

# Evidence Based Reading Interventions for Students With EBD: A Systematic Review and Quality Evaluation of Research Methods

**Argnue Chitiyo\***

*Ball State University*

**Seth King**

*University of Iowa*

**Jessica Wojton**

*Ball State University*

**Chaidamoyo Dzenga**

*Tennessee Technological University*

---

*Students identified with Emotional/Behavioral Disorders (EBD) often exhibit behavioral challenges that negatively affect their academic performance. One of the weaker academic areas for these students lies in reading. Although numerous research on interventions to improve their reading outcomes has been conducted, some students continue to lag behind their typical peers. The use of instructional practices not supported by strong empirical evidence may deter students' reading outcomes from improving. To address the reading performance gap, educators are encouraged to use evidence-based practices supported by strong empirical research. In order to facilitate this evidence base, the Council for Exceptional Children (CEC) provides research quality indicators to evaluate interventions' studies for validity and rigor. This study applied the CEC quality indicators to evaluate the methodological rigor of research on reading interventions for students with EBD published between 2000 and 2020. Findings suggest that most studies generally meet a lot of the rigor standards set forth by the CEC. Researchers using single case designs (SCDs), however, may need to include effect size measures in their results to illustrate the quantitative effects of the interventions on reading outcomes. Implications for future research and practice are discussed.*

---

**Keywords:** Emotional disorder, behavior disorder, reading, evidence-based practice

## INTRODUCTION

The number of students aged between 3 and 17 identified with emotional and behavioral disorders (EBD) continues to rise, with recent statistics ranging above 7 percent of the total student population (CDC, 2021). Approximately 3.2% (i.e., about 1.9 million) of the same age-range students have a diagnosis for depression (CDC, 2021). EBD is a mental condition associated with chronic feelings of anxiety, depression, and challenging behaviors that can negatively impact students' academic performance and social functioning (Reid et al., 2004; Stoutjesdijk, 2013). Typical behaviors characterizing the disorder include pervasive mood of unhappiness, hy-

---

\*Please send correspondence to: Argnue Chitiyo, Department of Special Education, Ball State University, 2000 W University Avenue, Muncie, IN 47306, USA, Email: achitiyo@bsu.edu.

peractivity, aggression, withdrawal, and immaturity (Kauffman, 2001). According to research, this group of students exhibit challenging behaviors that negatively affect their academic performance (Kremer et al., 2016; Wagner & Cameto, 2004) and social skills (Garwood et al., 2017; Sullivan & Sadeh, 2014). Relative to similar age typical peers, students with EBD have been shown to perform approximately 2.2 grade levels behind in elementary school, and this gap tends to widen as they progress to high school (Wagner et al., 2005). One of the bigger areas of their academic weakness lies in reading (Nelson et al., 2004). In a study examining academic profiles of students with disabilities, Wagner et al. (2005) reported that students with EBD read approximately 2.2 grades below grade level. Another study examining the academic achievement of k-12 students with EBD (Nelson et al., 2004) indicated that 83% of the study's sample of children with EBD scored below students in the typical group on a standardized measure of reading skills. This gap in reading tends to widen with progression to higher grades where content becomes denser and more complex (McKenna et al., 2018; Nelson et al., 2004). Given the co-occurrence of academic challenges and behavioral problems, delayed efforts to address the reading skills deficits for students with EBD may become less effective (Petersen et al., 2018; Utchel et al., 2015).

In order to address the reading deficits of students with EBD, numerous research studies examining different types of interventions have been conducted. Various strategies have emerged as promising practices for addressing reading outcomes of students with EBD. For example, a prior study examining the effectiveness of repeated readings, error correction, and performance feedback strategies on the reading fluency of middle scholars with behavioral problems, reported improvements in reading rate, literal, and inferential reading across all participants included in the study (Alber-Morgan et al., 2007). Other studies have also examined and reported positive effects of concept, text, and cognitive maps on the reading comprehensions skills among students with EBD (e.g., Blankenship et al., 2005; Stone et al., 2008). Previous reviews have also synthesized research on specific interventions. For example, Ryan and colleagues (2004) assessed literature on peer mediated interventions on the academic outcomes of students with EBD. They reported positive reading outcomes for all types of peer mediated strategies that included classwide peer tutoring, cooperative learning, cross-age tutoring, peer tutoring, and peer assisted learning. Berner and colleagues (2010) conducted a meta-analysis of literature on the effects of reading instruction on reading skills of students with or at risk for EBD, in which they found literacy instruction to have a positive effect on the reading outcomes of students with EBD. Other previous reviews have also reported potentially effective interventions on the reading outcomes of students with EBD (e.g., Burke et al., 2015; Dunn et al., 2017; Garwood, 2017; Mooney et al., 2005).

### ***Reforms in Educational Research and Move Towards EBP***

Despite years of research on reading interventions for students with EBD, there still exists a performance gap, indicating that either students are not responding to interventions as expected, or teachers continue to use practices that are not empirically demonstrated to be effective (Chitiyo et al., 2020; Garwood et al., 2014). An existence of potentially ineffective practices in this area can be attributed to challenges associated with the processes of developing evidence-bases on the part of

policy makers, or challenges of finding the EBPs on the part of teachers (Cook et al., 2009; Cook and Cook, 2011). High quality research that supports effective practices may be difficult to establish. Prior to the adoption of EBP approach to instructional practices, the quality of experimental research was rarely regulated and was mostly left to the determination of individual journals (Gersten et al., 2005; Horner et al., 2005). The absence of universally agreed research quality standards potentially results in publications that do not address certain important methodological features such as reliability of data collection, implementation fidelity, and validity of measurement methods (Cook et al., 2014; Horner et al., 2005). Some interventions may therefore result from research whose methodological rigor is compromised, and therefore research findings flawed or biased.

In order to address this research-practice gap, there is increased emphasis for teachers to use evidence-based practices. These are interventions that have been demonstrated to be effective through rigorous experimental research (Cook et al., 2008; Cook et al., 2009). To facilitate the identification of the EBP, the Department of Education, and other affiliated organizations (e.g., CEC, What Works Clearing House) have created legislations or guidelines outlining the procedures for identifying EBP. Generally, EBP constitute four different criteria which are (1) studies are conducted using experimental designs demonstrating functional relationships between interventions and outcomes, (2) studies are of high quality, (3) each practice is supported by a larger quantity of studies, and (4) effects are synthesized across multiple studies. Regarding the use of experimental design, EBP approach requires researchers to use experimental designs that demonstrate strong correlations between interventions and outcomes. In special education, these designs constitute group and single case designs (Horner et al., 2005). Regarding the quality of studies, EBP requires that methodologies address all potential threat to internal and external validity. Administrators in special education have come up with different types of quality indicators (e.g., CEC, WWC) to guide assessment of the methodological rigor of interventions' research. The CEC quality indicators consist of 8 methodological domains, which are *context and setting*, *participants*, *intervention agents*, *description of practice*, *implementation fidelity*, *internal validity*, *outcome measures/dependent variables*, and *data analysis* (Cook et al., 2015). The dimension of quantity of research requires that a single intervention be supported by at least 2 experimental studies of high quality. The aim is that intervention effectiveness is demonstrated by replication of the study and findings across more participants in different settings. Finally, studies should demonstrate, in quantitative terms, the magnitude of effects of the interventions on outcomes of interest using standard effect sizes. Systematic reviews and meta-analyses are commonly used to synthesize effects of interventions across multiple studies (Maggin et al., 2011; Talbott et al., 2018).

This study focuses on the quality dimension of the EBP criteria for identifying effective reading interventions. To ensure that researchers address all potential threats to validity, quality standards guiding research designs have been developed. A special issue in the journal of *Exceptional Children* features publications on some of the common standards that have been developed in special education (*Exceptional Children*, 2005, Special Issue 2). For example, Horner et al. (2005) proposed a set of quality indicators for assessing methods in SCDs. Gersten and colleagues (2005)

developed quality indicators to evaluate group and quasi-experimental research in special education. They suggest using a research study organizer template specifying “critical issues for consideration in research” (p. 149). The WWC developed a set of quality indicators to assess the rigor of interventions’ research in education. Using the WWC codes, McKenna and colleagues (2017) evaluated the methodological features of interventions’ research on reading skills of students with or at risk for EBD. In their findings, they reported a need for researchers to use “stronger designs and place a greater emphasis on investigating the effects of reading instructional practices in inclusive settings” (p. 868). Although WWC quality indicators have been widely applauded for addressing the important elements in primary research for quality improvement, they focus on a few aspects of research design and evidence.

The Council for Exceptional Children (CEC) also proposed a set of quality indicators for both group and SCDs specifically for special education research (Cook et al., 2014). The codes consist of 28 indicators divided into 8 domains: *context and setting*, *participants*, *intervention agents*, *description of practices*, *implementation fidelity*, *internal validity*, *outcome measures*, and *data analysis*. Eighteen codes apply to both group and SCDs, whereas six apply only to group designs. Four codes apply to SCD studies only. When using CEC indicators, studies are expected to address a minimum number of indicators in order for them to qualify as meeting EBP standards. Proper application of these quality indicators allows practitioners to determine if research outcomes are valid and findings can be trusted for the ultimate benefit of students with special needs.

Although previous reviews on reading interventions for EBD (e.g., Berner et al., 2010; Burke et al., 2015; Garwood et al., 2018; Joseph et al., 2009; Ryan et al., 2004) provide some useful insights on the interventions commonly used in this field, most have not evaluated the rigor of the methodological elements used in the primary studies, which leaves questions regarding the validity of the conclusions regarding the effectiveness of the interventions currently in place. The application of CEC quality standards to assess the methodological rigor of interventions research in reading for EBD will fill this gap. The purpose of this research was therefore to assess the methodological quality of the primary studies published between 2000 to present using the CEC quality standards. More specifically, the study addresses the following research questions: 1) What are the methodological characteristics of the primary studies included in the review (i.e., research designs, demographic characteristics, intervention types, and dependent variables)? 2) To what extent do researchers of reading interventions for EBD address CEC quality indicators? 3) What are the evidence-base classifications (i.e., EBP, Potentially EBP, Mixed Effects) of different types of interventions featured in the studies?

## METHOD

### **Literature Search**

To identify primary studies on reading interventions for EBD, a three steps process consisting of database search, ancestral search, and hand-search was conducted. In the first step, an electronic databases search consisting of PsycInfo, CINAHL, PsycArticles, and EBSCO Management Collection databases was run using 2000 to

2020-year range. The following search terms were entered into the search engine:

(AB (“emotion\* behavior\* dis\*” OR “ebd” OR “behavior\* dis\*” OR “emotion\* dis\*”)) AND (AB (“read\*” OR “litera\*” OR “comprehen\*” OR “phon\*” OR “alphabet\*” OR “fluen\*” OR “vocabulary\*”)).

A total of 5737 abstracts were returned consisting of 4211 journal articles. The rest were dissertations (895), books (564), magazines (24), CEUs (13), trade publications (4), and newspaper articles (n = 4). We limited analysis to peer reviewed journal articles. After removing duplicates, 3848 abstracts remained for screening. Two leading authors read the titles and abstracts of retrieved peer reviewed articles independently to determine the studies meeting the initial inclusion criteria. If criteria for inclusion/exclusion could not be ascertained from reading the title or abstract, a full text of the article was extracted. The two reviewers independently coded the titles/abstracts for inclusion and compared their findings upon completion. The two reviewers agreed on 98% of the titles/abstracts. Articles on which the reviewers differed were discussed and a final decision for inclusion/exclusion was made in consultation. A total of 21 articles met criteria for inclusion. All studies that did not meet the initial selection criteria were discarded.

Next, an ancestral search was conducted. The primary author went through the references lists of literature reviews, meta-analyses, and other primary studies identified in the database search phase to locate any primary studies that may have been missed in the database search. An additional 9 articles were identified. Finally, a hand search of three top journals from which most studies on reading interventions for EBD were located was conducted for the years 1996 to 2018. These were *Behavioral Disorders*, *Remedial and Special Education*, and *Exceptional Children*. The hand-search did not yield any additional articles. A spreadsheet listing all the articles identified across the three phases was prepared. Following application of inclusion criteria, a total of 30 peer-reviewed journal articles were retained.

### ***Inclusion Criteria***

Two independent coders reviewed the articles to make sure that they met the following inclusion criteria: (1) The studies evaluated interventions targeting reading outcomes of students with or at risk of EBD. If studies mentioned an academic intervention without specifically targeting reading outcomes for EBD, they were excluded. (2) Studies had to be published in peer reviewed journals. (3) The studies were published between 2000 and 2020. This period covers a time frame during which major reforms on educational policies and practices were instituted (e.g., NCLB, 2002; Race to Top, 2015), which resulted in significant changes in approaches to educational research. Educational Research Standards reforms put forth by the What Works Clearinghouse (Kratochwill et al., 2010) also suggest that research older than 20 years may be discarded from the EBP reviews owing to changes that have taken place over the years. (4) The participants in the studies were identified as having or at-risk for emotional and behavioral disabilities, behavioral disorders, or behavioral problems. (5) Studies used SCDs or group experimental designs. Studies using non-experimental designs (e.g., qualitative approaches) were excluded. (6) Studies were conducted in English.

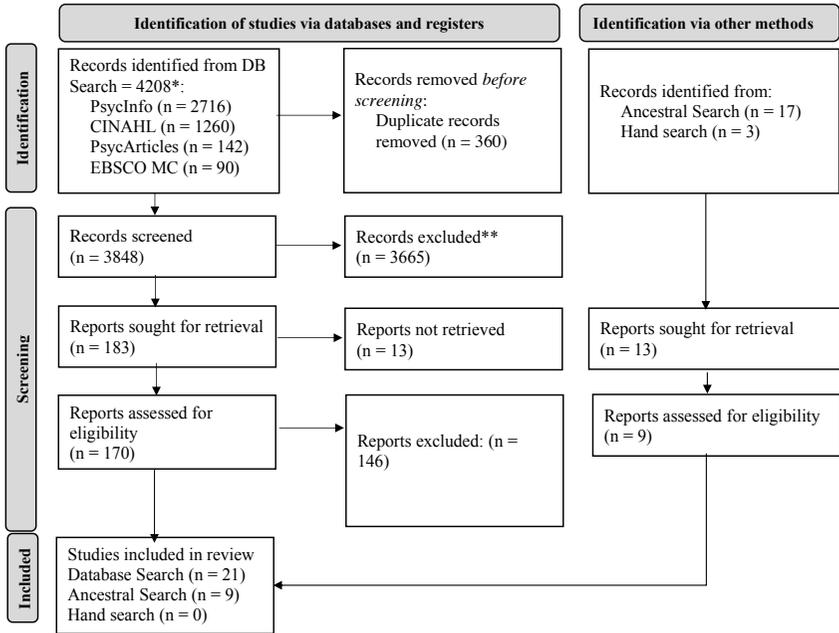


Figure 1. Literature Search

**Coder Training**

A graduate assistant was trained on the application of the CEC quality indicators for both group and SCDs. The primary author and the graduate student read the CEC quality indicators article (i.e., Cook et al., 2015), discussing the individual codes and their scoring criteria. Thereafter, the primary author and graduate assistant coded two practice articles independently for training purposes. The interrater agreement for the two articles was 85.7% (calculated by dividing the # of codes agreed on/total number of codes). Disagreements on the articles were discussed until consensus was reached. The two coders coded an additional 3 articles independently and results were compared upon completion of each article (IRA = 98%). Disagreements on any codes were discussed until consensus was reached. Finally, the two coders coded the first 16 articles independently (i.e., 53% of the articles), achieving an IRA of 98.7%. The remaining 14 articles were split evenly and coded separately.

**Coding Procedures**

Two separate codebooks denoting data for *narrative analysis* and *quality assessment* were created. The first codebook denoted descriptive characteristics of the primary studies, which include *research designs*, *participants’ demographic information*, *independent variables*, and *dependent variables*. The second part denoted the CEC quality indicators (Cook et al., 2014). The quality indicators (QIs) are divided into two sections consisting of (1) QIs to examine methodological soundness of the studies and (2) QIs for classifying the evidence base of practices on the basis of sound studies (Cook et al., 2014).

**Research design** codes pertained to the types of research designs used in the primary studies. Studies were coded as having used either SCD (i.e., reversal designs, multiple baselines across participants, multiple baselines across behaviors, and multiple probe designs) or group design. **Participants' demographics** pertained to the demographic characteristics of participants included in the studies. These included *number of participants, age, grade, diagnosis, gender, race, and socio-economic status (SES)* if available.

**Independent variables** codes pertained to the interventions featured in the studies. The subcategories included *types of interventions, instructional settings, and treatment fidelity*. **Dependent variables codes** referred to the specific outcomes targeted in the studies. The codes were split into four categories consisting of *reading fluency, reading comprehension, vocabulary, and phonemics*. The specific dependent variable codes were specifically created to facilitate further investigation for moderation analysis in a subsequent meta-analysis paper based on the interventions.

**Quality indicators** codes pertained to the CEC QIs (Cook et al., 2014). The QIs are divided into two sections consisting of (1) QIs to examine methodological soundness of the studies and (2) QIs for classifying the evidence base of practices on the basis of sound studies (Cook et al., 2014). The methodological soundness codes consist of 28 indicators divided into 8 domains consisting of *context and setting, participants, intervention agents, description of practices, implementation fidelity, internal validity, outcome measures, and data analysis*. Eighteen codes apply to both single case and group design studies, whereas six apply only to group designs. Four codes are specifically for SCD studies only. The codes are dummy coded (Yes and No) to indicate whether or not authors gave enough description of the specified elements. The classification codes for EBP consists of 5 categories namely 1) EBP, (2) potentially EBP, (3) practices with mixed effects, and (4) insufficient evidence, and (5) practices with negative effects (See Cook et al., 2014 for more detailed description of the codes). In order to classify the interventions across the 5 categories, we identified the interventions in each study, created categories for the specific types of interventions, and ascertained whether each intervention was supported by enough studies, participants, or effect sizes of the outcomes.

### **Data Analysis**

Data analyses involved reporting the percentages of studies meeting the respective QIs criteria. We entered all codes onto a spreadsheet and calculated descriptive statistics using Microsoft Excel. The descriptive statistics for SCDs were listed on a separate spreadsheet from the group designs. Missing information was noted as "not reported" with the exception of information pertaining to study participants. A combined spreadsheet for both group and SCDs was created to calculate the aggregate percentages for the quality indicators.

Interrater agreement (IRA) on abstract coding between two coders, calculated by dividing the number of agreements by the total number of abstracts, was 97.8%. Following coder training for CEC codes, the two coders coded the first 16 articles independently (i.e., 53% of the articles), comparing their findings after each article. The point-by-point IRA for the 16 articles totaled 98.7% ( $R = 94-100$ ). Any disagreements on the articles were discussed until consensus was reached. The remaining 14 articles were split evenly and coded separately.

## RESULTS

The literature search resulted in the identification of 30 studies meeting inclusion criteria. Most studies (73%;  $n = 22$ ) used SCDs, whereas 27% ( $n = 8$ ) used group designs. Researchers using SCDs used mostly multiple baseline design across participants (MBD-P; 95%,  $n = 21$ ) followed by multiple probe design (MPD; 5%,  $n = 1$ ).

**Participants' demographics.** There were on average, 4 participants across the 22 SCDs ( $R = 1 - 7$ ), and 48 participants in the group designs ( $R = 5 - 171$ ). In studies that reported participants' sex (83%;  $n = 25$ ), participants were dominantly male (80%,  $n = 121$ ) compared to females (20%,  $n = 38$ ). A fair number of the studies (27%,  $n = 8$ ) featured only male participants, whereas only one study featured female participants only. Information regarding participants' gender and age was not reported in 17% ( $n = 5$ ) and 20% ( $n = 6$ ) of the studies respectively. Information regarding participants' grade levels was reported in 70% ( $n = 21$ ) of the studies, and of these, 38% ( $n = 8$ ) targeted  $k - 3$ , whereas 57% ( $n = 12$ ) targeted grades 4 and above. Although this review primarily targeted studies involving students with or at risk for ED/EBD, some studies also featured participants with comorbid conditions. Studies that included participants with or at risk of ED/EBD only constituted 70% ( $n = 21$ ) of the studies, whereas 33% ( $n = 10$ ) of the studies featured comorbid disabilities. Comorbid diagnosis that featured more frequently included speech language impairment (SLI; 40%,  $n = 4$ ), ADHD (50%,  $n = 5$ ), OHI (30%,  $n = 3$ ), and LD (20%,  $n = 2$ ).

**Settings and intervention agents.** Interventions in 63% of the studies ( $n = 19$ ) were administered in self-contained or special education classrooms, whereas interventions in 13% of the studies ( $n = 4$ ) were conducted in general education classrooms. Interventions in 27% ( $n = 8$ ) of the studies were conducted in settings other than self-contained or special and general education classrooms. These included juvenile correctional facilities and specialized or residential treatment centers. Regarding intervention agents, teachers, tutors or instructors served as the treatment agents in 67% ( $n = 20$ ) of the studies, whereas researchers and graduate students delivered instruction in 23% ( $n = 7$ ) and 13% ( $n = 4$ ) of the studies respectively.

**Type/categories of interventions.** Curriculum-based interventions involve changes in reading materials, content, or scripted teaching procedures. These methods were used in 50% ( $n = 15$ ) of the studies. Practices used in curriculum-based methods included corrective reading, story/text/concept mapping, phonological awareness training for reading (PATR), stepping-stones to literacy, and other content-based reading programs (e.g., Scott Foresman Kindergarten, Orton Gillingham, and Great Leaps reading program). **Instruction-based** approaches involved methods of content administration or pedagogy, and these were used in 50% ( $n = 15$ ) of the studies. Instruction-based methods included repeated reading/error correction/performance feedback, direct instruction, PALS, choice antecedent instruction, self-graphing/self-regulated, and stimulus or reward-based methods.

**Dependent measures.** Most studies evaluated multiple types of reading skills. A preliminary analysis of the reading outcomes across the 30 studies revealed three main categories: phonics/phonological awareness, oral reading fluency, and comprehension. The most targeted reading outcomes were oral reading fluency, which featured in 77% ( $n = 23$ ) of the studies, followed by phonological awareness

(70%;  $n = 21$ ), and comprehension (37%;  $n = 11$ ) respectively. The most targeted phonological principles included phonemic blending or segmentation (48%,  $n = 10$ ), nonsense word fluency (48%,  $n = 10$ ), letter naming fluency (38%,  $n = 8$ ), initial sound fluency (38%,  $n = 8$ ), and vocabulary (19%,  $n = 4$ ). The skills acquisitions were measured as the number of phonemes, nonsense words, or letter sounds read per unit of time. Likewise, oral reading fluency in most studies was assessed by measuring the number of words and/or errors made per unit of time. Comprehension reading skills were assessed mostly through story retelling, summarization of passage main ideas, and responses to comprehension questions based on read passages. A summary of results for the narrative analysis are displayed in Table 1.1

### **Quality Assessment**

The 30 studies were evaluated for methodological rigor using the CEC quality indicators (Cook et al., 2015). Most studies (i.e., 93%; ( $n = 28$ ) provided detailed description of the study context and settings. Two studies that failed to address this indicator, and these stated the settings in which the participants were trained without describing the essential settings elements in which the interventions were administered. Although a few studies did not report the participants' ages or gender (i.e., 10% and 13% respectively), all studies stated at least two of the demographic characteristics indicated in the CEC guidelines. Authors in all studies (100%;  $n = 30$ ) described the participants' diagnosis. However, 3% ( $n = 1$ ) did not describe the diagnostic criteria used in ascertaining the participants' conditions.

**Intervention agent.** Most studies (93%;  $n = 28$ ) stated the roles of the intervention agents; however, 27% ( $n = 8$ ) did not describe the training procedures or necessary qualifications required for the implementation of interventions. In the studies addressing specific training needed to administer interventions (73%;  $n = 22$ ), most authors stated the qualifications of the interventionists (e.g., special education teachers, graduate students in special education or related fields), which was enough to indicate the training required to administer the interventions. Regarding description of independent variables, most studies (100%;  $n = 30$ ) described the implementation procedures for the interventions, including intervention components, instructional behaviors, manualized/scripted procedures, and dosages. Authors in all studies also described the materials used. The majority of the studies (40%;  $n = 12$ ), however, did not evaluate the social validity or consumer satisfaction with the interventions.

**Fidelity.** The CEC standards require that implementation fidelity for the interventions be assessed using three criteria: adherence, dosage, and scope. In order to meet the adherence criterion, researchers are expected to provide evidence about intervention agents' adherence to written down procedures (e.g., checklists), and this information was reported in 87% ( $n = 26$ ) of the studies. The code on dosage assessed implementation fidelity pertaining to duration or frequency of intervention (or participants' exposure to intervention). This was reported in 97% ( $n = 29$ ) of the studies. Scope assessed fidelity on two dimensions, namely consistency of fidelity assessment throughout the intervention period and assessment of each intervention agent's adherence to procedures. Authors reported quantity of sessions in which treatment fidelity was assessed in 97% ( $n = 29$ ) of the studies.

**Table 1.1. Summary of Studies**

Author	Participants N, Sex, Age	Grade	Race	Diagnosis	Setting	Intervention	Dependent Variables	Intervention Agent	Design
Alber-Morgan et al. (2007) *	4 (3M, 1F), 12-15	6, 7	2 AA, 2 Cauc	EBD, BP, LD	Self-Contained Classroom	RR, CR, & Performance Feedback	ARI: ORF, Comprehensions	Graduate Students	SCD: MBD-P
Allen-De Boer et al. (2006) *	4 (4M), 16-18	4, 5	3 AA, 1 Cauc	EBD, ADHD, CD, PPD	Juvenile Correctional Facility	CR	DIBELS ORF	Graduate Students	SCD: MBD-P
Babyak et al. (2000)	4 (4M), NM	4, 5	NM	BD	Self-Contained Classroom	Story Mapping Instruction	CBM: Comprehension	Teacher	SCD: MBD-P
Barton et al. (2005)	6 (4M, 2F), 8	3	NM	EBD	Self-Contained School	DI: Horizons Fast Track, PALS	CTOPP PA, WMRT-R: PB, NWF, ORF	Teachers	SCD: MBD-P
Bassette & Doughty (2013)	3 (2M, 1F), 7-11	2-5	2 White, 1 AA	EBD, LD, LSI	Special Education Class	Dog Assisted Reading Program	CBM: Reading Accuracy	Researcher	SCD: MBD-P
Blankenship et al. (2005)	3 (2M, 1F), 15	NM	NM	EBD	Self-Contained Classroom	Computer- Based Cognitive Mapping	CBM: Comprehension	Teachers	SCD: MBD-B
Daly et al. (2006)	2 (1M, 1F), 13	7	1 AA, 1 Cauc	BD	School Psychology Training Program	Choice Antecedent Instruction, Rewards, and Performance Feedback	CBM: ORF	Researcher	SCD: MPD-P
Devaney et al	6, 5M 2F, 11-13	6-7	2AA, 2 EA, 2 Biracial	EBD	Residential Treatment Center	AIMS Web Reading fluency and Comprehension instruction, MAZE	TOWRE ORF, Woodcock Johnson ORF, Comprehension	Teacher	SCD: MPD-P

Table 1.1. (Continued)

Author	Participants N, Sex, Age	Grade	Race	Diagnosis	Setting	Intervention	Dependent Variables	Intervention Agent	Design
Falk & Wehby (2001)	6 (6M), 5-6	K	NM	EBD, SLL, ADHD	Self-Contained Classroom	PALS	K-PALS: LNF, Letter Sound Identification, Letter Segmentation, Blending	Teacher & Peer directed	SCD: MBD-P
Gunter et al. (2003)	1 (1F), 9	3	NM	Serious EBD	Special Education Class	Self-Graphing Reading Performance	CBM: ORF	Teacher	SCD: ABAB
Hopewell et al. (2011)	2 (2M), 8	Pry	NM	Severe EBD	Self-Contained Classroom	Reading Racetrack & Token/ Response Cost System	WJ-III: SWRF	Teacher	SCD: MBD-P
Lane et al. (2001)	7 (5M, 2F), 6-7	1	2 AA, 1 Hispanic, 4 White	At Risk for EBD	General Ed Classroom	PA Training for Reading	DIBELS PA: NWF	Graduate students	SCD: MBD-P
Lane, Little et al. (2007)	7 (4M, 3F), 6-7	1	Cauc	Behavioral & Reading Problems	General Ed Classroom	Teacher Led Instruction & PALS	Correct Letter Sounds, ORF	Teacher	SCD: MBD-P
Lane & Fletcher (2007) *	24, 18M, 6F), 6	1	12 Europ, 3 AA, 9 Hisp	At Risk for EBD	Private Room	PA Training for Reading	PA, NWF	Paraprofessionals	Group
Lingo (2006)	7 (6M, 1F), 11-14	6-7	6 AA, 1 His	2 EBD, 4 OHI, 1 LD	Spec Ed Classroom	CR & RR	WRMT-Revised: ORF	Teachers	SCD: MPD-P
McDaniel et al. (2011) *	31 (27M, 4F), NM	4-8	0	E/BD	Self-Contained Classroom	CR	WJ-III: ORF	Teachers	Group

Table 1.1. (Continued)

Author	Participants N, Sex, Age	Grade	Race	Diagnosis	Setting	Intervention	Dependent Variables	Intervention Agent	Type of Design
Nelson et al. (2005) *	63 (47M, 16F), 5	K	9 AA, 47 Cauc, 6 Hisp, 1 Asian Amer	BD, Reading Difficulties	General Education Classroom	Stepping-Stones to Literacy ELI	WJ-III Reading: PA, Word Reading, LNS, & Rapid Naming	Researchers	Group
Nelson, Stage et al. (2005) *	36 (34M, 2F), 2	K	10 AA, 22 cauc, 3 His, 1 Asian Amer	ED, Reading Problems	General Education Classroom	Stepping-Stones to Literacy ELI	CTOPP, DIBELS: LNF, & WRMT-R: PA: ISF, PSF, NWF, RNF	Teachers	Group
Palmer et al. (2014)	4 (3M, 1F), 13-14	7	3 Cauc, 1 Hisp	2 OHI, 2 EBD	Resource Room	Concept Mapping, Dictionary Instruction	Vocabulary Acquisition Skills	Researcher	SCD: ABAB
Rogevich & Perrin (2005)	63 (63M), 13-16	13 to 16	35% Cauc, 41% AA, 24% Hisp	BD (31 had BD + ADHD)	Self-Contained Facility	TWA-WS: Reading Comprehension Intervention	CBM: Reading Comprehension	Instructor	Group
Scott & Lingo. (2002)	3 (3M), NM	Middle	NM	EBD	Self-Contained	Reading Fluency Instruction	CBM ORF	Teachers	SCD: MBD-P
Sanders	25, (19M, 6F), NM	6 to 12	NM	ED	Special Education Class	TWA-WS Reading Comprehension Intervention	Oral retell, Maze Vocabulary	Teacher	Group
Spilles	8, NM, 7-10	2,3	NM	BD	Inclusive School	Peer Tutoring	ORF	Teacher	MPD-P
Staubitz et al. (2005)	6 (4M, 2F), 9-11	4-5	85% AA, 13% Cauc, 2% Hisp/ Native Amer/ Asian	ED, ED + ADHD	Resource Classrooms for EBD	RR Instruction, PALS	WJ-III ORF, Reading Accuracy	Researcher	SCD: MBD-P

**Table 1.1. (Continued)**

Author	Participants N, Sex, Age	Grade	Race	Diagnosis	Setting	Intervention	Dependent Variables	Intervention Agent	Type of Design
Stone et al. (2008)	4 (NM), 15	9	NM	EBD	Self-Contained Classroom	Text Mapping	CBM Reading Comprehension Completed Text Map Scores	Instructor	SCD: MBD-P
Strong (2004) *	6 (6M), 12-14	7-8	4 AA, 2 Cauc	ED, ED + OHI + LD, ED +LSI	Self-Contained Classroom	CR & RR	CBM: ORF	Researcher	SCD: MBD-P
Sutherland & Snyder, (2007) *	4 (NM), 11-13	6-8	3 AA, 1 Cauc	ED, ED + SLI	Self-Contained Classroom	PALS	CBM Self Graphed ORF	Teacher	SCD: MBD-P
Trout et al. (2003)	18 (15M, 3F), 5	K	83% Cauc, 17% Hisp	At Risk for EBD	General Education Classroom	Reading Mastery, Great Leaps Reading Program	CTOPP Letter Sounds, Blending, Sight Words, ORF	Graduate Students	Group
Wehby et al. (2005)	4 (NM), 5-6	NM	2 AA, 2 EA	At Risk for EBD	Self-Contained Classroom	Scott Foresman Kindergarten Reading Instruction	CTOPP, WRMT-R, PA, NWF, Comprehension, ORF, Writing Grammar	Teachers	SCD: MBD-P
Wills et al. (2010)	171 (NM), NM	NM	32% W, 38% AA, 23% Hisp, 23% ELL	At Risk for EBD	Midwestern Metropolitan Area	Small Group Instruction, DI	DIBELS & WRMT-R: Grade NWF	Teachers	Group

*Notes:* M = Male, F = Female, NM = Not Mentioned, AA = African American, Cauc = Caucasian, Hisp = Hispanic, EA = European American, EBD = Emotional Behavioral Disorder, SLI = Speech Language Impairment, OHI = Other Health Impairment, LD = Learning Disability, ADHD = Attention Deficit Hyperactive Disorder, Dynamic Indicators of Basic Early Literacy Skills, CTOPP = Comprehensive Test of Phonological Processing, FAST = Formative Assessment System for Teachers, WRMT-R = Woodcock Reading Mastery Test Revised, CBM = Curriculum Based Measurement, GORT-4 = Gray Oral Reading Test 4th Edition, WJ-III = Woodcock Johnson Test of Achievement, CR = Corrective Reading, RR = Repeated Reading, DI = Direct Instruction, ELI = Early Literacy Instruction, PALS = Peer Assisted Learning Strategies ORF = Oral Reading Fluency, PA = Phonological Awareness, NWF = Nonsense Word Fluency, ISF = Initial Sound Fluency, LNF = Letter Naming Fluency, PSF = Phonemic Segmentation Fluency, PB = Phonemic Blending, SCD = Single Case Design, MBD-P = Multiple Baseline Design Across Participants, MBD-B = Multiple Baseline Design Across Behaviors, AID = Alternating Treatment Design, ABAB = Reversal Design, ABAB = Reversal Design. Studies with asterisk qualified for inclusion in meta-analysis.

**Internal validity codes** consisted of 5 elements to assess whether or not the independent variables were under the direct control of the researchers, and whether they were responsible for the observed change in the dependent variable(s). Evidence regarding **functional control** was reported in all studies (100%;  $n = 30$ ). **Baseline/control detail and integrity** pertaining the description of baseline and control procedures and integrity was sufficiently reported in 97% of the studies ( $n = 29$ )

The QI on **internal replication** pertained to SCD only and required authors to provide at least 3 demonstrations of experimental control at three different times. Graphs in 32% ( $n = 7$ ) of SCD studies did not adequately demonstrate internal replication in accordance with this standard. Of these studies, 5% ( $n = 1$ ) had some but not all graphs meeting this indicator. Twenty-seven percent ( $n = 6$ ) of the SCDs did not address this indicator in all of the graphs presented in the studies. Most such studies provided three tiered MBDs, but data collection across the baselines were not concurrent. Other studies (e.g., Lane et al. 2001; Lane, Little et al. 2007; Wehby et al., 2005) used multiple baseline designs with only two tiers, thereby lacking evidence for replication of effect.

SCD graphs in 73% of the studies ( $n = 16$ ) had **appropriate baselines** (i.e., stable trend, with at least three data points). The studies that failed to address this indicator had insufficient data points (i.e., less than 3 data points) in some tiers of the MTDs or had non-concurrent data collection sessions in different baselines. The QI on **adequate control** assessed whether authors addressed all common threats to internal validity (e.g., history, maturation, temporal precedence). In SCDs, this quality indicator was met if studies had well designed reversal, multiple baselines, or alternating treatment designs. Graphs demonstrated adequate control of the interventions in 59% ( $n = 13$ ) of the SCDs, whereas 41% ( $n = 9$ ) demonstrated partial control. The studies that partially addressed this indicator either had baselines with inadequate data or insufficient baseline phases to demonstrate predictability in future responding.

Internal validity for group designs was further assessed using 3 additional codes, namely random allocations, overall attrition, and differential attrition. Random allocations required researchers to demonstrate random assignment of participants between experimental and control groups. This indicator was addressed in all group designs (100%;  $n = 8$ ). Regarding attrition and differential attrition, the authors are required to report the number of participants that defaulted from the study before the end, and the procedures that were followed in replacing their data in analysis respectively. Authors disclosed information about attrition or completion rate of participation by participants as well as differential attrition in 75% ( $n = 6$ ) of group design studies respectively.

**Outcome measures** consist of 6 codes designed to measure the effect of the interventions on target outcomes. The 6 codes consist of importance, description, total effect, frequency, reliability, and instrument validity. Under importance, authors are expected to provide information regarding social significance of the outcome variables. Authors addressed social importance of dependent variables in all 30 studies. Most authors provided such information in the introduction sections, or they evaluated the dependent variables using social validity measures in the form of questionnaires or checklists. Authors in all the studies (i.e., 100%;  $n = 30$ ) also provided

detailed description of the outcome variables and the total effect of the interventions on the outcomes. The measures of effect featuring in group designs consisted of standardized mean differences (e.g., Cohen's  $d$ ) and Pearson's coefficient. SCDs illustrated effects by using visual analysis, which met the criteria for this indicator. Reliability procedures for the measurement of outcome variables were described in 97% ( $n = 29$ ) of the studies. **Effect sizes** were computed and reported in all group designs (i.e., 100%;  $n = 8$ ). Visual presentation of results through graphs was provided in all SCD studies. The summary of results for quality assessment are presented in table 1.2.

**Table 1.2. Summary of Quality Indicators**

	<b>Quality Indicator</b>	<b>SCD (%)</b>	<b>GD (%)</b>
1.1	<b>Context and Setting</b>	100	75
	<b>Participants</b>		
2.1	Description	100	100
2.2	Status	100	100
	<b>Intervention Agent</b>		
3.1	Role	91	100
3.2	Training	73	88
	<b>Description of IV</b>		
4.1	Procedures	100	100
4.2	Materials	100	100
4.3	Acceptability	68	50
	<b>Fidelity</b>		
5.1	Adherence	96	75
5.2	Dosage	100	88
5.3	Scope	100	88
	<b>Internal Validity</b>		
6.1	Manipulation	100	100
6.2	BL/control detail	95	100
6.3	BL/control integrity	95	100
6.4	Random Allocations		100
6.5	Internal replication <sup>2</sup>	68	
6.6	Appropriate baseline <sup>2</sup>	73	
6.7	Adequate control <sup>2</sup>	59	
6.8	Overall Attrition		75
6.9	Differential Attrition		75
	<b>Outcome Measures</b>		
7.1	Importance <sup>3</sup>	100	100
7.2	Description	100	100
7.3	Total effect	100	100
7.4	Frequency <sup>2</sup>	100	100
7.5	Reliability	100	100
7.6	Instrument Validity		100
8.1	Data analysis Methods		100
8.2	Graph	100	
8.3	Suitability of Effects		100

Evidence Based Practice Classifications. The interventions featuring in the 30 studies were analyzed and divided into categories consisting of interventions sharing similar features or characteristics. In total, 6 categories of interventions appeared. These are corrective reading, concept/story mapping, direct instruction, feedback based, peer mediated, and phonological awareness. All categories, except concept/story mapping met the classification for EBP. Concept/story mapping interventions were classified as potentially EBP. This was due to the absence of enough studies featuring the interventions. The summary results for the analysis are shown in table 1.3.

**Table 1.3. Classification for Evidence Base**

<b>Reading Strategy</b>	<b># Of sound GD Studies</b>	<b># Of sound SCD Studies</b>	<b>Studies with (-) ve Effects</b>	<b>EBP</b>	<b>Potentially EBP</b>	<b>Mixed Effects</b>
Corrective Reading	1	3	0		X	
Concept Mapping	0	3	0			X
Direct Instruction	5	14	0	X		
Feedback	1	3	0		X	
Peer Mediated	0	4	0			X
Phonological Awareness	2	2	0		X	

*Note:* All studies included in the Classification for EBP table addressed all CEC quality indicators. Studies that failed to address at least one CEC indicator were excluded.

## DISCUSSION AND CONCLUSION

Given the need to improve learning outcomes of children with special needs, there has been an increased emphasis for educators to use EBP. Children with EBD have consistently exhibited deficits in reading, partly owing to the use of instructional methods that are not fully proven to be effective. The purpose of this study was to assess the rigor in experimental methods used in reading interventions literature for students with EBD. Specifically, the study used the Council for Exceptional Children (CEC) quality indicators to assess the extent to which researchers addressed study elements that support evidence for practices. Thirty studies meeting inclusion criteria were scored and evaluated.

As shown in the results, most primary studies in this area are conducted using group and SCDs. The two methods render different advantages which strengthen the support for EBP. During the past two decades, SCDs have constituted a significant proportion of interventions' studies across the fields of education, nursing, and healthcare (Schlosser & Sigafoos, 2006). In this review, a larger number of studies were conducted using SDC relative to group designs. The growing prominence of SCD in special education can be attribute to its effectiveness in accounting for the unique characteristics in children with special needs (Horner et al., 2005; Kazdin, 2010). Furthermore, group designs require large amounts of resources, which may

prove costly for researchers (Handley et al., 2018). SCDs target limited participants, thereby limiting the amounts of resources needed to conduct the studies. SCDs also enable detailed analysis of both responders and non-responders to interventions, as well as offer cost-effective ways of identifying behavioral and educational interventions ideal for large scale analysis. Furthermore, although group designs have been credited for their strong inclination on using well established statistical methods to assess effect sizes of interventions, they tend to mask individual differences among participants, thereby limiting the important knowledge regarding the effectiveness of interventions on outcomes of students with unique needs (Horner et al., 2005). SCD on the other hand address such limitations by focusing analysis entirely on a few units/participants, targeting specific limited outcomes.

Quality assessment of the studies revealed some strengths and weaknesses across the literature. The description of study context (e.g., participants demographics and settings) is paramount in providing the evidence that the study was conducted in the context relevant to the investigation. Most studies provided the minimum necessary details for the description of the various study elements, including the participants' demographics, status, and study settings. Almost all studies described the important participants' demographic information like age, gender, and diagnosis. However, although the target for this study was EBD, the participants featured in some studies (32%;  $n = 9$ ) had comorbid conditions, which may complicate the interpretation of interventions effectiveness based on the wide range of participants' characteristics. However, the other conditions observed in participants (e.g., ADHD, learning disability) have been shown to frequently co-occur with EBD (Bunford et al., 2015; Rock et al., 1997; Spreen, 1989). Most interventions for EBD often address the other identified conditions (Bunford et al., 2015).

The types of reading interventions used across studies were generally well described. The practices most commonly used included repeated readings, corrective reading, peer mediated strategies, and teacher direct instruction. In most studies, authors used more than a single practice, often targeting different types of reading skills as well. None of the studies, however, conducted separate component analysis to show the disaggregated effectiveness of different intervention components, which may make it difficult to ascertain what parts of the interventions were associated with what proportion of the observed changes in outcomes. Results across all studies were clearly disaggregated for dependent variables. The effectiveness of the interventions on different specific outcomes in SCD were visually depicted, which made it easy to see which reading outcomes responded most to the interventions. The studies also generally described the implementation procedures for the interventions well enough, as well as the measures for implementation fidelity.

In order to establish EBP, one of the requirements pertains to researchers' ability to demonstrate experimental control of interventions (i.e., the evidence that the intervention was responsible for the observable change in the dependent variables). As shown on the indicators, authors in most studies described the internal validity procedures in sufficient details (i.e., manipulation, baseline control, baseline integrity). However, some SCD studies failed to adequately address the standards on internal replication, appropriate baseline, and adequate control. In most such studies, researchers did not allow for sufficient replication of the interventions following

baselines. In studies that used multiple baseline designs, some had two instead of three tiers, the minimum number required to demonstrate replication of the interventions. Furthermore, even though the CEC requires that baseline and treatment phases have at least 3 data points, more is always better to establish evidence that level of responding would not change unless the intervention is introduced. In group designs, reporting attrition is essential to provide evidence that researchers put in place procedures to combat this, thereby maintain confidence in the efficacy of the intervention. Of the 8 group design studies, only 2 did not report any information regarding attrition. The studies that reported attrition had 100% completion rates, thereby strengthening the evidence of the interventions' efficacy. Even if there are no participants leaving the study before completion, it is always necessary to report this for clarification's sake.

Regarding description of the study outcomes, authors in all studies described the procedures used in measurement of dependent variables as well as their social validity in sufficient details. However, although all studies provided enough data to compute effect sizes, only a few computed the effect sizes of the interventions on the outcomes. For a long time, there was no consensus amongst SCD researchers regarding the appropriate methods to use for measuring effects. Even though this has been the case, computation of mean percentages may provide some help in quantifying the effectiveness of interventions. Furthermore, some commonly applied methods of measuring effects in SCDs (e.g., PND and Tau U) do not require multiple participants. Researchers may improve their presentation of evidence by supplementing visual analyses with these methods.

When interventions were assessed for EBP, most interventions met the classification for EBP, whereas only one met classification for Potential EBP. Although there were not a lot of studies featuring the respective interventions, each intervention category had the minimum required number of studies supporting the interventions. Direct instruction strategies featured the most, and positive effects were reported across the studies featuring the methods. Direct instruction has a long-documented history of efficacy. However, teachers of students EBD often experience challenges administering instruction due to the problem behaviors exhibited by this group of students. The large volume of literature supporting direct instruction, however, demonstrates this method to be an effective approach in addressing academic needs of children with EBD. Peer mediated strategies also featured frequently and were classified as EBP. Again, peer mediated intervention strategies have a long history of effectiveness when used with typical students and students with other learning disabilities. The various formats of peer mediated strategies that featured across the studies provide a wide range of choices that instructors can use to fit their respective circumstances. The other strategies that featured more frequently include corrective reading and concept mapping. Although there were sufficient studies to support these strategies, the variations in the strategies across the studies may make their interpretation under one bracket complicated. Future research on these strategies may be necessary to further strengthen their efficacy.

### IMPLICATIONS AND DIRECTIONS FOR FUTURE RESEARCH

The application of quality indicators denoting methodological rigor in interventions research is important in guiding researchers on the best practices for carrying out research procedures to address any potential threats to reliability and validity, and to improve the confidence interventions research findings. The CEC quality indicators are very comprehensive as they cover the important aspects of experimental research. Addressing the quality indicators is therefore likely to improve the overall quality of the studies as well as the confidence with which the findings from the studies can be trusted. Judging from the findings on quality indicators applied in this review, several studies on reading interventions for students with EBD make an important contribution to the evidence base for literacy interventions. However, interpreting the efficacy of some interventions may be difficult if the interventions include many components without disaggregating the effects attributed to the different components. Future researchers may consider ways to narrow down their focus to specific interventions so that their effectiveness is not conflicted with more varied components. Researchers utilizing SCD methods can incorporate designs that examine individual intervention components separately. Research on component analysis suggests many different ways of achieving this (Krasny-Pacini & Evans, 2018; Ward-Horner & Sturmey, 2010). A few studies using SCDs also failed to adequately address quality indicators on internal replication, appropriate baseline, and adequate control either because their graphs had less than 3 tiers, insufficient datapoints in the respective phases, or concurrent data collections across tiers in multiple baseline designs. Future research can therefore improve their designs by incorporating these elements. SCDs meeting CEC standards should include at least three data sessions in each phase to establish a stable trend (Kratochwill et al., 2010). However, this is not a golden standard anymore since three data points are not enough, especially if there is high variability in the data (WWC, 2017). Future researchers may consider including more data sessions in each phase in order to establish more stable trends. Furthermore, future researchers using SCDs may improve their studies by computing effect sizes to supplement visual analysis (Barton et al., 2017).

Overall, the failure to identify sufficient studies fully supporting the evidence base for most interventions (see table 1.3) implies a need for more research in the future. Only direct instruction intervention had a sufficient number of studies supporting its efficacy as an EBP. However, direct instruction can be broad, involving several varied strategies. Teachers using this strategy may need to pay particular attention to specific components that can result in improved outcomes for students. Interventions that classified as potentially EBP (i.e., corrective reading, corrective feedback, and phonological awareness) may constitute the current best practices in supporting reading for students with EBD. More research targeting these strategies may help to strengthen their case.

## REFERENCES

- Alber-Morgan, A. R. (2007). Effects of repeated readings, error correction, and performance feedback on the fluency and comprehension of middle school students with behavior problems. *The Journal of Special Education, 41*(1), 17–30.
- Allen-De Boer, R. A., Malgren, K. W., & Glass, M. E. (2006). Reading instruction for youth with emotional and behavioral disorders in a juvenile correctional facility. *Behavioral Disorders, 32*(1), 18–28.
- Anderson, J. A., Kutash, A., Duchnowski, A. J. (2001). A comparison of the academic progress of students with EBD and students with LD. *Journal of Emotional and Behavioral Disorders, 9*(2), 106–115.
- Armstrong, K. H., Dedrick, R. F., & Greenbaum, P. E. (2003). Factors associated with community adjustment of young adults with serious emotional disturbance: A longitudinal analysis. *Journal of Emotional and Behavioral Disorders, 11*(2), 66–77. <https://doi.org/10.1177/106342660301100201>
- Babiyak, A. E., Koorland, M. A., & Mathes, P. G. (2000). The effects of story mapping instruction on the reading comprehension of students with behavioral disorders. *Behavioral Disorders, 25*(3), 239–258.
- Barton, 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.
- Bassette, L. A., & Doughty, T. (2013). The effects of a dog reading visitation program on academic engagement behavior in three elementary students with emotional and behavioral disabilities: A single case design. *Child Youth Care Forum, 42*(1), 239–256.
- Benner, G. J., Nelson, J. R., Ralston, N. C., & Mooney, P. (2010). Meta-analysis of the effects of reading on the reading skills of students with or at risk of behavioral disorders. *Behavioral Disorders, 35*(2), 86–102. <https://doi.org/10.1177/0741932513514856>.
- Blankenship, T. L., Ayres, K. M., & Langone, J. (2005). Effects of computer-based cognitive mapping on reading comprehension for students with emotional behavioral disorders. *Journal of Special Education Technology, 20*(2), 15–23.
- Bunford, N., Evans, S. W., Wymbs, F. (2015). ADHD and emotion dysregulation among children and adolescents. *Clinical Child and Family Psychology Review, 18*(3), 185–217.
- Burke, M. D., Boon, R. T., Hatton, H., & Bowman-Perrott, L. (2015). Reading interventions for middle and secondary students with emotional and behavioral disorders: A quantitative review of single-case studies. *Behavior Modification, 39*, 43–68.
- Chitiyo, A., King, S. A., Krizon, M., & Ablakwa, C., & Markelz, A. (2020). A methodological review of research syntheses involving reading interventions for students with EBD. *Behavioral Disorders. <https://doi.org/10.1177/0198742920919095>*
- Coleman, M., & Vaughn, S. (2000). Reading interventions for students with emotional/behavioral disorders. *Behavioral Disorders, 25*, 93–104.
- Cook, B. G. (2014). A call for examining replication and bias in special education research. *Remedial and Special Education, 35*(4), 233–246. <https://doi.org/10.1177/0741932514528995>.
- Cook, B. G., & Cook, S. C. (2011). *Thinking and communicating clearly about evidence-based practices in special education*. Division for Research, Council for Exceptional Children.
- Cook, B. G., Buysse, V., Klingner, J., Landrum, T. J., McWilliam, R. A., Tankersley, M., & Test, D. W. (2015). CEC's standards for classifying the evidence base of practices in special education. *Remedial and Special Education, 36*(4), 220–234. <https://doi.org/10.1177/0741932514557271>.

- Cook, B. G., Tankersley, M., Cook, L., & Landrum, T. J. (2008). Evidence based practices in special education: Some practical considerations. *Intervention in School and Clinic, 44*(2), 69–75. <https://doi.org/10.1177/001440290907500306>.
- Daly, E. J., Garbacz, A., Olson, S. C., Persampieri, M., & Ni, H. (2006). Improving oral reading fluency by influencing students' choice of instructional procedures: An experimental analysis with two students with behavioral disorders. *Behavioral Interventions, 21*(3), 13–20.
- Dunn, M. E., Shelnut, J., Ryan, J. B., & Katsiyannis, A. (2017). A systematic review of peer-mediated interventions on the academic achievement of students with emotional/behavioral disorders. *Education and Treatment of Children, 40*, 497–534.
- Falk, K. B., & Wehby, J. H. (2001). The effects of peer-assisted learning strategies on the beginning reading skills of young children with emotional or behavioral disorders. *Behavioral Disorders, 26*(4), 344.
- Garwood, J. D. (2017). Literacy interventions for secondary students formally identified with emotional and behavioral disorders: Trends and gaps in the research. *Journal of Behavioral Education, 27*(1), 23–52. <https://doi.org/10.1007/s10864-017-9278-3>
- Garwood, J. D., Brunsting, N. C., & Fox, L. C. (2014). Improving reading comprehension and fluency outcomes for adolescents with emotional-behavioral disorders: Recent research synthesized. *Remedial and Special Education, 35*(3), 181–194. <https://doi.org/10.1177/0741932513514856>.
- Gersten, R., Fuchs, L.S., Coyne, M., Greenwood, C., & Innocenti, M.S. (2005). Quality indicators for group experimental and quasi-experimental research in special education. *Exceptional Children, 71*(2), 49–164. <https://doi.org/10.1177/001440290507100>
- Gunter, P. L., Miller, K. A., & Venn, M. L. (2003). A case study of the effects of self-graphing reading performance data for a girl identified with emotional/behavioral disorder. *Preventing School Failure, 48*(1), 28–31. <https://doi.org/10.1080/1045988X.2003.10871076>
- Handley, M. A., Lyles, C. R., McCulloch, C., & Cattamanchi, A. (2018). Selecting and improving quasi experimental designs in effectiveness and implementation research. *Annual Review of Public Health, 39*, 5–25. <https://doi.org/10.1146/annurev-publ-health-040617-014128>
- Hill, J. W., & Coufal, K. L. (2005). Emotional/Behavioral Disorders: A retrospective examination of social skills, linguistics, and student outcomes. *Communication Disorders Quarterly, 27*(1), 33–46. <https://doi.org/10.1177/15257401050270010401>
- Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children, 71*(2), 165–179. <http://dx.doi.org/10.1177/001440290507100203>.
- Joseph, L. M., & Eveleigh, E. L. (2011). A review of the effects of self-monitoring on reading performance of students with disabilities. *The Journal of Special Education, 45*(1), 43–53. <https://doi.org/10.1177/0022466909349145>.
- Kazdin, A. E. (2010). *Single-case research designs: Methods for clinical and applied settings* (2nd ed.). Oxford University Press.
- Krasny-Pacini, A., & Evans, J. (2018). Single-case experimental designs to assess intervention effectiveness in rehabilitation: A practical guide. *Annals of Physical Rehabilitation Medicine, 61*(3), 164–179. <https://doi.org/10.1016/j.rehab.2017.12.002>
- Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M & Shadish, W. R. (2010). Single case designs technical documentation. In *What Works Clearinghouse: Procedures and standards handbook (version 2.0)*. Retrieved from <https://eric.ed.gov/?id=ED510743>.

- Kremer, K. P., Flower, A., Huang, J., & Vaughn, M. G. (2016). Behavior problems and children's academic achievement: A test of growth-curve models with gender and racial differences. *Children and Youth Services Review*, *67*, 95–104. <https://doi.org/10.1016/j.childyouth.2016.06.003>
- Lane, K. L. (2009). Young students at-risk for antisocial behavior: The utility of academic and social skills interventions. *Journal of Emotional and Behavioral Disorders*, *7*(4), 211–223.
- Lane, K. L., Barton-Arwood, S. M., Nelson, J. R., Wehby, J. H. (2008). Academic performance of students with emotional and behavioral disorders served in a self-contained setting. *Journal of Behavioral Education*, *17*(1), 43–62.
- Lane, K. L., Fletcher, T., Carter, E. W., & DeJud, C., & DeLorenzo, J. (2007). Paraprofessional-led phonological awareness training with youngsters at risk for reading and behavioral concerns. *Remedial and Special Education*, *28*(5), 266–276. <https://doi.org/10.1177/0741932507028005>
- Lane, K. L., Little, M. A., Redding-Rhodes, J., Phillips, A., & Welsh, M. T. (2007). Outcomes of a teacher-led reading intervention for elementary students at risk for behavioral disorders. *Exceptional Children*, *74*(1), 47–70. <https://doi.org/10.1177/001440290707400103>
- Lane, K. L., O'Shaughnessy, T., Lambros, K. M., Gresham, F. M., & Beebe-Frankenberger, M. (2001). The efficacy of phonological awareness training with first grade students who have behavior problems and reading difficulties. *Journal of Behavioral and Behavioral Disorders*, *9*(4), 219–231. <https://doi.org/10.1177/106342660100900402>
- Lingo, A. S. (2003). *Effects of corrective reading on the reading abilities and classroom behaviors of middle school students with reading deficits and challenging behavior*. (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses. (Accession Order No.
- Maggin, D. M., O'Keeffe, B. V., & Johnson, A. H. (2011). A quantitative synthesis of methodology in the meta-analysis of single-subject research for students with disabilities: 1985–2009. *Exceptionality*, *19*(2), 109–135. <https://doi.org/10.1080/09362835.2011.565725>
- McDaniel, S. C., Houchins, D. E., & Terry, N. P. (2010). Corrective reading as a supplementary curriculum for students with emotional and behavioral disorders. *Journal of Emotional and Behavioral Disorders*, *21*(4), 240–249.
- 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>
- Mooney, P., Ryan, J. B., Uhing, B. M., Reid, R., & Epstein, M. H. (2005). A review of self-management interventions targeting academic outcomes for students with emotional and behavioral disorders. *Journal of Behavioral Education*, *14*(3), 203–221. <https://doi.org/10.1007/s10864-005-6298-1>.
- Nelson, J. R., Benner, G. J., & Gonzalez, J. (2005). An investigation of the effects of a pre-reading intervention on the early literacy skills of children at risk of emotional disturbance and reading problems. *Journal of Emotional and Behavioral Disorders*, *13*(1), 3–12.
- Nelson, J. R., Benner, G. J., Lane, K., & Smith, B. W. (2004). Academic achievement of K-12 students with emotional and behavioral disorders. *Exceptional Children*, *71*(1), 59–73. <https://doi.org/10.1177/001440290407100104>.
- Nelson, J. R., Stage, S. A., Epstein, M. H., & Pierce, C. D. (2005). Effects of a pre-reading intervention on the literacy and social skills of children. *Exceptional Children*, *72*(1), 29–45.

- Palmer, J., Boon, R. T., & Spencer, V. G. (2014). Effects of concept mapping instruction on the vocabulary acquisition skills of seventh-graders with mild disabilities: A replication study. *Reading & Writing Quarterly*, 30(2), 165–182.
- Reid, R., Gonzalez, J. E., Nordness, P. D., Trout, A., & Epstein, M. H. (2004). A meta-analysis of the academic status of students with emotional/behavioral disturbance. *The Journal of Special Education*, 38(3), 130–143. <https://doi.org/10.1177/00224669040380030101>.
- Rock, E. E., Fessler, M. A., & Church, R. P. (1997). The Concomitance of Learning Disabilities and Emotional/Behavioral Disorders: A Conceptual Model. *Journal of Learning Disabilities*, 30(3), 245–263. <https://doi.org/10.1177/002221949703000302>
- Rogevich, M. E., & Perin, D. (2005). Effects on science summarization of a reading comprehension intervention for adolescents with behavior and attention disorders. *Exceptional Children*, 74(2), 135–154.
- Ryan, J. B., Reid, R., & Epstein, M. H. (2004). Peer-mediated interventions studies on academic achievement for students with EBD: A review. *Remedial and Special Education*, 25(6), 330–341. <https://doi.org/10.1177/07419325040250060101>.
- Schlosser, R. W. (2006). The role of systematic reviews and in evidence-based practice, research, and development. *Focus Technical Brief*, 15, 1–4.
- Schlosser, R., & Sigafoos, J. (2006). Augmentative and alternative communication interventions for persons with developmental disabilities: Narrative review of comparative single-subject experimental studies. *Research in Developmental Disabilities*, 27(1), 1–29.
- Scott, T. M., & Shearer, A. (2002). 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*, 46(4), 167–173.
- Spreen, O. (1989). The relationship between learning disability, emotional disorders, and neuropsychology; some results and observations. *Journal of Clinical and Experimental Neuropsychology*, 11(1), 117–140.
- Staubitz, J. E., Cartledge, G., Yurick, A. L., & Lo, Y. (2005). Repeated reading for students with emotional or behavioral disorders: Peer-and-trainer mediated instruction. *Behavioral Disorders*, 31(1), 51–64.
- Stone, R. H., Boon, R. T., Fore, C., Bender, W. N., & Spencer, V. G. (2008). Use of text maps to improve the reading comprehension skills among students in high school with emotional and behavioral disorders. *Behavioral Disorders*, 33(2), 87–98.
- Strong, A. C., Wehby, J. H., Falk, K. B., & Lane, K. L. (2004). The impact of a reading curriculum and repeated reading on the performance of junior high students with emotional and behavioral disorders. *School Psychology Review*, 33(4), 561–581.
- Sullivan, A. L., & Sadeh, S. S. (2014). Is there evidence to support the use of social skills interventions for students with emotional disabilities? *Journal of Applied School Psychology*, 30(2), 107–131. <https://doi.org/10.1080/15377903.2014.888528>
- Sutherland, K. S., & Snyder, A. (2007). Effects of reciprocal peer tutoring and self-graphing on reading fluency and classroom behavior of middle school students with emotional or behavioral disorders. *Journal of Emotional and Behavioral Disorders*, 15(2), 103–118. <https://doi.org/10.1177/10634266070150020101>.
- Talbott, E., Maggin, D. M., Van Acker, E. Y., & Kumm, S. (2018). Quality indicators for reviews of research in special education. *Exceptionality*, 26(4), 245–265.
- Trout, A. L., Epstein, M. H., Mickelson, W. T., Nelson, R., & Lewis, L. M. (2003). Effects of a reading intervention for kindergarten students at risk for emotional disturbance and reading deficits. *Behavioral Disorders*, 28(3), 313–326.
- U.S. Department of Education. (2012). *Students with learning disabilities intervention report: Peer-assisted learning strategies*. Washington, DC: US Department of Education.

- Wagner, M., & Cameto, R. (2004). *The Characteristics, Experiences, and Outcomes of Youth with Emotional Disturbances*. A Report from the National Longitudinal Transition Study-2. Washington, DC: National Center on Secondary Education and Transition
- Wagner, M., Kutash, K., Duchnowski, A. J., & Epstein, M. H. (2005). The special education elementary longitudinal study and the national longitudinal transition study: Study designs and implications for children and youth with emotional disturbance. *Journal of Emotional and Behavioral Disorders*, 13(1), 25-41.
- Wanzek, J., Otaiba, S. A., & Petscher, Y. (2014). Oral reading fluency development for children with emotional disturbance or learning disabilities. *Exceptional Children*, 80(2), 187-204.
- Ward-Horner, J., & Sturmey, P. (2010). Component analysis using single subject experimental designs: A review. *Journal of Applied Behavior Analysis*, 43(4), 685-704.
- Wehby, J., Lane, K., & Falk, K. (2005). An inclusive approach to improving early literacy skills of students with emotional and behavioral disorders. *Behavioral Disorders*, 30, 155-169.
- Wills, H., Kamps, D., Abbott, M., Bannister, H., & Kaufman, J. (2010). Classroom observations and effects of reading interventions for students at risk for emotional and behavioral