randomized Controlled Trial; DNMS = Does not meet WWC Standards; MS = Meets WWC Standards; MSWR = Meets WWC Standards With Reservations; WWC = What Works Clearinghouse.

Table 6

Study Determinants and Ratings (Single-Case Design).

Study	Clear manipulation of independent variable(s)?	3 attempts to demonstrate intervention effect?	Sufficient inter-assessor agreement reported?	Sufficient number of phases present based on design?	Sufficient number of data points per condition or phase?	Evaluation result
Beck, Burns, and Lau (2009)	Yes	Yes	Yes ^a	Yes	No; 3 points per condition	DNMS
Burke, Hagan-Burke, and Sugai (2003)	Yes	Yes	Yes	Yes	No; 3 points per condition	DNMS
Kamps, Barbetta, Leonard, and Delquadri (1994) ^b	Yes	Yes	Yes ^a	Yes	Yes for comprehension and duration	MS for comprehension and duration of social interaction
Lane, Little, Redding-Rhodes, Phillips, and Welsh (2007)	Yes	Yes	Yes ^a	Yes	Yes	MS
Lane et al. (2002) ^a	Visual analysis only available for 2 students; <i>M</i> , <i>SD</i> , and effect sizes given for all students on all dependent variables					
Lingo, Slaton, and Jolivet (2006)	Yes; instructional-level oral reading fluency and errors	Yes; instructional-level oral reading fluency and errors	Yes ^a ; instructional-level oral reading fluency and errors	Yes; instructional-level oral reading fluency and errors	Yes for four/seven students; instructional-level oral reading fluency and errors	MS, MSWR
	Yes; grade-level oral reading fluency and errors	Yes; grade-level oral reading fluency and errors	Yes ^a ; grade-level oral reading fluency and errors	Yes; grade-level oral reading fluency and errors	No; grade-level oral reading fluency and errors	MSWR
	Yes; appropriate behavior	Yes; for three/eight students; appropriate behavior	Yes ^a ; appropriate behavior	Yes; instructional-level fluency (four/seven students)	No	DNMS in gened setting; MSWR for three/eight students in Special Education setting
Scott and Shearer-Lingo (2002)	Yes	No	Yes ^a	Yes	No	DNMS
Wehby, Falk, Barton-Arwood, Lane, and Cooley (2003)	Yes	Yes	Yes ^a	Yes	No; NWF Yes for four/eight of students; sound blends No; letter ID No; sight words Yes for six/eight; time attending Yes for six/eight inappropriate behavior	MSWR for NWF MS and MSWR for blends MSWR for letter ID MSWR for sight word MS and MSWR for time attending MS and MSWR for inappropriate behavior

Note. DNMS = Does not meet WWC Standards; MS = Meets WWC Standards; MSWR = Meets WWC Standards With Reservations; NWF = nonsense word fluency; ID = Intellectual Disability; WWC = What Works Clearinghouse.

Percentage of interassessor agreement was not reported for each condition/phase.

^bVisual analysis not available for one or more dependent variables.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript