July 2018

Reduce Challenging Behaviors and Enhance Functioning in Youth with an Intellectual Disability: A Meta-Analysis of Behavioral Interventions Using Single Case Design

Brett A. Stone
University of South Florida, bastone2@mail.usf.edu

Reducing Challenging Behaviors and Enhancing Functioning in Youth with an Intellectual Disability: A Meta-Analysis of Behavioral Interventions Using Single Case Designs

by

Brett A. Stone

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctorate of Philosophy in School Psychology
Department of Psychological and Social Foundations
College of Education
University of South Florida

Co-Major Professor: Dr. Linda Raffaele-Mendez, Ph.D.
Co-Major Professor: Dr. John Ferron, Ph.D.
Julia Ogg, Ph.D.
Eric Storch, Ph.D.

Date of Approval:
June 25, 2018

Keywords: appropriate behavior, undesirable behavior, children, synthesis

Acknowledgements

I would like to thank the individuals who have given me guidance and support throughout my dissertation project. Specifically, I would like to thank my Co-Major Professors, Drs. Linda Raffaele-Mendez and John Ferron for their expertise and continuous support for this project. Their guidance helped me be able to conduct this type of a research design, enhanced the quality of the design and my writing skills, and greatly contributed to my overall understanding of the research process. I would also like to thank Drs. Julia Ogg and Dr. Eric Storch for their professional support through my graduate studies process, their honest and helpful feedback, and for their encouragement. All of the contributions of my committee members have helped me immensely to develop in my academic pursuit. Importantly, thank you to the graduate students who graciously gave me their time and effort to help me with the data collection. Also, I am very grateful to my family for their support and belief in me through this research process, and to Devon for her unconditional love. Finally, thank you to my friends and cohort members for providing much needed support.

Table of Contents

List of Tables

Abstract

Approximately 1.8% of students in the public school system have an intellectual disability or Autism Spectrum Disorder. These disabilities cause impairment in multiple domains of functioning. If these students also have challenging behaviors, such as noncompliance, aggression, and stereotypies, these behaviors have been found to cause impairment over and beyond those of the core symptoms associated with the disability. Challenging behaviors in youth with developmental disabilities do not typically subside on their own and need intervention. Thankfully, there are evidence-based behavioral interventions for individuals with developmental disabilities to reduce challenging behaviors and increase more functional behaviors including Applied Behavioral Analysis, Functional Behavioral Analysis, and School-Wide Positive Behavioral Support and Interventions (SWPBIS). There has been much research and positive effects found on the effectiveness of behavioral interventions for individuals with developmental disabilities, and there have been numerous meta-analyses conducted to synthesize these results. However, there have been only a few meta-analyses examining the effectiveness of school-based behavioral interventions for youth with developmental disabilities. A gap in the literature exists in understanding the effectiveness of behavioral interventions in schools from a SWPBIS perspective for youth with developmental disabilities. There also is a need to examine a wider range of dates and to examine the use of parametric statistical metrics. The current study addressed these issues by conducting a meta-analysis of single-case design studies over approximately the past 20 years to add to the current understanding of the effect of school-based behavioral interventions on behavioral outcomes of youth with developmental disabilities.

Additionally, moderator analyses were conducted on numerous participant, intervention, and study characteristics that have been deemed important in the literature. The effect size of behavioral interventions on youths' behavioral outcomes was determined through the use of a parametric statistical method, hierarchical linear modeling. The effect size was found to be large for a single case design synthesis of 3.31 and there were two moderating effects located, one being the type of classroom a participant was educated in and the other the type of specific outcome studied. The current study is important for decision makers in schools in terms of deciding on the specifics of behavioral interventions for youth with an intellectual disability. Additionally, the results of the study may be pertinent to other practitioners who work with youth is schools and their caregivers so that they can utilize school-based interventions to help increase the presentation of appropriate behaviors and reduction of challenging behaviors.

Chapter I: Introduction

Statement of the Problem

Intellectual disabilities (ID) and Autism Spectrum Disorder (ASD), both of which are considered developmental disabilities (DD), affect a small percentage of individuals in the general population and of students enrolled in the public school system. The effects of such disabilities can be very impairing (American Psychiatric Association, 2013). The American Psychiatric Association (APA; 2013) note that approximately 1% of the general population has a diagnosis of ID, and 1% has a diagnosis of ASD. Of the total percentage of youth who have a disability in the public school system (12%), 14.7% of youth have an ID, and 14% of youth with a disability have ASD (United States Department of Education, 2016). This totals 1.8% of the general student population with these types of DD. When educating students, there are many youth who are in need of services to accommodate the struggles that are typical when affected by a developmental disability (Harvey, Boer, Meyer, & Evans, 2009).

Due to the symptoms caused by these disabilities, there are several domains of functioning that are challenging for these individuals such as, communication, adaptive skills, relationships, other social skills, academics, and other occupational concerns (American Psychiatric Association, 2013). Furthermore, there is a large portion of this population that engage in challenging behaviors, specifically 20-30% of individuals with ID and an even higher percentage in individuals with ASD. The percent of individuals with ID or ASD that also have challenging behaviors is significantly more than the general population, of whom 4 to 10% engage in challenging behaviors (American Psychiatric Association, 2013; Dekker, Koot, Van

Der Ende, & Verhulst, 2002; Emerson & Hatton, 2007). It has been found that among youth with developmental disabilities, that if these youth engage in challenging behaviors beyond the symptoms of the disability, then these youth are functionally more impaired than those youth who do not engage in these behaviors (Emerson, 2003; Murphy et al., 2005). Furthermore, it has been found that unless there is intervention to remediate these concerns, then these challenging behaviors do not go away (Horner, Car, Strain, Todd, & Reed, 2002). It is very important to intervene when these challenging behaviors arise to help alleviate repercussions for the individual's quality of life and developmental outcomes (Emerson, 2003; Murphy et al., 2005). The Federal government is required through the Individuals with Disabilities Education Act (IDEA; 1975) to provide youth with disabilities free and appropriate education. This has enabled youth with developmental disabilities to be educated in the public school system with supports and services to accommodate the varying needs. To prevent and intervene with behavioral problems schools typically use a three-tiered approach to prevent and intervene called School Wide Positive Behavioral Intervention and Supports (SWPBIS). SWPBIS is applied at three levels of intervention in the school setting (Horner, Sugai, Todd, & Lewis-Palmer, 2005): (a) primary prevention, universal strategies applied to all students, across all settings; (b) secondary prevention, targeted strategies implemented to groups of students at risk for developing behavior problems; and (c) tertiary prevention, comprehensive supports applied to address the individual needs of students with challenging behaviors. However, for youth who are already placed in a special education exclusive setting, when challenging behaviors arise, there is not a federal mandate to use best practice when assessing and intervening, unless an individualized behavior support plan is developed (Goh & Bambara, 2010). It has been found that these best practices are significant indicators of intervention effectiveness, specifically using behavioral techniques

and conducting a functional behavioral analysis (FBA) to inform the intervention (Campbell, 2003; Carr, Horner, Turnbill, Marquis, Magito-McLaughlin, McAtee…Braddock, 1999; Denis, Van den Noortgate, & Maes, 2011; Didden, Duker, & Korzilius, 1997; Didden, Korzilius, van Oorsouw, & Sturmey, 2006; Goh & Bambara, 2010; Harvey et al., 2009; Gresham, McIntrye, Olson-Tinker, Dolstra, McLaughlin, & Van (2004) ; Heyvaert, Maes, Van den Noortgate, Kuppens, & Onghena, 2012; Heyvaert, Saenen, Campbell, Maes, & Onghena, 2014; Horner et al., 2002; Marquis, Horner, Carr, Turnbull, Thompson, & Behrens, 2000; Scotti, Evans, Meyer, & Walker 1991; Shogren, Faggella-Luby, Bae, & Wehmeyer, 2004). With the importance of intervening in challenging behaviors for this population, it is integral to understand what is working and what is not from an evidence-based approach.

Behaviorism has enabled the development of evidence-based treatments for challenging behaviors for youth with developmental disabilities including Applied Behavior Analysis, School-Wide Positive Behavior Support, and Functional Behavioral Analysis (Cooper, Heron, & Howard, 2007; Rosenwasser & Axelrod, 2001). There have been many quantitative reviews examining if behavioral interventions are effective in reducing undesirable behaviors and increasing desireable behaviors in this population, all finding a positive significant effect (Campbell, 2003; Denis et al., 2011; Didden et al., 1997; Didden et al., 2006; Goh & Bambara, 2010; Harvey et al., 2009; Gresham et al., 2004; Heyvaert et al., 2012; Heyvaert et al., 2014; Horner et al., 2002; Marquis et al, 2000; Scotti et al., 1991; Shogren et al., 2004). None of these reviews have included only children and only interventions in schools in the main analyses except two synthesis (Goh & Bambara, 2010; Gresham et al., 2004), and one included youth with ASD under the age of 8 in varied settings (Horner et al., 2002). However, no moderating effect for age has been found for the large majority of the syntheses (Campbell, 2003; Didden et

al., 1997; Didden et al., 2006; Denis et al., 2011; Harvey et al., 2009; Heyvaert et al., 2012; Heyvaert et al., 2014; Marquis et al., 2000; Shogren et al., 2004; Scotti et al., 1991). Heyvaert et al., (2012), did find a moderating effect of age, which results indicated behavioral interventions conducted with adults were more effective than with younger participants. Grade range was studied by Goh and Bambara (2010) and no moderating effect was found, while Gresham et al., (2004) did not conduct these moderator analyses. A more detailed review of the school-based syntheses is provided in the next section.

Effect of School-Based Behavioral Interventions on Individuals with Developmental Disabilities

There have been two SCD meta-analyses relevant to school-based behavioral interventions on challenging behaviors of youth with developmental disabilities (Goh & Bambara, 2010; Gresham et al., 2004) Both specifically examined interventions that utilized FBAs. However, Gresham et al., (2004) after collecting the data, decided to also conduct analyses on non-FBA based interventions due to the use of such in a large percent of studies. Both of these studies examined the effects on youth with various developmental disabilities. The year range between the two syntheses ranged from 1991-2008, whereas Gresham et al., (2004) included studies only published in The Journal of Applied Behavior Analysis (JABA). There were a variety of behaviors included in the analyses, for example off-task, disruptiveness, aggression, social behavior, and stereotypies. The main finding from both studies indicated that school-based FBA and in the case of Gresham et al., (2004), non-FBA based behavioral interventions, were effective in helping with behavioral challenges of youth with developmental disabilities (Goh & Bambara, 2010; Gresham et al., 2004). In summary, there is some

information pertaining to the effectiveness of school-based interventions for youth with developmental disabilities.

Single-Case Design Studies

The school-system has a chance to help youth with intellectual disabilities decrease challenging behaviors as well as enhance life functioning and this is often times studied within the literature through single-case design studies. Single-case designs are research experiments that include one participant or a small number of participants that experience an intervention and at least one dependent variable is repeatedly measured, typically through direct observation (Onghena, 2005). Single-case designs (SCDs) have gained popularity within education, as they are particularly useful in this field (Zhan & Ottenbacher, 2001). The What Works Clearinghouse (http://ies.ed.gov/ncee/wwc/), which is often considered in determining which educational interventions are efficacious, includes single-case studies as acceptable research designs for determining efficacy.

One concern that is often raised with regard to single-case design studies is that they are not viewed as reliable because their external validity is low. One way that this concern can be addressed is by integrating the findings of multiple single-case design studies through meta-analysis techniques (Riley-Tillman & Burns, 2009). The methodology exists for including SCDs in meta-analyses (Van den Noortgate & Onghena, 2008), and it is important to be able to synthesize single-case design studies to be able to further generalize the results. Also, it is important for researchers to continue to synthesize findings from individual studies through meta-analyses so that others can easily determine the "big ideas" or conclusions from a body of research (Glass, 1976). Conducting a meta-analysis with SCD allows for effect sizes of many different studies to be combined to determine the overall effect that behavioral interventions

have on behavioral outcomes in youth with developmental disabilities. Furthermore, conducting a meta-analysis provides a format for examining important variables that may moderate the effect of these interventions.

Present Study

To build on Goh and Bambara's (2010) and Gresham et al.'s, (2004) SCD meta-analyses, the present study addressed a gap in the literature by conducting a SCD meta-analysis that included both FBA-based and non-FBA-based interventions for youth with developmental disabilities (special education classification categories of intellectual disability, developmental disability, and Autism Spectrum Disorder with IQ lower than 70 or a report of an intellectual disability from an community evaluation). The study also collected data from a SWPBIS level, however only studies conducted at the tertiary level were found. The interventions that were included were implemented to reduce challenging behaviors and/or enhance areas of functioning. In addition, a large year range was searched for studies to include, specifically over approximately the past 19 years from 1997 (when IDEA mandated the use of PBS and FBAs in the schools) to June 2016. Furthermore, the parametric statistical method of hierarchical linear modeling was utilized to calculate effect sizes, as well as, to conduct a comprehensive moderator analysis.

Purpose of the Present Study

The purpose of the present study is to help shape future behavioral interventions in school settings for youth with developmental disabilities by elucidating the effects of such interventions on not only reducing challenging behaviors but also on enhancing functioning. In addition, the study is meant to help stakeholders understand any moderating effects of participant, intervention, or study characteristics to help enhance the effectiveness of intervention

selection and SWPBIS implementation for youth with developmental disabilities. Moreover, the particular dependent variable of interest, behavioral outcomes, are important to study, considering the contribution of alleviating these concerns for youth to experience school and life success. The results of the study may provide information to school psychologists and other stakeholders to help with their decision-making concerning how to utilize SWPBIS and other behavioral interventions for youth with developmental disabilities. Finally, another contribution is that the results of this study may further validate the utility of the results from single-case designs through aggregating the effects of single cases to obtain average treatment effects.

It seems that there is still a gap in the literature, as there is yet to be a SCD synthesis of all three levels of SWPBIS for youth with developmental disabilities that include both FBA and non-FBA based interventions, nor one that includes a wide range of year ranges searched for studies to be included in the analyses. Furthermore, nonparametric statistics were utilized in the two most relevant extant meta-analyses (Goh & Bambara, 2010; Gresham et al., 2004), whereas the present study used parametric statistical methods, specifically hierarchical linear modeling to synthesize the results (Van den Noorthgate & Onghena, 2003).

Research Questions

The present study addressed the following research questions:

1. On average, what is the effect size of behavioral interventions conducted in school settings on youth with developmental disabilities' behavioral functioning?
2. What participant characteristics moderate the relationship between behavioral interventions and youth with developmental disabilities' behavioral outcomes?

Specific participant characteristics that were examined included the following: (a) age range, (b) grade range, (c) gender, (d) specific disabilities (diagnoses of clinical disabilities such

as Autism Spectrum Disorder), (e) cognitive status (if participants were described as having a certain level of intellectual functioning), (f) level of verbal communication ability and (g) type of classroom setting the participant was educated in.

3. What intervention characteristics moderate the relationship between behavioral interventions and youth with developmental disabilities' behavioral outcomes?

Specific intervention characteristics that were examined included the following: (a) intervention type, (b) agent (who delivered the intervention), (c) setting (inclusive, exclusive classroom, therapy room, gym, etc.), (d) format (group or individual), (e) duration, (f) presence of a functional behavioral analysis (FBA), (g) if FBA data was used to inform the intervention, (h) FBA assessment agent, (i) FBA setting, (j) team decision -making during FBA, (k) techniques used to generalize behavior change, and (l) school-wide positive behavioral support tier (1, 2, or 3).

4. What study characteristics moderate the relationship between behavioral interventions and youth with developmental disabilities' behavioral outcomes?

Specific study characteristics that were examined included the following: (a) type of challenging behavior, (b) intervention fidelity, (c) social validity measures, (d) published/unpublished, (e) and inter-rater reliability data, and (f) type of single case design.

Definition of Key Terms

Antecedent intervention. A behavior change strategy that manipulates contingency-independent antecedent stimuli (Cooper et al., 1997).

Autism spectrum disorders. The DSM-5 describes autism spectrum disorder (ASD), as a clustering of symptoms that indicate deficits in social communication and social interaction in various settings, which include social reciprocity, nonverbal communication, and

social skills to develop, maintain, and understand relationships. In addition, the individual also engages in restricted, repetitive patterns of behavior, interests, or activities. Some examples of repetitive or stereotyped behaviors include motor stereotypies such as hand flapping, repetitive use of objects such as lining up toys, and repetitive speech, such as repeating words after someone else (American Psychiatric Association, 2013).

Behavioral interventions. The use of operant conditioning models (positive and negative reinforcement) as well as skill replacement and functional communication to modify undesired behaviors.

Behavioral outcomes. Behavior refers to any activity that living organisms can perform. As it relates to humans, this includes what we are able to do, what we think, and our feelings (Skinner, 1974). For this study we will be focused on "what we are able to do". Common problem behaviors include stereotypic behaviors, self-injury, aggression, and off-task verbal behaviors. Desirable behaviors may include on-task classroom behaviors, such as paying attention, writing when asked to write, and waiting quietly.

Contingent. Describes reinforcement that is delivered only after the target behavior has occurred (Cooper et al., 1997).

Developmental delays. This is a developmental disability educational label for children from birth to age three (under the Individuals with Disabilities Education Act; IDEA Part C) and children from ages three through nine (under IDEA Part B), the term developmental delay, as defined by each State, means a delay in one or more of the following areas: physical development; cognitive development; communication; social or emotional development; or adaptive/behavioral development (IDEA, 2004).

Functional behavioral Assessment (FBA) This is an assessment method that utilizes a specific process to identify challenging behaviors and the antecedent events that predict whether the behavior will or will not occur, and what consequential events will reinforce the behavior. . This assessment data is collected with the purpose of informing the development of behavioral interventions. (Sugai, 2000)

Hierarchical linear modeling (HLM). HLM is a parametric statistical tool that can be utilized for analyzing the results of a single-case design meta-analysis. HLM estimates linear equations that explain outcomes for members of groups as a function of the characteristics of the groups and the characteristics of the members (Van den Noortgate, 2012).

Individualized education plan (IEP). The federal law, IDEA, requires that public schools create an IEP for every child receiving special education services. Students from age 3 through high school graduation or a maximum age of 22 (whichever comes first) may be eligible for an IEP. The IEP is meant to address each child's unique learning issues and include specific educational goals. It is a legally binding document (United States Department of Education).

Intellectual disability. As defined by the Diagnostic Statistical Manual V (DSM-5), is when an individual has deficits in "general mental abilities, such as reasoning, problem solving, planning, and abstract thinking, judgment, academic learning, and learning from experience" (American Psychiatric Association, 2013 p. 33). It then goes on to state that these deficits cause impairments in adaptive functioning, which includes "personal independence and social responsibility in one or more aspects of daily life, including communication, social participation, academic or occupational functioning, and personal independence at home or in community settings" (American Psychiatric Association, 2013 p. 33).

Mean baseline level reduction (MBLR). MBLR is a statistical method used to compare data in baseline and treatment phases of a single case design study. MBLR is equated by calculating the mean treatment value and the mean baseline value, then subtracting these values respectively, followed by dividing by the mean baseline value (Kahng, Iwata, & Lewin 2002). When interpreting MBLR scores, 100% indicates the problem behavior has gone away completely, while 0% means that there was no change from baseline, and a negative score indicates that the problem behavior increased. (Heyvaert et al., 2014).

Meta-analysis. This statistical method was first introduced by Glass (1976) as a quantitative approach to summarize results of studies. Glass (1976) defined it as "the statistical analysis of a large collection of analysis results from individual studies for the purpose of integrating the findings" (p.3).

Moderators. A variable that changes the direction and/or significance of the correlation found between an independent and dependent variable (Baron & Kenny, 1986, p. 1174).

Neurodevelopmental disabilities. Within the DSM-5, there is a categorization of disorders called neurodevelopmental disorders (a.k.a. developmental disabilities), which have an onset during the developmental period. Some other characteristics of these disorders are that they cause developmental challenges and impairments in personal, social, academic, or occupational functioning. The various neurodevelopmental disorders that are classified in the DSM-5 include intellectual disability (intellectual developmental disorder; ID), autism spectrum disorder (ASD), communication disorders, attention deficit hyperactivity disorder (ADHD), neurodevelopmental motor disorders, and specific learning disorder (American Psychiatric Association, 2013). For the proposed study, there will be a focus on youth with ID and youth with both ID and ASD.

Percentage of all non-overlapping data (PAND). PAND is a statistical method used to compare data in baseline and treatment phases of a single case design study. PAND is a calculation of the percentage of data points that do not overlap between baseline and treatment phases. PAND is calculated by indicating the number of the overlapping data points, and dividing this by the total number of data points to obtain the percentage overlap, and then subtracing this percentage from 100% (Parker, Hagan-Burke, & Vannest, 2007). PAND has a scale of 50% to 100%, where 50% is chance level (Heyvaert et al., 2014, p. 2466).

Percentage of data points exceeding the mean (PEM). PEM is a statistical method used to compare data in baseline and treatment phases of a single case design study. To calculate PEM, first the median baseline point is determined as well as the amount of treatment data points that are greater than the median baseline point. Next the later number is divided by the former (Ma, 2006). A PEM score 90% or greater, indicates a highly effective treatment, a score between 90% and 70% indicates an effective treatment, a score between 70% and 50% indicates a questionable treatment, and a score less than 50% indicates an ineffective treatment (Heyvaert et al., 2014).

Percentage of non-overlapping data (PND). PND is a statistical method used to compare data in baseline and treatment phases of a single case design study PND is equated by identifying the amount of treatment data points that are greater than the highest baseline data point. This number is then divided by the total amount of data points in the treatment phase (Scruggs & Mastropieri, 1987). When interpreting PND scores, the same scale is used as the PEM statistic (Heyvaert et al., 2014)

Percentage of zero data (PZD). PZD is a statistical method used to compare data in baseline and treatment phases of a single case design study. This nonparametric is calculated by

identifying the initial treatment data point that reaches zero and then finding the percentage of treatment data points that remain at zero (Scotti et al., 1991). A PZD score greater than 80% shows a highly effective treatment, a score that falls between 80% and 55% indicates an effective treatment, a score between 55% and 18% indicates a questionable treatment, and a score less than 18% is labeled an ineffective treatment (Heyvaert et al., 2014).

Positive behavior supports (PBS). PBS refers to applying positive behavioral interventions and systems to promote socially appropriate and important behavior change. It was initially developed as a different approach compared to aversive interventions that were typically used with students with significant developmental disabilities who engaged in self-injury and aggression. Now the technique is applied to various populations of students, for a wide range of presenting concerns, to prevent challenging behaviors, and can be applied at the individual or school level (Sugai, 2000).

Single-case design. This type of research design involves one or multiple treatments at multiple time points, using the individual or a group as their own control (Kazdin, 2011).

Stereotypy. "Stereotypies are defined as involuntary, patterned, repetitive, coordinated, rhythmic, and non-reflexive behaviors that are suppressible by sensory stimuli or distraction (Freeman, Soltanifar, Baer, 2010)". These repetitive behaviors cause concern when they are atypically intense, have a long duration, are not present in the majority of a culture, cause self harm, or cause impairment in functioning (Freeman, Soltanifar, Baer, 2010).

Youth. The term refers to individuals from 3-22 years of age.

Chapter II: Review of the Literature

In this chapter, background information concerning prevalence and symptomology in youth with neurodevelopmental disabilities is provided and information concerning challenging behavior in this population, followed by a discussion of the theoretical underpinnings of behaviorism. Then there is information about school-based behavioral interventions for youth with neurodevelopmental disabilities. Next, a review ensues of the extant meta-analyses and a literature review related to the effects of behavioral interventions on individuals with neurodevelopmental disabilities' behavioral outcomes, and then meta-analyses are reviewed that examined only school-based behavioral interventions. Both main analyses and moderator analyses findings are reviewed. Finally, a discussion follows concerning the importance of single-case designs, integrating research findings through meta-analysis, and conducting meta-analyses of single-case design studies.

Youth with Neurodevelopmental Disabilities

Neurodevelopmental disabilities. The American Psychiatric Association, 2013 (APA) has a comprehensive book called the Diagnostic and Statistical Manual of Mental Disorders (DSM), which provides a way of classifying mental health disorders with criteria that have been established by a team of experts. One of the main intents of the DSM is to provide an objective assessment of symptoms that cluster together and form a disorder. The DSM is in its' 5th edition currently, with the 1st edition began being published in 1952 (American Psychiatric Association, 2013). Within the DSM-5, there is a categorization of disorders called neurodevelopmental disorders (aka developmental disabilities), which have an onset during the

developmental period. Some other characteristics of these disorders are that they cause developmental challenges and impairments in personal, social, academic, or occupational functioning. The various neurodevelopmental disorders that are classified in the DSM-5 include intellectual disability (intellectual developmental disorder; ID), autism spectrum disorder (ASD), communication disorders, attention deficit hyperactivity disorder (ADHD), neurodevelopmental motor disorders, and specific learning disorder. The terms neurodevelopmental disorder and developmental disorders will be used interchangeably throughout this proposal. Intellectual disability and autism spectrum disorders will be discussed in detail, as they are pertinent to the proposed study.

Intellectual disability. A description of intellectual disability disorder (ID) within the DSM-5, states that individuals have deficits in "general mental abilities, such as reasoning, problem solving, planning, and abstract thinking, judgment, academic learning, and learning from experience" (American Psychiatric Association, 2013 p. 33). The book then goes on to state that these deficits cause impairments in adaptive functioning, which includes "personal independence and social responsibility in one or more aspects of daily life, including communication, social participation, academic or occupational functioning, and personal independence at home or in community settings" (American Psychiatric Association, 2013 p. 33). Clinical assessment involves both deficits in intellectual and adaptive functioning. On standardized tests of intelligence, individuals with ID have scores two standard deviations or more below the population mean. They must also have deficits in adaptive functioning, which can be measured using standardized assessments as well. Adaptive functioning involves reasoning in conceptual, social, and practical domains. The conceptual domain, "involves competence in memory, language, reading, writing, math reasoning, acquisition of practical

knowledge, problem solving, and judgment in novel situations" (American Psychiatric Association, 2013 p. 33). The social domain, "involves awareness of others' thoughts, feelings, and experiences; empathy, interpersonal communication skills, friendship abilities, and social judgment" (American Psychiatric Association, 2013, p. 33). While the practical domain involves, "learning and self-management across life settings, including personal care, job responsibilities, money management, recreation, self-management of behavior, and school and work task organization"(American Psychiatric Association, 2013, p. 33). The individual must have deficits in the majority of one domain of adaptive functioning. Another related disorder is called global developmental delay, and is diagnosed when a person does not meet various developmental milestones in several domains of intellectual ability and who can not perform on standardized assessments of intelligence, which often times are children under the age of 5 (American Psychiatric Association, 2013). There are various levels of severity for this disorder, termed mild, moderate, severe, and profound. Previously, this disorder used to be called mental retardation (MR), however, a federal statue in the United States (Public Law 111-256, Rosa's Law) mandates this term be replaced with intellectual disability and that research journals also use the updated term. For the purposes of this proposed study, if an older study uses the term MR, the principal investigator will instead use the updated term ID.

Special education classification of youth with ID. Each school district follows state legislative for determining if a student will receive exceptional student education (ESE) for students between 3-22 years of age, who have disabilities. For example, in the state of Florida, following the Individuals with Disabilities Education Act (IDEA; 2004) there are 13 various ESE programs available to youth. The primary one relevant to the proposed study population of youth with intellectual disability (as labeled through the DSM-5) is also called intellectual

disability in the school system or InD (Florida Department of Education, 2015). To be diagnosed with an intellectual disability through the special education classification system, there are specific criteria established by IDEA, (2004), which are aligned with that of the DSM-5 criteria. Students must undergo school-based assessments to determine eligibility, which are detailed in state statutes. The youth scores on a standardized test of intellectual functioning must be two standard deviations below the mean, and the same for an assessment of adaptive functioning in two out of three domains of adaptive functioning (IDEA, 2004). The definition of adaptive functioning depends on state laws, but typically includes communication and social skills, independent living skills, personal care skills, employment/work skills, and practical academics (Florida Department of Education, 2015). Different than the DSM-5 criteria, student scores on a standardized test of academic or pre-academic performance must be consistent with that of a student with comparable intellectual functioning. A child can enter the public school system at the age of three if they have been found to have special needs, to provide early intervention services. When a child enters the school system at the age of three with a special education category of developmentally delayed, they are then reassessed at a later age (typically at age 6), and then given a different special education disability category out the 13 provided by IDEA (2004). It is typical for a child who will later be diagnosed as InD at age 6 to be diagnosed with developmental delay through the school system upon entering the system before the age of 6.

Prevalance of ID. One of the purposes of the DSM is to be a tool for collecting and detailing valid public health statistics on mental health disorder prevalence rates (American Psychiatric Association, 2013). Within the DSM-5 it states that the general population prevalence rate for intellectual disability is 1% and that it's .6% for the severely intellectually disabled. According to the National Center for Education Statistics the most recent data indicated

that in 2011/2012 out of the total enrollment of youth in the public school system 12.9% had a disability and were served under IDEA (2004). Of this percentage of public school youth, 0.9% were categorized as having an intellectual disability. Out of the percentage of total youth with a disability, 14.7% of these youth were diagnosed with InD (United States Department of Education, Institute of Education Sciences, 2016).

Autism spectrum disorder. Within the DSM-5 autism spectrum disorder (ASD), is described as a clustering of symptoms that indicates deficits in social communication and social interaction in various settings, which include social reciprocity, nonverbal communication, and social skills to develop, maintain, and understand relationships. In addition, the individual also engages in restricted, repetitive patterns of behavior, interests, or activities. Some examples of repetitive or stereotyped behaviors include motor stereotypies such as hand flapping, repetitive use of objects such as lining up toys, and repetitive speech, such as repeating words after someone else. Also, included in this category are resistance to change, and rituals, such as pacing a perimeter (American Psychiatric Association, 2013). Within the diagnoses of ASD, there are various specifiers that can be used and one of pertinence to the proposed study is, "with or without accompanying intellectual impairment" (American Psychiatric Association, 2013, p. 51). Furthermore, the DSM-5 format provides a way to state severity of ASD, by the level of support needed for deficits in both social communication and restricted, repetitive behavior domains. Intellectual disability is frequently found among individuals with ASD (American Psychiatric Association, 2013). In previous editions of the DSM, there was a separate diagnosis of Asperger's Disorder, which meant that the individual had autism but without intellectual impairment, and many researchers and clinicians are still referring to this as high functioning autism spectrum disorder (HF-ASD). Those individuals with a previous diagnoses of Asperger's

disorder, using the DSM-5 would now be given a diagnosis of autism spectrum disorder without the specifiers: language impairment and intellectual impairment. The DSM-5 ASD criteria also encompasses what used to be called pervasive developmental disorder not otherwise specified (PDD-NOS). Within the DSM-5 it states that individuals with ASD function with less impairment if they do not also have intellectual disability and/or language impairment (American Psychiatric Association, 2013).

Special education classification of youth with ASD. Just as with youth with intellectual disability, youth with ASD undergo assessment to determine the ESE disability category label to receive ESE services in the school system. To be given the ESE label of ASD, the criteria is aligned with that of the DSM-5. The specific assessment procedures are outlined in state statutes, and for example in Florida they include behavioral observations to evaluate social interaction, social communication skills, and restricted/repetitive behavior across settings, as well as social/developmental history, a psychological evaluation of academic, intellectual, social-emotional, and behavioral functioning, as well as a standardized measure for ASD, a language evaluation by a speech language pathologist, a standardized measure of adaptive behavior, and if behavioral concerns are present then a functional behavioral assessment to inform interventions on the youth's individualized education plan (IEP).

Prevalence of ASD. The DSM-5 indicates that the prevalence rate of ASD across the US and other countries is approximately 1% of children and the same for adults. The National Institute of Educational Statistics most recent data (2011/2012), also has found a similar rate, in that .9% of students have a ESE category label of ASD out of the total enrollment of youth in public school. Out of the percentage of total youth with a disability, 14% of these youth were diagnosed with ASD (United States Department of Education-Institute of Educational

Science, 2016). The most recent estimates by the Centers for Disease Control and Prevention from 2010, indicate that the prevalence rates of children with autism spectrum disorder (ASD) are 1 in 68 children. Among males it is five times more common (1 in 42) than in girls (1 in 189). The rate of youth with a diagnosis of ASD has risen dramatically in the past few decades. The prevalence rate of autism has increased 289.5% over the past 12 years (CDC, 2012). There is yet to be a consensus on why the rate of ASD has increased over the years, however, theories exist that it may be due to the expansion of the criteria in the DSM-5 vs the DSM-IV, more awareness of the disorder, research design differences, or a possibility that there are more individuals being born with the disorder (American Psychiatric Association, 2013).

Challenging Behavior in Youth with Neurodevelopmental Disabilities

Challenging behaviors are often developed by various influences from factors within the person and factors within the environment, and interactions of these factors. There have been many examples of these factors found in the literature including: age, gender, level of ID of an individual, and "poor adaptive skills, poor social skills, psychological stress, inadequate problem-solving skills, impaired language, socioeconomic deprivation, negative life events, secondary disabilities and psychiatric disorders (as cited in Heyvaert et al., 2012)". Children who have problem behaviors have a higher risk of being excluded from educational settings, being isolated, have difficulties with social relationships, excluded from typical home environments, and participating in community activities (Sprague & Rian, 1993). It has been shown that once a child with developmental disabilities exhibits challenging behaviors, the behavior will not typically decrease unless interventions are put in place (Horner, Car, Strain, Todd, & Reed, 2002).

Overall, regardless of disability status, research has found that children with limited communication and social skills have a higher risk of developing challenging behaviors (Borthwick-Duffy, 1996). Often individuals with developmental disabilities, such as intellectual disabilities, as well as ASD, have significant problems with communication (Sigafoos & Drasgow, 2001). These issues lead to impairments in communicating wants, needs, refusals, agreements, or social conversations (Sigafoos & Drasgow, 2001). Approximately 50% of individuals with ASD cannot express themselves verbally in a way that does not cause impairment in these above-mentioned areas (Koul, Schlosser, & Sancibrian, 2001). Severe communication issues can cause issues in education, employment, family, and community life (Beukleman & Mirenda, 2005). Students have difficulties requesting, asking for help, asking for breaks, or responding (Bondy & Frost, 2001). There is research suggesting that problem behaviors in youth with developmental disabilities negatively impacts quality of life and is a predictor of negative future outcomes (Emerson, 2003; Murphy et al., 2005). Also, the behaviors can serve as challenges to delivering interventions and educational programming (Harvey, Boer, Meyer, & Evans, 2009).

Intellectual disability. There is specific research concerning individuals with intellectual disabilities and challenging behaviors, such as aggression, noncompliance, disruptiveness, destructiveness, and self-injury, as well as, mental health disorders such as anxiety, depression, and mania, which are all found at high frequencies (Allen, 2013). Research has indicated that approximately 20 to 30% of youth with intellectual disability engage in behaviorally challenging behaviors (Dekker Koot, Van Der Ende, & Verhulst, 2002; Emerson & Hatton, 2007; Linna, Moilanen, Ebeling, Piha, Kumpulainen, Tamminen, & Almqvist, 1999). While 4 to 10% of youth without an intellectual disability engage in such behaviors (Emerson &

Hatton, 2007). Research has shown that the more severe the disability, then the higher the chance the individual will have challenging behaviors (Heyvaert, 2010). Challenging behaviors have been shown to become a lifelong struggle for people with ID, as well as for their family and service providers (Murphy et al., 2005). According to the National Institute of Education Statistics (2016) teachers report a high level of challenging behaviors amongst students with this educational classification label.

Autism spectrum disorder. There is also specific research concerning youth with ASD and challenging behaviors. Self-injury and disruptive behaviors are more common in youth with ASD, even more so than in youth with ID (American Psychiatric Association, 2013). Challenging behaviors of aggressive, stereotypies, and self-injury are found frequently in individuals with autism (Matson & LoVullo, 2008; Murphy & Leader, 2009). These problem behaviors can reduce the quality of life of the person with ASD (Walsh, Mulder, & Tudor, 2013). The DSM-5 reports that a minority of people with ASD are able to live and work independently as adults.

Theoretical Underpinnings of Behaviorism

The science of behaviorism has been determined to be valid through experiments and can explain the relationship between behavioral interventions/techniques and its effects on the challenging behaviors of youth with developmental disabilities. To begin, a brief history of behaviorism will be presented, followed by descriptions of specific behavioral techniques, and lastly a summary describing functional behavioral assessments.

B.F. Skinner brought respondent and operant behavior into the academic world with the publication of his book in 1938, which summarized his laboratory research from 1930 to 1937 (Cooper et al., 2007). Respondent behaviors are conceptualized as involuntary and present when

a stimuli is presented. Whereas operant behaviors are not present with the presence of antecedent stimuli, but are present when there are stimulus changes after the behavior. Skinner conducted and explained various experiments showing the relationship between behavior and environmental events. These experiments legitimized the concept of operant behavior and continue to be the cornerstones for behavioral interventions today (Cooper et al., 2007). Skinner's experiments involved animals, mainly rats and pigeons, whereas in 1949 Fuller published a study using the principles with a person. Throughout the 1950s and 1960s researchers conducted experiments to understand if the behavioral principles were relevant to humans. They were found to be effective. Next came applied behavioral analysis in the 1960s whereby researchers attempted to apply these principles in applied settings instead of in a laboratory. During this time many first attempts and successful findings were made in regards to behavioral principles and education, such as contingent teacher praise and attention. Then universities set up behavioral academic programs in the 60s and early 70s and in 1968, The Journal of Applied Behavior Analysis (JABA) was created. The definition of applied behavior analysis (ABA) from Cooper et al., (2007) is, "the science in which tactics derived from the principles of behavior are applied systematically to improve socially significant behavior and experimentation is used to identify the variables responsible for behavior change" (p. 20).

Behaviorism has hypothesized various functions of challenging behavior, which have been categorized into social positive reinforcement, social negative reinforcement, and sensory/automatic reinforcement (Lloyd & Kennedy, 2014). The social positive reinforcement hypothesis posits that the behavior is maintained by receiving social stimulus upon presenting the behavior. Specific examples include adult attention, physical attention, peer attention, tangible items, or preferred activities (Lloyd & Kennedy, 2014). The social negative

reinforcement hypothesis posits that the behavior is maintained by removing the social stimulus. Specific examples include removal of task demand, escape from aversive stimuli, or social avoidance (Lloyd & Kennedy, 2014). The treatment for socially mediated behaviors is changing the environmental contingencies, (e.g., if aggression is present for escape from a task, then escape from task is not allowed upon aggressive behavior presentations). The sensory or automatic reinforcement hypothesis posits that the behavior is maintained by internal reasons to the individual and is not dependent on the social environment (Lloyd & Kennedy, 2014). An example of a sensory reinforced behavior is a verbal stereotypy maintained due to enjoying the feeling from a vibration made from repeating a certain noise, such as "Mmmmmmmm". The treatment for sensory based behaviors is typically to reduce the value of the reinforcing consequences, so in the example provided above the treatment may be to replace the behavior with a more socially acceptable way of receiving that sensory input by providing a vibrating toy and allowing access to it at certain scheduled times of the day. It's important to note that often times the function of the behavior does not fall into only one of these categories of social positive reinforcement, negative reinforcement, or sensory reinforcement. Often times the behavior can be maintained by multiple functions, also the function can vary by environmental setting, and can change over time (Lloyd & Kennedy, 2014).

There is an assessment type that is utilized and evidence-based to hypothesize what maintains a behavior or what the function of the behavior is, called functional behavioral assessment (FBA). FBAs come from a body of literature that indicates that operant behavior is influenced by various components including: (a) the consequences a behavior has on an environment, (b) antecedents that trigger a behavior, (c) and events in a setting that change the value of the consequences that will ensue upon behavioral activation (Bijou & Baer, 1961).

Horner, Carr, Strain, Todd, & Reed (2002) describe that conducting a FBA involves sequential steps by first identifying the problem behavior/s, building hypotheses about what is maintaining the behavior/what the function is, then testing/confirming the hypotheses, and lastly designing an intervention based on the data from the FBA. When conducting a FBA there are experimental methods, descriptive methods, and combinations of these methods (Goh & Bambara, 2010). Experimental methods include setting up the environment to understand if the hypothesized function of the behavior is maintaining the behavior and analyzing the behavior during the experiment through hypothesis testing. Descriptive methods involve indirect methods, such as interviews, rating scales, archival record review, and observations. While often times experimental and descriptive methods are used in combination to conduct a FBA (Goh & Bambara, 2010). Many previous quantitative reviews have found that behavioral interventions that utilize functional behavioral assessments have a significantly higher effect as compared to studies that do not conduct FBAs (Campbell, 2003; Carr et al., 1999; Didden et al., 1997; Didden et al., 2006; Harvey et al., 2009; Heyvaert et al., 2014; Horner et al., 2002; Marquis et al., 2000; Scotti et al., 1991).

Behavioral Interventions

Youth with neurodevelopmental disabilities. There are different categories of positive behavioral interventions such as reinforcement-based strategies and antecedent interventions. Reinforcement-based strategies include examples such as differential reinforcement and extinction. Extinction is a process whereby the behavior is no longer reinforced. For example, if the function of biting was to gain access to verbal and physical attention, this attention would be withdrawn. Extinction is typically used along with differential reinforcement. Differential reinforcement is providing reinforcement contingent on an alternative behavior (DRA) or on

time intervals without the presentation of the undesirable behavior/other behavior (DRO). In the previous extinction example, this may be accompanied by using DRO, whereby if the individual does not bite the person for 30 seconds then verbal and physical attention is given (Lloyd & Kennedy, 2014). Due to theories that often times challenging behaviors in themselves are a form of communication for individuals with developmental disabilities functional communication training has been developed as a behavioral intervention (Durand, 1990). This intervention combines differential reinforcement and extinction of challenging behaviors with utilizing appropriate forms of communication as the replacement behavior. Noncontingent reinforcement (NCR) involves providing reinforcement on a certain time schedule so that the reinforcement is given separate from the presence of the behavior, however it is withheld if a challenging behavior occurs immediately before the time scheduled reinforcement. The mechanism by which it is proposed that noncontingent reinforcement is effective is by deprivation and satiation. There are cited limitations to NCR in that replacement behaviors are not taught and it may not be feasible to have such a schedule in a naturalistic environment (Lloyd & Kennedy, 2014).

Behavioral interventions also include preventative efforts instead of only dealing with a problem behavior once it has occurred. Antecedent interventions include changing the environment before the occurrence of a behavior, and two common interventions are curriculum changes and choice-making. Some examples of curriculum changes include adjusting the task type, duration, and/or difficulty. These aim at decreasing the aversiveness of tasks, and are utilized when the function of the behavior is to escape task demands. Choice-making is theorized to give environmental control (automatically reinforcing) to the individual and it is also theorized that they are effective in the same way as curriculum changes, due to reducing aversiveness (Lloyd & Kennedy, 2014). Other antecedent strategies include modifying the physical

characteristics of a setting, changing schedules, using reminders, and redesigning social groups (Horner et al., 2002). Interventions should be designed to address all of the functions of the behavior found in the FBA, therefore the various techniques can be used combined together to serve meeting the function of the behavior (Lloyd & Kennedy, 2014).

Some other examples of behavioral interventions include differential reinforcement of incompatible behavior, antecedent exercise, social stories, and picture exchange communication system (PECS) (Heyvaert et al., 2014). Positive behavior supports, do not include behavioral techniques that are aversive conditioning, such as restraint, aversive smells, aversive tastes, water misting, and aversive sounds as examples (Didden et al., 1997). Evidence indicates that behavior support is not an attempt to change individuals to fit environments, but to change environments to fit the individuals in that setting (Horner et al., 2002). There is evidence that interventions using behavioral techniques and interventions linked to the function of the behavior have a significant effect on behavioral outcomes for individuals with developmental disabilities (Campbell, 2003; Carr et al., 1999; Denis et al., 2011; Didden et al., 1997; Didden et al., 2006; Gage, Lewis, & Stichter, 2012; Goh & Bambara, 2010; Harvey et al., 2009; Gresham et al., 2004; Heyvaert et al., 2012; Heyvaert et al., 2014; Horner et al., 2002; Marquis et al., 2000; Scotti et al., 1991; Shogren et al., 2004).

School-based interventions. ABA is an evidence-based treatment for individuals with developmental disabilities in reducing challenging behaviors and enhancing areas of deficits (e.g., social skills, communication, adaptive skills) and is recommended by the Surgeon General of the United States (Rosenwasser & Axelrod, 2001). Furthermore, the federal education legislation, including the Individuals with Disabilities Education Act (IDEA) of 2004 and No Child Left Behind Act of 2001 (NCLB) has mandated the use of evidence-based practices for

ensuring all youth are able to be educated in the least restrictive environment (Odom, Brantlinger, Gersten, Thompson, & Harris 2005). It has been found that students with developmental disabilities with challenging behaviors need to receive interventions to help reduce these behaviors so that learning can occur (Demaray, Malecki, & DeLong, 2006). To remediate these challenging behaviors and to help youth succeed in the school setting, researchers recommend utilizing preventive interventions (Sprague & Horner, 2006), however traditionally punishment or exclusionary practices were the most common intervention (Goh & Bambara, 2010). These sorts of interventions have been found to work immediately but do not promote maintenance effects or generalization (Zhang Katsiyannis, & Herbst, 2004). The evidence-base for applied behavior analysis, which utilizes functional behavioral assessments and positive behavioral supports, has made large and important contributions to individuals with developmental disabilities and in the educational setting over the past 40 years (Gresham et al., 2004). The amendments to IDEA of 1997 and 2004 actually mandate the use of FBAs and positive behavioral supports for designing and implementing interventions for students with disabilities (Gresham et al., 2004). Prior to the use of FBAs becoming federal law, they had been considered best practices (Gresham et al., 2004).

School-Wide Positive Behavior Interventions and Supports (SWPBIS) are built upon the behavioral principles reviewed above. Research finds that SWPBIS is the method to help reduce problem behaviors in schools and with good maintenance of gains (Safran & Oswald, 2003). SWPBIS is founded on the theories based on applied behavior analysis, person-centered planning, inclusion, and systems change principles to impact challenging behaviors and enhance students' quality of life in the school setting (Carr, Dunlap, Horner, Koegel, Turnbill, Sailor…Fox, 2002). SWPBIS is applied at three levels of intervention in the school setting

(Horner, Sugai, Todd, & Lewis-Palmer, 2005): (a) primary prevention, universal strategies applied to all students, across all settings; (b) secondary prevention, targeted strategies implemented to groups of students at risk for developing behavior problems; and (c) tertiary prevention, comprehensive supports applied to address the individual needs of students with challenging behaviors.

The third level of support (individualized Positive Behavior Support or iPBS) has evidence-base to help reduce problem behaviors in youth with intellectual disabilities both in school and nonschool settings (Carr et al., 1999). Some specific components of iPBS include: the intervention is informed by a FBA, there is a focus on preventing challenging behaviors through changes in the environment and teaching alternative behaviors that they are easily generalizable to various agents implementing them, it is team based, and iPBS promotes socially acceptable and feasible interventions (Goh & Bambara, 2010). The Individuals with Disabilities Education Act of 1997 (IDEA, 1997) and its reauthorization in 2004 promoted the use of SWPBIS at this third level with youth with disabilities. IDEA states that school personal will use functional behavioral assessments (FBA) when students are at risk for a change of placement because of problem behaviors but a FBA is not mandated once they are already in an exclusive environment (Goh & Bambara, 2010). IDEA (2004) states that positive behavioral approaches should be used whenever an individualized behavior support plan is needed. Syntheses of the literature have shown that conducting an FBA before a behavioral intervention increases the success of the intervention. However, this is not always done within the school system, as sometimes (if the youth is already in an exclusion classroom and an individualized plan is not needed) it is not mandated or for other reasons (Gresham et al., 2004). Furthermore, there have been many SCD meta-analyses that have indicated the effectiveness of FBA-based interventions for decreasing

challenging behaviors in people (both adults and youth) with developmental disabilities (Campbell, 2003; Carr et al., 1999; Denis et al., 2011; Didden et al., 1997; Didden et al., 2006; Gage, Lewis, & Stichter, 2012; Goh & Bambara, 2010; Harvey et al., 2009; Gresham et al., 2004; Heyvaert et al., 2012; Heyvaert et al., 2014; Horner et al., 2002; Marquis et al., 2000; Scotti et al., 1991; Shogren et al., 2004).

While SWPBIS is meant to prevent challenging behaviors, it can also be used to help enhance functioning in youth with intellectual disabilities directly or indirectly (Harvey et al., 2009). It is important to not only look at using behavioral interventions for youth with intellectual disability to reduce challenging behaviors but also how these interventions target progression in their adaptive skills. As issues in these areas of functioning lead to a lower quality of life and poorer adult outcomes (Emerson, 2003; Murphy, Beadle-Brown, Wing, Gould, Shah, & Holmes, 2005). Gresham et al., (2004) in a meta-analysis that included 150 school-based intervention studies examined the effectiveness of behavioral interventions on youth with developmental disabilities, and found over half of the studies examined appropriate behavior acquisition (desirable behaviors).

Effects of Behavioral Interventions on Youth with Developmental Disabilities' Behavioral Outcomes: Syntheses

In this section, the literature on the effects of behavioral interventions on individuals with developmental disabilities' challenging behaviors is reviewed. Included is a review of the existing quantitative reviews (Campbell, 2003; Denis et al., 2011; Didden et al., 1997; Didden et al., 2006; Goh & Bambara, 2010; Harvey et al., 2009; Gresham et al., 2004; Heyvaert et al., 2012; Heyvaert et al., 2014; Horner et al., 2002; Marquis et al, 2000; Scotti et al., 1991; Shogren et al., 2004) and two literature reviews (Carr et al., 1999; Horner et al., 2002) that synthesize

single-case design studies. The purpose of this review is to provide a detailed account of the majority of extant studies on the effects of behavioral interventions on the aforementioned outcome and population. Moderator analyses were conducted in many of these syntheses and are summarized by study, and in a later part of this section they are summarized by characteristic type. The review of the syntheses is organized by delineating reviews of individuals with developmental disabilities (ID or a combination of diagnoses), those examining participants with solely a diagnosis of autism, and school-based intervention syntheses. Table 1 provides a summary of the main analyses descriptions of the syntheses that are reviewed below. Table 2 includes information to discern what indicates that a statistical metric is effective or not effective. While Table 3 provides a summary of the moderator analyses.

Individuals with intellectual disabilities or combined developmental disabilities.

There have been many quantitative reviews examining the effects of behavioral interventions on individuals with developmental disabilities (Denis et al., 2011; Didden et al., 1997; Didden et al., 2006; Heyvaert et al., 2012; Marquis et al., 2000; Scotti et al., 1991; Shogren et al., 2004). These reviews will be discussed in the following section.

Researchers cite Scotti et al., (1991) as the first comprehensive meta-analysis to examine the effect of behavioral interventions on challenging behavior in people with developmental disabilities. The researchers included studies from 1976-1987 and included 318 studies, and 403 participants. The analyses included both adults and youth, with 67% between 6 to 21 years old, 9% 5 or younger, and 24% adults. The researchers utilized percentage of nonoverlapping data (PND) and percentage of zero data (PZD) as the statistical tools to examine the effectiveness of the interventions, and found that behavioral interventions were effective. These researchers categorized the level of intervention intrusiveness by adopting an established categorical system

by Evans & Meyer, 1985 and Stephenson, Dempsey, & Scotti, 1983). This system (Levels 1, 2, and 3) ranged from least to most intrusive, restrictive, or aversive. Furthermore, behaviors were categorized into severity levels, from least to most severe (Level 1, 2, and 3). For PND approximately 33% were highly effective, 30% fairly effective, 17% in the questionable range and 20% as ineffective. For PZD, there were 25% of studies in each of the effectiveness ranges. In regards to the interaction effects between the level of behaviors, there were none found for PND but for PZD scores it was found that level 2 and 3 interventions were significantly more effective than level 1. Although, this paper argued that the delivery of the behavioral interventions to individuals with DD at that time was not ideal. Specifically, these researchers recommended that researchers improve the use of best practice assessment (FBA) and intervention (use of positive behavioral supports over aversive techniques) when working with this population.

The following participant characteristics were analyzed: age, gender, disability type, severity of disability (amount of impairments), level of verbal skills, and functional level. Age, gender, disability type, severity of disability, and level of verbal skills did not moderate the effect. The functional level as measured by mild to moderate and severe to profound intellectual disability was found to moderate the effect. Results indicated no effect for PND for intervention and follow-up but for PZD there was an effect during both phases, with higher mean scores for participants in the mild to moderate range.

The intervention characteristics that were analyzed included type of behavioral strategies, intervention setting, presence of an FBA, generalizability, and agent**.** Type of behavioral strategy moderated the effect, in only that medication was significantly less effective than the 13 other classes of behavioral strategies. Intervention setting, and agent did not moderate the effectiveness

of the intervention. FBA was found to moderate the effect positively if conducted for the PND statistic and no effect was found for PZD statistic. Generalizability moderated the effect in that if attempts were made to generalize the intervention effectiveness then these studies were significantly more effective on PND and PZD statistics.

The study characteristic "type of behavioral problem" moderated the effect. It was found that physically aggressive/tantrum behaviors and destructive/disruptive behaviors had the least treatment effect and the differences were significant for PND scores as compared to self-injury, stereotypies, and inappropriate social/other behaviors. For PZD, destructive/disruptive and stereotypic behaviors were affected the least and there were significant differences between these and the other behavioral classes (inappropriate social behavior/other, self-injury, physically aggressive/tantrums).

Didden et al., (1997) conducted a meta-analysis on the same variables as Scotti et al., (1991), to address the limitations these researchers found in the former meta-analysis. To accomplish this, Didden et al., (1997) searched more journals and categorized the data to include more delineation of the particular challenging behaviors studied and behavioral techniques utilized in the interventions. There were 482 included studies, and 1,451 comparisons between baseline and treatment conditions, with a search of journals between 1968 and 1994. The mean age of participants was 16.4 years old with a range of 1 to 66 years old. These researchers calculated the effect size using percentage of nonoverlapping data (PND) and found that response contingent behavioral interventions were more effective than the other types of treatment, which included antecedent control procedures, pharmacology, and response non-contingent procedures. Furthermore, externally destructive behaviors (e.g., property destruction) had significantly lower mean percentage of nonoverlapping data scores as compared to the

treatment of socially disruptive (i.e., self-injurious behavior [SIB], stereotypic behavior, pica) and internally maladaptive behaviors (i.e., public disrobing, inappropriate vocalizations) (Didden et al., 1997).

Didden et al., (1997) examined multiple variables for potential moderating effects, including the following participant characteristics: developmental level, age in years, disability, and secondary disabilities. In addition, the following intervention characteristics were examined: presence of FBA, intervention setting, and duration of sessions. None of these variables moderated the effect, with the exception of FBAs, which were found to positively affect the outcomes if conducted. The study characteristic, type of problem behavior was examined. It was found that externally destructive behaviors were rectified less than internally maladaptive or socially disruptive behaviors.

Harvey et al., (2009) also sought to replicate and update the Scotti et al., (1991) meta-analysis. The meta-analysis included 142 studies (316 participants) from 1988 to mid-2006, from birth to 21 years of age with a mean age of 9.7 years old. Specifically, 44% were diagnosed with an intellectual disability, 33% with ASD, and 17% with multiple developmental disorders. Scotti et al., (1991) used the study as the unit of analysis while Harvey et al., (2009) used the individual as the unit of analysis. The researchers utilized four types of metrics to calculate effect sizes, including PZD, PND, Allison-MT, and standardized mean difference (SMD). Interventions utilizing antecedents, skills replacement, and consequence-based (reinforcement-based) techniques produced fairly effective effect sizes for all four statistical metrics. Overall, this meta-analysis was in consensus with the Scotti et al., (1991) meta-analysis, in that the results indicate that behavioral treatments compared to no treatment reduce challenging behaviors in individuals with developmental disabilities.

Harvey et al., (2009) stated that they purposefully analyzed many of the same moderators as Scotti et al., (1991). The following participant characteristics were analyzed: age, gender, ethnicity and they did not have an effect, although ethnicity was rarely reported. Other participant characteristics that were examined and did moderate the effect included, disability type and severity of behaviors. The disability type moderated the effect; specifically youth with autism responded significantly more to antecedent interventions compared to youth with other developmental disabilities when analyzing two of the four metrics, while for the other two metrics, this was not found to be the case. If the behaviors were more severe, the lower the effectiveness of the treatment.

In Harvey et al., (2009) the intervention characteristics that were analyzed included intervention setting, duration, presence of FBA, and intervention type. Intervention setting did not moderate the effect, duration was found to moderate the effect with interventions between 3 to 20 weeks being most effective as compared to those less than 3 weeks or more than 20 weeks, and the presence of a FBA was associated with a higher effect. The type of intervention did moderate the effect. Intervention strategies when used alone were not *highly* effective, however, antecedent, skills replacement and consequences were found to be *fairly* effective when used alone. Teach replacement skills was consistently (across metrics) found to be the most effective, and even more effective when used with systems change or traditional antecedent and consequence manipulation. Yet, there was no single universal behavioral intervention that was effective for all types of challenging behavior, and no single behavioral strategy more effective than another. These findings highlight the importance for individualized interventions based off of the results of FBAs.

The study characteristic, type of challenging behavior was examined in Harvey et al., (2009) and it was found to moderate the effect. Specifically, self-injury, stereotypy, socially inappropriate, and destructive behavior responded more to interventions than disruptive and aggressive behavior.

Didden et al., (2006) conducted another meta-analysis specifically on individuals with mild intellectual disability, as they stated that previous meta-analyses in the field had conducted their research on individuals with moderate to severe intellectual disability (Campbell, 2003; Didden et al., 1997). Studies with an independent variable of either behavioral interventions or psychotherapeutic interventions were included. There were eighty studies included, with 133 participants, and studies ranged from a publication year of 1980 to 2005. The mean age of the participants was 14.5 years old and the age range was from 2 to 45 years old. Effect sizes were calculated by using PND and percentage of zero data (PZD). The main finding was that behavioral interventions had a significant effect on reducing challenging behaviors in individuals with mild intellectual disability (PND Ms = 75%; PZD Ms= 35%), and an effect, although less, was found for cognitive behavioral interventions, such as anger management. No other treatment methods such as counseling or psychotherapy were located.

Didden et al. (2006) examined various moderating variables. The participant variables that were studied included age, gender, and diagnosis and none moderated the effect. All of the intervention characteristics, FBA presence, FBA utilization, and generalization techniques utilized moderated the effect. Specifically, if an experimental FBA was conducted and the intervention was informed by the FBA data then the effect was greater, using PND and PZD metrics. The type of intervention also moderated the effect. It was found that behavioral

interventions were more effective than those using psychotherapeutic, specifically cognitive or self-management techniques.

The study characteristics that were examined included: year of publication, type of behavior, type of design, procedural reliability, and reliability of recording, and generalization**.** The type of design had a moderating effect in that AB designs as compared to reversal and multiple baseline designs had significantly lower effects using PND and PZD metrics. Using PZD, it was also found that the reliability of recording moderated the effect. If the reliability of recording was measured then these studies produced a greater effect (Didden et al., 2006).

Denis, Van den Noortgate, & Maes, (2011) conducted a SCD meta-analysis examining the effect of non-aversive and non-intrusive forms of reinforcement on self-injurious behavior of individuals with profound intellectual disability (IQ < 25). These researchers noted that previous syntheses had not focused specifically on these exact dependent and independent variables on this exact population. The researchers included 18 studies between 2000-2008, and the mean age was reported to be 27.5 years old. Hierarchical linear modeling was used to calculate the effect. The findings indicated that the treatment effect was significant and large, whereby problem behaviors were 2.54 standard deviations lower in treatment conditions as compared to baseline conditions.

Denis et al., (2011) found no moderating effects for the following participant characteristics: medication, motor impairment, age, and gender. Also no moderating effect was found for the following intervention characteristics: setting, matching of treatment with behavioral function, and contingency. A moderating effect was found for sensory impairment in that it indicated that if impairments were present then the treatment was significantly less effective.

Marquis et al., (2000) conducted a meta-analysis of the studies included in Carr et al., (1999). Carr et al., (1999) conducted a descriptive research synthesis of SCD studies (*n*=109) using positive behavioral supports between 1985-1996 with both adults and children with developmental disabilities and in varied settings. The mean age of the participants (*n*= 220) was 14.5 years old, and included both adults and children. The participants had various diagnoses pertaining to developmental disabilities, including: 50.9%with ID, 10.5% with ASD, and 12% with both ASD and ID, and 22.3% accounted for those with combined diagnoses of ASD and ID plus another disability. This synthesis was conducted upon the request from the United States Department of Education, Office of Special Education programs, to understand what the extant literature indicated in regards to the effect of positive behavior interventions. In the synthesis, Carr et al., (1999) found that PBS was effective in the reduction of challenging behaviors in 50-66% of participants and that the effectiveness was significantly greater if a FBA was conducted before and if the intervention was linked to this data. Marquis et al. (2000) used SMD, a nonparametric statistical tool to calculate the effect size of the studies included in Carr et al., (1999). Marquis found a positive significant effect size. Specifically it was reported that the SMD ES was 2.1 for single interventions using stimulus based (antecedent) and reinforcement strategies, and for those that used multiple interventions it was 3.1.

The participant characteristics that were analyzed in Marquis et al., (2000) included gender, age, diagnosis, and level of intellectual disability. Gender, age, and diagnosis were not found to moderate the effect. The level of intellectual disability was found to moderate the effect and specifically, the lower the intelligence of the participant the less effective the intervention.

In terms of intervention characteristics the following were examined: whether an FBA was conducted, whether the FBA data was utilized to inform intervention, whether the

intervention was stimulus-based or reinforcement based, the agent, the setting, and if it included a non-positive behavioral component. Conducting an FBA and using it to inform the intervention moderated the effect, indicating that doing this resulted in an increased effect. There were moderating effects found for the type of intervention, in that using both stimulus-based and reinforcement based interventions together increased the effect size by 1.0 units and was significant. Although the researchers state this should be interpreted with caution due to the low amount of data utilized to calculate this effect. The agent was found to moderate the effect, specifically that the intervention was more effective if administered by a typical agent, and when interventions were implemented in a typical setting then the treatment was more effective (ES increased from 1.7 to 2.6). However, it was found that the typical agent and typical setting were not independent of each other. The interventions that also included a non-positive behavioral component in addition to PBS did not moderate the effect.

In terms of study characteristics, Marquis et al., (2000) examined the following variables effect: type of data collected (whether the data was percentage or frequency count), the amount of data points collected in the baseline phase, slope, type of problem behavior, and function of the behavior (i.e., attention, escape, tangibles, and sensory). A moderating effect was found for the type of data collected. Specifically, if the study used frequency counts then the effect size decreased by .3 to .5 units. The number of baseline data points also moderated the effect, specifically, as the amount of data points increased the effect size decreased. The researchers did not report on what was found in regards to the slope. In terms of problem behaviors (i.e., aggression, self-injurious behaviors, property destruction, and tantrums), the overall effect size was 1.8 and for those with aggression as the target behavior it increased by .5 units. It was found

that PBS was effective for all problem behaviors, and significantly more so for aggression as the targeted outcome. The function of the behavior did not moderate the effect.

Heyvaert et al., (2012) conducted a SCD meta-analysis on the effect of behavioral interventions for reducing problem behaviors in individuals with intellectual disabilities. These researchers included both children and adults. There were 285 studies included (155 were SCD and 130 were *small-n* designs) with 598 participants and studies were published between 2000 to April 2011. The mean age of the participants was 18 with a range of 1 to 65 years old. The parametric statistical technique of hierarchical linear modeling was utilized and the treatment effect was statistically significant and large, specifically the level of challenging behavior was 2.96 standard deviations lower in the treatment conditions than baseline.

Heyvaert et al., (2012) conducted a comprehensive moderator analyses. This involved the following participant characteristics: age, gender, diagnosis of ASD, sensory impairment, motor impairment, and communicative impairment. Of these variables, the following moderated the effect, age and diagnosis of ASD. Specifically the results indicated that interventions conducted on adults were more effective than younger participants. The information could not be located within the study as to who the researchers considered younger versus older or how many participants fell into each category. In addition, it was found that the interventions were more effective for individuals with ASD as compared to those with other primary developmental disability diagnoses.

The following intervention characteristics were examined: FBA presence, setting, format (individual or group), family involvement, duration, agent, peer involvement, uni- vs. multicomponent intervention, and intervention components (i.e., teaching alternative replacement skills, reward, praise, attention, punishment, use of restraints, manipulating

antecedent factors, and extinction; social-contextual intervention, and environment factors of informing, educating, training the environment to the participant's needs). The presence of the behavioral component of manipulating antecedent factors of informing, educating, and training the environment was found to create a significantly larger effect. None of the other characteristics were found to moderate the effect (Hevaert et al., 2012).

The following study characteristics were analyzed: type of problem behavior, design, presence of interrater reliability, presence of follow up data, publication year, and study quality. None moderated the effect except type of problem behavior. Specifically, behavioral interventions for aggression and destructive behavior were less effective than other challenging behaviors (i.e., self-injury, stereotypies; Hevaert et al., 2012).

Shogren et al., (2004) conducted a SCD meta-analysis on the effects of choice-making interventions on challenging behaviors of individuals with intellectual disabilities on studies published before 2003. With the attention that self-determination received in the 1990s this led to educators providing individuals with disabilities choice-making opportunities (Shogren et al., 2004). At the time of this publication the mechanisms for which choice-making have a positive effect on challenging behaviors was not fully understood, although it was hypothesized through a literature review that it provided control over one's environment and provided an adaptive way of communicating needs before needs were not met and problem behaviors arose. The researchers utilized PND and PZD metrics to examine the efficacy of these interventions. There were 13 studies, with 30 participants that met inclusion criteria, the mean age was 11.1 years for females and 10.1 years for males with approximately 66% being male. The age range of participants spanned from 1 to 50 years old. The researchers report that 85% involved children between the ages of 5 to 21 years old. The main finding that the researchers discussed was that

choice interventions had a significant effect in reducing the amount of challenging behaviors, however based off of the PND and PZD scores these indicated questionable effects. The PND overall score was 65.7% and the overall mean PZD score was 42.3%.

Shogren et al., (2004) analyzed many potential moderating variables. The participant characteristics included in the analyses were gender, age, and diagnosis. It was found that gender moderated the effect, specifically that males had a higher level of reduction of problem behaviors than females. Age and diagnosis were found to not moderate the effect. The intervention characteristics that were studied included: the type of choice intervention used, the type of activity in which the choice procedure was embedded, if interventionist had training prior to implementation, setting, presence of FBA, and FBA data utilization. None of these variables were found to moderate the effect. Lastly, the two study characteristics, the type of study and type of behavior, were analyzed and no moderating effects were indicated.

In sum of the main analyses findings, all of the reviewed syntheses found that behavioral interventions had a positive effect on behavioral outcomes for individuals with developmental disabilities (Denis et al., 2011; Didden et al., 1997; Didden et al., 2006; Heyvaert et al., 2012; Marquis et al., 2000; Scotti et al., 1991; Shogren et al., 2004). Also all of these reviews included youth and adults combined in the main analyses. Two studies examined the effect of behavioral interventions on individuals with various developmental disorders, such as ID and ASD (Carr et al., 1999; Marquis et al., 2000; Scotti et al., 1991). While other studies examined the effects on individuals with intellectual disability (Didden et al., 1997; Heyvaert et al., 2012; Shogren et al., 2004), one study examined the effect on individuals with mild ID (Didden et al., 2006), and another on individuals with severe ID (Denis et al., 2011). Almost all studies examined many different topographies of behavior (Denis et al., 2011; Didden et al., 1997; Didden et al., 2006;

Heyvaert et al., 2012; Marquis et al., 2000; Scotti et al., 1991) such as self-injury, tantruming, aggression, stereotypies, and destructive behavior. Shogren et al., (2004) specifically examined the effect of one type of behavioral intervention, choice-making, on one outcome, self-injury. Also, overall there was a large representation of years studied, the amount of studies, the amount of participants, and all reviews included interventions conducted in varied settings. The findings from the moderator analyses will be reviewed in a later section. Next, three SCD meta-analyses examining the effect of behavioral interventions on the challenging behaviors of individuals with a sole diagnosis of autism spectrum disorder with intellectual impairment will be reviewed.

Individuals with ASD. There have been many meta-analyses that have examined the effectiveness of behavioral interventions on individuals with ASD (Campbell, 2003, Heyvaert et al., 2014; Horner et al., 2002). One such meta-analysis is Campbell (2003), whom included both children and adults (mean age 10.02 with a range of 2 to 31 years old, and studies published between 1966-1998. The meta-analysis included 117 studies with 181 participants. The researchers examined the effect of behavioral interventions on challenging behaviors of individuals with ASD (mean IQ of 42.2) and found that the interventions were effective. Three metrics were used to calculate the effectiveness of the interventions including PND, PZD, and mean baseline reduction (MBLR). It was found that on average treatment reduced problem behaviors by 75% as compared to baseline levels per the MBLR effect size, and the PND mean score was 84%, and the PZD mean score was 43%. The moderator analyses for Campbell (2003) are reviewed along side the section concerning the moderator findings in Heyvaert et al., (2014)

Heyvaert et al., (2014) updated the previous mentioned meta-analysis by including studies published between 1999-2012 and examined the same moderators as in Campbell (2013). Heyvaert et al., (2014) included 213 studies and 358 participants within the studies. The study

included individuals with ASD with and without intellectual disability, specifically 94.4% of participants had an IQ less than 70 and the mean age of the participants was 10.24, although both adults and children were included. The researchers utilized five ways of calculating the effect of the intervention at both the study and participant level, all but one statistical method indicated that behavioral interventions for individuals with ASD were effective in reducing the challenging behaviors examined. Specifically, across all participants the averages were 74.9% (PND), 44.7% (PZD), 70.2% (MBLR), 90% (PEM), and 91.9% (PAND). At the study level, the averages were 75.9%, 47.3%, 74.2%, 93.0%, and 92.3%, respectively. The PZD mean averages indicted the treatment was questionably effective, while the other 4 indicated on average interventions were effective in reducing challenging behaviors at both the participant and study levels.

Heyvaert et al., (2014) analyzed various participant, intervention, and study characteristics. With regards to participant characteristics, specifically, age, gender, intellectual disability level, and level of verbal communication ability were examined. No significant effects were found for any of the participant characteristics examined. They also analyzed various intervention characteristics including, type of intervention, FBA presence, parental involvement in the intervention, and presence of generalizability techniques. Heyvaert et al., (2014) found that positive combination interventions were statistically significantly better at reducing challenging behaviors as compared to antecedent control interventions. Also it was found that aversive and positive combinations, positive combinations, differential reinforcement of other behaviors only, antecedent control only, differential reinforcement of alternative behaviors only, noncontingent reinforcement only, and social stories only interventions were statistically more effective at impacting challenging behaviors as compared to PECS only interventions. However, it should be noted that there was a small sample size of interventions utilizing PECS only interventions.

Also, pretreatment FBA moderated the effectiveness of the interventions. Significant moderating effects were found for the remaining intervention characteristics examined including: parental involvement, generalizability, and presence of follow up data. However, further details could not be located about the moderating effect specifics.

The researchers also analyzed the following study characteristics: type of behavioral problem, criteria used for diagnosing autism, experimental design, amount of baseline data points, amount of treatment data points, publication year, presence of inter-rater reliability data, and the presence of follow up data. No effect was found for the type of behavioral problem examined or criteria used for diagnosing autism. However, significant moderating effects were found for the other study characteristics examined, although no further information could be located (Heyvaert et al., 2014).

Campbell (2003) did not find moderating effects on any of the participant characteristics examined, which were largely the same variables as in Heyvaert et al., (2014). However, different than Heyvaert et al., (2014), Campbell (2013) did not find moderating effects for any of the intervention characteristics examined. But Campbell did also find a moderating effect for pretreatment functional analysis, which indicated a higher effect if utilized. Consensus was found between the two meta-analyses that the same experimental characteristics moderated the effect, but no further details could be located.

Horner et al., (2002) conducted a literature review of past meta-analyses published between 1988 to 2000 pertinent to the study of behavioral interventions for individuals with ASD, as well as, a meta-analysis of the effect of behavioral interventions for youth 8 or younger with ASD. The publication years included in the meta-analysis were between 1996 and 2000. The literature review included six meta-analyses that examined the effects of behavioral

interventions on youth with ASD, and included youth below the age of 8 (Carr et al., 1999; Marquis et al., 2000; Lennox, Miltenberger, Spengler, & Erfanian, 1988; Scotti, Ulcich, Weigle, Holland, & Kirk, 1996; Didden et al., 1997; Scotti et al., 1991). Horner et al., (2002) stated that these reviews indicated that behavioral interventions are effective in reducing challenging behaviors in individuals with developmental disabilities, which included autism. They found that in 50-66% there were 80% reductions in problem behaviors. Horner et al., (2002) also conducted a meta-analysis of studies published between 1996 to 2000 that included youth with autism below the age of eight. The researchers chose to use such a small publication year range because they wanted to understand the most current published research, given the increased use of FBA and PBS. The meta-analysis included 9 studies, 24 participants, and 37 comparisons (baseline to treatment). The mean reduction in challenging behaviors was 85% (SD = 19), with a median reduction level of 93.2%. Furthermore, 59% indicated a reduction in challenging behaviors by 90% of greater.

Horner et al. (2002) reported various moderating effects through a literature review of the four extant meta-analyses on behavioral interventions of youth with autism below the age of 8. These researchers gathered that FBAs moderate the effectiveness of the interventions specifically that if a FBA informs the intervention, then the outcomes are significantly greater. In addition that typical agents (e.g., families, teachers) are correlated with greater effects, but the researchers state this may be because more difficult behavior is referred to atypical agents (e.g., hospitals, specialists).

All three of the located reviews on the effects of behavioral interventions for individuals with autism spectrum disorders included participants with ASD and intellectual impairments (Campbell, 2003; Heyvaert et al., 2014; Horner et al., 2002) and all three reviews found a

significant positive effect. Horner et al., (2002) was the only review out of the three that included only youth in their analyses, specifically youth 8 or younger, while the other reviews combined findings of youth and adults (Campbell, 2003; Heyvaert et al., 2014). Between Campbell (2003) and Heyvaert et al., (2014) these researchers attempted to examine the same variables however, Campbell examined studies from 1966 to 1998 and Heyvaert et al., (2014) examined them where the review had left off from 1999 to 2012. While Horner et al., (2002) examined studies for a 4-year time period between 1996 and 2000. As in the reviews that examined the variables with individuals with ID or combined diagnoses of developmental disorders, these researchers also covered a wide range of challenging behaviors, the settings were varied, and there were a large number of studies and participants included. Next, two SCD meta-analyses relevant to school-based behavioral interventions on challenging behaviors of youth with developmental disabilities will be reviewed.

School-based. Goh & Bambara, (2010) studied the effectiveness of school-based behavioral interventions that utilized FBAs prior to implementation at the individual positive behavioral support (iPBS) level, also known as the third tier of support. Goh & Bambara (2010) were focused on finding FBA based iPBS interventions for youth with any disability, including intellectual disability, or developmental disabilities, autism or pervasive developmental disorder, emotional and behavioral disorder (EBD), learning disabilities (LD), attention-deficit/ hyperactivity disorder (ADHD), other health impairment (OHI), physical disabilities, and other disabilities. The researchers included studies between 1997 and 2008. The highest frequency of studies were conducted with participants in elementary school at 69%, 21% in middle school, followed by 10% in high school. This meta-analysis included 83 studies with 145 participants and found positive significant effects. The PND effect was considered moderate overall for

interventions focused on reducing problem behavior and increasing appropriate behaviors. The median percentage of nonoverlapping data (PND) was 88%. Intervention studies that analyzed solely the reduction of problem behaviors (reduction behaviors), had a slightly lower median PND of 80%, this effect was still moderate. While those intervention studies that analyzed interventions that increased appropriate behaviors (acquisition behaviors) were slightly higher than the overall effect, with a median PND of 90%, still a moderate effect size. The maintenance effects, measured on 28 time series was large with a median PND of 100%, with the duration ranging from 1 week to 2 years. Overall these findings for school based FBA-based interventions for youth with various disability classifications, grade level, and classroom settings were moderately effective in reducing challenging behaviors and increasing functioning of youth.

Goh & Bambara (2010) conducted a comprehensive moderator analysis and included the following participant characteristics: gender, grade range, diagnosis, and classroom setting the participant was educated in. No moderating effects were found. There were also no moderating effects found for any of the intervention characteristics that were analyzed, including: assessment method used for FBA (e.g., observations, experimental, interviews), who conducted the assessments (assessment agent), assessment setting (e.g., exclusion classroom, therapy room, general education), presence of team decision making during assessment, intervention type, intervention agent, intervention setting, and length of treatment. The researchers also examined if there was a moderating effect of the study characteristic of whether social validity measures were or were not gathered, and no effect was found.

Another meta-analysis was conducted to understand the status of applied behavioral analytic techniques used in the schools, specifically FBAs and positive behavioral supports for youth with developmental disabilities. The purpose of this study was to assess the effect of these

interventions with antecedent functional behavioral assessments, for various behavioral outcomes (Gresham et al., 2004). These researchers specifically only reviewed studies in the Journal of Applied Behavioral Analysis from 1991-1999 and specifically FBA based school-based interventions conducted with youth less than 1 year to 18 years of age. The researchers included separate analyses for interventions that did not use FBA procedures due to finding that 52% of the located studies did not report use of such procedures. The article reviewed 150 school-based intervention studies. The researchers found that there were no differences between the effect of interventions that were linked to FBAs than to those that were not linked to FBAs. Two statistical methods were utilized to calculate effect sizes of interventions. These included the standardized difference effect size (Faith, Allison, & Gorman, 1997), as well as the percentage of non-overlapping data points (PND) between baseline and treatments phases (Scruggs & Mastropieri, 1998). It was found that the effect size of interventions that did not use FBAs was 6.77, and those that used FBAs was 4.60. The researchers posit several reasons that the non-FBA interventions may have had larger effect sizes such the as the legitimacy of the statistical methods used (Strain, Kohler, & Gresham, 1998), as well as that non-FBA studies published may have been a subset of effective non-FBA studies due to publication bias, and/or it may be that studies may have used an FBA but did not report it (Gresham, et al., 2004). It should be noted that the information could not be located within the Gresham and colleagues (2004) study, concerning what type of developmental disabilities were included and how disability status was determined when reviewing the studies. Gresham et al (2004) did not conduct moderator analyses.

There were two meta-analyses located that studied the effect of school-based behavioral interventions that utilized FBAs, on youth with developmental disabilities (Goh & Bambara,

2010; Gresham et al., 2004). However, Gresham et al., (2004) after collecting the data, decided to conduct analyses on non-FBA based interventions as well as FBA based interventions due to a large percent of studies including non-FBA interventions. Both of these studies examined the effects on youth with various developmental disabilities. The year range with these two studies ranged from 1991-2008, where Gresham et al., (2004) included studies specifically published in The Journal of Applied Behavior Analysis (JABA). There were a variety of behaviors included in the analyses, for example off-task, disruptiveness, aggression, social behavior, and stereotypies. The main findings from both studies indicated that school-based FBA and, in the case of Gresham et al., (2004), non-FBA based behavioral interventions were effective in helping with behavioral challenges of youth with developmental disabilities (Goh & Bambara, 2010; Gresham et al., 2004).

In summary, there have been numerous meta-analysis conducted examining the effectiveness of behavioral interventions for people with developmental disabilities, more narrowly focused meta-analyses for youth with specific diagnoses, and some information pertaining to the effectiveness of school-based interventions for youth with developmental disabilities.

Moderating effects of behavioral interventions. As reviewed above in the various syntheses, there have been a variety of moderators studied to examine the relationship between behavioral interventions and individuals with developmental disabilities. Table 3 provides this review categorized by participant, intervention, and study characteristics. Within this section, there is a description of this table to synthesize the moderator findings in the existing quantitative reviews by characteristic. It is important to refer back to Table 1 when interpreting these reviews, due to the various population differences studied among extant syntheses.

Table 1

Descriptive Information of Extant Syntheses

Author	Yr	N of Studies	N of Participants	Disability Status	Age	Setting	Independent Variables	Dependent Variables	Main Findings	Statistical Metric
Scotti et al. (1991)	1976-1987	318	403	DD	67% 6-21 years old, 9% 5 or younger, and 24% adults	Varied	Behavioral Interventions	Physically aggressive/tantrum, destructive/disruptive, self-injury, stereotypies, inappropriate social/other behaviors	Behavioral interventions effective. PND 33% highly effective, 30% fairly effective, 17% questionable, 20% ineffective. PZD: 25% in each range; Level of behaviors, PND: did not interact with effectiveness; PZD Level 2 and 3 more effective than 1 (but researchers say level 1 should be used due to least restrictive/aversive)	PND; PZD
Didden et al. (1997)	1968-1994	482	Did not report, but had 1,451 comparisons between baselines and treatments	ID	Combined: mean age 16.4, range 1 to 66	Varied	Behavioral Interventions	34 topographies of behavior, most frequent were self-injurious and stereotypic behaviors, then disruptiveness, aggression, and rumination	Response contingent behavioral interventions $t(991) = 4.10$, $p < .001$ were more effective than other types of treatment including: antecedent control procedures, , pharmacology, $t(991) = 6.68$, $p < .001$, and response noncontingent procedures, $t(991) = 5.92$, $p < .001$; Externally destructive behaviors significantly lower mean PND scores compared to socially disruptive & internally maladaptive behaviors.	PND

Table 1 Cont'd

Author	Yr	N of Studies	N of Participants	Disability Status	Age	Setting	Independent Variables	Dependent Variables	Main Findings	Statistical Metric
Harvey et al. (2009)	1988-2006	142	316	DD: 44% ID, 33% ASD, 17% multiple	Youth:0 to 21 years old, mean age 9.7	Varied	Behavioral Interventions, Educational, Psychotherapeutic Interventions	Self-injury, destructiveness, stereotypies, and aggression	Behavioral interventions effective compared to no treatment, intervention strategies used singly (e.g., an antecedent intervention only), none of the effect sizes indicated highly effective outcomes on any of the four statistics. However, all three treatments of antecedents, skills replacement, and consequences (recall that system change was never used alone) produced effect sizes in the fairly effective range. Refer to study for specific metrics, many reported.	PZD, PND, Allison-MT, SMD
Didden et al. (2006)	1980-2005	88	133 comparisons	Mild ID	Combined: mean age 14.5, range from 2 to 45	Varied	Behavioral and psychotherapeutic treatments	Most frequent: Physical aggression, disruptive behavior & a comb. of aggressive, disruptive, and destructiveness. Others:compulsive, stereotypies, stealing, self injury	Behavioral interventions effective compared to no treatment; Effective but less so were cognitive behavioral interventions PND Ms=75% and PZD 35%	PND; PZD

Table 1 Cont'd

Author	Yr	N of Studies	N of Participants	Disability Status	Age	Setting	Independent Variables	Dependent Variables	Main Findings	Statistical Metric
Denis et al. (2011)	2000-2008	18	Not given	Severe ID (IQ < 25)	Combined: Mean age 27.5 yrs old	Varied	Non-aversive and non-intrusive reinforcement	Self injury	Significant and large, 2.54 standard deviations lower in treatment conditions than baseline	HLM
Marquis et al. (2000)	1985-1996	109	220	DD (ID 50.9%; ASD 10.5%, ID and ASD 12.7%; ID, ASD and another disability & another disability 22.3%)	Combined: Mean age 14.5	Varied	Positive behavioral interventions	Aggression, self injury, property destruction, tantrums	Positive effect: reported the SMD ES = 2.1 for single interventions using stimulus-based and reinforcement based interventions and for combined interventions it was 3.1	Percentage reduction measure; SMD; HLM
Carr et al. * (1999)	1985-1996	109	220	DD	Combined: Mean age 14.5	Varied	Positive behavioral interventions	Aggression, self injury, property destruction, tantrums	Effective in 50-66% of participants, greater if FBA conducted	Literature review
Heyvaert et al. (2012)	2000-April 2011	285 (155 SCD, 130 small-*n*)	598	ID	Combined: Mean age a18, range 1 to 65	Varied	Behavioral Interventions	Self injury, stereotypy, aggression, destructive, disruptiveness	Significant and large, the level of challenging behavior is 2.96 standard deviations lower in the treatment conditions.	HLM

Table 1 Cont'd

Author	Yrs	N of Studies	N of Participants	Disability Status	Age	Setting	Independent Variables	Dependent Variables	Main Findings	Statistical Metric
Shogren et al. (2004)	unknown to 2003	13	30	ID	Combined: Mean age 11.1 for females, 10.1 males. 85% children between 5 to 21; overall range 1-50	Varied	Choice-Making Interventions	Aggression, noncompliance, leaving an area, off-task behavior, property destruction	Questionable effect, overall PND score was 65.7% (SD - 41.0) and overall mean PZD score was 42.3% (SD=42.2). Both indicating questionable effects.	PND; PZD
Campbell (2003)	1966-1998	117	181	ASD (mean IQ 42.2)	Combined; range 2 to 31, Mean 10.08	Varied	Behavioral Interventions	Self-injury, stereotypies, disruptiveness, aggression, property destruction; and combinations	Effective; on average treatment reduced problem behaviors by 75% from baseline levels (MBLR effect size); PND score averaged 84, PZD averaged 43	PZD; PND, MBLR
Heyvaert et al. (2014)	1999-2012	213	358	ASD (with & without ID): 94.4% had IQ less than 70	Combined: Mean 10.24	Varied	Behavioral Interventions	Self-injury, stereotypies, disruptiveness, aggression, property destruction	The averages were 74.9%, 44.7%, 70.2%, 91.4% and 91.9% for PND, PZD, MBLR, PEM, and PAND at the participant level. The averages were 75.9%, 47.3%, 74.2%, 93.0%, and 92.3% at the study level. PND, MBLR, PEM, and PAND indicate on average effective at both participant and study level. PZD the mean averages at participant and study level indicate questionable effects	PND, PZD, PEM, MBLR, PEM,

Table 1 Cont'd

Author	Yrs	N of Studies	N of Participants	Disability Status	Age	Setting	Independent Variables	Dependent Variables	Main Findings	Statistical Metric
Horner et al. * (2002)	1996-2000	9	24	ASD	Youth 8 or younger	Varied	Behavioral Interventions	34 different challenging behaviors but 76% examined tantrums, 59% aggression, stereotypy 14%, and self injury 11%	Mean reduction in problem behavior of 85% (SD 19), with a median reduction level of 93.2% and a mode of 100%. 59% recorded problem behavior reduction of 90% or greater, and 25 comparisons (68%) indicated problem behavior reduction of 80% or greater	Does not say explicitly
Goh & Bambara (2010)	1997-2008	83	145	DD	Elementary school students 69%), middle school students 21%, and high school students 10%.	School	School-based behavioral interventions that utilized FBAs	Off-task or disruptive behavior, aggressive or self-injurious behavior, and stereotypy; engagement and the increase in social or communication	Positive significant effects. PND moderate. Median PND was 88%; Maintenance effects large with a median PND of 100%	PND
Gresham et al. (2004)	JABA 1991-'99	150	Not given	DD	0 to 18	School	School based FBA behavioral treatments	Academic behavior and combined outcomes both	Interventions used FBAS 4.60 (PND =51.41); those did not 6.77 (PND = 66.15), both effective	SMD; PND

Table 1 Cont'd

Author	Yrs	N of Studies	N of Participants	Disability Status	Age	Setting	Independent Variables	Dependent Variables	Main Findings	Statistical Metric
Cont'd Gresham et al. (2004)								22.67%, social behavior 19.333%, stereotypical/destructive behavior 10%, disruptive behavior 8%, academic-related behavior 6.67%, daily living skills 6%, eating, 2%, other 2%		

Note: PND=percentage of nonoverlapping data; PZD=percentage of zero data; MBLR=mean baseline reduction; PEM = percentage of data points exceeding the median; PAND = percentage of all nonoverlapping data; HLM=hierarchical linear modeling; Allison-MT=; mean average trend; DD=developmental disability; ID=intellectual disability; ASD=autism spectrum disorder. * Indicates that a study includes a literature review, however Horner et al., (2002) also conducted a quantitative review, while Carr et al., 1999 is solely a literature review.

Table 2

Determinants of Effectiveness of Various Statistical Metrics

Metric	Highly Effective	Fairly Effective	Questionably Effective	Ineffective
PND *	> 90%	90 to 70%	70 to 50%	< 50%
PZD *	> 80%	80 – 55%	55- 18%	<18%
PEM *	> 90%	90 to 70%	70 to 50%	< 50%
PAND *	PAND is scaled from 50% to 100%, where 50% is chance level (cf. Parker et al., 2011).			
Allison-MT	>.47	.19-.46	.04-.18	<.04
SMD	>.80	.50-.79	.30-.49	<.30
HLM	There are statisticians who are working towards having a comparable effect size calculation for SCDs (Hedges, Pustejovsky, & Shadish, 2012; Shadish, Hedges, Pustejovsky, Boyajian, Sullivan, Andrade, & Barrientos, 2014), as compared to group design ES calculations.			
MBLR *	An MBLR score of 100% means total reduction of the challenging behavior, score of 0% indicates no change from baseline. A negative MBLR score reflects an increase in the behavior during treatment.			

Notes. *Definitions adopted from Heyvaert et al., (2014); Others from Harvey et al., (2009)

Table 3

Moderating Effects of Participant, Intervention, and Study Characteristics Between Behavioral Interventions and Individuals' with Developmental Disabilities Behavioral Outcomes Using Extant Quantitative Analyses

Type of Characteristic	Specific Characteristic	Moderator Findings
Participant	Grade range:	No effect (Goh & Bambara, 2010)
	Age range:	No effect (Campbell, 2003; Denis et al., 2011; Didden et al., 1997; Didden et al., 2006; Harvey et al., 2009; Heyvaert et al., 2014; Marquis et al., 2000; Scotti et al., 1991; Shogren et al., 2004); Interventions conducted with adults more effective than younger participants (Heyvaert et al., 2012)
	Gender:	No effect (Campbell, 2003; Denis et al., 2011; Didden et al., 1997; Goh & Bambara, 2010; Harvey et al., 2009; Heyvaert et al., 2012; Heyvaert et al., 2014; Marquis et al., 2000; Scotti et al., 1991); Moderating effect, males had a higher level of reduction of challenging behaviors (Shogren et al., 2004)
	Specific Disability:	No effect (Didden et al., 1997; Goh & Bambara, 2010; Marquis et al., 2000; Scotti et al., 1991; Shogren et al., 2004); No effect for motor impairment (Denis et al., 2011; Heyvaert et al., 2012) No effect for sensory impairment (Heyvaert et al., 2012); Effect found for sensory impairment, that if indicated then the effect was significantly lower (Denis et al., 2011) Moderated the effect, youth with autism as compared to youth with other developmental disabilities responded more on 2 of 4 statistical metrics, the other 2 metrics showed no effect (Harvey et al., 2009); Interventions conducted with individuals with ASD greater effect than with individuals with other developmental disabilities (Heyvaert et al., 2012)
	Verbal communication ability	No effect (Campbell, 2003; Heyvaert et al., 2014; Scotti et al., 1991) No effect for communicative impairment (Heyvaert et al., 2012)
	Cognitive Status	No effect (Campbell, 2003; Didden et al., 1997; Heyvaert et al., 2014); Moderated, lower the intelligence the less effective (Marquis et al., 2000); The more severe the developmental disability the less effective the intervention was found to be (Harvey et al., 2009) No effect for PND scores, PZD higher mean scores for participants in mild to moderate intellectual disability range (Scotti et al., 1991)
	Race/Ethnicity	No effect, and not much data (Harvey et al., 2009)
	Medication	No effect (Denis et al., 2011)
	Type of classroom	No effect, the type of classroom setting (Goh & Bambara, 2010)

Table 3 Cont'd

Type of Characteristic	Specific Characteristic	Moderator Findings
Intervention	Intervention type	No effect (Campbell, 2003; Shogren et al, 2004); Type of strategy moderated the effect in that medication was significantly less effective than the other 13 behavioral strategies (Scotti et al., 1991) No effect for uni vs. multicomponent interventions (Heyvaert et al., 2012); Strategies when used alone were not highly effective, but antecedent, skills replacement, and consequences fairly effective used alone. Replacement skills was most effective and more so when used with systems change, traditional antecedent, and consequence manipulation (Harvey et al., 2009); Using both stimulus-based and reinforcement based interventions together moderated the effect, but interpret with caution low amount of data (Marquis et al., 2000); Positive combination interventions more effective than antecedent control only interventions. Also it was found that aversive and positive combinations, positive combinations, differential reinforcement of other behaviors only, antecedent control only, differential reinforcement of alternative behaviors only, noncontingent reinforcement only, and social stories only interventions were more effective than PECS only interventions (but small sample of PECS only interventions; Heyvaert et al., 2014); The presence of manipulating antecedent factors (informing/educating/training environment) created larger effect (Heyvaert et al., 2012) Behavioral interventions moderated the effect more than psychotherapeutic interventions using cognitive or self-management techniques (Didden et al., 2006) No effect for PND scores for the level of intrusiveness of the interventions; for PZD level 2 and 3 more effective than level 1 (Scotti et al., 1991)
	Intervention agent	No effect (Heyvaert et al., 2012; Scotti et al., 1991) No effect found if the agent had training or not (Shogren et al., 2004) Moderated the effect positively if typical agent (Horner et al., 2002; Marquis et al., 2000), but not independent of typical setting (Marquis et al., 2000)
	Intervention setting	No effect (Denis et al., 2011; Didden et al., 1997; Harvey et al., 2009; Heyvaert et al., 2012; Shogren et al., 2004; Scotti et al., 1991) Moderated the effect positively if typical setting, but not independent of typical agent (Marquis et al., 2000)
	Format (individual or group)	No effect (Heyvaert et al., 2012)

Table 3 Cont'd

Type of Characteristic	Specific Characteristic	Moderator Findings
	Duration	No effect (Didden et al., 1997; Heyvaert et al., 2012) Moderated the effect, with 3 to 20 weeks being most effective compared to those less than 3 or more than 20 weeks (Harvey et al., 2009)
	FBA presence	No effect (Heyvaert et al., 2012; Shogren et al., 2004); Moderated the effect positively if present (Campbell, 2003; Carr et al., 1999; Didden et al., 1997; Didden et al., 2006; Harvey et al., 2009; Heyvaert et al., 2014; Horner et al., 2002; Marquis et al., 2000; Scotti et al., 1999)
	FBA data utilized to inform intervention	No effect (Shogren et al., 2004) No effect for matching treatment with behavioral function (not specifically an FBA; Denis et al., 2011) Moderated the effect positively if utilized (Carr et al., 1999; Didden et al., 2006; Marquis et al., 2000)
	FBA Assessment agent	No effect (Goh & Bambara, 2010)
	FBA Assessment setting	No effect (Goh & Bambara, 2010)
	Team decision making during FBA	No effect (Goh & Bambara, 2010)
	Parental/Family involvement	No effect (Campbell, 2003; Heyvaert et al., 2012) Significantly moderated effect (no details located; Heyvaert et al., 2014)
	Peer involvement	No effect (Heyvaert et al., 2012)
	Efforts to generalize behavior change	No effect (Campbell, 2003) Significantly moderated effect if techniques used to generalize (Didden et al., 2006; Heyvaert et al., 2014; Scotti et al., 1991)
Study	Type of problem behavior	No moderating effect (Didden et al., 2006; Heyvaert et al., 2014; Shogren et al., 2004) Moderated, all significant effects however, if the target was aggression, then the ES was significantly higher as compared to self-injury, destruction, and tantrums (Marquis et al., 2000) Interventions for self-injury and stereotypies more effective than those for aggression and destructive behavior (Heyvaert et al., 2012) Self-injury, stereotypy, socially inappropriate, and destructive behavior responded more than disruptive and aggressive behavior (Harvey et al., 2009); Externally destructive behavior (destruction of property), less effected than internally maladaptive (self-injury, stereotypies) or socially disruptive behavior (Didden et al., 1997) Moderated the effect, PND: physically aggressive/tantrum and destructive/disruptive least effective compared to self-injury, stereotypies and inappropriate social/other behaviors; PZD:

Table 3 Cont'd

Type of Characteristic	Specific Characteristic	Moderator Findings
		destructive/disruptive and stereotypic least effects as compared to inappropriate social behavior/other, self-injury, physically aggressive/tantrums (Scotti et al., 1991) Moderated the effect, details not located (Campbell, 2003)
	Behavior Severity	No effect (Scotti et al., 1991);
	Criteria utilized for ASD diagnosis	No effect (Heyvaert et al., 2014) No effect (Campbell, 2003)
	Function of the behavior	No effect (Marquis et al., 2000)
	Type of data collected (percentage/frequency)	Moderating effect-if frequency counts then effect size decreased between .3 to .5 units (Marquis et al., 2000);
	Amount of data points in baseline phase	Moderated the effect- as the amount of data points increased the effect size decreased (Marquis et al., 2000) Moderated the effect, details not located (Heyvaert et al., 2014) Moderated the effect (Campbell, 2003)
	Amount of treatment data points	Moderated the effect, details not located (Heyvaert et al., 2014) Moderated the effect, details not located (Campbell, 2003)
	Publication year	Moderated the effect, details not located (Heyvaert et al., 2014) Moderated the effect, details not located (Campbell, 2003) No effect (Didden et al., 2006)
	Presence of inter-rater reliability data	Moderated the effect, details not located (Heyvaert et al., 2014) Moderated the effect, details not located (Campbell, 2003) Moderated the effect if present (Didden et al., 2006)
	Procedural reliability (intervention fidelity)	No effect (Didden et al., 2006)
	Presence of follow up data	Significantly moderated effect if present (Heyvaert et al., 2014) No effect (Campbell, 2003)
	Social validity measures	No effect (Goh & Bambara, 2010)

Note: *Horner et al., (2002) is a literature review and not a quantitative synthesis. Goh & Bambara, 2010; Gresham et al., 2004; Horner et al., 2002 include only youth as participants, while the other reviews include youth and adults combined.

Participant characteristics. In regards to participant characteristics, the moderating effect of age range was studied in many of the previous reviews (Campbell, 2003; Didden et al., 1997; Didden et al., 2006; Denis et al., 2011; Harvey et al., 2009; Heyvaert et al., 2012; Heyvaert et al., 2014; Marquis et al., 2000; Shogren et al., 2004; Scotti et al., 1991). All of these reviews found no moderating effect of age range except Heyvaert et al., (2012), which results indicated behavioral interventions conducted with adults were more effective than with younger

participants. Grade range was studied by Goh and Bambara (2010) and no moderating effect was found.

Gender was studied in 9 extant quantitative reviews and with no effect found (Campbell, 2003, Didden et al., 1997; Denis et al., 2011; Goh & Bambara, 2010; Harvey et al., 2009; Heyvaert et al., 2012; Heyvaert et al., 2014; Marquis et al., 2000; Scotti et al., 1991) and one found a moderating effect in that males had a higher level of reduction of challenging behaviors (Shogren et al., 2004).

Specific developmental disability was not found to moderate the effect in most reviews (Didden et al., 1997; Goh & Bambara, 2010; Marquis et al., 2000; Scotti et al., 1991; Shogren et al., 2004). While two studies indicated that individuals with autism responded more on most statistical metrics as compared to individuals with other developmental disabilities (Harvey et al., 2009; Heyvaert et al., 2012). Some studies analyzed the effect of secondary disabilities and it was found that motor impairments did not moderate the effect in both reviews that examined this (Denis et al., 2011; Heyvaert et al., 2012). One study examined the moderating effect of sensory impairment on outcomes and found it did not have an effect (Heyvaert et al., 2012), while another found it did in that the effect was significantly lower if sensory impairment was present (Denis et al., 2011). While other studies examined the effect of verbal communication ability and all four reviews found no effect (Campbell, 2003; Heyvaert et al., 2012; Heyvaert et al., 2014; Scotti et al., 1991).

In regard to cognitive ability as measured typically by intelligence quotient, it was found that there was no effect in three reviews (Campbell, 2003; Didden et al., 1997; Heyvaert et al., 2014) and three found that treatment was less effective for individuals with lower scores of intelligence (Harvey et al., 2009; Marquis et al., 2000; Scotti et al., 1991).

Only one study collected data on race/ethnicity and it was found to not moderate the effect (Harvey et al., 2009). Medication usage was examined as a moderator in one study and it was found to not moderate the effect (Denis et al., 2011), although another study examined it as a behavioral strategy out of 14 total strategies and found that medication was significantly less effective than the other 13 behavioral strategies (Scotti et al., 1991). Also, classroom setting that a participant is educated in was examined in only one study and no effect was found (Goh & Bambara, 2010).

In review, many different participant characteristics have been examined in many extant reviews. Age range was not found to moderate the effect in 10 out of 11 syntheses. Gender did not moderate the effect in 9 out of 10 syntheses. The large majority found that the specific type of disability did not moderate the effect, however 2 found that treatment for those with ASD was more effective. It was also found that verbal ability did not moderate the effect in all reviews that examined this moderator. While, there was a split in whether analyses show that the level of intellectual impairment has an effect, in that half found it did not and half found that treatment for those with more intellectual impairment is less effective. And not much data has been collected on race/ethnicity, medication usage, or classroom setting and no studies examined social economic status.

Intervention characteristics. There was not a consistent pattern with how the moderating variable, *intervention type*, was analyzed throughout the various syntheses however, it is still possible to try to organize the findings. There was no effect found for two syntheses for the type of behavioral strategies used (Campbell, 2003; Shogren et al, 2004). One study examined the effect of behavioral strategies (n=14) and found that 1, medication, was significantly less effective (Scotti et al., 1991). There seemed to be no consensus on whether uni

versus multicomponent interventions were effective. While one found no effect (Heyvaert et al., 2012), another found that most strategies used alone were not as effective (Harvey et al., 2009), while yet others found that combinations of various strategies were most effective (Harvey et al., 2009; Heyvaert et al., 2014; Marquis et al., 2000). More details are provided in Table 4.

There were mixed findings as to the moderating effect of the intervention agent, in that it was not found to moderate the effect in two reviews (Heyvaert et al., 2012; Scotti et al., 1991), while in two others it did moderate the effect, specifically that if implemented by typical agents then the effect was greater (Horner et al., 2002; Marquis et al., 2000). The fairly consistent finding in terms of if intervention setting moderated the results was that it did not (Denis et al., 2011; Didden et al., 1997; Goh & Bambara, 2010; Harvey et al., 2009; Heyvaert et al., 2012; Shogren et al., 2004; Scotti et al., 1991). One study found that if the intervention was implemented in a typical setting (by a typical agent) this was more effective than an atypical setting (e.g., lab, facility that the individual did not attend regularly; Marquis et al., 2000). The format, whether the intervention was administered in a group or individual format was examined in one synthesis and no effect was found (Heyvaert et al., 2012). Also the duration of the intervention had mixed results and was not found to moderate the effect in two studies (Didden et al., 1997; Heyvaert et al., 2012) and did in another (Harvey et al., 2009). Specifically, that 3 to 20 weeks was most effective.

Most of the syntheses examined the moderating effect of whether a FBA was conducted or not. The majority found that the presence of a FBA significantly impacted the results positively if utilized (Campbell, 2003; Carr et al., 1999; Didden et al., 1997; Didden et al., 2006; Harvey et al., 2009; Heyvaert et al., 2014; Horner et al., 2002; Marquis et al., 2000; & Scotti et al., 1999), while two showed no effect (Heyvaert et al., 2012; Shogren et al., 2004). A few of

these syntheses then went on to analyze other moderating effects related to FBAs, if a FBA was conducted in a study. Specifically, the moderating effect of if the FBA data was utilized to inform intervention was conducted and two found it had no effect (Denis et al., 2011; Shogren et al., 2004), while three syntheses found it had a significant effect if utilized (Carr et al., 1999; Didden et al., 2006; Marquis et al., 2000). Goh & Bambara (2010), who specifically studied the variables of interest with youth in schools went on to analyze whether the FBA assessment agent, setting, and presence of team decision making during the FBA moderated the effect and none were found to do so.

A few other intervention characteristics have been examined in a few of the extant reviews. Parent/family involvement has been studied and mixed results have been indicated, with no effect being found (Campbell, 2003; Heyvaert et al., 2012), and a positive effect for family involvement (Heyvaert et al., 2014). Heyvaert et al., 2012 was unique in that they examined the effect of peer involvement and did not find an effect. One last intervention characteristic that has been examined was the presence of intervention techniques to generalize behavior change and the majority found that these efforts significantly moderated the effect positively if present (Didden et al., 2006; Heyvaert et al., 2014; Scotti et al., 1991) and one found no effect (Campbell, 2003).

In summary, many of the findings were mixed specifically the intervention type, agent, duration, and utilization of FBA data in the intervention development. Intervention type was not studied uniformly across any of the reviews. A consistently studied moderator was whether a FBA was conducted, and the large majority showed that the presence of a FBA moderated the effect in a positive direction. Furthermore, a few studies examined the effect of generalization techniques used in the intervention and the majority of these found that the presence of these

techniques significantly moderated the effect positively. Also, a popular moderator analysis and fairly consistent finding was that setting did not moderate the effect of the intervention. While, family involvement was studied by only a few syntheses, the majority found no effect for family involvement. Two uniquely studied moderators were peer involvement in the intervention and the format of the intervention group, both with no effect. The most closely related study to the proposed study, examined more variables related to FBA presence and found no effect for FBA agent, setting, or decision making utilization (Goh & Bambara, 2010).

Study characteristics. There have also been many moderator analyses conducted on various study characteristics in the extant reviews. The type of challenging behavior was studied in many studies. A few of the studies found no effect (Didden et al., 2006; Heyvaert et al., 2014; Shogren et al., 2004). While more studies found that there was a positive effect for all challenging behaviors but some challenging behaviors had more of an effect than others (Campbell, 2003; Didden et al., 1997; Harvey et al., 2009; Heyvaert et al., 2012; Scotti et al., 1991). The results were mixed in terms of which challenging behaviors moderated the effect. Please refer to Table 4 for a breakdown of these differences. To summarize very generally, the effect was found to be higher for aggression as compared to other specific challenging behaviors in 2 reviews (Marquis et al., 2000; Scotti et al., 1991), while it was found to be less effective as compared to some challenging behaviors in 3 reviews (Harvey et al., 2009; Heyvaert et al., 2012; Scotti et al., 1991). For destructiveness 5 reviews found interventions for this behavior to be less effective (Didden et al., 1997; Heyvaert et al., 2012; Marquis et al., 2000; Scotti et al., 1991). Note that Scotti et al., 1991 was counted twice as two different metrics were reported, with differing results. While one review found interventions to be more effective when destructiveness was the outcome (Harvey et al., 2009). Tantruming behaviors had mixed effects

for being more effective (Scotti et al. 1991) or less effective (Marquis et al., 2000; Scotti et al., 1991). Self-injury had the majority find it more effective (Didden et al., 1997; Harvey et al., 2009; Heyvaert et al., 2012; Scotti et al., 1991). Note that Scotti et al., 1991 found this to be the case for both metrics utilized. While one review found interventions used to reduce self-injury were less effective than other challenging behaviors (specifically, aggression) (Marquis et al., 2000). When stereotypies were the outcome of the intervention the interventions most often were found to be more effective than when used for other challenging behaviors (Didden et al., 1997; Harvey et al., 2009; Heyvaert et al., 2012; Scotti et al., 1991), and on one metric in Scotti et al., (1991) interventions focused on reducing stereotypies were found less effective. When socially inappropriateness was the focus of the intervention in all of the reviews the effect was found to be more effective in all of the reviews that studied this (Didden et al., 1997; Harvey et al., 1999; Scotti et al., 1991). Note that Scotti et al., 1991 found this to be the case for both metrics utilized. Also, disruptiveness was found to be less effective as compared to other challenging behaviors in two reviews (Harvey et al., 2009; Scotti et al., 1991).

A few study characteristics that were not analyzed by many studies and no effect was found included behavior severity (Scotti et al., 1991), criteria utilized for ASD diagnosis (Campbell, 2003; Heyvaert et al., 2014), function of the behavior (Marquis et al., 2000), intervention fidelity (Didden et al., 2006), and social validity measures (Goh & Bambara, 2010). Other moderators were examined by just a few reviews and found effects. The type of data collected (whether it was percentage or frequency) found that if frequency counts were used than the effect size decreased significantly (Marquis et al., 2000). Also, it was found that the more data points in the baseline phase the lower the effect size (Marquis et al., 2000), and the amount moderated the effect in two other studies but no further details were provided (Campbell, 2003;

Heyvaert et al., 2014). Also, the amount of data in the treatment phase was found to moderate the effect, but further details were not located (Campbell, 2003; Heyvaert et al., 2014). The publication status had mixed results in that it did not have an effect (Didden et al., 2006), and in two others it did moderate the effect, but further details could not be located (Campbell, 2003; Heyvaert et al., 2014). Mixed results were also found for the presence of follow up data in that it moderated the effect positively (Heyvaert et al., 2014) while another review found no effect (Campbell, 2003). The presence of interrater reliability data moderated the effect in all studies that analyzed this variable (Campbell, 2003; Didden et al., 2006; Heyvaert et al., 2014).

In sum, of the moderating effects of study characteristics in the extant quantitative reviews there were many variables that were studied only by a few researchers including: behavior severity, criteria used for ASD diagnosis, function of the behavior, intervention fidelity, social validity measures, type of data collected, amount of data in treatment or baseline phases, publication year, presence of follow up data, and the presence of interrelated reliability data. A consistently studied characteristic was the type of problem behavior and the studies found a positive effect for all challenging behaviors, while some challenging behaviors had more of an effect than others. Aggression, destructiveness, tantrums, self-injury, stereotypies, all had mixed results except socially inappropriateness, which the effect was always more effective and disruptiveness, which was always less effective than other specific behaviors.

Table 4

Investigation of the Moderator Findings from Extant Meta-Analyses Pertaining to Type of Behavioral Outcome

	More Effective		Less Effective	
Behavior	**Study**	**Comparison**	**Study**	**Comparison**
Aggression	Marquis et al., 2000	Self-injury, destruction, tantrums		
			Heyvaert et al., 2012	Self-Injury & stereotypies
			Harvey et al., 2009	Destructiveness, self-injury, stereotypies, & socially inappropriateness
			Scotti et al., 1991 (combined tantrums with aggression)	Self injury, stereotypies, & inappropriateness
	Scott et al., 1991 (combined tantrums with aggression)	Destructive/disruptiveness & stereotypies		
Destructiveness			Heyvaert et al., 2012	Self-Injury & stereotypies
			Didden et al., 1997	Self-injury, stereotypies, & socially inappropriateness
			Scotti et al., 1991 (combined with disruptiveness)	Self injury, stereotypies, & inappropriateness
			Scotti et al., 1991 (combined with disruptiveness)	Inappropriate social, self-injury, aggression/tantrum)
			Marquis et al., 2000	Aggression
Tantrums	Harvey et al., 2009	Aggression & Disruptiveness	Scotti et al., 1991 (combined tantrums with aggression)	Self injury, stereotypies, & inappropriateness
	Scott et al., 1991 (combined tantrums with aggression)	Destructive/disruptiveness & stereotypies		

Table 4 cont'd

More Effective			Less Effective	
Behavior	**Study**	**Comparison**	**Behavior**	**Study**
Self-injury	Heyvaert et al., 2012;	Aggression & destructiveness		
	Harvey et al., 2009	Aggression & Disruptiveness		
	Didden et al., 1997	Destructiveness		
	Scott et al., 1991	Destructive/disruptiveness & stereotypies		
	Scotti et al., 1991	Aggression/tantrum, & Disruptiveness/destructiveness		
			Marquis et al., 2000	Aggression
Stereotypies	Heyvaert et al., 2012;	Aggression & destructiveness		
	Harvey et al., 2009	Aggression & Disruptiveness		
	Didden et al., 1997	Destructiveness		
	Scotti et al., 1991	Aggression/tantrum, & Disruptiveness/desturctiveness		
			Scotti et al., 1991 (combined with disruptiveness)	Inappropriate social, self-injury, aggression/tantrum)
Socially inappropriateness	Harvey et al., 2009	Aggression & Disruptiveness		
	Didden et al., 1997	Destructiveness		
	Scotti et al., 1991	Aggression/tantrum, & Disruptiveness/desturctiveness		
	Scott et al., 1991	Destructive/disruptiveness & stereotypies		
Disruptiveness			Harvey et al., 2009	Destructiveness, self-injury, stereotypies, & socially inappropriateness
			Scotti et al., 1991 (combined with destructiveness)	Self injury, stereotypies, & inappropriateness

Integration of Research Findings

As demonstrated above, it is important to integrate findings across the literature in order to develop a more complete picture of the consistent themes across a body of research. An important way of integrating the findings of multiple studies examining the same variables is through meta-analyses. However, single-case designs have typically not been included in most meta-analyses, however due to the population of interest being low-incidence they have been often included in this body of literature.. To highlight the potential benefits of including SCD and to provide information on the state of research in this area, in this section, the following topics will be reviewed: (a) features and benefits of single-case designs; the (b) benefits of synthesizing research findings, particularly meta-analyses, (c) benefits of meta-analyses, and (d) the purpose and benefits of SCD meta-analyses.

Each single-case design begins with basic A-B (or baseline-intervention) phases, and then additional phases may be introduced through an A phase (no treatment) and then another B phase(treatment) (Riley-Tillman & Burns, 2009). During the A phase(s) of treatment, the DV is measured multiple times before the introduction of the intervention during the B phase(s) (Krysik & Finn, 2010). Then after the intervention has been implemented (B phase), the DV is measured on a regular basis. There are variations of these types of designs, creating a multitude of single-case design options (Owens, 2011). For example, there can be multiple participants or groups, and/or treatments. The purpose of this type of design is to understand if an intervention creates change (Krysik & Finn, 2010). This type of design has repeated data collection over time, showing small changes over time and the results are typically displayed graphically (Krysik & Finn, 2010).

SCDs have many benefits. One benefit of single-case design is that the documentation of the results of the treatment is systematic and there is frequent and repeated measurement of the DVs (Zhan & Ottenbacher, 2001). This allows the treatment effect to be analyzed using multiple observations, enabling the analysis of treatment effect changes over time (Owens, 2011). Moreover, this sort of design is more practical for practitioners, which shortens the distance between research and practice (Morgan & Morgan, 2001). Specifically in the school setting it is not usually appropriate to have a control group and this type of design does not call for randomization of participants (Riley-Tillman & Burns, 2009). In addition, replication of single- case design studies is easier to implement then group-based studies, which improves the generalization of findings. Zhan & Ottenbacher (2001) stated that a decision made concerning an individual student's educational decisions using evidence-based research that was conducted on many participants may cause problems when those findings are applied to individual cases of students. SCDs concentrate on the variation in the treatment effect at the individual level, which has been found to vanish when the focus is on the average treatment effect, as in group comparison designs (Barlow et al., 2009).

According to Owens (2011), the use of single-case designs has become more prolific with researchers in varying fields, such as school psychology, education, special education, and behavioral intervention studies, and it is important for researchers to synthesize these results through meta-analytic techniques. Quantitatively integrating the results of multiple studies for a particular population or a specific DV, through a meta-analysis, is a useful way to combine the findings so that research is organized in a way that is useful for practitioners, other researchers, and decision makers (Owens, 2011). Meta-analysis, as a statistical method, was first introduced by Glass (1976), as a quantitative approach to summarize results of studies. Glass (1976)

defined it as, “the statistical analysis of a large collection of analysis results from individual studies for the purpose of integrating the findings” (p.3). Meta-analyses have multiple purposes, including the following: (a) identification of variables that may influence outcome variables, (b) summarizing the overall effectiveness of the treatment that is being analyzed, (c) and describing the body of research as a whole (Blimling, 1988; Busk & Serline, 1992). Meta analyses allow others to access the literature by integrating the findings of multiple studies using a systematic approach to analyzing the research and generating conclusions (Owens, 2011). Kavale and Glass (1981) stated that research integration is needed to help legitimize the work of multiple researchers by allowing similar studies to be synthesized.

Using meta-analysis research design to analyze research from SCDs is a relatively new practice in the fields of psychology and education (Miller & Lee, 2013). There has been considerable debate over the best way to calculate effect sizes for this type of study (Maggin, O’Keeffe, & Johnson 2011). However, according to Schlosser (2005), "while there is still some debate about what 'effect size' is most appropriate, the question of whether or not to synthesize single-subject experimental designs using meta-analytic techniques is no longer in question” (p. 376). Meta-analyses of single-case designs should be performed more frequently, considering (a) the validity of findings of well-designed single-case research, (b) increase in the use of such designs in the past few decades, and (c) single-case designs to deem interventions as evidence-based (Miller & Lee, 2013). When multiple SCD findings are aggregated together, then the overall treatment effect, as well as the individual treatment effect can be estimated (Van den Noortgate & Onghena, 2003). By integrating the findings of multiple single-case design studies, theoretically, the generalizability of the results of the individual cases increases (Riley-Tillman & Burns, 2009). In addition, it has been found that many studies with youth with disabilities or in

a nonclinical setting are conducted utilizing single-case design because it is harder to have large numbers of participants when studying low incidence and small populations (Parker, Vannest, & Brown 2009). It is important to be able to synthesize single-case design studies for these populations and to analyze any potential moderating variables.

Conclusion

It seems that there was a gap in the literature, as there had yet to be a SCD synthesis of all three levels of SWPBIS for youth with developmental disabilities that include both FBA and non-FBA based interventions. Furthermore, nonparametric statistics were utilized in the two most relevant meta-analyses (Goh & Bambara, 2010; Gresham et al., 2004), whereas the present study utilized hierarchical linear modeling, a parametric statistical method to conduct a SCD meta-analyses (Van den Noorthgate & Onghena, 2003). The publication years that have been included have been limited to 1997 to 2008 for Goh & Bambara (2010), and for Gresham et al. (2004) from 1991-1999 and only studies from the Journal of Applied Behavioral Analysis were included.

The purpose of this study was to conduct a meta-analysis of single-case design studies that examined the effect of school-based behavioral interventions on the behavioral outcomes of youth with developmental disabilities at any tier of PBS support. This body of literature was in need of a quantitative synthesis including a large publication year range and of specifically school-based interventions (FBA and non-FBA) for youth with developmental disabilities. The majority of the previous meta-analyses grouped adults and youth together and of those that have examined school-based interventions they have been limited in various ways. The present studies' main analyses, along with a comprehensive moderator analysis, may provide a more accurate and detailed understanding of the effect of school-based behavioral interventions for

this population. In addition, the importance of intervening in behavioral problems is highlighted in its impact on optimal childhood and later, adult functioning. The present study has important implications for youth with developmental disabilities and challenging behaviors and the school-based practitioners that work with these youth.

Chapter III: Method

In this chapter, a detailed account is provided of the methods utilized in the study. The chapter highlights the various search strategies, as well as the inclusion and exclusion criteria for a study to be included in the present meta-analysis. These are followed by a description of the various processes that were used to establish if a study meets the inclusion criteria and then information on the organization of the eligibility phases is presented. After this, there is a detailed description of the system used to code the outcome variables and moderators, as well as how the data was extracted from the studies. Next there is a description of how graduate students assisted with the data collection. This chapter concludes with a description of the statistical analyses used to analyze the data.

Search Strategies

The present study utilized different search methods to locate studies. A comprehensive search was performed on relevant databases. The databases that were searched include: (a) PsychINFO (b) ERIC and (c) Proquest Dissertation and Thesis Abstracts. All three of these databases were searched simultaneously using EBSCO. The first two databases were chosen because they are the same databases searched in Goh and Bambara (2010) and the third was chosen, because it includes both unpublished and published studies. The electronic searches of the above mentioned databases involved combining search strings for the independent variable and dependent variable. Specifically the following keywords were searched on each database concerning the independent variable (intervention and treatment terms): (a) *positive behavior* support,* (b) *response to intervention,* (c) *functional analysis,* (d) *functional assessment,* and (e)

behavior modification. The following keywords were searched on each database concerning the dependent variable (symptom terms): (a) *behavior* problem* and (b) *disruptive* behavior**. The keywords were chosen by reviewing the prior school-based meta-analyses and scanning the results section for keywords these meta-analyses used and are relevant to the present study. The logic for use of these keywords was due to referring to the interventions or outcomes of interest (Littell, Corcoran, & Pillal 2008). Furthermore, the symbol * was utilized, because it expanded the keywords so that the database also searched for different versions of the root of the word (Littell et al., 2008). For example, *behavior* problems* searched for both behavioral and behavior. The way each search was performed was by utilizing a keyword from each of the two categories, independent and dependent variables, mentioned above. The starting search year was 1997, when IDEA enacted the first legislation that required PBS and FBAs to be conducted in the schools and the search continued through January 2017. A secondary search method, called "foot chasing" (White, 2009), was utilized by searching the reference list of the previous meta-analyses (Campbell, 2003; Denis et al., 2011; Didden et al., 1997; Didden et al., 2006; Goh & Bambara, 2010; Harvey et al., 2009; Gresham et al., 2004; Heyvaert et al., 2012; Heyvaert et al., 2014; Horner et al., 2002; Marquis et al, 2000; Scotti et al., 1991; Shogren et al., 2004) and literature reviews (Carr et al., 1999; Horner et al., 2009). An additional method of hand searching relevant journals was conducted to locate articles that did not emerge from the other search methods. The following journals were hand searched beginning with the publication year of 1997 through January 2017: (a) *Behavioral Disorders*, (b) *Journal of Applied Behavior Analysis*; (c) *Journal of Positive Behavior Interventions*, and (d) *Education and Treatment of Children.* The rationale for searching these particular journals was these are the journals that produced at least five percent of the included articles out of the total hand searching method in Goh & Bambara

(2010). Goh & Bambara searched 15 journals and the large majority of these did not produce many articles for study inclusion (n =11) and the current study had similar inclusion criteria to that of Goh & Bambara (2010).

A summary of the literature search methods is shown in Table 5. After all of the potential studies were gathered they then underwent eligibility review rounds based on inclusion and exclusion criteria, which is described in the next section.

Inclusion Criteria

1. The study independent variable (IV) is a school-based behavioral intervention.
2. The study was conducted with school-aged children between the ages of 3 and 22.
3. The study was conducted in the period from 1997 through January 2017.
4. The researcher measured at the DV of behavioral outcomes in relation to a school-based behavioral intervention. Some examples of behavioral outcomes include aggression, on task behavior, compliance, functional communication, initiating conversations, eye contact during conversation, and showing interest in other.
5. Articles must use a single-case design. This can include A-B-A-B designs, multiple baseline designs across subjects, A-B designs, multi-element designs, and multi-treatment designs.
6. Articles published in languages other than English were acceptable provided that a translation could be found. If a translation cannot be found, this study was ruled out.
7. Theses and dissertations were acceptable, provided they met the other criteria.
8. The study provided enough quantitative data to allow a calculation of a stable effect size, which is defined as at least three data points assigned to the baseline phase as well as to the treatment phase (Swanson, 2000).

9. The study provided data to permit the calculation of effect sizes or it was obtained from the lead researchers.
10. The study met the What Works Clearinghouse standards for a well-designed SCD, falling within the categories of "Meeting Standards" or "Meeting Standards with Reservations".

Table 5

Search Strategies

Search Strategy	Details
Database search	Variations of a keyword from each of the keyword categories including: independent variable, dependent variable, population of interest age, and population of interest disability type
Footchasing	Foot-Chasing Methods: checked the citation lists and included study reference lists of all extant meta-analysis and literature reviews for single-case design studies
Hand-searching Journals	Went through the table of contents of each issue of the 4 journals for publication years 1997- January 2017.

Exclusion Criteria

Studies that did not meet the inclusion criteria described above were excluded. This included studies that examined the effects of behavioral interventions of youth with developmental disabilities that were community or home-based.

Study Eligibility Process

There were six phases of review with set criteria for inclusion in each phase. There was a need to have another individual help in the review process other than the principal investigator.

A graduate student with a background in education helped as a research assistant with the collection of the data. The research assistant was recruited by sending out an email to the various relevant graduate programs to see if anyone was interested in helping with data collection. The research assistant was trained on the inclusion criteria and eligibility phase requirements. Inter-rater reliability was gathered during each eligibility phase for 10% of the identified studies in that review round/phase. If there was disagreement among the raters, then the particular study was brought to the principal investigator's attention and was reviewed together until consensus was met. The calculation of inter-rater agreement was conducted through the following formula: agreements divided by the number of agreements plus disagreements, multiplied by 100.

During the first phase of the eligibility review, the two eligibility criteria that were determined is if the study involved: (a) a behavioral intervention and (b) individuals with developmental disabilities (based on diagnoses/special education classification). Just as in Heyvaert et al., (2014) individuals who are labeled as "autistic-like" or engaging in "autistic-like behavior" will be excluded, unless a formal diagnosis/education classification has been given to the participant. At this stage, just the abstract of the article was reviewed. If the criteria could not be determined by only reviewing the abstract, then the reviewer read the entire article to make a determination. The decision to use these two criteria first was because this information was, for the most part, available in reviewing just the abstract and it allowed for a time effective way of exclusion of a large portion of the studies. During the second phase of eligibility review, the two criteria used to determine eligibility included whether or not the study was conducted on: (a) youth between 3 to 22 years of age and (b) in the school setting. School settings was defined as in Gresham et al., (2004) to include all educational institutions (i.e., public, private, hospital

schools, residential schools). The third phase involved a review of the abstract or article to see if the study measured behavioral outcomes (e.g., aggression, attention, and/or social behavior).

The fourth review round criteria was: *Is the study design a single-case design,* to determine eligibility into the next review round. For this phase of the review, the abstract or article was reviewed to be able to determine if the study met the inclusion criteria for Round 4. For studies that were determined to be SCDs, then the results section of the study was used to determine if the study met Round 5 inclusion criteria. These criteria examined whether the researchers of the final Round 4 studies, provided sufficient data for the proposed analyses, which was determined by the presence of three or more data points for a baseline phase as well as a treatment phase (Swanson, 2012). The final and 6th review phase was conducted by reviewing the study to determine if the IQ of the individual was given and is < 70, for those labeled as having Autism Spectrum Disorder or PDD-NOS, if not, then these were excluded. It was assumed if the youth were given a diagnosis of intellectually disabled or developmentally delayed then they have an intellectual impairment. Those studies that made it through Phase 6 will be included in the data analysis. Table 6, provides a summary of the proposed eligibility process. The final criteria, if the studies met the standards for a well-designed SCD, was determined a priori to be coded as the first variable when coding the final studies, due to the extensive nature of determining if it met the criteria of a well-designed SCD or met the standards with reservations, and all other studies were excluded if they didn't meet these two levels of standards.

Organization of Eligibility Phases

Online technologies were utilized to help the organization and extraction of data necessary for the current study. To be exact, the citation and resource management system,

RefWorks, was utilized for all of the online database searches. The lists of the studies that were found through the various eligibility phases were kept in separate folders within RefWorks. This organizational system allowed for the researcher to stay organized and enhance accurate reporting of data. This system also allowed the researcher to indicate how many studies were included or excluded at each phase.

If an article needed to reviewed in whole, then the researcher downloaded the full article and saved them into a DropBox folder for ease of locating the studies for other review rounds/data coding. The articles were located through the University of South Florida's library services. If a study was unavailable through the USF database system, then a request to the Interlibrary loan services was made. If after two weeks the study was still unavailable then the study was excluded.

Also, GoogleDocs was utilized to serve as a way for the principal investigator and research students helping with data collection to communicate about delegated responsibilities, track if a study met or did not meet criteria, and for data coding.

Coding System

Next the final studies were coded that met all six eligibility review round criteria (final studies). A list of operational definitions and coding criteria for each term and category was created for reliability of coding, refer to Table 8 for operational definitions and to Table 9 for an example of the coding key. Each category was assigned a numerical number to help with data analysis. Articles were coded for the dependent variables of behavioral outcomes. Also, studies were coded to indicate whether the study shows a positive effect if the baseline to treatment data declines or whether a treatment effect is indicated if the baseline to treatment data increases. This step was important to analyze the data validly. Each article was additionally coded for an

Table 6

Eligibility Review Process

Review Round #	Inclusion Criteria	Review Type	Inter-rater Reliability
1.	(a) Is the intervention a behavioral intervention? And (b) is it conducted with individuals with developmental disabilities (based on diagnoses and not IQ at this stage)?	Abstract review unless a full article review was needed to locate the information.	10% of studies with > 80% IRR
2.	(a) Is the intervention conducted with youth between 3 to 22 years of age? And (b) is it conducted in a school setting?	Abstract review unless a full article review was needed to locate the information.	10% of studies with > 80% IRR
3.	Did the study measure behavioral outcomes (e.g., aggression, attention, and/or social behavior)?	Abstract review unless a full article review was needed to locate the information.	10% of studies with > 80% IRR
4.	Is the study's design a single-case design?	Abstract review unless a full article review was needed to locate the information.	10% of studies with > 80% IRR
5.	If there sufficient data for the proposed analyses, which will be determined by the presence of three or more data points for a baseline phase as well as a treatment phase?	Review the results section of a study	10% of studies with > 80% IRR
6.	Is the IQ of the individuals 70 or below? Or has the individual been labeled as intellectually disabled?	Review the article.	10% of studies with > 80% IRR
Final Studies	Will be coded	Review the article	10% of the studies with > 80% IRR ; 10% of the graphs if DataThief II is needed to extract data

extensive list of characteristics including participant characteristics, intervention characteristics, and study design characteristics. This allowed for the examination of potential moderating variables. The particular participant characteristics that were coded included: age range, grade range, gender, specific disabilities, cognitive status, level of verbal communication ability, and classroom setting the participant is educated in. The intervention characteristics that were coded included: intervention type, agent, setting, format, duration, presence of a functional behavioral analysis (FBA), techniques used to generalize behavior change, and school-wide positive behavioral support tier (1, 2, or 3). If an FBA was conducted then the following was coded: FBA method, assessment agent, assessment setting, and team decision-making during assessment (Goh & Bambara, 2010). The study design characteristics that were coded included: type of challenging behavior, intervention fidelity, social validity measures, published/unpublished, inter-rater reliability data, type of single case design, and if they met the criteria for being a well-designed SCD. Many of these variables were chosen to be analyzed due to these examinations in past meta-analyses (Goh & Bambara, 2010). Table 7 provides details about the variables that were coded.

A coding database was developed in Google Docs that allowed for the data to be entered into an online database, so that the research assistant and the principal investigator had access to the data collection and the document saved simultaneously. The GoogleDoc was used during interrater reliability checks, of which 10% of the studies were coded by a second coder. The database from GoogleDocs is compatible with Excel and was exported to the Excel software program for later use for statistical analyses.

Outcome Data Extraction

Baseline and treatment raw data points were extracted from the studies. In order to extract the data, the following order of methods was used: (1) obtaining raw data from studies; (2) through the use of the DataThief III (2006) computer software; this software precisely extracts the data from the graphs provided in studies through importing the graphs in .JPEG file format; 3) if the graph or data were not provided in a study then the authors of the study were contacted. If the authors were unable to send the data within two weeks, then those cases were excluded. If, after exhausting all of these methods, the researcher was unable to extract the data, then these cases w excluded.

Table 7

Description of Coding of Study Moderators

Type of Characteristic	Specific Characteristic	Definition	Coding Categories
Participant	Grade range	The school grade(s) of the participants.	To be coded as *N*, for each category: preschool to pre-K, kindergarten to 1st, 2nd-3rd, 4th-5th, 6th-8th, 9th-12th, not provided
	Age range	The age of the participants.	To be coded as *N* for each cate
	Gender	The gender of the participants.	Male or female, or data not provided.
	Specific Disability	The special education diagnostic label given to participants (i.e., ASD, PDD-NOS, InD)	PDD-NOS, ASD, intellectual disability, combinations of disabilities, developmental disability, other, not provided
	Cognitive Status	Whether the participants are developmentally delayed or intellectually disabled or not.	Mild (IQ between 70-55), moderate (54-40), severe (39-below), profound labeled as intellectually disabled but no IQ specified, not provided
	Verbal Communication Ability	What the participant's verbal communication ability is	Nonverbal/mute, minimally verbal, echolalic, average language skills, repetitive speech, other (functional sign language), not reported (Campbell, 2003)
	Classroom	What type of classroom the participant is educated in	General education (i.e., participant received instruction in a general education classroom only), (b) special education (i.e., participant received instruction in a special education classroom only), and (c) combination of general and special education (i.e., participant received instruction in a both general and special education class- rooms) (Goh & Bambara, 2010)
Intervention	Duration	The total duration of treatment as defined by the total amount of treatment data points across all treatment phases (Goh & Bambara, 2010)	"Long (i.e., 21 or more data points), or short (i.e., 20 or fewer data points). The criterion for the length of treatment was based on Snell et al. (2005). (Goh & Bambara, 2010)"
	Type	The type of behavioral intervention	"(a) Skills training, interventions that targeted skill acquisition (e.g., self-management, functional communication training); (b) antecedent-based intervention,

Table 7 Cont'd

Type of Characteristic	Specific Characteristic	Definition	Coding Categories
			interventions that made environmental modifications before problem behavior occurred (e.g., curricular modification, noncontingent reinforcement, choice making); (c) consequence- based intervention, interventions that were implemented following the occurrence of a targeted behavior (e.g., positive reinforcement, differential reinforcement, extinction); and (d) multicomponent interventions (i.e., combinations of two or more intervention categories) (Goh & Bambara, 2010).
	Agent	The person that delivers the intervention.	Typical (involving at least one school employee), atypical (i.e., researcher or research assistant) (Goh & Bambara, 2010)
	Setting	The specific place that the intervention is implemented.	Typical (participant's classroom), Atypical (i.e., pull-out setting, such as an empty classroom)
	Format	How many participants are included in the intervention at the same time.	Individual-based, small group (< 10), medium group (10-30), large group (+30), whole class, other, not provided
	Generalization	Were there efforts to generalize the behavioral intervention results?	New situation or setting only, new behavior only, combination, with new person only (Campbell, 2003)
	SWPBS Tier	Whether the intervention was delivered school-wide, at Tier 2, or individualized at Tier 3	Tier 1, Tier 2, Tier 3, not provided
	Presence of FBA	If a FBA was conducted or not	Yes, no (if the study does not explicitly say a FBA was conducted then this will be coded as no)
	FBA method	"Method used to assess the environmental influences of the participant's behavior" (Goh & Bambara, 2010)	Experimental only, descriptive only, or combination of experimental and descriptive methods, or not conducted/provided

Table 7 Cont'd

Type of Characteristic	Specific Characteristic	Definition	Coding Categories
	FBA agent	Individuals who conducted or provided assessment data	Atypical (if the only person involved was a researcher), typical (involving at least one school employee) (Goh & Bambara, 2010)
	FBA setting	Where the assessment was conducted	Atypical (pull-out classroom foreign to student's educational setting), typical (typical classroom), not provided (Goh & Bambara, 2010)
Study	Social Validity Measures	Inclusion of measures that evaluated the acceptability of intervention goals, procedures, and/or outcomes by stakeholders.	Published, unpublished, not provided
	Inter-rater Reliability Data	Did the researchers collect inter-rater reliability when collecting data?	Yes or no (if does not explicitly state these measures were collected then this will be coded as no) Yes or no, not provided

Table 8

Example of the Coding Key

Participant Characteristics									
Grade Range:	Preschool to Pre-K	K to 1st	1st to 2nd	2nd to 3rd	4th to 5th	6th to 8th	9th to 12th	12th plus	Not provided
Code	1	2	3	4	5	6	7	8	9
Age Range:	3 to 5	6 - 8	9 to 11	12 to 14	15 to 18	19-21	Not provided		
Code	1	2	3	4	5	6	7		
Gender:	Male	Female	Not provided	Other					
Code	1	2	3	4					
Specific Disability:	ASD	ID	DD	PDD-NOS	No a priori (fill in combinations)				
Code	1	2	3	4	5	6	7	8	9

Inter-rater reliability as gathered for the outcome data extraction methods. To assess the reliability of data extraction through use of DataThief III, 10% of the studies were randomly selected and a second reviewer extracted the data. Agreement rates between the researcher and the coder were analyzed by the following method: the number of agreements divided by the total number of comparisons and then multiplied by 100. Agreement was operationalized as the value of two data points being identical or one unit apart. All disagreements in data were reconciled by going back to the original article. These data were input into the data coding GoogleDoc.

Team Involvement

Research assistant involvement in the data collection was as follows:

Literature Search (as outlined in Table 5): The principal investigator asked for help to complete the initial search strategies. This involved putting the keywords into the databases to gather the studies to be reviewed, footchasing, and handsearching journals to locate all potential studies.

Eligibility Review Rounds: The researcher and research assistant assisted in the review phases. Each person was given either the role of primary reviewer or a secondary reviewer for particular studies, which was randomly decided, using an online randomizer software.

Eligibility Review Rounds Reliability: For each eligibility review round, in order to calculate IRR, two raters reviewed 10% of the studies.

Literature Search and Eligibility Review Rounds Training: The researcher conducted a two-hour training on the literature search methods, in particular how to conduct the literature search for databases, handsearching, and footchasing, how to utilize the pertinent online technologies for the literature search, and what to do if they are unclear on processes/decisions. The second part of the training concered inclusion criteria, each review round phases' criteria, inter-rater reliability (IRR) methods related to inclusion criteria, and training on usage of RefWorks, DropBox, and GoogleDocs for organization of eligibility rounds.

Data Coding Reliability: At this phase of data collection, 10% of the studies were reviewed by two reviewers. During the coding of the studies, if coding disagreements occurred, then discrepancies were resolved through discussion. The way that IRR data was collected for the coding of the data was adapted from Carmago (2012) using a worksheet. The coder referenced the article and the primary evaluator's coding of the article on GoogleDocs to answer the worksheet (e.g., Is this an accurate coding of age?). Disagreements on the worksheet were handled by discussion until consensus was met. The worksheet can be found in Appendix A.

Data Coding Training: Once all of the studies were identified for inclusion in the proposed study, then a second two-hour training was held on how to code variables, and use of GoogleDocs for coding of data. Part of this training included a practice coding session. Specifically, each person coded the same article utilizing a specific set of directions, and then

interrater reliability was calculated and any discrepancies discussed, as well as any questions/concerns were addressed.

Outcome Data Extraction Training: At the training listed above for coding the data, the team members also coded the outcome data from the practice article using DataThief III and questions were addressed.

Outcome Data Extraction Reliability: For raw data that was extracted without DataThief III (2006), 10% of the studies was reviewed by a second coder. When DataTheif III was used then 10% of the studies were randomly selected to be coded by a second rater.

Analyses

To answer the proposed research questions, hierarchical linear modeling was used. There have been multiple studies that provide evidence that hierarchical linear modeling (HLM) is a valid statistical tool to combine and analyze the data among cases in a study and across studies (Moeyaert, Ugille, Ferron, Beretvas, & Van den Noortgate, 2013). The use of hierarchical linear models is a way to summarize the findings of multiple cases examined in the same or several studies. It is important to synthesize the results to understand the generalizability of the findings to see if the same effect will be found across studies and how large of an effect one may expect from a given intervention (Van den Noortgate & Onghena, 2007).

Another advantage of HLM is that it is easy to account for autocorrelation even when there are few observations per case (Van den Noortgate & Onghena, 2008). In other words, HLM can address the fact that measurements closer in time to one another may be more related compared to later measurements in time. In addition, HLM can provide information on linear or nonlinear time trends within phases of the design, and variances within cases, across cases, and across studies (Moeyaert, Ugille, Ferron, Beretvas, & Van den Noortgate, 2015). Given that

these issues are key in single-case designs, HLM is particularly well-suited to synthesize SCD studies.

Standardization of Data

Prior to running the analyses, each DV in a study was standardized per case, since many different scales of measurement were used across studies. There was a focus on analyzing the data from the first phase change or AB transition phases within the same time series. Also another focus was on examining the change in level between phases versus change in trend. The method to do this was proposed by Van den Noortgate and Onghena (2008). Then an ordinary least squares (OLS) regression for each subject from a study was performed separately (i.e., by using Equation 2, described further below), which provided an estimate of the residual within-subject standard deviation ($\hat{\sigma}_{ejk}$). Then the individual score (Y'_{ijk}) was divided by the estimated residual within-subject standard deviation ($\hat{\sigma}_{ejk}$).

$$Y'_{ijk} = \frac{Y_{ijk}}{\hat{\sigma}_{ejk}} \tag{1}$$

By using this method to standardize scores, the scores were not impacted by the size of the treatment effect and therefore the treatment effect estimates were not biased. There were not cases where there was no variability in both baseline and treatment phases. Then the data that was extracted was exported and imported into a data file in Statistics Analysis Software (SAS).

Hierarchical Model to Aggregate the Single-Case Data

After the data were standardized, then the effect sizes were calculated using the hierarchical model proposed by Van den Noortgate and Onghena (2003, 2008). This model has been validated through numerous studies (Ferron, Farmer, & Owens, 2010; Moeyaert et al.,2013; Owens & Ferron, 2012; Shadish, Rindskopf, & Hedges, 2007; Van den Noortgate & Onghena, 2003, 2008).

The use of the restricted maximum likelihood procedure in SAS proc MIXED was utilized to estimate the model parameters (Littell, Milliken, Stroup, Wolfinger, & Schabenberger, 2006). The Satterthwaite method to get an estimate of the degrees of freedom was used (Satterthwaite, 1941). This method was used because it has been found to give accurate confidence intervals for estimates of the average treatment effect for the analysis of two-levels of multiple-baseline data (Ferron, Bell, Hess, Rendina-Gobloff, & Hibbard, 2009).

A four-level HLM was utilized for all outcomes. The four-level structure was as follows: level one measurements were grouped by dependent variable (DV; level 2), which will be grouped within cases (level 3), which will be grouped within studies (level 4).

At the first level of the model, the regression equation shows the within-subject variability (Equation 1). Y_{ijkl} is the observed score on the i^{th} measurement occasion (i = 1,2, . . . I), for the j^{th} DV (j = 0,1, . . . J), for the k^{th} case (k = 0,1, . . . K), and for the l^{th} study (l = 0,1,…L) and was modeled as a function of D, a dummy coded variable that describes if the measurement occasion i from the j^{th} DV, of the k^{th} case, in the l^{th} study is part of the baseline phase (D_{ijkl} = 0) or the treatment phase (D_{ijkl} = 1).

$$Y_{ijkl} = \beta_{0jkl} + \beta_{1jkl}D_{ijkl} + e_{ijkl} \text{ with } e_{ijkl} \sim N(0, \hat{\sigma}^2) \quad (2)$$

The coefficient β_{1jkl} is then interpreted as the immediate effect of the treatment on the j^{th} DV, for the k^{th} case, in the l^{th} study, whereas coefficient β_{0jkl} is the baseline level on the j^{th} DV, for the k^{th} case, in the l^{th} study.

At the second level of the model, the variation across DVs within a case is described using two equations:

$$\begin{aligned} \beta_{0\,jkl} &= \theta_{00kl} + u_{ojkl} \\ \beta_{1\,jkl} &= \theta_{10kl} + u_{1\,jkl} \end{aligned} \qquad \text{with} \quad \begin{bmatrix} u_{0\,jkl} \\ u_{1\,jkl} \end{bmatrix} \sim N(0, \Sigma_u) \tag{3}$$

Overall, these equations show that the β coefficients from Equation 2 equate to a case specific baseline level (θ_{00kl}) with random error to account for variation across DVs, and a case specific treatment effect (θ_{10kl}) with random error to account for variation across DVs.

At the third level, the case specific regression coefficients were modeled as random errors from the study average baseline level (γ_{000l}) and the study average treatment effect (γ_{100l}) as follows:

$$\begin{aligned} \theta_{00kl} &= \gamma_{000l} + \nu_{00kl} \\ \theta_{10kl} &= \gamma_{100l} + \nu_{10kl} \end{aligned} \qquad \text{with} \quad \begin{bmatrix} \nu_{00kl} \\ \nu_{10kl} \end{bmatrix} \sim N(0, \Sigma_\nu) \tag{4}$$

At the fourth level, the study level regression coefficients were modeled as random errors from the overall average baseline level (δ_{0000}) and the overall average treatment effect (δ_{1000}) as follows:

$$\gamma_{000l} = \delta_{0000} + \omega_{000l} \qquad \gamma_{100l} = \delta_{1000} + \omega_{100l} \qquad \text{with} \begin{vmatrix} \omega_{000l} \\ \omega_{100l} \end{vmatrix} \sim N(0, \Sigma_{\omega}) \tag{5}$$

Residuals at each of the four levels were presumed to be multivariate normally distributed (Moeyaert, Ugille, Ferron, Beretvas, & Van den Noortgate, 2015). The δ*'s* are the fixed effects referring to the mean regression coefficients. δ_{1000} represent the overall treatment effect (i.e., the immediate treatment effect averaged across DVs, cases, and studies).

Moderator Analysis

Hierarchical linear modeling provides for an approach to systematically examine moderator variables. The variety of procedures, interventions, and subject characteristics in single-case studies allows for a source of information to identify variables that moderate the effect (Van den Noortgate, & Onghena, 2007). The moderators listed above in the following section, *Categorization of Variables* were analyzed. More specifically, moderator analyses were conducted if there were at least five units at each level of the moderator variable, of which there were for every variable at every level. The moderators were added to the four-level model in order to investigate if they have an impact on the effectiveness of the treatment. They were set as fixed effects to minimize the iterations and add to the reliability in the analysis (Wang, Cui, & Parrila, 2011), and added in at the appropriate level (i.e., case level moderators were added in at level 3, whereas study level moderators were added in at level 4). For example, to examine the

potential moderation of a study characteristic, Y, Equation 5 was altered by adding Y as a predictor:

$$\gamma_{000l} = \delta_{0000} + \delta_{0001} Y + \omega_{000l}$$
$$\gamma_{100l} = \delta_{1000} + \delta_{1001} Y + \omega_{100l} \qquad \text{with} \quad \begin{bmatrix} \omega_{000l} \\ \omega_{100l} \end{bmatrix} \sim N(0, \Sigma_{\omega}) \tag{6}$$

Significance of the Current Study

With regard to the significance of this study for youth with developmental disabilities there may be information gleaned that will further validate and possibly enhance behavioral interventions being used in schools for youth with developmental disabilities. School psychologists and other school-based practitioners may be better informed when developing the behavioral interventions to help with particular youth at multiple levels of tiered services in schools. Furthermore, the particular dependent variable of interest, behavioral outcomes, on youth with developmental disabilities and challenging behaviors is important to study, considering the positive contribution that the lack of challenging behaviors has on youths' school and life success. This study also served to fill a gap in the literature in terms of providing a comprehensive (approximately 20 years were examined for relevant studies) search, including both FBA and non-FBA based interventions, attempted to include studies in various tiers of PBS support, and used parametric statistical analyses.

Chapter 4: Results

This chapter presents the results of the statistical analyses that answer the research questions within the current study. Descriptive analyses are provided first, including the literature search methods descriptives, reasons for study exclusion during data coding, interrater agreement, study characteristics, participant characteristics, and intervention characteristics. Results from the hierarchical linear modeling for the effect of school-based behavioral interventions on youths' with intellectual disabilities behavioral outcomes are presented next. Subsequently, results of the moderator analyses follow. It should be noted that upon consultation with a statistician that no time series that included a second or third intervention type were included, due to the nature of these subsequent interventions, as they built upon the learnings of the first intervention phase to enhance the results. For example, if an intervention only included communication training in the first intervention phase and the researchers did not see a desirable effect then during the second phase they added a visual component to the communication system, if they wanted to enhance the results further in the third phase they then added a reward system. Given that there were less than 5% of the observations that included phases such as those just described, these phases were excluded. Out of 6235 observations, there were 315 that had either a second treatment phase that built on the learnings of the first phase and these were in 7 cases, across 4 studies. Furthermore, only 12 observations had a third treatment phase that was developed as a learning from not only the first intervention phase but the second intervention phase.

Descriptive Statistics

Literature search method descriptives. The literature search resulted in 74 studies that met all of the study inclusion criteria. Table 9 shows that there were numerous studies identified for each search method. The database search method ended up including 119 studies that were included through the final review round. The type of search method that comprised the database search method in reference to Table 9, were a result of adding the database final studies and the duplicate final studies, because RefWorks separated the duplicates into a separate folder, thus that's why the chart is delineated in this manner. The handsearching method located 20 studies that made it into the dataset and the footchasing method located 31 studies. All search methods proved to be useful, as various different studies were located using each search methods. When analyzing the final studies 29 duplicated individual studies were located amongst the various search methods. Furthermore, there were seven studies that were requested through the Interlibrary Loan or the first authors of the studies and these studies were not provided.

There were 59 studies excluded during the data coding stage of the study, which accounted for 34% of the studies that had made it into the final round being excluded. Refer to Table 10 for the detailed exclusion reasons. The main reason for exclusion was due to the study not providing information as to whether the participant had an intellectual disability, although often implied, there was not information that specifically stated this information within 44% of the excluded studies. The second reason, accounting for 22% of the excluded studies, was due to the study being an FBA to analyze the function of the behavior however, not to conduct interventions. The rest of the exclusion reasons included: 12% did not meet criteria due to data concerns (e.g., no baseline data, raw data not available, not enough data points in a phase), 8% of

the studies were not SCD's, 7% the interventions were not conducted in the schools, 5% the dependent variable did not meet inclusion criteria, and 2% of the excluded studies were not included due to not meeting the age inclusion criteria.

Interrater reliability for review rounds and data coding. The IRR score for each search method is shown in Table 11 and the IRR score for the data coding stage and use of the software, DataThief III (2006) is shown in Table 12. IOA for each stage ranged from 87.5 % to 100% with most IOA above 90% (average IOA was 97.64% across all search methods and review rounds), which suggested that it was appropriate to proceed with analyzing the data to determine the effectiveness of school based interventions overall and across different moderators.

Characteristics of the included studies. One hundred and fifty-five participants were included across the 74 studies, there were 424 time series across studies and participants, and 13 specific behavioral outcomes studied across all cases. The information regarding the various variables that were coded for the main and moderator analyses for study characteristics are included in Table 13.

As shown in Table 13, the most frequent type of study design was alternating treatments (27.83%), followed by multiple baseline across subjects (25.71%), and then multiple baseline across settings (16.04%). Combining all of the variants of multiple baseline designs, these type were actually the most frequent type of design at 43.87%. The other three located study designs, including AB, ABAB, and multi-element accounted for approximately 9% each of the type of design. For specific outcome, which described the exact type of behavioral dependent variable, there were many different specific outcomes in the various time series, coded into 13 categories. Pro-social behaviors accounted for 27.83%, followed by drooling/mouthing/spitting at 15%, self-

stimulatory behaviors at 11.79%, off-task behaviors at 11.08% and a combination of challenging behaviors at 9.10%, while the remaining variables were studied in less than 6% of the times series. These included: disruptive behavior, daily living skills, academic behavior, work completion, compliance, aggression along, noncompliance, and happiness. Using the What Works Clearinghouse guidelines for what is considered a good quality SCD, the following categories were established to be coded for the quality of the study design. Please see Appendix B for specific criteria to meet standards. The various categories were, meets standards, meets standards with reservations, and does not meet standards. Of the 424 time series (74 studies), 68.87% met criteria and 21.13% met with reservations. Intervention fidelity was also coded, and all time series included intervention fidelity within the design of the study. Social validity was measured in 49.06% of the time series and was not measured in 50.94%. All of the included time series were published and none were unpublished, please note that if a dissertation was found and then it was published, this researcher used the published version of the study.

The information regarding the frequency of the various variables that were coded for the moderator analyses for participant characteristics are included in Table 14. The age ranges of the participants in the time series were found to be 3.77% in the 3 to 5 year old range, 13.68% in the 6 to 8, 26.18% in the 9 to 11, 16.89% in the 12 to 14, 33.49% in the 15 to 18, and 4.72% in the 19-22 year old range. Only .24% (*n*= 1) did not provide this information, although the participant was clearly a child and 4 time series (.94%) were coded as the participants being between the ages of 6 to 12 years old. In terms of grade range, the frequency of pre-school to pre-kindergarten was (1.65%), elementary aged youth (5.66%), middle school aged youth (.94%), and high school aged youth (.24%); however, 81.37% did not provided this information. There was not enough data to do moderator analyses. Of the time series, 71.93% were conducted with

males and 28.07% with females. The specific disability of the participants in the time series were 43.40% were conducted with youth with an Autism Spectrum Disorder with an Intellectual Disability, and 56.6% on youth with solely an Intellectual Disability. The frequency of cognitive status was found to be 17.69% of times series had participants with a mild cognitive impairment, 21.23% a moderate impairment, 34.43% a severe impairment, and 26.65% indicated that they had an intellectual disability but did not give a specific range of the impairment. In terms of the verbal ability of the participants per time series, 39.39% did not provide this information, while 31.84% were minimally verbal, 15.33% were nonverbal, and then similar frequencies were found for the categories of echolaliac (5.90%), use of sign language or the picture exchange communication system (4.48%), and average language skills were reported in 3.07% of the time series. In terms of the classroom setting of the participants, in 95.08% of the time series, the youth was taught in a special education setting, 3.30% in a combination of special education and general education, and 1.42% in general education.

Table 15 also provides the frequency information of the intervention characteristics by total time series. The majority of the interventions in the time series were considered short (less than 20 data points across intervention phases) at 60.14% and 39.86% were considered to have a long intervention phase. The type of intervention that was found the most frequently was multicomponent at 41.04%, followed by consequence-based at 26.89%, then skills training at 18.16%, and lastly, antecedent-based at 13.92%. In terms of the agent, who conducted the intervention, 80.42% had a typical agent, while 19.58% were conducted by an atypical agent. For the setting of the interventions, 71.93% were in a typical setting, while 28.07% were in an atypical. The large majority, 98.82% were conducted in an individual setting, while 1.19% in a small group setting. Generalization of treatment effects were also coded, and 62.97% did not

collect this information, while 28.54% included generalization data for a new situation or setting, while 5.19% collected data on a combination of situation, setting, or behaviors, lastly, 3.30% collected this information utilizing a new agent only. Data was collected on the Tier level of support that the intervention was implemented, however all 424 (100% of the time series were implemented at the Tier 3 level. Data was collected on the presence or absence of a FBA and it was almost equal in that 43.16% did conduct an FBA, while 42.69% did not, and 14.15% used data from a FBA conducted shortly before the study of the time series occurred. For the following remaining variables, 242 or 57.08% of the time series did not have this information, because the data was only applicable if an FBA was conducted. For the FBA method, 22.64% had a combination of experimental and descriptive FBA methods, while 16.04% used experimental methods only and 4.25% used descriptive methods only. The agent of the FBA was more frequently coded as typical at 24.29%, and atypical was found in 18.87%. The FBA setting was found to be in a typical setting 29.48% of the time series and 13.92% of the time in atypical settings. A team decision was utilized in determining the function of the behavior in 2.59% of the time series, and not in 31.37%. In sum, there was much variability among the case, study, and participant characteristics.

Inferential Analyses

There were 74 included studies in the data analysis and multiple time series per study for a total of $N = 424$ time series, and a total of $N = 155$ cases. After coding the data points of each dependent variable, the data were transformed into standardized scores. It was observed that the hierarchical linear model had four levels. These levels included observations nested within specific outcomes, nested within cases, nested within studies. For the analysis of the effect of school based behavioral interventions on the behavioral outcomes of youth with an

intellectual disability, a total of 6235 individual observations were nested within the data set and less than 6% (*n*= 371) could not be used leaving a total of 5864 observations in the analyses.

Behavioral outcomes. The four-level hierarchical linear model without moderators is presented in Table 16. This analysis shows that on average school based behavioral interventions are significantly effective in comparison to the baseline conditions for changing youths' behavioral outcomes. Specifically it was found that the level of desirable behaviors is 3.31 (95% CL 3.21 to 3.41, p = < 0.001) standard deviations higher in the treatment conditions, which is statistically significant. Looking at the covariance parameter estimates in Table 17, the intervention effects vary significantly over time series within a case (i.e., across the multiple dependent variables within a case), with an estimated variance of 0.98, $Z = 7.19$, $p = <.0001$. The intervention effect did not vary significantly for the cases, with an estimated variance of 0, and they vary significantly over the studies, with an estimated variance of .69, $Z = 3.65$ $p = .0001$. The residual within participants' variance is .9998, which means the standard deviation within a time series is about 1.0, which was expected because the data had been standardized within time series. Calculating the Interclass Correlation, then this indicates that at baseline 26% of the variation in behavioral outcomes exists between studies, 0% between cases, 37% exists between specific dependent variables. Leaving 37% of the variance in behavioral outcomes existing within time series.

Moderator analyses for behavioral outcomes. In order to examine the research questions related to which variables moderate the relationship between the effect of school-based behavioral interventions on the behavioral outcomes of youth with an intellectual disability refer to Table 18. Table 18 shows a statistically significant moderating effect of the variable *Type of Classroom* ($F (2, 5488) = 421.97$, $p = <0001$). The specific type of classroom that had a

moderating effect was special education only classrooms in comparison to the reference group, general education classrooms (t (5488) = 1.97, *p* = .0493). It was found that there was a statistically significant difference, with interventions conducted in special education classrooms having the larger effect, however please note that there were only six time series coded for general education and 404 for special education. Each level of a moderator only required at least five time series, so this variable did meet the criteria. The type of intervention in comparison to the reference group seemed to moderate the effect (F, (3, 5490)= 2.81, p = .0382, however It was not found to be significant when comparing the reference group to the specific groups, with the largest difference being greater than > .05, please refer to Table 18. None of the other study, intervention, or participant variables moderated the main effect analyzed. The variable, SWBSPBS was not able to be analyzed due to all time series being conducted at the Tier 3 level. Note that for the variable, type of SCD, there was enough data to calculate the moderating effects, however, in order to do so the various types of multiple baseline designs were combined into one category due to low amounts of time series in some of these categories.

Table 9

Literature Search Methods Descriptives

Eligibility Review Round	Database	Duplicated	Handsearch	Footchasing	Other	Total	
Initial	8859	466	12	26	0		
Round 1	602	257	10	26	0		
Round 2	373	178	5	26	0		
Round 3	138	62	4	21	0		
Round 4	76	5	0	16	0		
Round 5	61	4	0	16	0		
Round 6	61	58	26	16	0	0	17
Duplicates							29
Excluded During Data Coding							59
Interlibrary Loan Did not Locate							7
Total Studies							74

Table 10

Reasons for Study Exclusion During Coding of Data

Author	Study Title	Reason for Exclusion
Agosta (2004)	"Treatment of Self-Injurious Behavior through Overcorrection Procedures"	Did not specify Intellectual Disability (IND)
Asmus, Wacher, Harding, Berg, Derby, & Kocis (2013)	"Evaluation of Antecedent Stimulus Parameters for the Treatment of Escape-Maintained Aberrant Behavior"	Interventions not conducted in schools

Table 10 Cont'd

Author	Study Title	Reason for Exclusion
Butler (2009)	"Wetting and Soiling"	Did not specify Intellectual Disability (IND)
Camp, Iwata, Hammond, & Bloom (2009)	"Antecedent versus Consequent Events as Predictors of Problem Behavior"	Not a Single Case Design (SCD) for purpose of treatment, but for an experimental functional analysis
Carbone, Morgenstern, Zecchin-Tirri, & Kolberg (2010)	"The Role of the Reflexive-conditioned Motivating Operation (CMO-R) During Discrete Trial Instruction of Children with Autism"	Did not specify Intellectual Disability (IND)
Carison, Luiselli, Slyman, & Markowski (2008)	"Choice-Making as Intervention for Public Disrobing in Children with Developmental Disabilities"	Did not specify Intellectual Disability (IND)
Carnahan, Musti-Rao, & Bailey (2009)	"Promoting Active Engagement in Small Group Learning Experiences for Students with Autism and Significant Learning Needs"	Did not specify Intellectual Disability (IND)
Coleman & Holmes (1998)	"The Use of Noncontingent Escape to Reduce Disruptive Behaviors in Children with Speech Delays"	Did not specify Intellectual Disability (IND)
Cooper (2014)	"Response to Interventions (RtI): A Mixed Methods Study Evaluating the Effects of Behavior Training Software on Behavior of In-School Suspension Students"	Study not a SCD

Table 10 Cont'd

Author	Study Title	Reason for Exclusion
Dewein & Miller (2008)	"The Effect of a Teacher Report on the Sustainability of an Intervention to Facilitate Engagement by a Child with Developmental Delays"	Did not specify Intellectual Disability (IND)
Didde, Prinsen, & Sigafoos (2000)	"The Blocking Effect of Pictorial Prompts on Sight-Word Reading"	Dependent variable not related to current study inclusion criteria
Downs, Downs, Johansen, & Fossum (2007)	"Using Discrete Trial Teaching within a Public Preschool Program to Facilitate Skill Development in Students with Developmental Disabilities"	Not a SCD
Ellingston, Miltenberger, & Long (1999)	"A Survey of the Use of Functional Assessment Procedures in Agencies Serving Individuals with Developmental Disabilities"	Not a SCD
Ganz, Bourgeois, Flores, & Campos (2008)	"Implementing Visually Cued Imitation Training with Children with Autism Spectrum Disorders and Developmental Delays"	Did not specify Intellectual Disability (IND)
Garfinkle & Schwartz (2002)	"Peer Imitation: Increasing Social Interactions in Children with Autism and Other Developmental Disabilities in Inclusive Preschool Classrooms"	Did not specify Intellectual Disability (IND)

Table 10 Continued

Author	Study Title	Reason for Exclusion
Garbutt & Furniss (2007)	"Context Sampling Descriptive Assessment: A Pilot Study of a Further Approach to Functional Assessment"	Not a SCD
Heinicke, Carr, Mozzoni, & Roane (2009)	"Using Differential Reinforcement to Decrease Academic Response Latencies of an Adolescent with Acquired Brain Injury"	Did not specify Intellectual Disability (IND)
Hetzroni (2004)	"Effects of a Computer-Based Intervention Program on the Communicative Functions of Children with Autism"	Not a SCD
Howell, Rueda, & Rutherford (1983)	"A Procedure for Teaching Self-Recording to Moderately Retarded Students"	Data not reported like SCD: median and ranges given instead
Kee, Hill, Weist (1999)	"School-Based Behavior Management on Cursing, Hitting, and Spitting in a Girl with Profound Retardation"	Not enough data points in phases to meet inclusion criteria
Kelley, Shillingsburg, Castro, Addison, & LaRue (2007)	"Further Evaluation of Emerging Speech in Children with Developmental Disabilities: Training Verbal Behavior"	Did not specify Intellectual Disability (IND)
Kennedy & Meyer (1996)	"Sleep Deprivation, Allergy Symptoms, and Negatively Reinforced Problem Behavior"	Not a SCD
Kern, Childs, Dunlap, Clarke, & Falk (1994)	"Using Assessment-Based Curricular Intervention to Improve the Classroom Behavior of a Student with Emotional and Behavioral Challenges"	Did not specify Intellectual Disability (IND)
Lalli, Livezey, & Kates (1996)	"Functional Analysis and Treatment of Eye Poking with Response Blocking"	Did not specify Intellectual Disability (IND)

Table 10 Continued

Author	Study Title	Reason for Exclusion
Lane, Harris, Graham, Weisenbach, Brindle, & Morphy (2008)	"The Effects of Self-Regulated Strategy Development on the Writing Performance of Second-Grade Students with Behavioral and Writing Difficulties"	Did not specify Intellectual Disability (IND)
Lang, O'Reilly, Machalicek, Lancioni, Rispoli, & Chan (2008)	"A Preliminary Comparison of Functional Behavior Results when Conducted in Contrived versus Natural Settings"	Did not specify Intellectual Disability (IND)
Magee & Ellis (2001)	"The Detrimental Effects of Physical Restraint as a Consequence for Inappropriate Classroom Behavior"	No baseline data
Marcus & Vollmer (1995)	"Effects of Differential Negative Reinforcement on Disruption and Compliance"	Purpose of study did not align with the current study's purpose
Martens & Houk (1989)	"The Application of Herrnstein's Law of Effect to Disruptive and On-task Behavior of a Retarded Adolescent Girl"	Only an FBA
May & Howe (2013)	"Evaluating Competing Reinforcement Contingencies on Off-task Behavior in a Preschooler with Intellectual Disability: A Data-Based Case Study"	Did not specify Intellectual Disability (IND)
McComas, Hoch, Paone, & El-Roy (2000)	"Escape Behavior During Academic Tasks: A Preliminary Analysis of Idiosyncratic Establishing Operations"	Did not specify Intellectual Disability (IND)
Mcentee & Saunders (1997)	"A Response-Restriction Analysis of Stereotypy in Adolescents with Mental Retardation: Implications for Applied Behavior Analysis"	No Baseline data

Table 10 Continued

Author	Study Title	Reason for Exclusion
Mueller, Wilcynzski, Moore, Fusilier, & Trahant (2001)	"Antecedent Manipulations in a Tangible Condition: Effects of Stimulus Preference on Aggression"	Only an FBA
Mullins & Christian (2001)	"The Effects of Progressive Relaxation Training on the Disruptive Behavior of a Boy with Autism"	Interventions not conducted in schools
Napolitano, Smith, Zarcone, Goodkin, & McAdam (2010)	"Increasing Response Diversity in Children with Autism"	Dependent variable not related to current study inclusion criteria
Nikopoulas, Canavan, & Nikopoulou-Smyrni (2009)	"Generalized Effects of Video Modeling on Establishing Instructional Stimulus Control in Children with Autism: Results of a Preliminary Study"	Did not specify Intellectual Disability (IND)
Northup, Wacker, Berg, Kelly, Sasso & DeRaad (1994)	"The Treatment of Severe Behavior Problems in School Settings Using a Technical Assistance Model"	Did not specify Intellectual Disability (IND)
Parry-Cruwyes, Neal, Ahern, Wheeler, Permchander, Lobe, & Dube (2011)	"Resistance to Disruption in a Classroom Setting"	Did not specify Intellectual Disability (IND)
Peters-Schiffer, Didden, Mulders, & Korzilius (2010)	"Low Intensity Behavioral Treatment Supplementing Preschool Services for Young Children with Autism Spectrum Disorders and Severe to Mild Intellectual disability"	Not a SCD
Plavnick & Ferreri, (2011)	"Establishing Verbal Repertoires in Children with Autism Using Function-Based Video Modeling"	Did not specify Intellectual Disability (IND)
Potoczak, Carr, & Michael (2007)	"The Effects of Consequence Manipulation During Functional Analysis of Problem Behavior Maintained by Negative Reinforcement"	Purpose of study did not align with the current study's purpose

Table 10 Continued

Author	Study Title	Reason for Exclusion
Rispoli, Davis, Goodwyn, & Carmago (2013)	"The Use of Trial-Based Functional Analysis in Public School Classrooms for Two Students With Developmental Disabilities"	Only an FBA
Querim, Iwata, Roscoe, Schlichenmeyer, Ortega, & Hurl (2013)	"Functional Analysis Screening for Problem Behavior Maintained by Automatic Reinforcement"	Only an FBA
Robertson, Simon, Pachman, & Drabman (1979)	"Self Control and Generalization Procedures in a Classroom of Disruptive Retarded Children"	Data not reported for interpretation and author did not respond (all subjects data combined)
Sarakoff, Taylor, & Poulson (2001)	"Teaching Children with Autism to Engage in Conversational Exchanges: Script Fading with Embedded Textual Stimuli"	Interventions not conducted in schools
Sasso, Reimers, Cooper, Wacker, Berg, Steege, Kelly, & Allaire (1992)	"Use of Descriptive and Experimental Analyses to Identify the Functional Properties of Aberrant Behavior in School Setting"	Did not specify Intellectual Disability (IND)
Spitalnik & Drabman (1976)	"A Classroom Timeout Procedure for Retarded Children"	No baseline data
Taylor, Sisson, McKlivey, & Trefelner (1993)	"Situation Specificity in Attention-Seeking Problem Behavior-A Case Study"	Only an FBA

Table 10 Cont'd

Author	Study Title	Reason for Exclusion
Taylor & Romancyzk (1994)	"Generating Hypotheses about the Function of Student Problem Behavior by Observing Teacher Behavior"	Only an FBA
Thiemann & Goldstein (2001)	"Social Stories Written Text Cues, and Video Feedback: Effects on Social Communication of Children with Autism"	Did not specify Intellectual Disability (IND)
Tomlin & Reed (2012)	"Effects of Fixed-Time Reinforcement Delivered by Teachers for Reducing Problem Behavior in Special Education Classrooms"	Did not specify Intellectual Disability (IND)
Van Houton & Rolider (1988)	"Recreating the Scene: An Effective Way to Provide Delayed Punishment for Inappropriate Motor Behavior"	Interventions not conducted in schools
Vaughn, Clark, & Dunlap (1997)	"Assessment-Based Intervention for Severe Behavior Problems in a Natural Family Context"	No baseline data
Venn, Wolery, & Greco (1996)	"Effects of Every-Day and Every-Other-Day Instruction"	Did not specify Intellectual Disability (IND)
Vollmer & Northup (1996)	"Some Current Themes in Functional Analysis Research"	Only an FBA

Table 10 Cont'd

Author	Study Title	Reason for Exclusion
Vollmer, Marcus, & Ringdahl (1995)	"Progressing from Brief Assessments to Extended Experimental Analyses in the Evaluation of Aberrant Behavior"	Only an FBA
Vaughn, Clark, & Dunlap (1997)	"Assessment-Based Intervention for Severe Behavior Problems in a Natural Family Context"	No baseline data

Table 11

Interrater Reliability Calculations Per Review Round

Review Round	**Database IRR**	**Footchasing IRR**	**Handsearching IRR**	**Duplicates in Database**	**Average IRR**
1	97.7%	88%	87.5%	96%	92.8%
2	97.8%	100%	93.8%	98.7%	97.03%
3	95%	100%	100%	93.3%	97%
4	92.9%	100%	100%	90.9%	96.76%
5	100%	100%	100%	100%	100%
6	100%	100%	100%	100%	100%
7	100%	100%	100%	100%	100%
					97.64%

* IRR is an abbreviation for interrater reliability

Table 12

Interrater Reliability Calculations During Data Coding

	Percent of Studies Calculated	**IRR**
Data Coding	10%	97%
Graphs	50%	93%

Table 13

Study Characteristics Frequency by Time Series

Study Characteristic	**Total Number of Time Series**	**% For each subcategory**	**Enough Data For Analyses**
Type of SCD			Yes
Alternating Treatments	118	27.83%	
AB	41	9.67%	
ABAB	39	9.20%	
Multiple Baseline Across Subjects	109	25.71	
Multielement	40	9.43%	
Multiple Baseline with Reversal	5	1.18%	
Multiple Baseline Across Settings	68	16.04%	
Multiple Baseline Across Behaviors	2	0.47%	
Multiple Baseline Across Tasks	2	0.47%	
Quality of SCD			
Meets	292	68.87%	
Meets with Reservations	132	31.13%	

Table 13 Cont'd

Study Characteristic	Total Number of Time Series	% For each subcategory	Enough Data For Analyses
Intervention Fidelity			
Yes	424	100%	Yes
Social Validity			
Yes	208	49.06%	
No	216	50.94%	
Publication Status			
Yes	424	100%	
Specific Outcome			
Pro-Social Behaviors (appropriate touching, communication, obeying, waiting)	118	27.83%	
Drool/Mouthing/Spitting	65	15.33%	
Self Stimulatory	50	11.79%	
Off Task	47	11.08%	
Challenging Behaviors (self injury, aggression, tantrum)	39	9.20%	Yes
Disruptive Behaviors (out of seat, talking out, throwing)	23	5.42%	
Self Injury	15	3.54%	
Daily Living Skills	15	3.54%	
Academic Achievement	10	2.36%	
Work Completion	10	2.36%	
Compliance to Teacher	9	2.12%	
Aggression Towards Other	8	1.89%	
Noncompliance	7	1.65%	
Happiness	4	0.94%	

Table 14

Participant Characteristics Frequency by Time Series

Variable	Total Number of Time Series	% For each subcategory	Enough Data For Analyses
Grade Range			Yes (not High School)
Preschool to Pre-K	7	1.65%	
Elementary	24	5.66%	
Middle School	4	0.94%	
High School	1	0.24%	
Not Provided	345	81.37%	
Age Range (years old)			Yes
3 to 5	16	3.77%	
6 to 8	58	13.68%	
9 to 11	111	26.18%	
12 to 14	72	16.98%	
15 to 18	142	33.49%	
19-22	20	4.72%	
Not Provided	1	.24%	
6 to 12	4	.94%	
Gender			Yes
Male	305	71.93%	
Female	119	28.07%	
Specific Disability			Yes
Autism Spectrum Disorder with an Intellectual Disability	184 240	43.40% 56.60%	
Cognitive Status			Yes
Mild	75	17.69%	
Moderate	90	21.23%	
Severe	146	34.43%	
IND, no IQ	113	26.65%	

Table 14 Cont'd

Variable	Number of Time Series	Frequency Percentage	Enough Data for Analyses (> 5 per category)
Verbal Ability			
Nonverbal	65	15.33%	
Minimally Verbal	135	31.84%	
Echolaliac	25	5.90%	
Average Language Skills	13	3.07%	
Sign Language/PEC	19	4.48%	
Not provided	167	39.39%	
Classroom			
General Education	6	1.42%	
Special Education	404	95.08%	
Combination	14	3.30%	

Table 15

Intervention Characteristics Frequency by Time Series

Variable	Number of Time Series	Frequency Percentage	Enough Data for Analyses (> 5 per category)
Duration			Yes
Long	169	39.86%	
Short	255	60.14%	
Type of Intervention			Yes
Skills Training	77	18.16%	
Antecedent-Based	59	13.92%	
Consequence-Based	114	26.89%	
Multicomponent	174	41.04%	
Agent			
Typical	341	80.42%	Yes
Atypical	83	19.58%	
Setting			Yes
Typical	305	71.93%	
Atypical	119	28.07%	

Table 15 Continued

Variable	Number of Time Series	Frequency Percentage	Enough Data?
Format			Yes
Individual	419	98.82%	
Small Group	5	1.18%	
Generalization			Yes
New Situation or Setting Only	121	28.54%	
Combination of Situation,Setting, or Behavior	22	5.19%	
New Person Only	14	3.30%	
No Generalization	267		
SWPBS Tier			Yes
Tier 3	424		
Presence of FBA			Yes
No	181		
Yes	183		
Prior FBA	60		
FBA Method			
Experimental Only	68		Yes
Descriptive Only	18		
Combination	96		
No FBA	242		
FBA Agent			
Typical	103		Yes
Atypical	80		
Not Applicable	242		
FBA Setting			
Typical	125		Yes
Atypical	59		
Not Applicable	242		
Team Decision			
Yes	11		Yes
No	133		
Not Reported	38		
Not Applicable	242		

Table 16

Results of the 4-level HLM Final Estimation of Fixed Effects for Behavioral Outcomes

Fixed effect	Coefficient	SE	*T*-Value	Approx. d.f.	*p*-Value
Intercept	1.0774	.1290	8.35	39	<.001
Tx	3.3092	.3038	10.89	39	<.001

Table 17

Covariance Parameter Estimates

Variance Parameter Estimates				
Parameter	Estimate	SE	Z	*p*-value
Variance in Treatment Effects				
Between Time Series	7.4553			
Between Cases	0			
Between Studies	3.1209			
Variance in Baseline Levels				
Between Time Series	.9791			
Between Cases	0			
Between Studies	.6895			
Variance Within Time Series	.9998			

Table 18

Moderator Effects Statistics on the Effect of Behavioral Interventions on Youths' Behavioral Outcomes

Study Characteristics					
Type of SCD					
Moderator	**Estimate**	**Standard Error**	**DF**	***t* Value**	**Pr > |t|**
Alternating Treatments	1.9166	1.2350	5492	1.55	.1207
AB	.7463	1.2470	5492	.60	.5496
ABAB	.9218	1.2378	5492	.74	.4565
Multiple Baseline Across Subjects	.9251	1.9645	5492	.47	.6377
Multielement	.04187	1.2629	5492	.03	.9736
Multiple Baseline with Reversal	3.3034	2.0560	5492	1.61	.1082
Multiple Baseline Across Settings, Behaviors, Tasks	0	-	-	-	-
Quality of SCD					
Moderator	**Estimate**	**Standard Error**	**DF**	***t* Value**	**Pr > |t|**
Meets	.2660	.5378	5491	.49	.6209
Meets with Reservations	0	-	-	-	-
Generalization					
Moderator	**Estimate**	**Standard Error**	**DF**	***t* Value**	**Pr > |t|**
New Situation or Setting Only	2.0773	.8669	5492	2.40	.0166
Combination of Situation, Setting, or Behavior	1.2355	1.1799	5492	1.05	.2951
New Person Only	1.4160	1.1167	5492	1.27	.2049

No Generalization	0	-	-	-	-

Presence FBA

Moderator	Estimate	Standard Error	DF	*t* Value	Pr > |t|
No	-.3727	1.0315	5492	-.36	.7179
Yes	.1376	1.0281	5492	.13	.8935
Prior FBA	0	-	-	-	-

FBA Method

Moderator	Estimate	Standard Error	DF	*t* Value	Pr > |t|
Experimental Only	-2.6712	1.2144	5492	-2.20	.0279
Descriptive Only	-.8901	1.0426	5492	-.85	.3933
Combination	-1.2166	.7013	5492	-1.73	.0828
No FBA	0	-	-	-	-

FBA Setting

Moderator	Estimate	Standard Error	DF	*t* Value	Pr > |t|
Typical	.4904	.6153	5492	.80	.4254
Atypical	-1.1691	1.8121	5492	-.65	.5189
Not Applicable	0	-	-	-	-

FBA Agent

Moderator	Estimate	Standard Error	DF	*t* Value	Pr > |t|
Typical	.8575	.6441	5492	1.33	.1831
Atypical	-.9789	.9777	5492	-1.00	.3167
Not Applicable	0	-	-	-	-

Social Validity

Moderator	Estimate	Standard Error	DF	*t* Value	Pr > |t|
Yes	1.0324	.6222	5492	1.66	.0971
No	0	-	-	-	-

Specific Outcome

Moderator	Estimate	Standard Error	DF	*t* Value	Pr > |t|
Academic Achievement	2.8531	1.8135	34	1.57	.1249
Pro-Social Behavior-appropriate touching, communication, appropriate waiting	2.5363	1.4935	34	1.70	.0986
Drooling/Mouthing/Spitting	-1.2802	2.0470	34	-.63	.5359
Challenging Behavior- self injury, aggression, tantrums	-.1480	1.8497	34	-.08	.9367
Self Injury Alone	-.4906	1.8175	34	-.27	.7889
Academic Achievement	.2737	1.8023	34	.15	.8802
Work Completion	-.4320	1.5082	34	-.29	.7763
Compliance	.09532	1.5032	34	.06	.9498

Aggression Alone	.7361	1.6284	34	.45	.6541
Self Stimulatory	-.3811	1.6360	34	-.23	.8172
Noncompliance	1.0874	1.8689	34	.58	.5645
Other- pica, incontinence, behaviors during toileting	-.1186	2.0401	34	-.06	.9540
Disruptiveness- out of seat, talking out, throwing	.3205	2.1357	34	.15	.8816
Happiness	.4303	2.7906	34	.15	.8784
Off Task	1.2150	2.0444	34	.59	.5562
Daily Living Skills	0	-	-	-	-

Participant Characteristics

Grade Range

Moderator	Estimate	Standard Error	DF	*t* Value	Pr > |t|
Preschool to Pre-K	-.9589	1.5587	5492	-.62	.5385
Elementary	2.5056	.8490	5492	2.95	.0032
Middle	.1262	1.6868	5492	.07	.9404
High School	.8684	3.1944	5492	.27	.7858
Not Provided	0	-	-	-	-

Age Range (years old)

3 to 5	3.8585	2.4049	5492	1.60	.1087
6 to 8	2.5382	2.2253	5492	1.14	.2541
9 to 11	2.0303	2.2222	5492	.91	.3609
12 to 14	1.5883	2.2688	5492	.70	.4839
15 to 18	2.0253	2.2544	5492	.90	.3690

19 to 22	3.0036	2.7359	5492	1.10	.2723
Not Provided	4.0396	3.3752	5492	1.20	.2314
6 to 12	0	-	-	-	-
Gender					
Female	.07356	.4968	5492	.15	.8823
Male	0	-	-	-	-
Specific Disability					
Intellectual Disability	.4094	.5707	5492	.72	.4732
Autism Spectrum Disorder & Intellectual Disability	0	-	-	-	-
Verbal Ability					
Nonverbal	1.9113	1.1613	5492	1.65	.0999
Minimally Verbal	-.8910	1.0778	5492	-.83	.4085
Echolaliac	6.4074	3.2032	5492	2.00	.0455
Average Language Skills	.4672	.9558	5492	.49	.6250
Sign Language/Pec	0	-	-	-	-
Cognitive Status					

Mild	-.1743	1.0514	5492	-.17	.8683
Moderate	.4799	.9654	5492	.50	.6191
Severe	.2689	.9340	5492	.29	.7734
IND, no IQ	0	-	-	-	

Intervention Characteristics

Setting

Moderator	**Estimate**	**Standard Error**	**DF**	***t* Value**	**Pr > |t|**
Typical	1.6893	1.0308	5492	1.64	.1013
Atypical	0	-	-	-	-

Duration

Long	.7869	.6703	5492	1.17	.2405
Short	0	-	-	-	-

Type of Intervention

Skills Training	-0.1661	1.7493	866	-0.09	0.9244
Antecedent-Based	-.03797	.6962	5490	-.05	.9565
Consequence-Based	-.7207	.8801	5490	-.82	.4129
Multicomponent	0	-	-	-	-

Classroom

General Education	.9257	.4708	5488	1.97	.0493
Special Education	-2.3076	.4858	5488	-4.75	<.0001

	Combination	0	-	-	-	-
Agent						
	Typical	.1136	.8732	5492	.13	.8965
	Atypical	0	-	-	-	-
Unit						
	Individual	0.07669	1.5913	865	0.05	0.9616
	Small Group (< 10)	0.9986	1.5997	865	0.62	0.5326
	Large Group (> 30)	0	-	-	-	-

*p < 0.005 (statistically significant effect)

Chapter 5: Discussion

The current study investigated the effects of school-based behavioral interventions on youths' behavioral outcomes by conducting a meta-analysis of single-case design studies for a 20-year timeframe from 1997 to 2017. Comprehensive search methods were utilized to locate single-case design studies that met inclusion criteria. The primary purpose of this study was to understand the effect that school-based behavioral interventions have on youth with an intellectual disability', behavioral outcomes by synthesizing the results of single case design studies. The importance of synthesizing these types of designs is highlighted by the fact that usually SCDs are conducted on low-incidence populations and by combining the effects of many studies this gives an overall effect size for the research that meets the inclusion criteria. Another purpose of this study was to conduct a comprehensive analysis of any moderating effects of study, intervention, or participant characteristics to help guide school-based practitioners in the use of behavioral interventions as an intervention to help promote desirable outcomes with this particular population of youth. This chapter summarizes the results of the current study, relates these findings to existing literature, discusses alternative explanations for the results and limitations of this research, and suggests implications for practice, policy, and for research.

Descriptive Analyses

There were 74 studies that met inclusion criteria, and 424 time series across all of the studies, giving an adequate sample size to conduct the meta-analysis. It is important to note for future researchers who wish to synthesize the results of studies, that although there may be a

fewer number of participants in SCD studies in a body of literature than group design studies, one SCD study often has multiple time series to synthesize. An often cited limitation of SCDs are that they may not be as reliable as group design studies since the external validity is low, but by synthesizing the results of multiple SCDs this helps to generalize the results (Riley-Tillman & Burns, 2009). It is important to synthesize SCDs for this particular body of literature, as all of the studies were conducted on youth with a clinically diagnosable disorder of an intellectual disability and some youth had both an Autism Spectrum Disorder and an intellectual disability. This study serves to provide pertinent information regarding the effect of school-based behavioral interventions for lower incidence populations of youth, whom have a high rate of experiencing challenging behaviors.

The current study used multiple types of search methods to locate studies for the meta-analysis and from analyzing these different methods there are important findings to discuss. It was found that handsearching accounted for 15% of the total studies located and foot-chasing accounted for 9% of the total studies, before duplicates and any studies were excluded for not meeting inclusion criteria. The database search method provided 76% of the studies. Please note that the database studies and deleted studies are all from the database search, however RefWorks deleted all of the duplicates and put them into a separate folder, and this researcher then had to review those studies separately from the other database studies. This finding highlights the importance of having multiple methods of searching the literature when conducting a meta-analysis. Another finding gleaned from analyzing the search methods was that 36% of the studies were excluded when undergoing the data coding phase, with the most common reason being due to not meeting criteria that specifically states that the participant has an intellectual disability as reported (review round 6 criteria), at 45% of the overall exclusion reason. Upon

further analysis, it was hypothesized that this large percent of studies (*n*= 26), were included initially during the 6th review round because in most studies it was implied that the participants had an intellectual disability but it was not explicitly stated, and the term developmental disability was commonly used. The next most common reason for exclusion was 17% of the studies were FBA's, so the study presented as SCDs, however, the purpose of the SCD was to find out the function of a behavior. The third most common reason was due to issues with the data (12%), for example, not having enough data in the phases (review round 5 criteria), or it was reported in a way that was unusable and data was not able to be obtained in another format. Then there were three reasons that accounted for approximately 9% each of the exclusion reasons. These reasons included, the intervention not being conducted in the schools, the variables were not related to the purpose of the current study, and the study was not a SCD. Additionally, there were 17% of the studies duplicated across the various search methods. So out of the 170 final studies, there were actually 74 studies viable for study inclusion.

Study characteristics descriptive findings. A descriptive statistic related to study characteristics to highlight was that there were a variety of behavioral outcomes studied, in fact 13 different types. Initially there were 26 types, however to be able to run moderator analyses some had to be combined and construct-wise it made sense to do so, for example one code was for communication, however "prosocial behavior" also included communication, so these were combined. Pro-social behaviors, included appropriate touching, communication, listening to directives, and waiting calmly. This category accounted for the highest frequency of behaviors at 27.83%. This is very uplifting, as behaviorists are taught to teach replacement behaviors that are worded positively to replace undesirable behaviors. It is best practice not just to work on reducing an undesirable behavior, but to replace it with a behavior that enhances functioning and

matches the same function the undesirable behavior was serving. A limitation of this study is that it would be useful to recode all this data into individual pro-social behaviors rather than have them grouped all together. Many reductive behaviors (behaviors wished to be reduced) were coded separately and the most common reduction behavior coded was drooling/mouthing/spitting at 15.33%, followed by self-stimulatory behaviors 11.79%, off-task behaviors 11.08%, challenging behaviors (self-injury, aggression, tantrum combinations) 9.20%, and disruptive classroom behaviors (out of seat, talking out, throwing items combinations) at 5.42%. The remaining coded behaviors were all under 5% and included, self-injury alone, daily living skills (also a pro-social behavior), academic achievement, (pro-social) work completion (pro-social), compliance (pro-social), aggression alone, noncompliance, and happiness (prosocial). However, if you think of it as pro-social or acquisition behaviors versus behaviors to reduce or reduction behaviors, that means these accounted for 34.13% of the data collected while data on reducing challenging behaviors accounted for 65.77%. In one of the most relevant meta-analyses, Gresham et al. (2004) they reported the outcomes in categories of specific outcomes as well and included, academic behavior, academic related behaviors, social behaviors, disruptive behaviors, stereotypies/destructive behavior, daily living skills, eating, combined, and other. In retrospect, it may have been helpful to code the outcomes in these same categories as much as possible.

Another study characteristic to highlight is that all of the studies met the criteria of being a good quality SCD: Meets with Reservation (31%), while 69% meet the criteria Meets, as described in Chapter 4 and in Appendix B. This may be attributed to the fact that 100% of the studies were published, and possibly in order to be published journals are utilizing the criteria established to determine the quality of SCD design. Furthermore, it was found that a variety of

SCD types were included, 44% were multiple baseline designs, 28% were alternating treatments, and 9% each were AB designs, ABAB, and multi-element. This data was not reported in the similar extant meta-analyses (Goh & Bambara, 2010; Gresham et al., 2004), and this data adds to this body of research.

Lastly, it was encouraging that 100% of the studies included intervention fidelity as part of their design, and that 50% of studies included a social validity measure. This is best practice when doing behavioral interventions. This was not examined in Gresham et al. (2004), however Goh and Bambara (2010) did collect data on social validity and found a less frequent rate of inclusion of this information, finding social validity was collected in 39% of the data.

In sum, the main descriptive discussion points for study characteristics were derived from the finding that it was important to include various methods for searching the literature, that upon further analysis there were a large percentage of final studies that were excluded for various reasons, that the included studies met criteria for being quality designs by established standards, and that the current study is a novel meta-analysis in that it included descriptive data on a few variables that had yet to be examined in the most similar meta-analyses (Goh & Bambara, 2010; Gresham et al., 2004), including the type of SCD, the quality of the SCD, and intervention fidelity.

Intervention characteristics descriptive findings. Findings from descriptive analyses of the intervention characteristics indicate that the majority of the interventions were conducted for short duration (60.14%), meaning there were less than 20 data points across all treatment and the remaining portion (39.86%) were conducted for a long duration. Similarly, Goh and Bambara (2010) found that 72% of the studies have a short length of treatment. Another finding related to intervention characteristics was that the person that implemented the intervention, or

agent, was mainly a person in the participant's environment like the teacher, (80.42%), while 19.58% were researchers. In line with this, it was found that 71.93% of the time series were conducted in a typical setting, such as the classroom while 28.07% were conducted in a pull out classroom or other atypical setting. This indicates that the data is generalizable, as teachers mainly implemented the interventions in classrooms, which is what would most likely happen if the intervention was not part of a research study. Goh and Bambara (2010) included this information as well and found in line with the current study, that the large majority of interventions were conducted by typical agents (81%) and typical settings (81%).

It was thought that a unique addition the current meta-analysis would add to the research was examining the interventions from a SWBPBS perspective, and it was unique, however, 100% were conducted at the Tier 3 level, and in line with this 98.82% at the individual level (not in a group), one participant at a time. Before a behavioral intervention is implemented it is best practice to conduct an FBA. It was found that 43.16% of the time series did so, while 42.69% did not, and 14.15% used data from a FBA conducted shortly before the study and not included as part of the study itself. Interestingly, in Gresham et al., (2004), the researchers had sought to only included school based studies that included an FBA, yet they found that 52% of the 150 located studies did not include an FBA, so they did separate analyses for these studies. This finding was similar to that of the current study, in that close to half (43% in the current study) also did not include an FBA.

In the prior meta-analysis Goh & Bambara (2010) the following coding occurred related to FBAs and was followed in the current research project. Please note that for all of the following FBA related variables, that 242 time series or 57% did not conduct an FBA, so the total time series included in this section is 182. This was similar in Gresham et al., (2004) in that 52% did

not conduct an FBA, whereas Goh & Bambara, (2010) only included studies with an FBA. So of 182 time series in the current study, 68 or 37.4% used an experimental method for the FBA, while 9.9% used only descriptives, and 52.85% used a combination of experimental and descriptive. In Goh & Bambara (2010) a direct comparison can be made in that 21% used experimental, 41% descriptive, and 38% a combination, which was much different than in the current study. In the following summary of the remaining descriptives related to the FBAs, the percentage obtained in Goh & Bambara, (2010) will be in parenthesis to allow for an easier comparison. The agent of the FBA in the current study was similar in that 56.6% (81%) were coded as a typical agent and 44.06% (19%) as an atypical agent. For the FBA setting, 68.68% (81%) were coded as a typical setting while 32.42% (19%) an atypical setting. And although it is best practice to use a team decision process in determining the function of a behavior, 73% (68%) did not use a team decision process and 6.04% did (32%), while 20.88% (0%) did not report on this. In sum there was variation in these variables among the current and extant meta-analyses.

Another characteristic to highlight is that the most frequent type of intervention found was the multicomponent at 41.04% (Goh and Bambara, 2010 reported 46%) which used at least two from the other categories, and these were "consequence-based" accounting for 26.89% (15%), skills training 18.16% (17%) and antecedent-based accounted for 14% (23%). So while both studies found multicomponent as the most frequent, each of the other types were close to the same percentage across the remaining three categories.

It was also found that 62.97% of the time series, were not conducted in studies that had a part of the design where the researchers attempted to generalize the intervention effect. While the remaining percent did. It is best practice to generalize the results to new situation, setting, or

behavior or a combination of such, so it is encouraging that 38% did try to generalize the results. No descriptive information was provided in the pertinent extant meta-analyses concerning this intervention characteristic (Goh & Bambara; Gresham et al., 2004).

In sum, the main descriptive discussion points for intervention characteristics were derived from the findings indicating that short durations of the intervention were most frequent and they were most commonly conducted by a typical adult in the student's natural setting as well in a typical setting. Furthermore, the current study was similar to that of Gresham et al., (2004) whereby 43% in the current study and 52% in Goh and Bambara (2010) found that there was not an FBA conducted. Although, not much analysis could be conducted on the different tiers of behavioral support in the schools for this population, it was highlighted that 100% of these studies included interventions at the Tier 3 level, and it is hypothesized because they need to be so individualized, furthermore conducting an FBA is always an individualized assessment and 57% did include this. Lastly, due to the very individualized approach behavioral interventions have with this population, the finding that 41% of the interventions were of a multicomponent type, makes sense, using a combination of skill training, antecedent manipulation and consequence-based approaches and in line with Goh and Bambara's (2010) finding. There were more similarities found in the intervention characteristics among the current and extant meta-analyses than not, however the type of FBA method did seem quite different, in that the current study found the most frequent type is combination of experimental and descriptive and the least common type was descriptive. It is best practice to not only do a descriptive FBA assessment method, so possibly due to the current study having more recent studies, researchers have been heeding best practices in the type of FBA method used.

Participant characteristics descriptive findings. Findings from descriptive analyses of the participant characteristics show that a larger percentage of the participants were males (71.93%), while females made up 28.07% of the participants. As compared to the current study, Goh and Bambara (2010) also found that there were more male than female participants (74%), while Gresham et al., (2004) did not report this information.

Additionally, it was revealed that children with a diagnosis of an Intellectual Disability (without ASD) accounted for 56.60% of the participants while 43.40% had both ASD and an Intellectual Disability. A difference between the current study and the most similar meta-analyses, Goh and Bambara (2010) included any disability, and Gresham et al., (2004) included only youth with a developmental disability however, it could not be determined through reading the study how this label was determined and anything more specific. Cognitive status was not studied in either of the relevant extant meta-analyses. It was found that 34.43% of the participants had a diagnosis of a severe ID, 21.23% a moderate, and 17.69% mild ID, while 26.65% did not report the specifics of the severity of the intellectual disability. Another characteristic that was a novelty coding, as compared to the two most relevant meta-analyses was the verbal ability of the participants. It was found in the current study that 39.39% did not provide this information however the rest did and that 31.84% were minimally verbal, 15.33% nonverbal, and then between 3-6% were separately coded as echolaliac, having average language skills, or using gestures or sign language or pecs (one category). Age range and grade range were coded, yet 81.37% did not include the grade range, so the focus on the age range is indicated as better variable to examine. It was found that 33.49% were between the ages of 15-18, 26.18% between ages of 9 to 11, 16.98% 12 to 14, and 13.68% between 6 to 8 years old. The age groups of 3 to 5 (3.77%) and 19-22 (4.72%) did not account for much of the data, and therefore the

results are less generalizable to these two age groups. Interestingly, Goh and Bambara (2010), were able to code by grade range and had the largest percentage (69%) in elementary school, 21% in middle, and 7% in high school. Gresham et al., (2004) did not include this information, however did state they included youth ages 1 to 18 years old.

The classroom setting was also coded just as in Goh and Bambara (2010) and the percentages of the various categories from that study are in parenthesis after the current study's findings. It was found that 95% of the participants were in special education classrooms, while Goh and Bambara found 45.5%, and 3% (19%) were in special education and some general education, and 1.42% in general education solely (32.4%). This discrepancy is most likely due to Goh and Bambara (2010) including students with any disability as a participant, and not specifically those with an intellectual disability.

In sum, the main descriptive discussion points for participant characteristics indicated that as in the past most similar meta-analyses there were more male than female participants, and the other participant characteristics were dissimilar due to meta-analyses inclusion criteria differences. Additionally, new descriptive information was collected in this body of literature, in that the current study collected data on the verbal ability and the cognitive status of the participants, as noted was done in past meta-analyses that examined the effect of non-school based behavioral interventions on people with an intellectual disability (Campbell, 2003; Harvey et al., 2009; Heyvaert et al., 2014; Marquis et al., 2000; & Scotti et al., 1991).

Inferential Statistics

The results of the current study indicate that school-based behavioral interventions are significantly effective in helping youth with an intellectual disability increase desirable behaviors and decrease undesirable behaviors.

Behavioral outcomes. A large effect size of 3.31 was found for the main effect by synthesizing 424 time series. In comparison to the most relevant meta-analyses, Goh and Bambara (2010) found a moderate effect size using PND, at 88% PND. When these researchers separated the data by reduction behaviors it was an 80% PND and acquisition behaviors indicated a higher effect at 90%. Whereas, Gresham et al., (2004) found an even higher effect size, also large like the current study. These research conducted separate analyses for studies that included an FBA and those that did not. The effect size for the studies that did not include a FBA was 6.77 and those that did conduct an FBA was 4.60 (see Chapter 2 for reasons these researchers hypothesized that this was the outcome). The current study adds information to this body of literature that is similar to that of the previous meta-analyses examining the effect of school-based interventions on youths' behavioral outcomes. However, adds to it in that it's the only study that includes the past 20 years of studies, multiple journal sources, and specifically studies that examine the effect on students with an intellectual disability. Remember that search methods and inclusion criteria differed between the two extant studies and within the current study. Mainly this is the study that provides data from the largest time frame from 1997 to 2017 (20 years), whereas Goh and Bambara included studies from 1997 to 2008, so this added 9 years of data. Gresham et al., (2004) only included studies from JABA and from 1991-1999. Furthermore, Goh and Bambara (2010) included participants with any disability, and Gresham et al., (2004) didn't go into great detail other than that participants had a developmental disability. Furthermore, the current study had the age range that aligns with the age range that students with an intellectual disability are able to be included in public school, ages 3 to 22, and Gresham et al., (2004) included 1 to 18 years old, while Goh and Bambara (2010) included elementary through high school students. This was the first meta-analysis to duplicate and build on Goh

and Bambara's comprehensive moderator analyses, and those results will be highlighted below. This should continue to be explored in future meta-analyses, until consensus can be determined with enough replication of data. Goh and Bambara (2010) is the main comparison meta-analyses as Gresham et al., (2004) did not conduct moderator analyses.

Moderator analyses. Of the participant characteristic moderators that were examined in the current study and Goh and Bambara, (2010) there is a consensus that the participant's gender, grade range, and diagnosis do not make a difference on the effect of behavioral interventions on youths' behavioral outcomes. Also, although the present study was the only meta-analysis out of the syntheses that are most related (Goh & Bambara, 2010; Gresham et al., 2004), to conduct moderator analyses on age range, cognitive status, and verbal ability, these variables did not impact the effectiveness of the intervention either. The only variable that had a moderating effect was the type of classroom the participant attended while at school and it was found that there was a moderating effect in that interventions were more effective for youth who attend special education classes as compared to those that are in general education setting or a combination of the two types of classrooms. Whereas, in Goh and Bambara, (2010) they did not find a moderating effect, and they included the same categories within this variable (special education, general education, and a combination of both). It should be noted that the prior meta-analysis included youth with all disabilities and possibly this impacted the results of the classroom setting, as the current study only included youth with an intellectual disability or whom also have an ASD. These special education classrooms can qualitatively be experienced very differently. In sum, there were three new variables examined, age, cognitive status, and verbal ability, and a consensus was met on all of the other variables that have been examined for a moderating effect (gender, grade range, and diagnosis), and a difference found in the effect on

the classroom setting of the participants, particularly that the interventions were more effective for students whom are educated exclusively in special education, while Goh & Bambara (2010) found no difference.

In terms of the moderating effects of intervention characteristics, all of the following variables were analyzed in both the current study and Goh and Bambara (2010), FBA presence, FBA method, FBA agent, FBA setting, intervention type, agent, setting, and duration of intervention. They do not seem to make a difference in the effectiveness of the intervention on youths' behavioral outcomes. This is very interesting considering FBAs are the best practice when conducting behavioral interventions. The current study utilized the same categories within each variable as in Goh and Bambara (2010) to help with consistency of research in this field. The remaining intervention characteristics, format of intervention (group or individual) and generalization, were only analyzed in the current study and neither were found to moderate the effect. It seems as if in the study that included participants with a variety of diagnoses, these researchers did not find any intervention characteristics that moderate the effect (Goh & Bambara, 2010), nor in the current study specifically for youth with an ID or ASD and ID. However, the overall effect is large in the case of the current study, meaning intervening is important and it would be interesting to understand the effectiveness if instead of researchers, who are highly trained in best practices of behavioral interventions weren't the designers of the interventions but rather the typical adults in the child's life. Possibly, this is why it's not being captured that FBA's makes a significant difference in the effectiveness of interventions as to those that do not have FBAs. There seems to need to be data collected on who designs the intervention and who designs the FBA and understand if this moderates the effect.

Almost all of the study characteristic moderators that were analyzed in the current study were not examined in the past meta-analyses, including type of SCD, the quality of the SCD, and intervention fidelity. None of these variables moderated the effect, nor did the presence of a social validity, which was not measured in Goh and Bambara (2010).

In sum, we have learned about the moderating effects of variables on the effectiveness of school-based behavioral interventions on youths' behavioral outcomes, specifically youth with an intellectual disability. Furthermore, we have learnings by comparing the current study results to past meta-analyses results. We have learned that some moderators seem to be consistently showing that they do not moderate the effect, including gender, grade range, diagnosis, everything related to FBAs, intervention type, agent, setting, and duration, as well as the presence of a social validity measure. We have learned that there was a difference found in the study that included youth with varying diagnoses in terms of the classroom they attended and in the current study, in that the first found no effect and the later found a moderating effect favoring those attending special education classrooms. We have also learned that some variables have only been analyzed in the current study, in relation to the most similar meta-analyses to the current one, and these include, age range, cognitive status, verbal ability, generalization of the intervention, the type of SCD, the quality of the SCD, and intervention fidelity, all of which did not moderate the effect.

Based on the findings from the current study indicating a large effect size, it is important to encourage the use of behavioral interventions for challenging behaviors and to enhance functioning of youth with an intellectual disability in school settings. The knowledge gained concerning the moderating effects is important to highlight so that future researchers continue to collect this data to help better inform researchers and practitioners of these effects, however it is

posited that data be collected on who conducts the FBA, if there is or is not training for individuals to learn how to conduct an FBA, and who then designs the intervention. Al beit the findings, it is recommended to continue the best practice use of FBAs to inform the development of interventions.

Generalizability of Conclusions

The findings in the current study can be generalized to youth with an intellectual disability with or without autism, having varying severity levels of cognitive disability, and mainly those youth between the ages of 6 to 22. Students ages 19-22 were a smaller percentage of the sample population so it is suggested to use caution in interpreting the effect size for this population. Furthermore, most of the participants were male (71.93%).

The findings in regards to intervention characteristics should be generalized to behavioral interventions conducted at schools, utilizing various types of interventions, mainly conducted within a typical setting within a school, and given to an individual student, versus a group. Furthermore, the results are generalizable to a wide variety of behavioral outcomes, comprising of both reduction and acquisition behaviors.

Limitations

One limitation of the current study is that, although this researcher coded various categories for reduction behaviors, much less delineation of separate acquisition behaviors were conducted and they were grouped mainly as “pro-social behaviors”. However, no moderating effect of the specific outcome studied was found, the descriptive information would have been useful. It may have been helpful to try to use the same categories for the specific outcomes studied as in Gresham et al., (2004) for the purpose of replication in research within this field of study. Another limitation of the current study is that an effect size was not calculated to examine

the maintenance effect of the interventions, although many studies did collect maintenance data and Gresham et al., (2004) found 100% PND for the maintenance effect in their meta-analysis. A limitation of the current study was that there was a strict criteria on how it was determined if a participant had an intellectual disability, in that there needed to be a norm referenced test score or a diagnosis given, which resulted in the exclusion of 26 studies after the final review phase. The implication of this decision for the findings is that possibly studies that did include youth with an intellectual disability were excluded and this could have changed the effect size with the addition of almost a third more total studies. There is a limitation of the methodology choice of only including A and B phases in the analyses instead of also including C or D phases was because a small percentage had a C or D phase and they usually were included a priori, as a result of learnings from phase B. The implication is that the effect size may have been higher had this data been included, since C and D phases were typically implemented when it was noted that the B phase could be implemented in a way that would change behavior better. However, it could also be posited that a limitation of the meta-analysis itself is that it included only single case design studies and not also group design studies, and can only compare baseline to treatment and not control groups to treatment groups. In single case design studies, the standardization is within person variability versus when you have a control group it is between person variability.

Another limitation of the study is that the results are generalizable only to the particular settings, participants, and interventions that were examined in the meta-analysis. An additional limitation is that the search methods may not have located all of the feasible studies. Publication bias is a commonly cited limitation and there were 100% published versus unpublished studies. Additionally, there are chances of data entry and calculation errors but interrater reliability

checks were conducted at various stages of the data collection, extraction, and coding methods with acceptable percentages of agreement.

Implications and Interpretation for Theory, Policy, and Practice

In this section the implications and interpretation of the results will be discussed for theory, policy, and practice. In terms of implications for theory, this study did attempt to examine the independent variable that is based on behaviorism and the constructs of respondent and operant conditioning. Specifically, studies were included that were behavioral by design, were applied in a natural setting, which is the field of applied behavior analysis, and studies were included that conducted an FBA as well as those that did not. It was interesting that there was no moderating effect for those studies that did or did not include a FBA before the development of the intervention. It is hypothesized that this could be because the researchers designed the interventions and were very well versed in thinking through a behavioral theory lens and collecting information to decide what the function or functions of a behavior were, although possibly informally doing so, therefore not showing a difference. In Gresham et al., (2004), these researchers found a higher effect size in studies that did not include an FBA, which would be counter to the theory of behaviorism and applied behavior analysis. However, they posited that this possibly was due to studies that had a higher effect size being published, therefore the studies that did not conduct an FBA and were not effective, were not as likely to be published, causing a publication bias. The large effect size found in the current meta-analysis adds further to the strong evidence that behavioral interventions are useful and specifically, within schools with students with an intellectual disability and with or without autism.

The results of this meta-analysis provides implications for policy-makers, to help stakeholders advocate to make or keep policies that protect youth with an intellectual disability

to receive evidence-based behavioral interventions to help these youth succeed behaviorally within the school setting and potentially beyond. It also implies that schools should intervene not only when there are challenging behaviors but also to increase functioning, such as in Daily Living Skills. So instead of simply teaching the skills through a curriculum, utilizing evidence based behavioral principles to do so.

Implications for practitioners can also be posited. This meta-analysis provides information that even short durations of a behavioral intervention, as long as based in strong behavioral principles, can have a positive effect on youths' behavioral outcomes. This information is important for educators to know so that when challenging behaviors arise, the adults in the child's life know that there is a strong basis in the literature to conduct behavioral interventions for these youth and that change can occur. Moreover the data indicates that behavioral interventions can increase functioning and in many domains such as communication, social skills, and daily living skills. The lack of data at the Tier 1 and Tier 2 level, are slightly related to the type of design included in the current study, however, SCDs can be used grouping students together working towards the same outcomes. Furthermore, it may be beneficial during graduate training of the future school based practitioners, specifically school psychologists to receive more training in applied behavior analysis to be able to design, implement, and evaluate these types of interventions among youth with an intellectual disability, or other disabilities.

Guidelines for Future Research

Further research using SCDs could be conducted surrounding educator training on how to conduct a FBA and how to design an intervention based on this assessment, to understand if a similar effect is found if educators design the interventions instead of researchers. In future meta-analyses, it may be useful to collect data on the same moderators and categories,

to aide in replication of research, and therefore generalizability of the results. It is recommended that the maintenance effect also be included in future meta-analyses, as many studies included a maintenance phase. It would be interesting to understand if students with an intellectual disability are receiving Tier 1 and Tier 2 supports to help prevent and modify behavioral challenges and to increase functioning, before a Tier 3 level intervention is needed. The finding regarding that around 40% of the studies included generalizability methods, calls for this percentage to be higher and closer to 100% to help solidify behavior change across contexts and as best practice in research. In addition, upon review of the included studies, although not specifically coded and analyzed there seemed to be a dearth of studies using mixed methods, with a lack of qualitative research being included with the quantitative research. This is an area for future direction in research in this field.

Conclusion

The results of this study have important implications for this specific population of youth and those who work with or care for these youth. Also, the results are important for policy-makers and practitioners to advocate for the use of sound behavioral interventions as a way to help promote appropriate behaviors and decrease challenging behaviors within schools. Furthermore, all studies in this body of literature should collect data on potential moderating variables as well as encouraging researchers of individual studies to collect data on generalizability. The large effect size of the current study is very promising to indicate the evidence-base of utilizing behavioral interventions for youth with an intellectual disability at schools, and it is a hope that this research will encourage the use of such well-designed interventions to not only decrease challenging behaviors, increase appropriate behaviors, and to also enhance functioning through skill teaching using behavioral techniques. How many

diagnoses exist that have interventions that have small effect sizes, whereas for this population of students there seems to be a very specific theory that allows for, on average, a large effect on various different behaviors to change in a desirable way? Let's use them. Let's not only work on eliminating challenging behaviors but having high expectations and using these principles to reach multiple domains of functioning. Let's make sure the natural adults in these children's school lives know how to design, implement, and understand without a doubt if they are working, and furthermore let's figure out how to make these techniques instilled and feasible within the school setting.

References

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

Allen, D. (2013). Towards the prevention of behavioral and psychiatric disorders in people with intellectual disabilities. *Journal of Applied Research in Intellectual Disabilities, 26*(6), 501.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC.

Barlow, D. H., Nock, M. K., & Hersen, M. (2009). Single case experimental designs: Strategies for studying behavior change (3rd ed.). Boston, MA: Pearson.

*Barry, L. M. (2000). Making it work at school and home: A need based collaborative, across settings, behavioral intervention. *California School Psychologist*, 5, 43-51.

Beukleman, D. R., & Mirenda, P. (2005). *Augmentative and alternative communication: Management of severe communication disorders in children and adults* (2nd ed.). Baltimore: Brookes.

Bijou, S., & Baer, D. M. (1961). *Child development. Vol. I: A systematic and empirical theory.* New York, NY: Appleton- Century-Crofts.

*Bilias Lolis, E. (2012). Exploring the Utility of Self-Modeling in Decreasing Disruptive Behavior in Students with Intellectual Disability. *Psychology in the Schools*, 49(1), 82-92.

*Blair, K., Liaupsin, C. J., Umbreit, J., Kweon, G. (2006). Function-based intervention to

support the inclusive placements of young children in Korea. *Education and Training in Developmental Disabilities,* 41, 48–57.

Blimling, G. S. (1988). Meta-Analysis: A Statistical Method for Integrating the Results of Empirical Studies. *Journal of College Student Development, 29*(6), 543-49.

Bondy, A., & Frost, L. (2001). The picture exchange communication system. *Behavior Modification, 25*(5), 725-744.

Borthwick-Duffy, S. A. (1996). Evaluation and measurement of quality of life: Special considerations for persons with mental retardation. *Quality of life, 1,* 105-119.

*Braithwaite, K. L., Richdale, A. L. (2000). Functional communication training to replace challenging behaviors across two behavioral outcomes. *Behavioral Interventions,* 15, 21–36.

*Brooks, A., Todd, A. W., Tofflemoyer, S., & Horner, R. H. (2003). Use of functional assessment and a self-management system to increase academic engagement and work completion. *Journal of Positive Behavior Interventions, 5*(3), 144.

*Bruhn, A. (2015). Self-Monitoring Interventions for Students with Behavior Problems: A Systematic Review of Current Research. *Behavioral Disorders,* 40(2), 102-121.

Busk, P. L., & Serline, R. C. (1992). Meta-analysis for single-case research. In T. Kratochwill, & J. Levin (Eds.). *Single-case research design and analysis: New directions for psychology and education* (pp. 187-212). Hillsdale, NJ: Lawrence Erlbaum Associates.

*****Cale, S. I., Carr, E. G., Blakeley-Smith, A., & Owen-DeSchryver, J. (2009). Context-based assessment and intervention for problem behavior in children with autism spectrum disorder. *Behavior Modification, 33*(6), 707-742.

Campbell, J. M. (2003). Efficacy of behavioral interventions for reducing problem behavior in persons with autism: A quantitative synthesis of single-subject research. *Research in Developmental Disabilities, 24*, 120–138. doi:10.1016/ S0891-4222(03)00014-3.

*Cannella-Malone, H., Fleming, C., Chung, Y., Wheeler, G. M., Basbagill, A. R., & Singh, A. H. (2011). Teaching daily living skills to seven individuals with severe intellectual disabilities: A comparison of video prompting to video modeling. *Journal of Positive Behavior Interventions, 13*(3), 144-153.

Carmago, S., H. (2012). *Behaviorally based interventions for improving social interaction skills of children with ASD in inclusive settings: a systematic review and meta-analysis*. (Doctoral dissertation). Retrieved from Proquest Theses and Dissertations.

Carr, E. G., Horner, R. H., Turnbull, A. P., Marquis, J. G., Magito- McLaughlin, D., McAtee, M. L.,. Braddock, D. (1999). *Positive behavior support for people with developmental disabilities: A research synthesis.* Washington, DC: American Association on Mental Retardation.

Carr, E. G., Dunlap, G., Horner, R. H., Koegel, R. L., Turnbull, A. P., Sailor, W., . . . Fox, L. (2002). Positive behavior support: Evolution of an applied science. *Journal of Positive Behavior Interventions*, *4*, 4–16. doi:10.1177/109830070200400102

*Carey, Y. A., Halle, J. W. (2002). The effect of an idiosyncratic stimulus on self-injurious behavior during task demands. *Education and Treatment of Children*, 25, 131–141.

Center for Disease Control. (2012). Prevalence of autism spectrum disorders-autism and developmental disabilities monitoring network, 14 Sites, United States, 2008. MMWR

Surveillance Summaries, 61(3), 1–24.

Civic Impulse. (2016). S. 2781 — 111th Congress: Rosa's Law. Retrieved from https://www.govtrack.us/congress/bills/111/s2781

*Clemons, L. L., Mason, B. A., Garrison-Kane, L., & Wills, H. P. (2016). Self-monitoring for high school students with disabilities. *Journal of Positive Behavior Interventions, 18*(3), 145-155.

Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). Applied behavior analysis.

*Craft, M. A., Alber, S. R., & Heward, W. L. (1998). Teaching elementary students with developmental disabilities to recruit teacher attention in a general education classroom: Effects on teacher praise and academic productivity. *Journal of Applied Behavior Analysis, 31*(3), 399-415.

Dekker M. C., Koot H. M., Van Der Ende J. & Verhulst F. C. (2002). Emotional and behavioral problems in children and adolescents with and without intellectual disability. *Journal of Child Psychology and Psychiatry*, 43, 1087–1098.

Demaray, M. K., Malecki, C. K., & DeLong, L. K. (2006). Support in the lives of aggressive students, their victims, and their peers. In S. R. Jimerson, & M. J. Furlong, (Eds.), *Handbook of school violence and school safety: From research to practice* (pp. 21–29). Mahwah, NJ: Lawrence Erlbaum.

Denis, J., Van Den Noortgate, W., & Maes, B. (2011). Self-injurious behavior in people with profound intellectual disabilities: A meta-analysis of single-case studies. *Research in Developmental Disabilities*, 32, 911–923.

*Dibley, S., Lim, L. (1999). Providing choice making opportunities within and between daily

school routines. *Journal of Behavioral Education*, 9, 117–132.

Didden, R., Duker, P. C., & Korzilius, H. (1997). Meta-analytic study on treatment effectiveness for problem behaviors with individuals who have mental retardation. *American Journal on Mental Retardation*, 101, 387-399.

Didden, R., Korzilius, H., van Oorsouw, W., & Sturmey, P. (2006). Behavioral treatment of challenging behaviors in individuals with mild mental retardation: Meta-analysis of single-subject research. *American Journal of Mental Retardation*, *111*, 290–298. doi:10.1352/0895-8017

*DiGennaro, F. D., Martens, B. K., & Kleinmann, A. E. (2007). A omparison of performance feedback procedures on teachers' treatment implementation integrity and students? Inappropriate behavior in special education classrooms. *Journal of Applied Behavior Analysis*, *40*(3), 447-461.

*Ducharme, J. M., & DiAdamo, C. (2005). An errorless approach to management of child noncompliance in a special education setting. Research brief. *School Psychology Review, 34*(1), 107-115.

*Dunlap, G. (1991) "Functional assessment, curricular revision, and severe behavior problems." *Journal of Applied Behavior Analysis*, 24(2), 387-397.

Durand, V. M. (1990). *Severe behavior problems: A functional communication training approach*. New York: Guilford Press.

*Durand, V. M. (1999). Functional communication training using assistive devices: Recruiting natural communities of reinforcement. *Journal of Applied Behavior Analysis, 32*(3), 247-267.

*Durand, V. M., & Carr, E. G. (1991). Functional communication training to reduce challenging behavior: Maintenance and application in new settings. Journal of Applied Behavior Analysis, 24(2), 251-264.

*Dyer, K., Dunlap, G., & Winterling, V. (1990). Effects of choice making on the serious problem behaviors of students with severe handicaps. *Journal of Applied Behavior Analysis, 23*(4), 515-524.

*Egan, P. J., Zlomke, L. C., & Bush, B. R. (1993). Utilizing functional assessment, behavioral consultation and videotape review of treatment to reduce aggression: A case study. *Special Services in the Schools*, 7(1), 27-37.

Emerson, E. (2003). Prevalence of psychiatric disorders in children and adolescents with and without intellectual disability. *Journal of Intellectual Disability Research,* 47, 51–58.

Emerson E. & Hatton C. (2007) Mental health of children and adolescents with intellectual disabilities in Britain. *British Journal of Psychiatry,* 191, 493–499.

Evans, I., & Meyer, L., (1985). An educative approach to behavior problem: A practical decision model for interventions with severely handicapped learners. Baltimore: Brookes.

Faith, M. S., Allison, D. B., & Gorman, B. S. (1997). Meta-analysis of single-case research. In R. D. Franklin, D. B. Allison, et al. (Eds.), *Design and analysis of single-case research* (pp. 245–277). Mahwah, NH: Erlbaum.

Ferron, J. M., Farmer, J. L., & Owens, C. M. (2010). Estimating individual treatment effects from multiple-baseline data: A Monte Carlo study of multilevel-modeling approaches. *Behavior Research Methods*, *42*(4), 930-943.

Ferron, J. M., Bell, B. A., Hess, M. R., Rendina-Gobioff, G., & Hibbard, S. T. (2009). Making treatment effect inferences from multiple-baseline data: The utility of multilevel modeling approaches. *Behavior Research Methods, 41*(2), 372-384.

Florida Department of Education. (2015). *Exceptional Education and Student Services.* Retrieved from http://www.fldoe.org/academics/exceptional-student-edu/ese-eligibility

*Foster-Johnson, L., Ferro, J., & Dunlap, G. (1994). Preferred curricular activities and reduced problem behaviors in students with intellectual disabilities. *Journal of Applied Behavior Analysis*, 27, 493-504.

*Foxx, R. M., & Shapiro, S. T. (1978). The timeout ribbon: A nonexclusionary timeout procedure. *Journal of Applied Behavior Analysis, 11*(1), 125-136.

*Francisco, M. T., & Hanley, G. P. (2012). An evaluation of progressively increasing intertrial intervals on the acquisition and generalization of three social skills. *Journal of Applied Behavior Analysis, 45*(1), 137-142.

*Frea, W. D., Arnold, C. L., Vittimberga, G. L., & Koegel, R. L. (2001). A demonstration of the effects of augmentative communication on the extreme aggressive behavior of a child with autism within an integrated preschool setting. *Journal of Positive Behavior Interventions, 3*(4), 194.

Freeman, R. D., Soltanifar, A., & Baer, S. (2010). Stereotypic movement disorder: easily missed. *Developmental Medicine & Child Neurology*, *52*(8), 733-738.

Gage, N. A., Lewis, T. J., & Stichter, J. P. (2012). Functional behavioral assessment-based

interventions for students with or at risk for emotional and/or behavioral disorders in school: A hierarchical linear modeling meta-analysis. *Behavioral Disorders,* 55-77.

*Ganz, J. B., Heath, A. K., Lund, E. M., Camargo, S. P. H., Rispoli, M. J., Boles, M., et al. (2011). Effects of peer-mediated implementation of visual scripts in middle school. *Behavior Modification, 36*(3), 378-398.

Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher, 5*(10) 3-8.

Goh, A. E. & Bambara, L. M. (2010). Individualized positive behavior support in school settings: A meta-analysis. *Remedial and Special Education,* XX, X, 1-16. doi: 10.1177/0741932510383990

*Graetz, J. E., Mastropieri, M. A., & Scruggs, T. E. (2009). Decreasing inappropriate behaviors for adolescents with autism spectrum disorders using modified social stories. *Education and Training in Developmental Disabilities, 44*(1), 91-104.

*Graff, R. B. (1999). Functional analysis and treatment of screaming in a young girl with severe disabilities. Behavioral Interventions, 14(4), 233-239.

Gresham, F. M., Mclntyre, L. L., Olson-Tinker, H., Dolstra, L., McLaughlin, V., & Van, M. (2004). Relevance of functional behavioral assessment research for school-based intervention and positive behavioral support. *Research in Developmental Disabilities,* 25(1), 19-37, doi:10.1016/ j.ridd.2003.04.003.

*Grey, I., Healy, O., Leader, G., & Hayes, D. (2009). Using a time timer[TM] to increase appropriate waiting behavior in a child with developmental disabilities. *Research in Developmental Disabilities: A Multidisciplinary Journal, 30*(2), 359-366.

*Hall, A., Neuharth-Prichett, S., & Belfiore, P. J. (1997). Reduction of aggressive behaviors with changes in activity: Linking descriptive and experimental analyses. *Education and Training in Mental Retardation and Developmental Disabilities, 32*(4), 331-39.

*Hall, R. V., Fox, R., Willard, D., Goldsmith, L., Emerson, M., Owen, M., et al. (1971). The teacher as observer and experimenter in the modification of disputing and talking-out behaviors. *Journal of Applied Behavior Analysis, 4*(2), 141-149.

*Hansen, B. D. (2014). The effects of goal setting, contingent reward, and instruction on writing skills. *Journal of Applied Behavior Analysis*, 47(1), 171-175.

*Haring, T. G. (1990). Contextual control of problem behavior in students with severe disabilities. *Journal of Applied Behavior Analysis*, 23(2), 235-243.

Harvey, S. T., Boer, D., Meyer, L. H., & Evans, I. M. (2009). Updating a meta-analysis of intervention research with challenging behavior: Treatment validity and standards of practice. *Journal of Intellectual & Developmental Disability, 34*, 67–80. doi:10.1080/13668250802690922

Hedges, L.V., Pustejovsky, J., & Shadish, W.R. (2012). A standardized mean difference effect size for single-case designs. *Research Synthesis Methods*, 3, 224-239.

Heyvaert, M. (2010). A meta-analysis of intervention effects on challenging behavior among persons with intellectual disabilities. *Journal of Intellectual Disability Research, 54*(7), 634.

Heyvaert, M. (2012). A multilevel meta-analysis of single-case and small-n research on interventions for reducing challenging behavior in persons with intellectual disabilities.

Research in Developmental Disabilities, 33(2), 766.

Heyvaert, M., Saenen, L., Campbell, J. M., Maes, B., & Onghena, P. (2014). Efficacy of behavioral interventions for reducing problem behavior in persons with autism: An updated quantitative synthesis of single-subject research. *Research in Developmental Disabilities, 35*(10), 2463-2476.

*Hirsch, N., & Myles, B. S. (1996). The use of a pica box in reducing pica behavior in a student with autism. *Focus on Autism and Other Developmental Disabilities, 11*(4), 222-225.

Horner, R. H., Carr, E. G., Strain, P. S., Todd, A. W., & Reed, H. K. (2002). Problem behavior interventions for young children with autism: A research synthesis. *Journal of Autism and Developmental Disorders, 32*, 423–446. doi:10.1023/A:1020593922901

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*, 165–179.

*Hughes, C. (1997). Functional Analysis and Treatment of Social-Communicative Behavior of Adolescents with Developmental Disabilities. *Journal of Applied Behavior Analysis*, 30(4), 701-04.

*Hunt, P., Alwell, M., Goetz, L., & Sailor, W. (1990). Generalized effects of conversation skill training. Journal of The Association for Persons With Severe Handicaps, 15, 250-260.

Individuals with Disabilities Education Act, Pub. L. No. 105-17, 111 Stat. 37 (1997).

Individuals with Disabilities Education Improvement Act, Pub. L. No. 108-446, 118 Stat. 2647 (2004).

*Jackson, J. L., & Calhoun, K. S. (1977). Effects of two variable-ratio schedules of timeout:

Changes in target and non-target behaviors. *Journal of Behavior Therapy and Experimental Psychiatry, 8*(2), 195-199.

*Johnson, J. W., Van Laarhoven, T., & Repp, A. C. (2002). Effects on stereotypy and other challenging behavior of matching rates of instruction to free-operant rates of responding. *Research in Developmental Disabilities, 23*(4), 266-84.

Kahng, S., Iwata, B. A., & Lewin, A. B. (2002). The impact of functional analysis on the treatment of self-injurious behavior. In S. R. Schroeder, M. L. Oster-Granite, & T. Thompson (Eds.), *Self-injurious behavior: Gene-brain-behavior relationships* (pp. 119–131). Washington, DC: American Psychological Association.

*Karsh, K. G., Repp, A. C., Dahlquist, C. M., & Munk, D. (1995). In vivo functional assessment and multi-element interventions for problem behaviors of students with disabilities in classroom settings. *Journal of Behavioral Education*, 5, 189-210.

*Kay, S., Harchik, A. E., & Luiselli, J. K. (2006). Elimination of drooling by an adolescent student with autism attending public high school. *Journal of Positive Behavior Interventions, 8*(1), 24-28.

Kazdin, A. E. (2011). *Single-Case Research Designs: Methods for Clinical and Applied Settings* (2nd ed.). New York: Oxford University Press.

*Keeter, D., & Bucholz, J. L. (2012). Group delivered literacy-based behavioral interventions for children with intellectual disability. *Education and Training in Autism and Developmental Disabilities, 47*(3), 293-301.

*Kelley, M. E., Lerman, D. C., & Van Camp, C. M. (2002). The effects of competing

reinforcement schedules on the acquisition of functional communication. *Journal of Applied Behavior Analysis, 35*(1), 59-63.

*Kennedy, C. H., & Itkonen, T. (1993). Effects of setting events on the problem behavior of students with severe disabilities. *Journal of Applied Behavior Analysis, 26*(3), 321-27.

*Koegel, L. K. (1998). Reducing aggression in children with autism toward infant or toddler siblings. *Journal of the Association for the Severely Handicapped*, 23(2), 111-118.

Koul, R. K., Schlosser, R. W., & Sancibrian, S. (2001). Effects of symbol, referent, and instructional variables on the acquisition of aided and unaided symbols by individuals with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities*, *16*(3), 162-169.

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. Retrieved from What Works Clearinghouse website: http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf.

Krysik, J. L., & Finn, J. (2010). *Research for Effective Social Work Practice* (2nd ed.). New York: Rouledge.

*Lalli, J. S., & Others, A. (1993). Teacher use of descriptive analysis data to implement interventions to decrease students' problem behaviors. *Journal of Applied Behavior Analysis, 26*(2), 227-38.

*Langdon, N. A. (2008). Functional Analysis of Precursors for Serious Problem Behavior and Related Intervention. Behavior modification, 32(6), 804-827.

Linna S. L., Moilanen I., Ebeling H., Piha J., Kumpulainen K., Tamminen T. & Almqvist F. (1999). Psychiatric symptoms in children with intellectual disability. *European Child and*

Adolescent Psychiatry 8, 77–82.

Littell, J. H., Corcoran, J., & Pillai, V. (2008). *Systematic reviews and meta-analysis.* New York: Oxford University Press.

Littell, R. C., Milliken, G. A., Stroup, W. W., Wolfinger, R. D., & Schabenberger, O. (2006). SAS for mixed models 2nd ed. SAS Inst., Cary, NC.

Lloyd, B. P., & Kennedy, C. H. (2014). Assessment and treatment of challenging behavior for individuals with intellectual disability: A research review. *Journal of Applied Research in Intellectual Disabilities*, *27*(3), 187-199.

*Lohrmann-O'Rourke, S., & Yurman, B. (2001). Naturalistic assessment of and intervention for mouthing behaviors influenced by establishing operations. *Journal of Positive Behavior Interventions, 3*(1), 19-27.

*Luiselli, J. K., Suskin, L., & Slocumb, P. R. (1984). Application of immobilization time-out in management programming with developmentally disabled children. *Child & Family Behavior Therapy, 6*(1), 1-15.

Ma, H. H. (2006). An alternative method for quantitative synthesis of single-subject researches: Percentage of data points exceeding the median. *Behavior Modification*, 30, 598–617.

*Mace, A. B., Shapiro, E. S., & Mace, F. C. (1998). Effects of warning stimuli for reinforcer withdrawal and task onset on self-injury. *Journal of Applied Behavior Analysis, 31*(4), 679-682.

*Mancine, C., Tankersley, M., Kamps, D., Kravits, T., & Parrett, J. (2000). Reduction of inappropriate vocalizations for a child with autism using a self-management treatment program. *Journal of Autism and Developmental Disorders, 30*(6), 599-606.

Marquis, J. G., Horner, R. H., Carr, E. G., Turnbull, A. P., Thompson, M., Behrens, G. A., et al. (2000). A meta-analysis of positive behavior support. In R. Gersten, E. P. Schiller, & S. Vaughn (Eds.), *Contemporary special education research: Synthesis of knowledge base on critical instructional issues* (pp. 137–178). Mahwah, NJ: Lawrence Erlbaum Associates.

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,* 109–135.

Matson, J. L., & LoVullo, S. V. (2008). A review of behavioral treatments for self-injurious behaviors of persons with autism spectrum disorders. *Behavior Modification*, 32, 61–76.

Marquis, J., Horner, R., Carr, E., Turnull, A., Thompson, M., & Behrens, G., et al. (2000). A meta-analysis of positive behavior support. In R/. Gersten, E.P. Schiller, & B.J. Vaugn (Eds.), *Contemporary special education research: Syntheses of the knowledge base on critical instructional issues* (pp.17-178). Mahwah, NJ: Lawrence Erlbaum Associates.

Miller, F. G., & Lee, D. L. (2013). Do functional behavioral assessments improve intervention effectiveness for students diagnosed with ADHD? A single-subject meta-analysis. *Journal of Behavioral Education, 22*(3), 253-282.

Moeyaert, M., Ugille, M., Ferron, J. M., Beretvas, S. N., & Van den Noortgate, W. (2013). Modeling external events in the three-level analysis of multiple-baseline across-participants designs: A simulation study. *Behavior Research Methods, 45*(2), 547-559.

Moeyaert, M., Ugille, M., Ferron, J. M., Beretvas, S. N., & Van den Noortgate, W. (2015). The misspecification of the covariance structures in multilevel models for single-case data: A Monte Carlo simulation study. *The Journal of Experimental Education,* 1-37.

*Moore, T. C., Robinson, C. C., Coleman, M. B., Cihak, D. F., & Park, Y. (2016). Noncontingent reinforcement to improve classroom behavior of a student with developmental disability. *Behavior Modification, 40*(4), 640-657.

Morgan, D. L. & Morgan, R. K. (2001). Single-participant research design: Bringing science to managed care. *American Psychologist, 56*, 119-127.

Murphy, G. H., Beadle-Brown, J., Wing, L., Gould, J., Shah, A., & Holmes, N. (2005). Chronicity of challenging behaviors in people with severe intellectual disabilities and/or autism: A total population sample. *Journal of Autism and Developmental Disorders*, 35, 267–280.

Murphy, O., Healy, O., & Leader, G. (2009). Risk factors for challenging behaviors among 157 children with autism spectrum disorder in Ireland. *Research in Autism Spectrum Disorders*, 3, 474–482.

*Newman, B., Needelman, M., Reinecke, D. R., & Robek, A. (2002). The effect of providing choices on skill acquisition and competing behavior of children with autism during discrete trial instruction. *Behavioral Interventions, 17*(1), 31-41.

*Nunes, D. L., Murphy, R. J., & Ruprecht, M. L. (1977). Reducing self-injurious behavior of severely retarded individuals through withdrawal of reinforcement procedures. *Behavior Modification, 1*(4), 499-516.

Odom, S., Brantlinger, E., Gersten, R., Horner, R., Thompson, B., & Harris, K. (2005). Research in special education: Scientific methods and evidence-based practices. *Exceptional Children, 71*(2), 137-148.

*Ohtake, Y., Takeuchi, A., & Watanabe, K. (2014). Effects of video self-modeling on eliminating public undressing by elementary-aged students with developmental

disabilities during urination. *Education and Training in Autism and Developmental Disabilities, 49*(1), 32-44.

Onghena, P. (2005). Single-case designs. In Everitt, B., Encyclopedia of statistics in behavioral science (Vol. 4, pp. 1850–1854). Chichester: Wiley.

*O'Reilly, M. F., Tiernan, R., Lancioni, G. (2002). Use of self-monitoring and delayed feedback to increase on-task behavior in a post-institutionalized child within regular classroom settings. Education and Treatment of Children, 25, 91–102.

Owens, C. M. (2011). Meta-analysis of single-case data: A Monte Carlo investigation of a three level model (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. UMI No. 3449468

Owens, C. M., & Ferron, J. M. (2012). Synthesizing single-case studies: A Monte Carlo examination of a three-level meta-analytic model. *Behavior Research Methods, 44(*3), 795-805.

*Park, S., Singer, G. H. S., Gibson, M., & Koegel, R. L. (2005). The functional effect of teacher positive and neutral affect on task performance of students with significant disabilities. *Journal of Positive Behavior Interventions, 7*(4), 237-246.

Parker, R. I., Hagan-Burke, S., & Vannest, K. (2007). Percentage of all non-overlapping data (PAND): An alternative to PND. *Journal of Special Education*, 40, 194–204.

Parker, R., Vannest, K., & Brown, L. (2009). The improvement rate difference for single-case research. *Exceptional Children, 75*, 135-150.

*Payne, L. D., Scott, T. M., & Conroy, M. (2007). A school-based examination of the efficacy of function-based intervention. *Behavioral Disorders, 32*(3), 158-174.

*Peterson, S. M., Frieder, J. E., Smith, S. L., Quigley, S. P., & Van Norman, R. K. (2009). The effects of varying quality and duration of reinforcement on mands to work, mands for break, and problem behavior. *Education and Treatment of Children, 32*(4), 605-630.

*Repp, A. C., Barton, L. E., & Brulle, A. R. (1983). A comparison of two procedures for programming the differential reinforcement of other behaviors. *Journal of Applied Behavior Analysis, 16*(4), 435-445.

*Repp, A. C., & Karsh, K. G. (1994). Hypothesis-based interventions for tantrum behaviors of persons with developmental disabilities in school settings. *Journal of Applied Behavior Analysis, 27*(1), 21-31.

Riley-Tillman, T. C., & Burns, M. K. (2009). *Single case design for measuring response to educational intervention.* New York: Guilford.

*Roane, H. S., Kelly, M. L., & Fisher, W. W. (2003). The effects of noncontingent access to food on the rate of object mouthing across three settings. *Journal of Applied Behavior Analysis, 36*(4), 579-582.

*Romaniuk, C. (2002). The influence of activity choice on problem behaviors maintained by escape versus attention. Journal of Applied Behavior Analysis, 35(4), 349-362.

Rosenwasser, B., & Axelrod, S. (2001). The contributions of applied behavior analysis to the education of people with autism. *Behavior Modification*, 25(5), 671-677.

Safran, S. P., & Oswald, K. (2003). Positive behavior supports: Can schools reshape disciplinary practices? *Exceptional Children*, 69, 361–373.

Satterthwaite, F. E. (1941). Synthesis of variance. *Psychometrika, 6,* 309-16.

*Scattone, D. (2002). Decreasing disruptive behaviors of children with autism using social stories. Journal of autism and developmental disorders, 32(6), 535-543.

Schlosser, R. W. (2005). Meta-analysis of single-subject research: How should it be done? *International Journal of Language & Communication Disorders, 40*, 375-377. doi:10.1080/13682820500057327

*Schmidt, J. D., Luiselli, J. K., Rue, H., & Whalley, K. (2013). Graduated exposure and positive reinforcement to overcome setting and activity avoidance in an adolescent with autism. *Behavior Modification, 37*(1), 128-142.

Scruggs, T. E., & Mastropieri, M. A. (1998). Summarizing single subject research: Issues and applications. Behavior Modification, 22(3), 221–242.

Scotti, J. R., Evans, I. M., Meyer, L. H., & Walker, P. (1991). A meta-analysis of intervention research with problem behavior: Treatment validity and standards of practice. *American Journal on Mental Retardation*, 96(3), 233-256.

*Secan, K. E., & Egel, A. L. (1986). The effects of a negative practice procedure on the self-stimulatory behavior of developmentally disabled students. *Education and Treatment of Children, 9*(1), 30-39.

* Seybert, S., Dunlap, G., & Ferro, J. (1996). The effects of choice-making on the problem behaviors of high school students with intellectual disabilities, *Journal of Behavioral Education*, 6(1), 49-65.

*Sigafoos, J. (1996). Functional communication training for the treatment of multiply determined challenging behavior in two boys with autism. Behavior modification, 20(1), 60-84.

*Sigafoos, J. (1996b). Naturalistic Assessment Leading to Effective Treatment of Self-Injury in a Young Boy with Multiple Disabilities. Education and Treatment of Children, 19(2), 101-23.

Sigafoos, J., & Drasgow, E. (2001). Conditional use of aided and unaided AAC a review and clinical case demonstration. *Focus on Autism and Other Developmental Disabilities*, 16(3), 152-161.

*Sugai, G. (1984). The effect of self-recording on out-of-seat behavior of an EMR student. Education and Training of the Mentally Retarded, 19(1), 23-28.

Shadish, W. R., Rindskopf, D. M., & Hedges, L. V. (2007). The state of the science in the meta-analysis of single-case experimental designs. *Evidence-Based Communication Assessment and Intervention, 2*(3), 188-196.

Shadish, W.R., Hedges, L. V., Pustejovsky, J. E., Boyajian, J. G., Sullivan, K. J., Andrade, A., & Barrientos, J. L. (2014). A d-statistic for single-case designs that is equivalent to the usual between-groups d-statistic. *Neuropsychological Rehabilitation*, 24, 528-553.

Shogren, K. A., Faggella-Luby, M. N., Bae, S. J., & Wehmeyer, M. L. (2004). The effect of choice-making as an intervention for problem behavior: A meta-analysis. *Journal of Positive Behavior Interventions*, 6, 228–237.

Skinner, B. F. (1974). *About Behaviorism*. New York: Knopf.

Snell, M. E., Voorhees, M. D., & Chen, L. (2005). Team involvement in assessment-based interventions with problem behavior: 1997–2002. *Journal of Positive Behavior Interventions, 7*, 140–152. doi:10.1177/10983007050070030301

*So, H. L., Odom, S. L., & Loftin, R. (2007). Social engagement with peers and stereotypic

behavior of children with autism. *Journal of Positive Behavior Interventions, 9*(2), 67-79.

Sprague, J. R., & Rian, V. (1993). *Support systems for students with severe problem behaviors in* Indiana: A descriptive analysis of school structure and student demographics. Unpublished manuscript, Indiana University, Bloomington. Indiana.

Sprague, J. R., & Horner, R. H. (2006). School wide positive behavioral support. In S. R. Jimerson & M. J. Furlong (Eds.), *Handbook of school violence and school safety: From research to practice* (pp. 413–428). Mahwah, NJ: Lawrence Erlbaum.

Stephenson, A., Dempsey, L., Scotti, J. (1983). *Guidelines and standards for habilitative interventions utilizing aversive or unpleasant consequences*. Binghamton, NY: Broome Developmental Services.

*St. Lawrence, J. S., & Drabman, R. S. (1984). Suppression of chronic high frequency spitting in a multiply handicapped and mentally retarded adolescent. *Child & Family Behavior Therapy, 6*(1), 45-55.

Strain, P., Kohler, F., & Gresham, F. M. (1998). Problems in logic and interpretation with quantitative syntheses of single-case research: Mathur and colleagues (1998) as a case in point. *Behavioral Disorders*, 24, 74–85.

Sugai, G. (2000). Applying positive behavior support and functional behavioral assessment in schools. *Journal of Positive Behavior Interventions*, 2(3), 131.

Swanson, H. L., & Sachse-Lee, C. (2012). A meta-analysis of single-subject-design intervention research for students with learning disabilities. *Journal of Learning Disabilities, 33*(2), 114-136.

Tummers, B. (2006). DataThief III [Computer software]. Retrieved from <http://datathief.org/

United States Department of Education. Building the legacy: IDEA 2004. Retrieved from

http://idea.ed.gov/explore/view/p/,root,regs,300,D,300%252E320

United States Department of Education, Institute of Education Sciences, National Center for Educational Statistics (2016). *Fast Facts, Students with Disabilities.* Retrieved from https://nces.ed.gov/fastfacts/display.asp?id=64

Van den Noortgate, W., & Onghena, P. (2003). Combining single-case experimental data using hierarchical linear models. *School Psychology Quarterly, 18*(3), 325-346.

Van den Noortgate, W., & Onghena, P. (2007). The aggregation of single-case results using hierarchical linear models. *The Behavior Analyst Today*, 8, 196–209.

Van den Noortgate, W., & Onghena, P. (2008). A multilevel meta-analysis of single-subject experimental design studies. *Evidence-Based Communication Assessment and Intervention, 2*(3), 142-151.

*Wadsworth, J. P. (2015). Increasing Compliance in Students with Intellectual Disabilities Using Functional Behavioral Assessment and Self-Monitoring. Remedial and Special Education, 36(4), 195-207.

Walsh, C. E., Mulder, E., & Tudor, M. E. (2013). Predictors of parent stress in a sample of children with ASD: Pain, problem behavior, and parental coping. *Research in Autism Spectrum Disorders*, 7, 256–264. http://dx.doi.org/10.1016/j.rasd.2012.08.010

Wang, S. Y., Cui, Y., & Parrila, R. (2011). Examining the effectiveness of peer-mediated and

video-modeling social skills interventions for children with autism spectrum disorders: A meta-analysis in single-case research using HLM. *Research in Autism Spectrum Disorders, 5*(1), 562-569.

*Wehman, P., & Marchant, J. A. (1978). Reducing multiple problem behaviours in a profoundly retarded child. *British Journal of Social & Clinical Psychology, 17*(2), 149-152.

*Weeks, M., & Gaylord-Ross, R. (1981). Task difficulty and aberrant behavior in severely handicapped students. *Journal of Applied Behavior Analysis, 14*(4), 449-463.

White, H. D. (2009). *The handbook of research synthesis and meta-analysis.* New York: Russell Sage Foundation.

Zhan, S., & Ottenbacher, K. J. (2001). Single-subject research designs for disability research. *Disability and Rehabilitation, 23*, 1-8.

Zhang, D., Katsiyannis, A., & Herbst, M. (2004). Disciplinary exclusions in special education: A 4-year analysis. *Behavioral Disorders, 29*, 337–347.

Appendix A.

DATA EXTRACTION RELIABILITY CHECK

Directions: Please, read the summary provided on the GoogleDoc for the identified time series and the article and highlight yes or no. If the answer is no, then write in the column (new description).

Time Series Identifying Information for Secondary Coding:

Study code:_____ Secondary Evaluator:_____

Time Series _____ (to be filled in by Primary Evaluator)

Behavior _____ (to be filled in by Primary Evaluator)

Is this accurate information for:	**Answer:**	**New Code if, answer no:**
1. Age?	Yes No	
2. Grade range?		
2. Gender?	Yes No	
3. Specific disability?	Yes No	
4. Cognitive status?	Yes No	
5. Level of verbal communication ability?	Yes No	
6. Classroom setting of participant?	Yes No	
7. Intervention type?	Yes No	
8. Agent?	Yes No	
9. Setting?	Yes No	
10. Format?	Yes No	
11. Duration?	Yes No	
12. Presence of FBA?	Yes No	

Data Points fill in with numbers:

1st Baseline:

1st Treatment:

2nd Baseline:

2nd Treatment:

3rd Baseline:

3rd Treatment:

4th Baseline:

IF YES TO NUMBER 12 answer 13-16, if not skip to 17		
13. FBA method?	Yes No	
14. FBA agent?	Yes No	
15. FBA setting?	Yes No	
16. FBA team decision-making?	Yes No	
17. Techniques for generalization?	Yes No	
18. School-wide positive behavioral support tier?	Yes No	
19. Type of challenging behavior	Yes No	
20. Intervention fidelity measures?	Yes No	
21. Social validity measures?	Yes No	
22.Published/unpublished	Yes No	
23. Inter-rater reliability data?	Yes No	
24. Type of SCD?	Yes No	

Appendix B.

QUALITY INDICATORS FOR DECISION MAKING FOR THE DESIGN QUALITY OF A SCD

*information from Kratochwill, et al., 2010

	Meet	With Reservation	Does not meet	Notes
1.The independent variable (i.e., the intervention) must be systematically manipulated, with the researcher determining when and how the independent variable conditions change		N/A		
2. Each outcome variable must be measured systematically over time by more than one assessor, and the study needs to collect inter-assessor agreement in each phase and on at least twenty percent of the data points in each condition (e.g., baseline, intervention) and the inter-assessor agreement must meet minimal thresholds.				
a. measured by more than one assessor?		n/a		any one of these are a no, then study does not meet
b. IRR each phase?		n/a		
c. IRR 20% of each condition?		n/a		
d. meet minimum thresholds (.8-.9 for percentage, .6 for cohen's kappa)		n/a		

3. The study must include at least three attempts to demonstrate an intervention effect at three different points in time or with three different phase repetitio		n/a		Examples of designs meeting this standard include **ABAB designs, multiple baseline** designs with at least three baseline conditions, **alternating/simultaneous treatment** designs with either at least three alternating treatments compared with a baseline condition or two alternating treatments compared with each other, **changing criterion designs** with at least three different criteria, and more complex variants of these designs. Examples of designs not meeting this standard include AB, ABA, and BAB designs.10
4.) For a phase to qualify as an attempt to demonstrate an effect, the phase must have a minimum of three data points		n/a		4.1 Meet Standards a reversal /withdrawal (e.g., ABAB) design must have a minimum of four phases per case with at least 5 data points per phase.
	n/a			4.2 To Meet Standards with Reservations a **reversal /withdrawal** (e.g., ABAB) design must have a minimum of four phases per case with at least 3 data points per phase. Any phases based on fewer than three data points cannot be used to demonstrate existence or lack of an effect
		n/a		4.1 To Meet Standards a **multiple baseline design** must have a minimum of six phases with at least 5 data points per phase.
	n/a			4.2 To Meet Standards with Reservations a **multiple baseline design** must have a minimum of six phases with at least 3 data points per phase. Any phases based on fewer than three data points cannot be used to demonstrate existence or lack of an effect
		n/a		4.1 An **alternating treatment design** needs five repetitions of the alternating sequence to Meet Standards. Designs such as ABABBABAABBA, BCBCBCBCBC, and AABBAABBAABB would qualify, even though randomization or brief functional assessment may lead to one or two data points in a phase.

	n/a			**4.2 alternating treatment design** :A design with four repetitions would Meet Standards with Reservations, and a design with fewer than four repetitions Does Not Meet Standard

Quality Indicator									
Study	Type of SCD	1	2a.	2b	2c	2d	3	4.1	4.2