2019-03-01

A Systematic Review of Interventions for Implementation Fidelity for Academic Interventions

Emily Morgan Beecher
Brigham Young University

A Systematic Review of Interventions for Implementation Fidelity for Academic Interventions

Emily Morgan Beecher

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Educational Specialist

Christian Sabey, Chair
Ellie Young
Timothy Smith

Department of Counseling Psychology and Special Education

Brigham Young University

ABSTRACT

A Systematic Review of Interventions for Implementation Fidelity
for Academic Interventions

Emily Morgan Beecher
Department of Counseling Psychology and Special Education, BYU
Educational Specialist

To address students' academic and behavioral needs, schools are held accountable for implementing effective evidence–based interventions. An important relationship exists between implementation fidelity and the effectiveness of interventions. The purpose of this systematic review was to identify and evaluate the evidence of interventions to improve the implementation fidelity of academic interventions and to evaluate the quality of the existing research with a focus on the quality of the research on the most successful interventions. A total of 13 studies met the inclusion criteria. Each study was coded based on a quality of evidence coding protocol and the findings were then reviewed and synthesized.

The results show that performance feedback was the most used and successful intervention for increasing implementation fidelity of an academic intervention. Professional development and teacher training were other interventions that were implemented to improve implementation fidelity. These results are summarized and implications for school-based practice are discussed. With such few studies that met the inclusion criteria, there is a need for more research in this area.

Keywords: implementation fidelity, evidence-based interventions, performance feedback

ACKNOWLEDGEMENTS

I don’t know if I have ever been happier to complete something. Research does not come easily to me, and this took a lot of mental and emotional endurance. I could not have done it without my supportive husband, Cameron, who handled all of my meltdowns, my parents for always believing in me, my advisor Christian for answering all my questions, and my Heavenly Father for everything.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Introduction

Advocating for and implementing interventions to enhance academic performance is essential in the school system (Gilbertson, Witt, Singletary, & VanDerHeyden, 2007). Academic interventions are key to deciding if a student is improving and learning or if they need more individualized instruction, such as special education (Gilbertson et al., 2007). It is not enough for interventions to be implemented. They must be implemented with fidelity, which refers to the degree to which an intervention is delivered as intended (Perepletchikova, Treat, & Kazdin, 2007). According to Greenberg, Domitrovich, Graczyk, and Zins (2005), intervention can both reduce academic deficits and problem behaviors while promoting competence. With such a great need for effective treatment for students with academic deficits and problem behaviors, schools have an obligation to adapt to meet the needs of these students.

In a study of academic interventions (Duhon, Mesmer, Gregerson, & Witt, 2009), the students' academic performance was identified, as well as their baseline levels of performance and response to current specific interventions. Then the authors administered a pre-intervention assessment to help identify the target skill deficit and find an intervention to improve that skill. The teachers were trained one day before the intervention was implemented. The following day the intervention called Drill Sandwich was implemented to the students' (Duhon et al., 2009). When the teachers received feedback, they had high levels of implementation fidelity but when feedback was removed implementation dropped off immediately and the intervention was no longer being conducted (Duhon et al., 2009). This is one example of how important implementation fidelity is to the effectiveness of an intervention. Additionally, it demonstrates the ways that implementation fidelity can be addressed and improved to make a difference in an

intervention being conducted. If we want students' academic performance to improve, we need to give them the assistance to make that happen. Another study of academic prereferral interventions showed that, the effectiveness of the intervention is linked to implementation fidelity of the intervention (Noell, Witt, Gilbertson, Ranier, & Freeland, 1997). Mortenson and Witt (1998) found that students' academic performance improved with high levels of implementation fidelity during the feedback and assistance phase and students' academic performance decreased when implementation fidelity was low due to the no feedback or assistance phases.

In the recent past there has been an increased emphasis on the steps of creating and applying research-based knowledge to practice in both education and psychology (Kratochwill & Shernoff, 2003). This has led to an increased surge of interest in evidence-based practice within the mental health and educational fields (Walker, 2004). The Individuals with Disabilities Education Improvement Act (IDEIA, 2004) explains that school-based interventions are more sophisticated now than ever before. There are new federal mandates for pre-referral intervention, documentation of student's response to intervention, and intervention plans based on functional behavior assessments (IDEIA, 2004). This need for evidence-based interventions also leads to the importance of implementation fidelity.

Implementation Fidelity

There is an important relationship that exists between intervention effectiveness and fidelity of implementation (Noell, Gresham, & Gansle, 2002). Educators are expected to implement interventions to improve academic performance and improve the behavior of struggling students but when interventions are not implemented properly (i.e., with fidelity) they do not usually produce the results they would like or were expecting. Thus, the implementer

may believe that the intervention is ineffective but really it is ineffective because it was not implemented properly. Research shows that failure to implement with fidelity may result in an intervention being less efficient, less effective, and producing less predictable responses (Noell et al., 2002). It is crucial for the educator implementing the intervention to understand how to implement it as it was designed (The Iris Center, 2016). Fidelity of implementation is an important link between the use and effectiveness of interventions both in school settings and scientific investigation (Wilkinson, 2007). It is critical that fidelity be measured when implementing interventions in order to distinguish between ineffective interventions and potentially effective interventions that are being implemented incorrectly (Wilkinson, 2007). When implementation fidelity is not observed then the potential benefit of interventions can be overlooked and disregarded.

Martens and McIntyre (2009), explained that when implementation fidelity is not measured, it is not possible to draw conclusions on the efficacy of the intervention. In order to help school personnel who are implementing interventions they must know what is expected of them and what will help them implement with fidelity. Measuring implementation fidelity can also contribute to understanding if the intervention is a match for the setting and resources available. Low implementation fidelity can show that teachers may need more support and coaching to get implementation fidelity to effective levels (Martens & McIntyre, 2009).

Implementation fidelity reflects the accuracy and consistency that each aspect of the treatment or intervention plan in implemented (Wilkinson, 2007). There are three important factors that contribute to implementation fidelity. They are adherence, meaning the instructional procedures were followed and in the correct order (Wilkinson, 2007). Next exposure/duration, meaning the intervention is given for the correct amount of time, over the correct number of

weeks or months, and the frequency of the intervention is correct (Wilkinson, 2007). Lastly, is the quality of deliverance, meaning good teacher practices such as enthusiasm, allowing students to ask questions and give feedback and managing transitions (Wilkinson, 2007). When interventions are implemented with fidelity the results that are expected are far more likely to happen (Noell et al., 2002).

The definition of fidelity of implementation is the extent to which an intervention is implemented as it was intended (Noell et al., 2002). Fidelity of implementation is also commonly referred to as treatment integrity, procedural fidelity, intervention integrity, procedural reliability, or procedural adherence. The terms that will primarily be used throughout this paper will be fidelity of implementation and implementation fidelity.

Problem Statement and Research Questions

Implementation fidelity is a critical feature of evaluating the effectiveness of an intervention. For this reason, researchers have developed interventions to improve implementation fidelity. At present, it is unclear how effective these interventions are for improving the implementation fidelity of academic interventions. Furthermore, existing research has focused on program and organizational characteristics that contribute to but not directly alter implementation fidelity (Wilkinson, 2007). The purpose of this review is to evaluate the existing research on implementation fidelity interventions for academic interventions. Specifically, we will evaluate the quantity, quality, and strength of the evidence. Additionally, we will evaluate differential effects based on possible moderator variables (e.g., interventionist, sample). Specific research questions are as follows:

1. To what extent have interventions been effective in helping individuals, such as school teachers or other school personnel implement academic treatments with fidelity?
2. Is there a specific intervention that has been notably more successful in terms of helping individuals implement academic treatments with fidelity?
3. What is the quality of the research that has been conducted on interventions to improve implementation fidelity of academic treatments?
4. To what extent does existing research address interventionist (school teacher or other school personnel) characteristics as a moderating factor of implementation fidelity?

CHAPTER 2

Literature Review

If we are to bridge the gap between research and practice of proper implementation of interventions, then consumers of research must be able to guarantee that the treatment outcomes are due to the intervention (Lane, Beebe-Frankenberger, Lambros, & Pierson, 2001). Some of the necessary components to guarantee treatment outcomes are due to an intervention include the following: social validity, implementation fidelity, generalization, and maintenance (Lane et al., 2001). This study will analyze the effectiveness of different interventions to improve the implementation fidelity of academic interventions.

Assessment of Implementation Fidelity

It is essential that fidelity of implementation data is collected when conducting school-based interventions in order to draw proper conclusions about the effectiveness of the intervention (Lane, Bocian, MacMillan, & Gresham, 2004). This is important because if the intervention is not implemented with integrity, then the internal and external validity of the process are threatened (Gresham, Gansle, & Noell, 1993). When interventions are altered during implementation, the alterations are seldom documented in a way that allows anyone to evaluate the components that were changed, resulting in unknown effectiveness of the intervention (Lane et al., 2004).

There are a number of ways school personnel can evaluate implementation fidelity. Elliott and Busse (1993) explained that school personnel can assess fidelity of implementation by (a) direct observation, (b) feedback from consultants, (c) self-monitoring, self-reporting, and behavioral interview techniques, (d) permanent products, and (e) manualized treatments and intervention scripts. Some researchers and practitioners may video the implementation and then

rate the fidelity of the implementation. Both the interventionist and the coach can observe the implementation and problem solve when implementation is insufficient. Fixsen, Naoom, Blase, and Friedman (2017) found through review of the literature, that feedback on implementation fidelity, provided in a timely fashion (short feedback loops recurring), and delivered personally by a respected source was most effective when accompanied by written material. Dusenbury, Brannigan, Falco, and Hansen (2003), explained that in order for research to successfully be translated into practice, programs and policies must be understood and implemented so that quality is maintained, and the programmatic objectives intended by program developers are achieved.

Dane and Schneider (1998), also explain five areas of implementation fidelity to consider, which include the following: (a) adherence, which means the extent to which specific program components were implemented as intended by program developers; (b) exposure refers to the number of sessions implemented, length of each session, or frequency with which program aspects were implemented; (c) quality of delivery refers to the aspects of program delivery such as implementers enthusiasm, leader qualities and attitude; (d) participant responsiveness, which is the participant's participation and excitement for program sessions, and (e) program differentiation is used to ensure that the program being implemented is different from one that is currently happening. Most importantly, the intervention procedures and guidelines must be clearly and explicitly specified and understood by teachers and paraprofessional so that they can implement the intervention as it is intended (Lane et al., 2004).

Predictors of Fidelity

In a review of research on implementation fidelity by Dusenbury and colleagues (2003), they found that key elements of high fidelity include, teacher training, program characteristics,

teacher characteristics, and organizational characteristics. Research has mainly been concerned with program characteristics and organizational characteristics, rather than teacher training and the actual individual implementing. Simons-Morton, Parcel, Baranowski, Forthofer, and O'Hara (1991), in a study of Teenage Health Teaching Modules, found that teachers who received training were far more likely to implement the curriculum with fidelity than teachers who had no training. Multiple studies have found that live teacher training results in greater fidelity than video-based training alone, although video-based training is still better than no training (Basen-Engquist et al., 1994, Brock & Carter, 2015).

Another factor discovered by Schoenwald, Sheidow, Letourneau, and Liao (2003) to be associated with practitioner fidelity is organizational culture and climate. Experts have referred to organizational culture as "essentially a social indoctrination of unwritten rules that people learn as they try to fit in a particular group" (Schein, 1993). Schools want their culture to be positive and one that is healthy where students can adapt to new challenges. Climate in schools can be defined as quality and character of school life (National School Climate Center, 2017). In a review of implementation fidelity in school-wide programs by Bruhn, Hirsch, and Lloyd (2015), they found that fewer than half of the studies that reported implementation fidelity included a quantifiable index of implementation fidelity. They also found that in the majority of studies that reported implementation fidelity, it was only assessed once or annually (Bruhn et al., 2015). Implementation fidelity helps show if an intervention is effective or if something else is influencing the results and outcomes. Without implementation fidelity being reported there is no way to figure out which of those is occurring.

Lane and colleagues (2004) found that practitioner beliefs, motivation, and investment in the program have the potential to be important to implementation fidelity. Previous studies have

concluded that higher quality of program implementation can be attributed to greater self-efficacy, meaning one's belief of his or her ability to accomplish a specific task. For example, the Bierman and Conduct Problems Prevention Research Group found that young children at high risk for long-term antisocial behavior were given a universal level classroom program that included social skills training, academic tutoring, parent training, and home visits (McMahon et al., 1999). There were high levels of self-efficacy, which is one's belief in their ability to succeed in specific situations, amongst the implementers in this study (McMahon et al., 1999). Other studies have also found that teachers who had higher self-efficacy, enthusiasm, preparedness, as well as more positive beliefs about the effectiveness of the program were found to influence implementation quality, quantity, and sustainability (Ringwalt et al., 2003).

Factors that Contribute to Implementation Fidelity

Existing research shows that potential factors that contribute to implementation fidelity are intervention complexity, strategies to facilitate implementation, quality of delivery, and participant responsiveness (Carroll et al., 2007). The complexity of the intervention and description of the intervention has been found to affect implementation fidelity. For example, detailed and specific interventions have been found to be implemented with high fidelity compared to vaguely described interventions. A strategy that is beneficial to implementation fidelity is feedback (Faw, Hogue, & Liddle, 2005). This allows any deviations from the original intervention to be addressed and corrected throughout the intervention. The idea of feedback also influences the quality of delivery of an intervention (Carroll et al., 2007). Participant responsiveness is also a key factor in the effectiveness of an intervention and level of implementation fidelity. Multiple studies have found that the implementation fidelity of prescribed drug interventions for elderly individuals can be low because patients deliberately

failed to comply with their prescribed regimens. Patients chose to not to take the prescribed drug because of side effects and other reasons (Maidment, Livingston, & Katona, 2003).

After reviewing existing research, it is unclear whether interventions targeting implementation fidelity of academic treatments consistently improve their fidelity of implementation. These studies show predictors of implementation fidelity, aspects of implementation fidelity, and how to assess implementation fidelity. They do not however discuss quality indicators of implementation and how implementation fidelity can be addressed and improved. This present study will analyze the literature to synthesize what interventions improve implementation fidelity and what the key factors of implementation fidelity are in regards to academic interventions.

CHAPTER 3

Method

Study Design

A systematic review was conducted to gather and synthesize the current research on interventions to improve implementation fidelity of academic interventions. A systematic review was chosen over a meta-analysis because of the limited number of articles that were found on this topic. A systematic review allows for a starting point of what is in the literature right now, what has been done, and what are the next steps that should be taken to improve interventions for implementation fidelity of academic interventions.

Inclusion Criteria

The purpose of this review was to synthesize research findings across studies to identify what interventions are effective for increasing implementation fidelity in school based academic interventions. The following criteria were used to determine eligibility for inclusion in the present review. First, the article needed to report on the implementation of an academic intervention. Second, the intervention had to be conducted in the school setting. Interventions that were implemented in a clinic, home, or hospital setting were excluded. Third, the study needed to evaluate implementation fidelity as the primary dependent variable, meaning fidelity was the purpose of the study. Fourth, the study needed to actively manipulate some independent variable to influence fidelity of implementation. Finally, the intervention for which implementation fidelity data were collected, needed to target academic outcomes. Essentially, the inclusion criteria for this review limited studies to those addressing interventions to improve a teacher or other school personnel's implementation fidelity when implementing an academic intervention.

Study Identification

To identify studies for this review we searched the following databases: Academic Search Premier, Educational Resources Information Center (ERIC), and PSYCHINFO. The search terms used were: *Treatment fidelity* OR *treatment integrity* OR t*reatment adherence* OR *treatment reliability* OR *treatment implementation* OR *procedural fidelity* OR *procedural integrity* OR *procedural adherence* OR *procedural implementation* OR i*ntervention fidelity* OR *intervention integrity* OR *intervention adherence* OR *intervention reliability* OR *intervention implementation* OR *treatment plan integrity* OR *treatment plan fidelity* OR *treatment plan adherence* OR *treatment plan reliability* OR *treatment plan implementation* OR *fidelity of implementation* OR *implementation fidelity* OR *implementation integrity* OR *implementation adherence* OR *implementation reliability* OR *program fidelity* OR *program integrity* OR *program adherence* OR *program reliability* OR *program implementation* AND *school* OR *class* OR *classroom* OR *teacher* OR *education* OR *k-12* NOT *medicine* OR *medical* OR *drug* OR *prescription* OR *addiction*. This initial search produced 9,476 articles. To obtain a more manageable sample of articles, we applied the following limiters, school based intervention, intervention, behavior problems, treatment, attention deficit disorders, treatment effectiveness evaluation, classroom behavior, cognitive behavior therapy, treatment outcomes, prevention, evidence based practice, health behavior, schools, mental disorders, teachers, coping behavior, physical activity, childhood development, teaching methods, educational programs, special education, treatment compliance, at risk population, decision making, racial and ethnic difference, behavior change, measurement, behavioral assessment, mental health services, social skills, behavior disorders, and motivational interviewing. There were no restrictions made on publication date. This search resulted in 517 studies. These studies were imported into Zotero

(zotero.org) for further review. In Zotero, the first author reviewed the titles of each article and excluded studies that clearly did not meet the inclusion criteria. This first round of review resulted in excluding 234 articles. Articles that were excluded were moved to a folder entitled "Title Out". Next, the first author and a research assistant reviewed the abstracts of the remaining studies based on the inclusion criteria to identify articles appropriate for this review. Studies that did not meet the criteria of this study were rejected (N=132). These studies were moved to a folder entitled "Abstract Out". In the final step, the first author and research assistant reviewed the full text of the remaining studies. When studies did not meet the criteria of this review they were rejected (N=82). In the end, six relevant studies were used from this search.

In order to include all relevant articles a second search was conducted and to ensure interrater reliability a second research assistant completed the search as well. Multiple trainings were done with the second researcher. The following databases were searched: Academic Search Premier, Educational Resources Information Center (ERIC), and PSYCHINFO. The search terms used were (treatment* OR procedur* OR interven* OR implement* OR program* OR plan) AND (integrity OR fidelity OR adhere* OR reliab*) AND (school* OR class* OR educat* OR teach*) AND (academic OR read* OR Math* OR writi*) NOT (medic* OR drug* OR prescription* OR addict* OR clinic* OR health*). The limiters that were used were school based intervention, intervention, behavior problems, autism spectrum disorders, treatment, attention deficit disorders, treatment effectiveness evaluation, classroom behavior, cognitive behavior therapy, treatment outcomes, prevention, evidence based practice, health behavior, schools, mental disorders, teachers, coping behavior, physical activity, childhood development, teaching methods, educational programs, special education, treatment compliance, at risk population, decision making, racial and ethnic difference, behavior change, measurement,

elementary school students, behavioral assessment, mental health services, social skills, behavior disorders, and motivational interviewing. From this search four more articles were included.

In an effort to identify articles that may not have been captured in our electronic search, we conducted an ancestral or backward search of the 10 identified articles. To complete this search, the first author reviewed the reference section for each article and identified any titles that appeared to be relevant to this review. This search produced 63 potentially relevant articles. The first author retrieved and reviewed the text for each article. Two of these articles met criteria for inclusion in this review, bringing the total number of articles included in this review to 12. One article included two independent studies, so, although the total number of articles included in this review was 12, the total number of studies analyzed was 13 (see Figure 1).

The study identification information included the type of publication (e.g., journal article, dissertation, thesis, report), the year of publication and the names of the author(s). It also included the name of the journal. This information was intended to establish broad demographic characteristics of the publication sources.

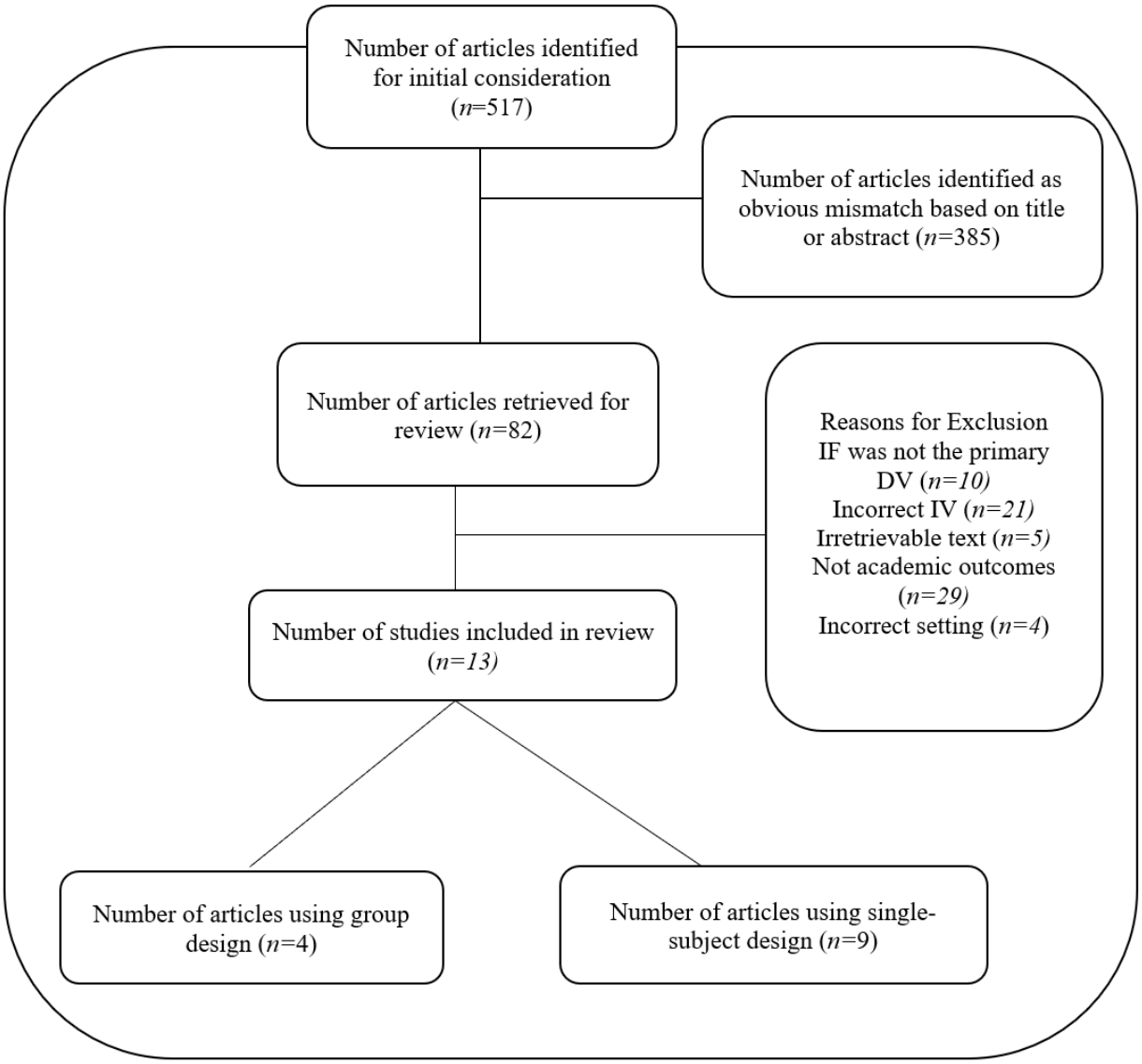

Figure 1. Schematic summary of the study identification process. IV= independent variable; DV= dependent variable; IF= implementation fidelity.

Coding Procedures

A coding protocol was developed to capture the relevant information from the 13 studies that were identified for review. To collect the information on the coding protocol, the primary researcher developed an online survey using Qualtrics online survey software (www.qualtrics.com). Once all studies were coded, the responses were exported into an Excel spreadsheet for analysis. The protocol grouped study information into ten categories: (a) study

information, (b) participant (educator) information, (c) setting information, (d) number of dependent measures, (e) dependent variables, (f) study design, (g) number of independent variables (interventions), (h) independent variable (intervention), and (i) quality indicators Council for Exceptional Children (CEC). These categories made it possible for the researchers to analyze the characteristics of each study in order to recommend conditions under which the studied outcomes might be possible.

Participant Characteristics

Coders reported the demographic characteristics of the participants, L1= the teacher or educator who tried to change student academic performance and L2= the person that tried to help L1 implement with fidelity. This included job title, intervention that was implemented, interventionist to improve the implementation fidelity, training or qualifications of the interventionist (L2), total sample size, number of male participants, number of female participants, race/ethnicity of participants, average age of participants, the grade the educator (L1) worked with, and the disability or disorder of the children with whom the educator worked. Participant characteristics provide a detailed understanding of the population from which meaningful findings may be generalized.

Setting Information

The setting information included where the study took place (public school, private school, charter school etc.). As well as, where the intervention was implemented for which fidelity was measured.

Independent Variable

The information recorded in this section included the intervention the researchers used, the length of the intervention (e.g., number of days from when treatment began to when it

ended), the number of treatment sessions, the duration of the sessions, and the identity of the interventionist (e.g., researcher, teacher).

Dependent Variable

To enable comparisons across studies, specific features of the dependent variable were identified, including what dependent variables the researchers reported, how the researchers measured the dependent variable (e.g., direct observation, questionnaire, survey, or rating scale), how direct observation data was collected (e.g., frequency count, interval recording, duration recording, latency recording etc.), what questionnaires (rating scale) the researchers used, what standardized assessment the researchers used, whether or not the researchers measured the maintenance of the intervention effect (follow-up measures), whether or not the effect of the intervention was maintained, if adherence (degree to which the program was implemented as prescribed) was measured, if exposure (the number of sessions that participants received) was measured, if quality of delivery (aspects of program delivery such as implementer enthusiasm, preparedness, and attitude) was measured, if participant responsiveness was measured, and if program differentiation was measured.

Design

Each study was categorized according to features of the research design so that comparisons across and within designs could be completed. Aspects included the type of study (single-subject or N), the unit of analysis (individual or group), and the specific type of design (e.g., quasi-experimental, multiple baseline across participants).

Quality of Evidence

The quality of evidence was assessed using a quality indicator tool we created called "Quality Indicator Tool for Studies of Implementation Fidelity Interventions." It was adapted

from the Council for Exceptional Children Standards for Evidence-Based Practices in Special Education (Council for Exceptional Children, 2014) and "A Conceptual Framework for Implementation Fidelity" (Carroll et al., 2007). We started with the Council for Exceptional Children Standards (CEC), then we modified that based on four supplemental elements of implementation fidelity (Content, Coverage, Frequency, and Duration) and four moderators of implementation fidelity (Intervention Complexity, Facilitation Strategies, Quality of Delivery, and Participant Responsiveness) that Carroll and colleagues (2007) discusses in "A Conceptual Framework for Implementation Fidelity" (Carroll et al., 2007). Next, we sent out the tool to experts in the field to get their feedback on what had been created and adapted. We collaborated and made the necessary edits. Then an article was used as a practice code and any issues were worked out based on the feedback and results from the practice code.

Coding

There were four individuals that coded the articles included in this study. Two school psychology masters students, the thesis advisor who is a faculty member, and a psychology undergraduate student. Each article was coded by two coders to ensure interrater reliability. After the articles were coded by an individual coder, the two coders met and discussed the articles they coded and resolved any differences on the codes.

Interrater Reliability

All raters had completed a semester-long research methods course, that explained the basics of coding and interrater reliability. Additionally, three team coding meetings were held with all four members to learn the coding procedures for this study and went over an article together. The team agreed at 90% or better in order for there to be high interrater reliability and to consider the article coded. After training on how to use the coding protocols, each rater

completed a practice protocol prior to coding studies for this review. Additionally, all articles were double coded to assess interrater reliability. One of the four raters computed the Tau-U scores for all single-subject studies.

In order to train the second coder for article sorting, the first author met with her to train her on sorting the articles. The inclusion and exclusion criteria were explained, and a copy of both criteria were provided, and any questions were answered. They also met to discuss the articles that met criteria resolve any discrepancies. The second coder (psychology undergrad) and the first author met to train on how to code the selected articles. The primary researcher also met with the school psychology master's student to train her on coding as well. The coding protocol was provided and explained with examples from an article. Next all four coders met to do a practice code together and address any questions with the coding survey. Together the article was read and coded and questions were answered, and each code was explained with examples.

Raters met to resolve any disagreements on coded responses for articles they coded. During this time, they identified any discrepancies in the coding and then discussed as a group while referencing the article itself what the correct answer would be and then came to a unanimous consensus on what the correct code was for each article.

CHAPTER 4

Results

Study Features

There were 12 articles included in this review, with one of the articles having two separate studies included in the article. A total of 13 studies were analyzed for this review. Two of the 12 articles had been published as theses, which go through a different process for approval than studies published in peer reviewed journals.

More than half of the studies included in this review were single subject research designs (*n*=9). Four of the 13 studies included used a group research design (*n*=4), and of those studies one used a randomized control trial experimental design, whereas the remaining three studies used a quasi-experimental design.

Participant Characteristics

Student participants. All together the studies in this review included 205 participants. In almost half (*n*=6) of the studies, researchers did not report the ethnicity or race of the participants. Caucasian/White participants were included in 86% of the studies in which the researchers reported race/ethnicity of the participants. African American/Black participants were included in 43% of the studies and Latino/Hispanic and Asian participants were included in 14% of the studies.

Six of the studies included second grade students. Students in Kindergarten, first grade and third grade were included in five of the studies. Four studies included students in the fourth grade and three studies included students in the fifth grade. Only one study included grades 9-12. The ages that were reported ranged from 6-10 for the elementary aged students. The majority of the students included had academic performance deficits or academic concerns (e.g.,

reading and math). There were also struggling readers, deficits in math, deficits in letter naming skills, and at-risk failure in writing. The majority of consultants were school psychology doctorate students or practicum students (n=10).

Teacher participants. All together 39 teachers were involved in the studies. Each teacher was certified to teach as a general education teacher. Only four of the studies reported on the ethnicity of the teachers. Of those reported, one was African American and the rest were Caucasian. Of the genders reported all of the teachers were female except for one male teacher. One teacher had a Master's in Education and had been teaching for 22 years. Another teacher taught a program for gifted children. It was reported that some of the teachers had experience with student referrals and working with school psychologists. Three of the teachers only had one year of teacher experience but the majority had at least seven years of teaching experience (*n*=29). In some of the studies it was explicitly reported that participation in the study was voluntary.

Setting

The studies included four rural communities, four urban communities, and two suburban communities. The other three studies did not report on the setting. Six of the studies mentioned that the population in the schools that were included had high levels of economically disadvantaged students and received free and reduced lunch. Of the six schools reported to be low socioeconomic status, the range of economically disadvantaged students within the school was 43% to 83% with a median of 68%. All 13 studies took place in public schools in the United States. Only one study was done in a high school setting (grades 9-12), while the other 12 studies took place in elementary schools.

The inclusion criteria for this review limited studies to those addressing interventions to improve a teacher or other school personnel's implementation fidelity when implementing an academic intervention. The majority of studies evaluated the implementation of reading interventions (n=6), while other studies evaluated math interventions (n=4), and a few studies evaluated combined math and reading interventions (n=2), with one study evaluating a writing intervention. As a result, we can have greater confidence in the findings of the reading studies because there were more studies addressing reading interventions.

Interventions

Across all studies performance feedback for the teacher implementing the intervention was the most common feature of the treatment, either as a standalone intervention or as part of an intervention package. Performance feedback was the intervention most used in the single-subject studies. Performance feedback can be given as correction both during implementation and after, and it can also include praise for proper implementation. Out of the 13 studies, 4 used performance feedback alone as the intervention. Two studies used teacher and in-class trainings as the intervention to improve teacher/school personnel implementation fidelity. The intervention used in three studies was a combination intervention of performance feedback and some form of consultation or training. Professional development was used as the intervention in three studies. Professional development included district, school, and grade-level teams meeting together to be trained by someone who is knowledgeable of a certain topic.

In nine studies the researchers used permanent products to evaluate the effect of the intervention. Permanent products are physical or digital evidence of participants' performance. A few examples of permanent products are checklists, intervention records, and charts. In the Sanetti and Kratochwill (2009) study they used graphs of the student's progress on a math

assignment and the student feedback form as permanent products. Permanent products were used because they are more practical for assessment of fidelity of interventions given that they do not require the presence of the consultant (Gilbertson et al., 2007). The majority of the studies (67%) used a form of a checklist as the permanent product. The permanent products were completed by the researcher based on the steps that were completed by the teacher implementing the intervention.

Effects

The general aspects of each single-subject study were coded to identify the intervention, measure used to evaluate the dependent variable, and the Tau-U effect size. The general aspects of each group study were coded to identify the intervention, the variables, and the Cohen's *d* effect size (see Table 2). The underlying metric in the group studies was different measures of implementation fidelity. The Tau-U effect sizes of performance feedback interventions ranged from 0.50-0.92 (see Table 1). The intervention with the lowest Tau-U effect size was teacher training (0.41). The comparison being made for single-subject studies was from baseline to treatment.

Table 1

Implementation Fidelity Effect Size Estimation for Single-Subject Studies

Design Type	Intervention	Study	Measure	Tau-U
Single Subject	In Class Training & Team Meetings	Duhon et al. (2009) Exp. 1	Permanent Products: Intervention Records	1.0
	Performance Feedback	Duhon et al. (2009) Exp.2	Permanent Products: Intervention Records	0.92
	Faded Classroom Training & Performance Feedback	Gilbertson et al. (2007)	Permanent Products: Chart, Flashcards, & Math Probe	0.78
	Performance Feedback	Gross (2013)	Permanent Products: Checklist	0.77
	Teacher Training	Mocco (2014)	Permanent Products: Intervention Checklist & Direct Observation	0.41
	Performance Feedback	Mortenson & Witt (1998)	Permanent Products: Checklist	0.67
	Follow Up Meeting and Performance Feedback	Noell et al. (2000)	Permanent Products: Checklist	0.63
	Consultation and Performance Feedback	Noell et al. (1997)	Permanent Products: Treatment Steps Checklist	0.76
	Performance Feedback (TIPP)	Sanetti & Kratochwill (2009)	Permanent Products: Checklist	0.50

Table 2

Implementation Fidelity Effect Size Estimations for Group Design Studies

Design Type	Intervention	Study	Variable	Cohen's *d*
Randomized-controlled trial	Professional Development	Harris et al. (2015)	Story Elements	*d*= 1.53
			Quality	*d=0.*98
			Personal Narrative Elements	*d*=1.26
			Motivation (Intrinsic)	*d=1.3*
			Effort	*d*=1.26
Quasi-experiment	Follow-up & integrated support	Noell et al. (2017)	Weekly vs. Integrated Consultant Evaluation Form	*d*=1.3
			Consultant Rating Scale	*d*=0.49
			Intervention Implementation	*d*=0.45
			Intervention Effective	*d*=0.33
			Satisfaction with Effect	*d*=0.21
				d=0.065
	Professional Development	Capraro et al. (2016)	Mathematics (longitudinal)	*d*=1.41
			Science (longitudinal)	*d*=2.03
			Reading (longitudinal)	*d*=0.36
			Matched mathematics	*d*=0.62
			Matched science	*d*=0.92
			Matched reading	*d=-0.8*
	Professional Development	Vernon-Feagans et al. (2015)	Implementation Quality	*d*=0.95
			Teacher Efficacy	*d=0.50*
			Spelling of Sounds	*d=0.46*

Out of the four group design studies three used a quasi-experiment design and one used a randomized controlled trial design. Professional Development was the intervention used to improve teacher/school personnel implementation fidelity in three out of the four studies. Follow-up and integrated support were the other interventions used in the studies.

Quality of Evidence

The quality of evidence was assessed using an adaptation of the Council for Exceptional Children Standards for Evidence-Based Practices (EBPs) in Special Education, The CEC (2014) Standards for EBPs. These standards work as a quality appraisal tool to support the categorization of evidence-based practices in special education (Lane, Common, Royer, & Muller, 2014). The Standards for EBP are different than the CEC standards alone because they include eight categories of quality indicators and include weighted Quality Indicator coding, meaning partial credit can be given. Absolute coding and weighted coding can be done, which captures a wider variety of methodologically sound practices and studies. The purpose of this analysis was to determine the degree of confidence that one might have in the findings reported by the various researchers. Each study was evaluated on primary and secondary quality indicators.

We developed a separate QI evaluation tool specifically to address implementation fidelity. We sent out the tool to experts in the field to get their feedback on what had been created and adapted. We collaborated and made the necessary edits. Then an article was used as a practice code and any issues were worked out based on the feedback and results from the practice code. This project was the pilot for this tool.

This tool included secondary quality indicator items targeting implementation fidelity items adapted from *A Conceptual Framework for Implementation Fidelity* (Carroll et al., 2007). The article includes four supplemental elements of implementation fidelity (Content, Coverage, Frequency, and Duration) and four moderators of implementation fidelity (Intervention Complexity, Facilitation Strategies, Quality of Delivery, and Participant Responsiveness that should be evaluated when looking at implementation fidelity.

CEC Quality Indicators

Four out of the nine single-subject studies met 21 out of the 22 CEC quality indicators, (Duhon et al., 2009; Gilbertson et al., 2007; Gross, 2013; Noell et al., 1997). Two studies, Mocco (2014) and Noell and colleagues (2002) met 20 out of the 22 quality indicators. Three studies met at least 17 of the quality indicators (Duhon et al., 2009; Mortenson & Witt, 1998; Sanetti & Kratochwill, 2009).

On individual items, 13 out of the 22 quality indicators were met by every single study (see Table 3). Whereas item 3.2 was not included in any of the studies. Item 3.2 said, "The study describes any specific training (L2) (e.g., amount of training, training to a criterion) or qualifications (e.g., professional credential) required to implement the intervention, and indicates that the interventionist has achieved them. Only six out of the nine studies met Item 3.1, which says, "The study describes the role of the intervention agent (L2; e.g., researcher, coach, teacher, school psychologist, consultant) and, as relevant to the study, background variables (e.g., race/ethnicity, educational background/licensure)." All other CEC quality indicators items were reported by all nine studies or eight out of the nine studies.

Table 3

CEC Quality Indicators Single-Subject Studies

	1	**2**		**3**		**4**		**5**			**6**						**7**					**8**	Totals (22)
	1	1	2	1	2	1	2	1	2	3	1	2	3	5	6	7	1	2	3	4	5	2	
Study																							
Gilbertson et al. (2007)	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	21
Gross (2013)	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	21
Mocco (2014)	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	20
Mortenson & Witt (1998)	1	1	1	0	0	1	1	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	18
Noell et al. (1997)	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	21
Noell et al. (2000)	1	1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	20
Sanetti & Kratochwill (2009)	1	1	1	0	0	1	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	17
Duhon et al. (2009) Exp. 1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	21
Duhon et al. (2009) Exp. 2	1	1	1	1	0	0	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	18
Totals (9)	9	9	9	6	0	8	8	8	6	8	9	9	9	8	9	9	8	9	9	9	9	9	

Note. The numbers on the top horizontal axis are the CEC Quality Indicators. For example, 1.1, 2.1, 2.2. each correspond to a quality indicator. The 1s and 0s represent whether each study included that quality indicator. A 1 means it was included in the study and the 0 means it was not included in the study. The totals on the vertical axis is the total number of quality indicators each study included. The totals on the horizontal axis are the totals for each quality indicator.

Table 4

CEC Quality Indicators Group Design Studies

	1	2		3		4		5			6								7						8		Totals (24)
Study	1	1	2	1	2	1	2	1	2	3	1	2	3	4	8	9	1	2	3	4	5	6	1	3			
Harris et al. (2015)	1	1	1	0	0	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1			20
Noell et al. (2017)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1			24
Capraro et al. (2016)	1	1	1	0	0	1	1	1	0	0	1	0	1	1	1	1	1	1	1	1	1	1	0	0			17
Vernon-Feagans et al. (2015)	1	1	1	1	1	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1			19
Totals (4)	4	4	4	2	2	3	2	2	2	2	4	3	4	4	4	4	4	4	4	4	4	4	3	3			

Note. The numbers on the top horizontal axis are the CEC Quality Indicators. For example, 1.1, 2.1, 2.2. each correspond to a quality indicator. The 1s and 0s represent whether each study included that quality indicator. A 1 means it was included in the study and the 0 means it was not included in the study. The totals on the vertical axis is the total number of quality indicators each study included. The totals on the horizontal axis are the totals for each quality indicator.

One of the four group design studies included all of the CEC quality indicators (24). The only study was Noell and colleagues (2017). The next two articles included 19 and 20 of the quality indicators (Harris et al., 2015; Vernon-Faegans et al., 2015). The last group design study included 17 of the quality indicators (Capraro et al., 2016).

On individual items, 14 out of the 24 items were met by every single study. Four of the items were included by three out of the four studies. The remaining items six items were met by half of the studies (see Table 4).

Secondary Quality Indicators

Secondary quality indicators were adapted based on four supplemental elements of implementation fidelity (Content, Coverage, Frequency, and Duration) and four moderators of implementation fidelity (Intervention Complexity, Facilitation Strategies, Quality of Delivery, and Participant Responsiveness) that Carroll and colleagues (2007) discusses in A Conceptual Framework for Implementation Fidelity (Carroll et al., 2007). Two studies, Noell and colleagues (2017) and Duhon and colleagues (2009) Experiment 1, met 14 of the secondary quality indicators. Two studies met 13 out of the 15 secondary quality indicators: Gilbertson and colleagues (2007) and Gross (2013). Two studies included 12 out of the 15 secondary quality indicators, Mocco (2014) and Noell and colleagues (1997). Mortenson and Witt (1998) included 11 out of the 15 secondary quality indicators in the study. Noell and colleagues (2000) included 10 of the 15 secondary quality indicators. Harris et al. (2015) scored 9 on the secondary quality indicators of the 15. Two studies, Sanetti and Kratochwill (2009) and Capraro et al. (2016) included eight of the 15 secondary quality indicators. Duhon and colleagues (2009), Exp. 2 included six of the 15 secondary quality indicators. Vernon-Feagans and colleagues (2015) scored 5 out of 15 on the secondary quality indicators.

On individual secondary indicators, one item, item 2.2, was reported by all 13 studies. Three items were reported on by 12 of the 13 studies (1.2, 9.3, 10.4). Item 2.3 was included in 11 of the 13 studies. Ten studies included items 5.1 and 5.3, which are also items that are on the CEC quality indicators. Items 9.4 and 10.2 were included in 9 of the 13 studies and items 5.2 and 9.1 were included in 8 of the 13 studies. Six studies included item 10.1 and five studies included 9.5 (see Table 5). No studies included 10.3, which said, "The researchers evaluate and report on the quality of the delivery of the intervention (e.g., enthusiasm of the interventionist)."

Table 5

Secondary Quality Indicators

Study	1.2	2.2	2.3	**5.1**	**5.2**	**5.3**	9.1	9.2	9.3	9.4	9.5	10.1	10.2	10.3	10.4	Totals (15)
Single-Case																
Gilbertson et al. (2007)	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1	13
Gross (2013)	1	1	0	1	1	1	1	1	1	1	1	1	1	0	1	13
Mocco (2014)	0	1	1	1	1	1	1	0	1	1	1	1	1	0	1	12
Mortenson & Witt (1998)	1	1	1	1	1	1	1	1	1	1	0	0	0	0	1	11
Noell et al. (1997)	1	1	1	1	1	1	1	1	1	1	0	0	1	0	1	12
Noell et al. (2000)	1	1	1	1	0	1	1	1	1	0	0	0	1	0	1	10
Sanetti & Kratochwill (2009)	1	1	1	0	0	0	0	1	1	0	0	1	1	0	1	8
Duhon et al. (2009) Exp. 1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	14
Duhon et al. (2009) Exp. 2	1	1	0	1	0	1	0	0	1	0	0	0	0	0	1	6
Group																
Harris et al. (2015)	1	1	1	0	1	1	0	1	1	1	0	0	0	0	1	9
Noell et al. (2017)	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	14
Capraro et al. (2016)	1	1	1	1	0	0	0	0	1	1	0	1	1	0	0	8
Vernon-Feagans et al. (2015)	1	1	1	0	0	0	0	1	0	0	0	0	0	0	1	5
Totals (13)	12	13	11	10	8	10	8	10	12	9	5	6	9	0	12	

Note. Bolded items were taken from the CEC standards for implementation fidelity.

CHAPTER 5

Discussion

The purpose of this review was to identify and evaluate the evidence of interventions to improve the implementation fidelity of academic interventions and to evaluate the quality of the existing research with a focus on the quality of the research on the most successful interventions.

Interventionists

In all of the reviewed studies the consultant that was implementing the intervention to improve the teacher's implementation fidelity was a researcher, doctoral, or practicum student. None of the studies reviewed included typically available school personnel (e.g., teacher, school psychologist) as the consultant. This shows that a researcher-implemented intervention may be successful, but it does not show that the same intervention would work when administered by typical school personnel. Consequently, it is unclear whether these interventions would be effective when administered under typical school circumstances. Now that we know these implementation fidelity interventions are effective when done by a researcher, we need researchers to look at using these interventions in a school context with the consultant being a typical school personnel. In order for these concepts to be a broadly adopted approach the interventions need to be simple enough and have enough detail in order for schools without universities and researchers nearby to still be able to benefit from them. The majority of schools do not have access to research personnel, so it is critical for research to be done that using typical school personnel as the consultants to see if what has been found is possible under normal school circumstances.

Quality Indicators

There are two quality indicators items that were hardly reported on among the articles. Item 3.2 of the CEC quality indicators ("The study describes any specific training (L2) (e.g., amount of training, training to a criterion) or qualifications (e.g., professional credential) required to implement the intervention, and indicates that the interventionist has achieved them") was only reported in two articles (15%). We believe this is a critical component for that should be included in studies because it allows the reader to understand what training is needed in order to incorporate a similar intervention. The secondary quality indicator 10.3 was not reported by any of the articles ("The researchers evaluate and report on the quality of the delivery of the intervention (e.g., enthusiasm of the interventionist)."). Quality of delivery is defined as "the manner in which a teacher, volunteer, or staff member delivers a program." (Carroll et al., 2007). We believe that the quality of the delivery of the intervention could be a critical feature of implementation fidelity because of the research that has been done, specifically by Carroll and colleagues (2007). Carrol and colleagues (2007) reported the idea of feedback also influences the quality of delivery of an intervention. Interventions can be delivered, but delivered poorly, which in return the implementation fidelity could be adversely affected. We believe that the delivery makes a difference in the implementation of the intervention. The delivery of the intervention should be reported so that better conclusions can be drawn based off the results of the interventions and the implementation fidelity.

CEC Quality Indicators

When looking at all of the studies and whether or not they met the CEC quality indicators, we found that only one of the studies met all of the CEC quality indicators (Noell et al., 2017). According to CEC, only one of these studies would be included in an Evidence Based

Practice review (Council for Exceptional Children, 2014). It was also a group design study. The question we need to ask ourselves is how we can efficiently include critical components of implementation fidelity in our research studies and practice. We know that implementation fidelity is critical in understanding if an intervention is working, so we must be better at monitoring it and giving enough details in order to maintain fidelity. The questions that need to be asked and future research should address are, what do we do with the studies that don't meet all of the CEC quality indicators? Do we examine them more closely? If it is determined that all the quality indicators are essential then, future researchers should work to address the CEC quality indicators in their studies in order to be able to compare studies and see effects when there is a universal standard.

The results from this review confirm what is found in the current literature on implementation fidelity (Carroll et al., 2007). There needs to be a focus on the implementation of an intervention before it can be deemed ineffective. The research shows that in the past we have been more concerned with program characteristics, but we are moving more towards teacher characteristics and the training that teachers are given (Carroll et al., 2007). In the studies included in this review teacher training was an important component to the interventions to improve teacher's implementation fidelity. One thing that research suggests but was not always reported in the studies we looked at was the complexity of the intervention. It is important for the complexity of the intervention to be included with great detail in order to ensure high levels of implementation fidelity.

Secondary Quality Indicators

Overall, the articles did fairly well on the secondary quality indicators as 8 of the 15 studies included at least 10 of the 15 indicators. The article that included the least secondary

quality indicators was also one of the articles that included the least CEC quality indicators. At a quick glance the articles did similarly on the CEC quality indicators and the secondary quality indicators that focused on implementation fidelity. Articles that scored high on the CEC quality indicators also scored high on the secondary quality indicators. When comparing just the CEC quality indicators that involved implementation fidelity (5.1, 5.2, and 5.3) to the secondary quality indicators focusing on implementation fidelity all of the articles scored similarly on both. This shows that the secondary quality indicator tool that was created has some validity and measured what it was intended to measure, and it also tells us that the CEC quality indicators focused on implementation fidelity may be enough as they seem to capture if an article is including implementation fidelity measures or not. It is possible that the secondary quality indicators are not necessary but could be referenced when needing to take a deeper look at the implementation fidelity of an intervention.

There were three items that were not included in the majority of articles. One of the items was 9.5 which said, “The researchers directly measured and reported the duration of intervention from initiation to termination and indicated that the duration was consistent with the prescribed intervention procedures.” Also 10.1 which says, “The researchers include some evaluation of the complexity of the intervention” and 10.3 which says, “The researchers evaluate and report on the quality of the delivery of the intervention.” Even though these items were not reported by many of these articles, we still believe they should be considered in the future and more research should be done to see if they are critical components. We believe these items as well as the other secondary quality indicators are important because they show the validity of the intervention and allow for easier replication.

When we can understand what has been done before and the quality indicators that were included in a study we are able to make better decisions in the classroom with students. That is what it is all about. What we found opens new doors to what is important when it comes to implementation fidelity and what needs to be done in the future.

Single Case vs. Group Design

None of the group studies included performance feedback as the intervention. It is possible that none of the group studies used performance feedback to improve implementation fidelity because performance feedback takes a lot of work and money to give the support of performance feedback. The group studies that were included had many participants which made other interventions more convenient for a large scale. In order to make more comparisons and draw better conclusions, we believe it would be insightful to have a study done on performance feedback in a group design. The group studies did not do as well on the secondary quality indicators as the single case studies. The pooled mean of the effect size of performance feedback was 0.72 which shows that it is a moderately successful intervention for improving implementation fidelity. If performance feedback is the frontrunner for implementation fidelity interventions, then we need to find a way to implement it and measure it in group studies.

It appears in the literature that there is much more research on reading interventions and implementation fidelity interventions to improve the delivery of reading interventions than any other subject (math, writing, science, social studies etc.). We have greater confidence in the reading interventions and the implementation fidelity interventions for reading. There is a gap between implementation fidelity intervention for reading interventions and other subjects. We need more research on implementation fidelity interventions for math, science, social studies,

writing. Researchers should conduct studies that will give more information and insight on these subjects.

Limitations and Future Research

One of the limitations of this study was the lack of studies on interventions to improve implementation fidelity of academic interventions. There were only 13 studies that fit the criteria. There should be more studies done on this topic in order to have more studies to compare and draw further conclusions. Along with few studies, there was a gap between reading interventions and all other subjects. Future research should focus their efforts on other subjects in order to know what works for those subjects as well, especially math and writing. There is also a need to conduct these studies with special education teachers to see if the effects are similar. All of the included studies for this review included only general education teachers. Another idea for future research is to include more studies in the high school setting. It is unclear how effective these interventions are in the high school setting because of lack of research in that specific setting.

Other limitations of this study were that we created the secondary quality indicator tool, and this was the first time it was being used. This tool should be used in more studies in order to gain validation of the tool before it is continued to be used. We did find it to be consistent with the CEC implementation quality indicators, but more test runs of the tool may help determine its validity and reliability. Overall, as a field there should be a consensus on what items (CEC quality indicators) matter or not. It would be beneficial to decide if some need to be weighted more than others and whether some are mandatory, and others are supplementary. If this was decided there would be better results and more of a consensus when looking across studies.

Another limitation of this study was that we were not able to compare our results from the single subject studies to the group studies. In the future, the BC-SMD statistic could be used to compare group studies to single subject studies (Pustejovsky & Ferron, 2017). Quality of delivery and complexity of the intervention were items did not clear descriptions or good accepted measures to use to determine if they were included in a study. Future research could create good accepted measures and descriptions of quality of delivery and complexity of the intervention for future studies.

Conclusion

An important relationship exists between implementation fidelity and the effectiveness of interventions. Performance feedback was the most used intervention to improve the implementation fidelity of academic interventions in single case studies and professional development was the most used intervention used in the group design studies. More studies should be done specifically with group design studies in order to draw further conclusions. There is also a discrepancy between the amount of studies on reading interventions compared to other subjects. It would be beneficial to have these studies replicated but with typical school personnel instead of researchers and graduate students. With such few studies that met the inclusion criteria, there is a need for more research in this crucial area to benefit students and their education.

References

Basen-Engquist, K., O'Hara-Tompkins, N., Lovato, C. Y., Lewis, M. J., Parcel, G. S., & Gingiss, P. (1994). The effect of two types of teacher training on implementation of smart choices: A tobacco prevention curriculum. *Journal of School Health, 64*(8), 334–339.

Brock, M. E., & Erik, W. C. (2015). Effects of a professional development package to prepare special education paraprofessionals to implement evidence-based practice. *Journal of Special Education, 49*(1), 39–51.

Bruhn, A. L., Hirsch, S. E., & Lloyd, J. W. (2015). Treatment integrity in school-wide programs: A review of the literature (1993–2012). *The Journal of Primary Prevention, 36*(5), 335–349.

Capraro, R. M., Capraro, M. M., Scheurich, J. J., Jones, M., Morgan, J., Huggins, K. S., . . . Han, S. (2016). Impact of sustained professional development in STEM on outcome measures in a diverse urban district. *Journal of Educational Research, 109*(2), 181–196. https://doi.org/10.1080/00220671.2014.936997

Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., & Balain, S. (2007). A conceptual framework for implementation fidelity. *Implementation Science, 2*(1), 40-49.

Council for Exceptional Children. (2014). *Council for Exceptional Children standards for evidence-based practices in special education*. Arlington, VA: Author. Retrieved from http://www.cec.sped.org/~/media/Files/Standards/Evidence%20based%20Practices%20and%20Practice/EBP%20FINAL.pdf

Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of control? *Clinical Psychology Review, 18*(1), 23–45.

Duhon, G. J., Mesmer, E. M., Gregerson, L., & Witt, J. C. (2009). Effects of public feedback during RTI team meetings on teacher implementation integrity and student academic performance. *Journal of School Psychology, 47*(1), 19–37. https://doi.org/10.1016/j.jsp.2008.09.002

Dusenbury, L., Brannigan, R., Falco, M., & Hansen, W. B. (2003). A review of research on fidelity of implementation: Implications for drug abuse prevention in school settings. *Health Education Research, 18*(2), 237–256.

Elliott, S. N., & Busse, R. T. (1993). Effective treatments with behavioral consultation. In J. E. Zins, T. R. Kratochwill, & S. N. Elliott (Eds.), *The Jossey-Bass social and behavioral science series. Handbook of consultation services for children: Applications in educational and clinical settings* (pp. 179-203). San Francisco, CA: Jossey-Bass.

Faw, L., Hogue, A., & Liddle, H. A. (2005). Multidimensional implementation evaluation of a residential treatment program for adolescent substance abuse. *American Journal of Evaluation, 26*(1), 77–94. doi:10.1177/1098214004273183

Fixsen, D. L., Naoom, S. F., Blase, K. A., & Friedman, R. M. (2017). Implementing effective educational practices at scales of social importance. *Clinical Child and Family Psychological Review, 20*(1), 25–35.

Gilbertson, D., Witt, J. C., Singletary, L. L., & VanDerHeyden, A. (2007). Supporting teacher use of interventions: Effects of response dependent performance feedback on teacher implementation of a math intervention. *Journal of Behavioral Education, 16*(4), 311–326. https://doi.org/10.1007/s10864-007-9043-0

Greenberg, M. T., Domitrovich, C. E., Graczyk, P. A., & Zins, J. E. (2005). *The study of implementation in school-based preventive interventions: Theory, research, and practice*

(Vol. 3). Rockville, MD: Center for Mental Health Services, Substance Abuse and Mental Health Services Administration.

Gresham, F. M., Gansle, K. A., & Noell, G. H. (1993). Treatment integrity in applied behavior analysis with children. *Journal of Applied Behavior Analysis, 26*(2), 257–263. doi:10.1901/jaba.1993.26-257

Gross, T. J. (2013). *Enhancing treatment integrity through systematic fading with indiscriminable contingencies* (Doctoral dissertation). Retrieved from ProQuest. (99161-061)

Harris, K. R., Graham, S., & Adkins, M. (2015). Practice-based professional development and self-regulated strategy development for Tier 2, at-risk writers in second grade. *Contemporary Educational Psychology, 40*(4), 5–16. https://doi.org/10.1016/j.cedpsych.2014.02.003

Individuals with Disabilities Education Act, 20 U.S.C. § 1400 (2004).

Kratochwill, T. R., & Shernoff, E. S. (2003). Evidence-based practice: Promoting evidence-based interventions in school psychology. *School Psychology Quarterly, 18*(4), 389-408.

Lane, K. L., Beebe-Frankenberger, M. E., Lambros, K. M., & Pierson, M. (2001). Designing effective interventions for children at-risk for antisocial behavior: An integrated model of components necessary for making valid inferences. *Psychology in the Schools, 38*(4), 365–379. doi:10.1002/pits.1025

Lane, K. L., Bocian, K. M., MacMillan, D. L., & Gresham, F. M. (2004). Treatment integrity: An essential—but often forgotten—component of school-based interventions. *Preventing School Failure: Alternative Education for Children and Youth, 48*(3), 36–43.

Lane, K. L., Common, E. A., Royer, D. J., & Muller, K. (2014). *Group comparison and single-case research design quality indicator matrix using Council for Exceptional Children 2014 standards*. Retrieved from http://www.ci3t.org/practice

Maidment, R., Livingston, G., & Katona, C. (2003). "Just keep taking the tablets": Adherence to antidepressant treatment in older people in primary care. *International Journal of Geriatric Psychiatry, 17*(8), 752–757. doi: 10.1002/gps.688

Martens, B. K., & McIntyre, L. L. (2009). The importance of treatment integrity in school-based behavioral intervention. In A. Akin-Little, S. G. Little, M. A. Bray, & T. J. Kehle (Eds.), *Behavioral interventions in schools: Evidence-based positive strategies* (pp. 59–71). Washington, DC: American Psychological Association.

McMahon, R. J., Bierman, K. L., Coie, J. D., Dodge, K. A., Greenberg, M. T., Lochman, J. E., & Pinderhughes, E. E. (1999). Initial impact of the fast track prevention trial for conduct problems: I. the high-risk sample." *Journal of Consulting and Clinical Psychology, 67*(5), 631–647.

Mocco, C. (2014). *Evaluating teacher responsiveness to antecedent based intervention for increasing treatment regularity* (Doctoral dissertation). Retrieved from ProQuest. (99150-224)

Mortenson, B. P., & Witt, J. C. (1998). The use of weekly performance feedback to increase teacher implementation of a prereferral academic intervention. *School Psychology Review,* 27(4), 613–627.

National School Climate Center. Retrieved April 1, 2017, from http://www.schoolclimate.org/

Noell, G. H., Gresham, F. M., & Gansle, K. A. (2002). Does treatment integrity matter? A preliminary investigation of instructional implementation and mathematics performance.

Journal of Behavioral Education, 11(1), 51–67. https://doi.org/10.1023/A:1014385321849.v

Noell, G. H., Volz, J. R., Henderson, M. Y., & Williams, K. L. (2017). Evaluating an integrated support model for increasing treatment plan implementation following consultation in schools. *School Psychology Quarterly*, *32*(4), 525–538. https://doi.org/10.1037/spq0000195

Noell, G. H., Witt, J. C., Gilbertson, D. N., Ranier, D. D., & Freeland, J. T. (1997). Increasing teacher intervention implementation in general education settings through consultation and performance feedback. *School Psychology Quarterly, 12*(1), 77-88.

Noell, G. H., Witt, J. C., LaFleur, L. H., Mortenson, B. P., Ranier, D. D., & LeVelle, J. (2000). Increasing intervention implementation in general education following consultation: a comparison of two follow-up strategies. *Journal of Applied Behavior Analysis*, *33*(3), 271-284. https://doi.org/10.1901/jaba.2000.33-271

Perepletchikova, F., Treat, T. A., & Kazdin, A. E. (2007). Treatment integrity in psychotherapy research: Analysis of the studies and examination of the associated factors. *Journal of Consulting and Clinical Psychology, 75*(6), 829–841. http://dx.doi.org/10.1037/0022-006X.75.6.829

Pustejovsky, J. E., & Ferron, J. M. (2017). Research synthesis and meta-analysis of single-case designs. In J. M. Kaufmann, D. P. Hallahan, & P. C. Pullen (Eds.), *Handbook of special education* (2nd ed.) New York, NY: Routledge.

Ringwalt, C. L., Ennett, S., Johnson, R., Rohrbach, L. A., Simons-Rudolph, A., Vincus, A., & Thorne, J. (2003). Factors associated with fidelity to substance use prevention curriculum guides in the nation's middle schools. *Health Education and Behavior, 30*(3), 375–391.

Sanetti, L. M. H., & Kratochwill, T. R. (2009). Treatment integrity assessment in the schools: An evaluation of the Treatment Integrity Planning Protocol. *School Psychology Quarterly, 24*(1), 24-35. http://dx.doi.org/10.1037/a0015431

Schein, E. H. (1993, January 15). How can organizations learn faster? The challenge of entering the green room. *MIT Sloan Management Review*, *34*(2), 85-92. Retrieved from https://sloanreview.mit.edu/article/how-can-organizations-learn-faster-the-challenge-of-entering-the-green-room/

Schoenwald, S. K., Sheidow, A. J., Letourneau, E. J., & Liao, J. G. (2003). Transportability of multisystemic therapy: Evidence for multilevel influences. *Mental Health Services Research, 5*(4), 223–239. doi:10.1023/A:1026229102151

Simons-Morton, B. G., Parcel, G. S., Baranowski, T., Forthofer, R., & O'Hara, N. M. (1991). Promoting physical activity and a healthful diet among children: Results of a school-based intervention study. *American Journal of Public Health, 81*(8), 986–991

The Iris Center. (2016). Evidence-based practices (part 2): Implementing a practice or program with fidelity. Vanderbilt Peabody College. Retrieved March 8, 2017. https://iris.peabody.vanderbilt.edu/module/ebp_02/

Vernon-Faegans, L., Bratsch-Hines, M., Varghese, C., Bean, A., & Hedrick, A. (2015). The targeted reading intervention: Face-to-face vs. webcam literacy coaching of classroom teachers. *Learning Disabilities Research and Practice, 30*(3), 135–147. https://doi.org/10.1111/ldrp.12062

Walker, H. M. (2004). Commentary: Use of evidence-based interventions in schools: Where we've been, where we are, and where we need to go. *School Psychology Review, 33*(3), 398–408.

Wilkinson, L. A. (2007). Assessing treatment integrity in behavioral consultation. *International Journal of Behavioral Consultation and Therapy, 3*(3), 420–432. doi:10.1037/h0100816