ASSESSMENT AND TREATMENT OF MALADAPTIVE BEHAVIORS WITHIN THE DEVELOPMENTAL DISABILITY LITERATURE: IMPACT ON OUTCOME VARIABLES

by

CAITLIN V. HERZINGER

(Under the Direction of Jonathan M. Campbell)

ABSTRACT

Several meta-analyses relating to the assessment and treatment of severe maladaptive behaviors within the developmental disability research were reviewed. Following a synthesis of the meta-analyses results, three common themes were found and discussed: (a) behavioral treatment constitutes the large majority of published treatment outcomes, (b) treatments employing different functional assessments produced different outcomes, and (c) few characteristics of individuals were identified as mediating or moderating factors in treatment effectiveness. This current quantitative review of developmental disability research aims to make comparisons of different functional behavioral assessment methodologies, both across and within diagnostic categories. Quantitative synthesis data were used to answer questions regarding behavioral function, assessment type, differences based upon diagnostic category, and treatment effectiveness. Results indicate that assessment methodology does not impact treatment effectiveness, but both identified functions and treatment effectiveness are impacted by diagnosis. Implications for clinicians as well as future research directions are also discussed.

INDEX WORDS: developmental disability, maladaptive behaviors, functional assessment, functional analysis, meta-analysis

ASSESSMENT AND TREATMENT OF MALADAPTIVE BEHAVIORS WITHIN THE DEVELOPMENTAL DISABILITY LITERATURE: IMPACT ON OUTCOME VARIABLES

by

CAITLIN V. HERZINGER

B.S, University of Florida, 2001

M.A., University of Georgia, 2005

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2008

ASSESSMENT AND TREATMENT OF MALADAPTIVE BEHAVIORS WITHIN THE DEVELOPMENTAL DISABILITY LITERATURE: IMPACT ON OUTCOME VARIABLES

by

CAITLIN V. HERZINGER

Major Professor: Jonathan M. Campbell

Committee: David L. Gast
Shawn M. Glynn
A. Michele Lease

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
May 2008

ACKNOWLEDGEMENTS

I would like to acknowledge many people for helping me throughout this process. I would especially like to thank my advisor, Jon Campbell, for his generous time and commitment. Throughout my graduate school experience he has supported and encouraged independent thinking, following new lines of research, and, despite my rash dismissals, a career in academia.

I am also grateful for having an exceptional doctoral committee and wish to thank David Gast, Michele Lease, and Shawn Glynn for their continual support. I am also indebted to Alice Shillingsburg and Nate Call, faculty at the Marcus Institute, for taking the time to provide thoughtful advice and encouragement.

I would like to thank my family and friends for their encouragement throughout the dissertation process. Thanks to both my mother and father for their helpful critiques and firm belief in me. Finally, I owe a special thanks to Jeff Delfs for his understanding, humor, and patience.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

GENERAL INTRODUCTION

The assessment and treatment of maladaptive behaviors, specifically within the developmental disability population, is an important area of research. Maladaptive behaviors, which can include self-injurious behavior [SIB] (e.g., self-hitting, head banging, eye poking), stereotypic behaviors (e.g., body rocking, hand flapping), property destruction, aggression towards others, and severe disruptions (e.g., tantrums), are often difficult to treat. Aside from tissue damage and property destruction, maladaptive behaviors can interfere with new learning, prevent acquisition of adaptive behavior and compete with socially acceptable behaviors. In addition, severe behavior problems often lead to high levels of stress for caregivers (Lecavalier, Leone & Wiltz, 2006). These factors, in turn, will have a negative impact on social relationships, academic performance, and overall personal growth for individuals who engage in maladaptive behaviors. For example, caregiver stress associated with parenting adolescents with autism has been shown to predict increased levels of maladaptive behavior and more severe symptoms of autism above and beyond prior levels of maladaptive behavior (Greenberg, Seltzer, Hong, & Orsmond, 2006). Epidemiological studies suggest that 13-30% of children with autism engage in problematic behaviors so severe that intervention is warranted (Horner, Carr, Strain, Todd & Reed, 2002). According to Wicks-Nelson and Israel (1999), 15% of developmentally disabled children engage in SIB.

One component of the assessment and treatment of maladaptive behaviors is functional assessment. The goal of functional assessment is to establish the purpose or function of the

behavior in order to generate hypotheses about how to approach treatment. There are several different functional assessment methodologies, but they can be categorized as being experimental or non-experimental in nature. Several researchers have looked at the differences in these methodologies and their effects on (a) treatment effectiveness (Campbell, 2003; Didden, Korzilius, van Oorsouw, & Sturmey, 2006; Newcomer & Lewis, 2004; Herzinger & Campbell, 2007), (b) cost-benefit analysis (Applegate, Matson, & Cherry, 1999; Sturmey, 1995), (c) experimental design (Sturmey, 1995) and (d) time and training required (Northup et al., 1991; Kern, Hilt, & Gresham, 2004). Other research foci include participant and study characteristics that may mediate or moderate treatment effects.

Meta-analyses have become the standard method for summarizing research findings in many scientific areas. “Meta-analysis refers to the analysis of analyses...the statistical analysis of a large collection of analysis results from individual studies for the purpose of integrating the findings” (Glass, 1976, p. 3). A meta-analysis is the process of compiling the results of multiple research studies and combining the findings into one summative product. Typically, one or more statistical techniques are applied to the new dataset to integrate the findings of the individual studies and establish common findings throughout the literature. Meta-analyses are used to make comparisons across several studies that examine similar constructs or relationships. Quantitative research methods such as statistical significance and effect size are often used to make summative comments about the larger, compiled datasets. By using standardized statistical measurement across primary studies, meta-analyses present a more replicable and dependable outcome than narrative reviews of the literature.

However, the summative collections must be assessed with a critical eye, not simply accepted at face value. Mostert (2003) noted that variation in the amount of reported detail in

each study may affect the judgments of face validity and decisions for replication. Because meta-analyses are a collection of independent, empirical research studies, the results are highly dependent on the clarity and specificity of other authors. Several researchers have previously commented on this issue. For example, Fisher, Piazza and Hanley (1998) noted some issues with overlapping data without proper notations in subsequent publications. Multiple presentations of the same data set in different articles without proper notations are especially of interest when that data is used in meta-analysis.

A quantitative synthesis of existing related meta-analyses or "mega-analysis" is another way to glean information from a large database. Reviewing related meta-analyses in a mega-analysis can be used to summarize findings across different research areas and knowledge domains (Mostert, 2003). Several mega-analyses exist in the general area of special education, specifically in learning disabilities (Mostert, 1996), emotional and behavioral disorders (Mostert, 2001), and treatment in special education (Lipsey & Wilson, 1993). The current review includes meta-analyses that focus on severe problem behaviors within the developmental disabilities literature.

I have chosen the two-paper option for my dissertation. In the first paper, I have reviewed research in the area of assessment and treatment of problem behaviors within the developmental disability research. Also, I reviewed relevant, published meta-analyses to determine prominent themes within the literature. The review of meta-analyses takes the form of a mega-analysis so that, when possible, quantifiable comparisons can be made between different studies. Gaps in the existing literature discovered during the literature review guide the meta-analytic review.

In the second paper, following a brief review of the overarching focus (i.e., assessment and treatment of problem behaviors within the developmental disabilities population), I have

attempted to address some of the gaps in the literature. The second paper includes a description of the experimental design, methodology, and results of the meta-analytic review. Limitations of the current research as well as future directions for research and current implications for practitioners and researchers are also included.

REFERENCES

Applegate, H., Matson, J.L., & Cherry, K.E. (1999). An evaluation of functional variables affecting severe problem behaviors in adults with mental retardation by using the Questions about Behavioral Function scale (QABF). *Research in Developmental Disabilities, 20,* 229-237.

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.

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 on Mental Retardation, 111,* 290-298.

Fisher, W.W., Piazza, C.C., & Hanley, G.P. (1998). Informing readers of the presence of data common to multiple investigations. *Journal of Applied Behavior Analysis, 31,* 703-704.

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

Greenberg, J.S., Seltzer, M.M., Hong, J., & Orsmond, G.I. (2006). Bidirectional effects of expressed emotion and behavior problems and symptoms in adolescents and adults with autism. *American Journal on Mental Retardation, 111,* 229-249.

Herzinger, C.V., & Campbell, J.M. (2007). Comparing functional assessment methodologies: A quantitative synthesis. *Journal of Autism and Developmental Disorders, 37,* 1430-1445.

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 Disabilities, 32,* 423-446.

Kern, L., Hilt, A.M. & Gresham, F. (2004). An evaluation of the functional behavioral assessment process used with students with or at risk for emotional and behavioral disorders (2004). *Education and Treatment of Children, 27,* 440-451.

Lecavalier, L., Leone, S., & Wiltz, J. (2006). The impact of behaviour problems on caregiver stress in young people with autism spectrum disorders. *Journal of Intellectual Disability Research, 50,* 172-183.

Lipsey, M.W. & Wilson, D.B. (1993). The efficacy of psychological, educational, and behavioral treatment: Confirmation from meta-analysis. *American Psychologist, 48,* 1181-1209.

Mostert, M.P. (1996). Reporting meta-analyses in learning disabilities. *Learning Disabilities Research and Practice, 11,* 2-14.

Mostert, M.P. (2001). Characteristics of meta-analyses reported in mental retardation, learning disabilities, and emotional and behavioral disorders. *Exceptionality, 9,* 199-225.

Mostert, M.P. (2003). Meta-analyses in mental retardation. *Education and Training in Developmental Disabilities, 38,* 229-249.

Newcomer, L.L. & Lewis, T.J. (2004). Functional behavioral assessment: An investigation of assessment reliability and effectiveness of function-based interventions. *Journal of Emotional and Behavioral Disorders, 12,* 168-181.

Northup, J., Wacker, D., Sasso, G., Steege, M., Cigrand, K., Cook, J., & DeRaad, A. (1991). A brief functional analysis of aggressive and alternative behavior in an outclinic setting. *Journal of Applied Behavior Analysis, 24,* 509-522.

Sturmey, P. (1995). Analog baselines: A critical review of the methodology. *Research in Developmental Disabilities, 16,* 269-284.

Wicks-Nelson, R., & Israel, A.C. (1999). Autism and schizophrenia. In *Behavior Disorders of Early Childhood* (pp. 330-363). Upper Saddle River, N.J.: Prentice Hall.

CHAPTER 2

ASSESSMENT AND TREATMENT OF MALADAPTIVE BEHAVIORS WITHIN THE DEVELOPMENTAL DISABILITIES LITERATURE: A MEGA-ANALYSIS

Herzinger, C.V. and J.M. Campbell. To be submitted to *Research in Developmental Disabilities*.

Abstract

Several meta-analyses relating to the assessment and treatment of severe maladaptive behaviors within the developmental disability research were reviewed. Following a synthesis of the meta-analyses results, three common themes were found and discussed: (a) behavioral treatment constitutes the large majority of published treatment outcomes, (b) treatments employing different functional assessments produced different outcomes, and (c) few characteristics of individuals were identified as mediating or moderating factors in treatment effectiveness. A quantitative synthesis of the literature focused on the congruence of data within and among the separate empirical studies. The usefulness of the mega-analysis as a way to quantify related research was also discussed.

KEY WORDS: developmental disabilities, maladaptive behaviors, functional assessment methodology, mega-analysis, meta-analysis, quantitative synthesis

A Mega-Analysis of Assessment and Treatment of Maladaptive Behaviors

Within the Developmental Disability Population

INTRODUCTION

Autism Definition and Prevalence

Autism is a pervasive developmental disorder that is characterized by qualitative social impairment, communication delays, and restricted patterns of behavior. For persons with autism, social impairments are often typified by marked impairment in using and understanding nonverbal social cues, such as facial expressions and body posture. Communication delays noted in persons with autism can range from a total lack of developmentally appropriate verbal or nonverbal language to the inability to initiate or sustain a conversation with others. The restricted patterns of behavior associated with autism include those behaviors that are abnormal in intensity or focus as well as an inflexible adherence to nonfunctional routines and rituals. The onset of autism is prior to 3 years of age. The prevalence rate of autism is currently estimated at 10-16 cases per 10,000 with a male to female ratio of 1.95 to 5.5:1 (Fombonne, 2005). Often the cognitive profile of an individual with autism is skewed with verbal skills being weaker than nonverbal skills. In addition to the core features of autism, individuals with autism often show behavioral symptoms such as deficits in attention, hyperactivity, sensory integration difficulties, and temper tantrums (APA, 2000). One of the most consistently noted concerns for individuals with autism is their propensity to engage in problem behaviors, such as aggression towards self and others.

Intellectual Disability Definition and Prevalence

Intellectual Disability (ID), previously known as Mental Retardation (MR), is typified by "significantly subaverage intellectual functioning (an IQ of approximately 70 or below) with

onset before age 18 years and concurrent deficits or impairments in adaptive functioning" (DSM-IV-R, p.39). Adaptive functioning refers to an individual's daily life skills and their personal independence. ID occurs in approximately 1-3% of the population, depending on the criteria used to determine the prevalence rates. Four degrees of severity have been established to better explain the variability of ID: Mild, Moderate, Severe, and Profound. Mild ID refers to individuals whose IQ scores range from approximately 50-55 to 70-75 and accounts for 85% of the ID population. Moderate ID, with IQ levels ranging from 35-40 to 50-55, makes up 10% of the ID population. Severe ID, ranging from 20-25 to 35-40, accounts for 3-4% of the ID population. Profound ID refers to those individuals whose IQ scores are below 20-25 and occurs in only 1-2% of the ID population.

For approximately 25% of individuals with ID, the etiology of the disorder is known. Known etiologies can typically be grouped in one of the following categories: genetics (e.g., chromosomal abnormalities, heredity); environmental issues (e.g., prenatal/early environmental insult or deprivation); general medical conditions; and mental disorders (e.g., autism). No organic cause can be identified for almost 75% of the ID population. These cases of unknown origin tend to involve milder forms of retardation (Weiten, 1995). Comorbidity of psychiatric problems for individuals functioning in the range of ID is estimated to be three to four times greater than in the general population.

Prevalence of Maladaptive Behaviors for Individuals with Autism and ID

Individuals with autism may demonstrate a wide range of problematic behaviors, such as self-injurious behavior [SIB] (e.g., self-hitting, head banging, eye poking), stereotypic behaviors (e.g., body rocking, hand flapping), property destruction, aggression towards others, and severe disruptions (e.g., tantrums). Epidemiological studies suggest that 13-30% of children with autism

engage in problematic behaviors so severe that intervention is warranted (Horner, Carr, Strain, Todd, & Reed, 2002). According to Wicks-Nelson and Israel (1999), 15% of developmentally disabled children engage in SIB. Some research has indicated that a diagnosis of autism is a risk marker for several forms of challenging behaviors within the larger population of individuals with ID (McClintock, Hall, & Oliver, 2003). For example, individuals diagnosed with autism are significantly more likely to engage in self injury than individuals without an autism diagnosis (McClintock et al.). These data, although not comprehensive, highlight the importance of assessing and treating maladaptive behaviors. Aside from tissue damage and property destruction, maladaptive behaviors can interfere with new learning, prevent acquisition of adaptive behavior, and compete with socially acceptable behaviors. These factors in turn typically have a negative impact on social relationships, academic performance, and overall personal growth for individuals who engage in maladaptive behaviors.

Many specific types of ID are also associated with a varying array of severe problem behaviors. Cri Du Chat, a genetic abnormality that affects 1 in 50,000, is associated with self injury and repetitive movements (Cornish & Pigram, 1996). Prader-Willi, another chromosome deletion disorder that is typified by obesity and food obsessions, is also associated with self injury, as well as temper outbursts and repetitive speech (Clarke & Boer, 1998; Greaves et al., 2006). Self injury is also associated with Smith-Magenis syndrome (Clarke & Boer, 1998), Lesch-Nyhan disease (Dykens, Hodapp, & Finucane, 2000), Coffin-Lowery syndrome (Sivagamasundari et al., 1994) and Cornelia de Lange syndrome (Hyman, Oliver, & Hall, 2002).

ASSESSMENT AND TREATMENT OF MALADAPTIVE BEHAVIORS

There are several different approaches for assessing and treating maladaptive behaviors. The most commonly reported methods include those based on operant conditioning within a

behavioral framework. Other approaches include those based in psychoanalysis, sensorimotor integration, and psychopharmacology. Treatment based on the psychoanalytic and sensorimotor approaches were once popular, however, they were not found in the current review of the literature. Thus, only a description of the two most common (i.e., psychopharmacology and behavioral) approaches are discussed.

Psychopharmacological Approach

The psychopharmacological approach to the assessment and treatment of maladaptive behaviors in people with developmental disabilities targets specific behaviors for treatment, such as "resistance to change, ritualistic/compulsive behaviours, hyperactivity, aggressive behaviours and sleep problems" (Gringas, 2000, p. 229). Overactive dopaminergic activity is thought to be the cause of several problem behaviors, notably stereotypies and general overactivity, and is often treated with dopamine receptor blockers such as haloperidol, a traditional antipsychotic (Volkmar, 2001). There is also research regarding the role of serotoninergic dysfunction in the etiology of autism (Schultz & Anderson, 2004). Like dopamine, altering the levels of serotonin in the body has been shown to improve some symptoms of autism. Schultz and Anderson report that serotonin reuptake inhibitors have had a positive affect on reducing autistic symptoms. They also note that symptoms of autism have been exacerbated by depletion of tryptophan, a precursor to serotonin production. Stimulants, antidepressants, and melatonin have also been used to treat problem behaviors. According to Bryson, Rogers, and Fombonne, (2003) "There is no curative treatment for autism, and psychotropic drugs have only a minimal role to play in its management. As a rule, drugs should be used sparingly and only when other strategies to reduce

maladaptive behaviours have been properly tried and have failed to bring about the desired changes" (p. 512). In any case, careful clinical and laboratory monitoring is recommended when psychotropic drugs are part of the treatment package.

Behavioral Approaches

The behavioral approach to treating aberrant behaviors has consistently focused on identifying the antecedents and consequences of the target behavior and typically uses the operant conditioning approach during treatment. Several assessment methodologies are focused on assessing antecedents of behavior, such as that involved in structural analysis. In the case of structural analysis, antecedent variables, such as task difficulty and attention to task, are manipulated to alter target behavior rates (Repp & Horner, 1999). Other behavioral assessment and intervention methodologies are consequence driven (i.e., reinforcement or punishment based treatments). The interventions that stem from this type of assessment focus on altering the immediate consequences to aberrant behaviors in order to effect change upon the target behavior. Prior to the 1980's, aberrant behaviors were often treated with punishment, without concern for the contingencies maintaining the behavior. In an extensive review of the literature, Kahng, Iwata, and Lewin (2002) found that interventions were often implemented arbitrarily without regard to the contingencies that maintained the target behavior. Although punishment was a successful treatment strategy (i.e., reduced maladaptive behaviors), treatments based on punishment are controversial. Punishment-based interventions are thought to be effective because their aversive properties are so salient. This is in contrast to reinforcement-based procedures which aim to alter the existing reinforcement contingencies that maintain behavior (Pelios, Morren, Tesch & Axelrod, 1999). Also, in contrast to reinforcement-based interventions, punishment procedures do not typically produce changes in positive behavior. Many researchers

and clinicians are concerned that punishment is intrusive, unethical, and inhumane (Laski, 1987). Still others are concerned with the ability of punishment interventions to generalize and maintain positive behavior changes across settings and over time (Matson & Taras, 1989).

With the growth of research defining and describing functional assessment of aberrant behavior, the behavior analytic perspective has quickly become the "gold standard" for the assessment and treatment of maladaptive behaviors. Carr (1977) and Johnson and Baumeister (1978) were some of the first to document the usefulness of identifying the environmental events that maintain aberrant behaviors. Behavior analysts hold the view that all behavior responses have a function and are observable demonstrations of cause-effect relationships (Skinner, 1953). To date, a few functions of maladaptive behavior have been identified: (a) *escape* from academic tasks or disliked events; (b) *attention* from peers, parents, or teachers; (c) *tangible*–characterized as leading to a physical or food reinforcer (e.g., toy or cracker); or (d) *automatic* function, in which it is assumed that the behavior is nonsocially mediated or maintained independent of social consequences (e.g., self stimulation). The most current taxonomy of behavioral function focuses on three types of reinforcement as the major behavioral mechanisms maintaining behavior: (a) positive reinforcement, which refers to contingent access to tangibles, attention, and sensory stimuli, (b) negative reinforcement, which refers to escape from tangibles, attention, and sensory stimuli, and (c) automatic reinforcement, in which environmental variables do not affect behavior. The adoption of at least part of the behavioral approach by the federal government with the amendments to the Individuals with Disabilities Education Act [IDEA] (P.L. 105-117, 1997) emphasizes the importance of the behavioral approach to assessment and treatment of problematic behaviors (Gresham et al., 2004).

FUNCTIONAL BEHAVIORAL ASSESSMENT

Functional Behavioral Assessment (FBA) is an umbrella term used to describe any methodology in which the goal is to document the "function" of a behavior. Function can be defined as the "purpose" of any behavior and describes the contingencies that are antecedent and consequential to the behavior. Once function is determined, the results of the FBA are to be used in treatment selection for reducing maladaptive behaviors, rather than selecting treatments arbitrarily. For example, if the FBA indicates that an individual is engaging in an aberrant behavior (e.g., aggression) to gain attention, the treatment may include the introduction of an alternative way to gain attention (e.g., saying "Help") and extinction to decrease the likelihood that the aggressive behaviors are reinforced with attention. Kahng et al. (2002) compared the effectiveness of treatments based on FBAs to arbitrarily chosen ones and found treatments based on FBAs to be more effective. Newcomer and Lewis (2004) also noted therapeutic, decreasing trends in rates of maladaptive behaviors in 100% of participants during function-based interventions, as compared to decreases in behavior in only 1 of 3 participants during non-function-based interventions.

FBA encompasses *indirect assessments* (e.g., interviews, questionnaires, rating scales), *descriptive assessments* (e.g., A-B-C sheets, direct observation with no variable/environment manipulation); and *functional analyses* (FA), also known as experimental, "true," or traditional functional assessments (e.g., analogue conditions in which variables are systematically manipulated within an experimental design). For the purposes of this paper, we are using the term FA to describe all experimental analyses. The term Behavioral Assessment (BA) refers to those assessments which are non-experimental in nature and includes both indirect and descriptive assessments. FBAs have led to an increase in the precision and resulting effectiveness

of reinforcement-based interventions (Kahng et al., 2002). The selection of instructional practices such as functional communication training (FCT), an intervention based on differential reinforcement, has also increased following the introduction of behavioral function identification through FBAs.

FBAs have become foundational in the study and treatment of maladaptive behaviors. For example, Kahng et al. report that there has been a dramatic increase in the number of published data sets investigating FBAs since the early 1980's. The increase coincides with and is presumed to be a result of the increase of data sets that included a FA. Reports have documented the increased use of FBA methodologies in the assessment of maladaptive behavior across time in both clinical and research settings (Pelios, Tesch, Morren, & Axelrod, 1999; Horner, Carr, Strain, Todd, & Reed, 2002). Based on the ability of FBAs to ascribe function to aberrant behaviors and thus improve treatment effectiveness of those behaviors through treatment selection (Kahng et al., 2003; Campbell, 2003), FBAs have become an integral part of the entire treatment process. For example, the positive behavior support (PBS) literature focuses on the importance of strategies intended to prevent aberrant behaviors and concurrently teach appropriate, alternative behaviors (Mesibov, Browder, & Kirkland, 2002). As such, FBAs are considered a major component of all positive behavior support approaches (Kincaid, Knoster, Harrower, Shannon, & Bustamante, 2002). In 1997, the amendments to IDEA required the use of FBA and positive behavioral interventions based on these assessments when student placement may be changed due to problem behavior. Prior to this, these assessments were considered "best practices" but were not mandated by federal law (Horner & Carr, 1997; Sugai, Horner, & Sprague, 1999). IDEA, and IDEIA (P.L. 108-446, 2004) do not, however, define what constitutes a valid FBA. Due to the lack of detailed description, school districts and other

professionals use a wide variety of FBA methods. The current quantitative review of the literature aims to delineate the differences in accuracy and effectiveness of different FBA methodologies.

Functional Analysis (FA) and Behavioral Assessment (BA)

FA represents simulation of natural environments and is the primary tool for demonstrating causal relationships between observed behavior and environmental consequences (Carr, Langdon, & Yarborough, 1999). With FA, the goal is to isolate the function or purpose of the maladaptive behavior. In order to determine the functional properties of maladaptive behavior, the therapist directly and systematically alters the social and physical environment in a way that is likely to alter the frequency of the maladaptive behavior. Iwata, Dorsey, Slifer, Bauman, and Richman (1982/1994) defined the first standardized and comprehensive method for conducting a FA. As outlined by Iwata et al., there are four different conditions in a FA: social disapproval, academic demand, unstructured play, and alone. In the social disapproval condition, the client receives therapist attention contingent on the target behavior. The social disapproval condition is thought to mimic naturalistic occurrences where individuals may respond with attention when the target behavior is exhibited. Attention may include one or more of the following: emotional behavior, physical contact, verbal reprimand, and consolatory attention. All other behaviors, even those that are socially acceptable, are ignored. In the academic demand condition, the client is presented several tasks and permitted escape contingent on targeted behavior. The unstructured play condition serves as a control for the experimenter's presence and is conducted in a room filled with toys and other possibly stimulating materials. In the alone condition, the therapist is assessing the possibility that the target behavior is self-reinforcing. Through each condition of the FA, frequency data are collected. Typically, conditions are

presented more than once and in a random order to prevent order bias although modifications to Iwata et al.'s (1982/1994) original methodology exist. The frequency data are graphed and rate of behavior comparisons are made across condition types. If the rate of behavior is significantly higher in one condition, then it is assumed that the problem behavior is a function of the variable assessed in that condition.

BA represents the other two categories mentioned under the umbrella term FBA: indirect and descriptive assessments. Also, BAs can be considered formal (i.e., structured) or informal depending on how systematically the information is collected. BAs range from brief interviews with parents in which they are asked questions such as, "Why do you think the problem behavior is occurring?" to long, extended observations in which the subject is observed in a variety of settings. The Motivation Assessment Scale (MAS; Durand & Crimmins, 1988), Questions About Behavioral Function (QABF; Matson & Vollmer, 1995) and Motivation Analysis Rating Scale (MARS; Wieseler et al., 1985) are three recognized indirect methods of assessing behavioral function. The Functional Assessment Observation Form (FAOF; O'Neill et al., 1997) is a direct observation method of behavioral assessment. Some research has been reported comparing the comparability of these methods of assessment to FA. For example, MAS ratings have been correlated with FA analogue data and the relationship between the two sources was found to be highly significant ($r = .99$, $p < .001$) (Durand & Crimmins). Thus, "the teacher ratings on the MAS predicted their student's behavior in the experimental condition" (Durand & Crimmins, p. 112). In contrast, Hall (2005) found that descriptive and experimental methods of FBA agreed only 25% of the time. In almost all published accounts of comparison data, the FA represented the gold standard for validity tests of other types of assessment.

Due to the variability in BA methods (e.g., rating scales, observations, interviews), assessment outcomes and treatment recommendations may differ. Some BAs are simply used to gain qualitative information that is later combined with other sources of information to identify the function of maladaptive behavior. Other BAs, such as the rating scales, use cut scores or relative rankings to identify the functions of the targeted behavior. For example, the authors of the MAS, a 16-item questionnaire, note that if one condition (e.g., attention) has clearly received the highest score it is assumed that this is the function of the behavior. The relative rankings of conditions are used to determine the most important influence on the behavior (Durand & Crimmins, 1992). For the MAS, like all other measures of behavioral function, more than one possible function may be identified. The authors provide a guide to interpret most outcome results, including decision rules when two functions are scored within .25 to .50 points of each other. In this scenario, both categories would be considered influences that may be causing the maladaptive behavior to continue (Durand & Crimmins, 1992). Regardless of the method of interpretation, the results of BAs are used in the same manner as those from FAs. The qualitative data from interviews, rating scale results, and direct observations are often combined and, if possible, a single function of behavior is identified. Treatment selection is then based on the outcome of the gathered information.

Benefits of FA when compared to BA

Much research has been published regarding the effectiveness of FAs. FA has provided methods to test the notion that each behavior has a function that can be observed. FA is able to identify causal and maintaining factors of problem behavior and has led to the development of effective treatments for severe problem behaviors. FA is a specific and structured assessment method, in which the analogue conditions and the contingencies set in place for each condition

are defined a priori. The structure and conditional quality of the FA makes replication possible. FA allows the clinician/ practitioner to base treatment selection on an observable, recorded (not simply reported) function of behavior. Thus, the primary benefit of FA is that the methodology allows for causal inferences to be made regarding the purpose of the targeted maladaptive behavior.

In contrast, BA does not allow for systematic manipulation of environmental variables (antecedents or consequences) and involves a recording of what is seen (or noticed) in the interaction. Therefore, BA methodologies are correlational in nature. BA is less thorough and is more susceptible to data recorder biases. For example, a teacher completing a functional rating scale may be more likely to recall instances in which her attention was drawn from a target activity to a maladaptive behavior than an instance in which a child engaged in maladaptive behaviors followed by a break from difficult academic tasks.

In 1994, Horner wrote that although different situations may require different functional assessment procedures, experimental FA "will remain the expectation within published research" (p. 402). Although other procedures for FBA should be created, considered, and researched, the clinical standard should include the following four standards: (a) problem behaviors are operationally defined, (b) antecedent behaviors are identified, (c) hypotheses are developed concerning variables maintaining problem behaviors, and (d) direct observation data are collected (Horner).

Benefits of BA when compared to FA

According to research by Sturmey (1995), the experimental design of FA presents a number of procedural and psychometric problems. FAs are time consuming, require extensive training, and are not cost-effective. Matson, Bamburg, Cherry, & Paclawskyj (1999) report that a

single FA can require several hours daily over a two- to three-week-period. Iwata et al. (1994) summarized data from over 150 FAs and reported that the mean length of assessment was 26 sessions, which represented approximately 6.5 hours of direct observation, not including the time spent preparing for each condition, analyzing data, and creating a visual representation in the form of graphs. When assessing a severe behavior problem that is life threatening (e.g., head banging), this time constraint may have vital consequences and immediate attention may be necessary.

Thorough training is necessary for the service provider conducting the FA sessions as well as the behavior data specialists who record the antecedents, behaviors, and consequences of the interactions during the analogue sessions. Limitations in mental health funding have exposed the monetary constraints on using FA as an automatic prelude to behavioral treatment (Applegate, Matson, & Cherry, 1999). Also, the FA methodology described previously neglects to identify establishing operations and minimizes the relevance of understanding antecedent manipulations. For example, FAs do not identify antecedent events that are functional, yet temporally distant from the behavior (Horner, Day, & Day, 1997). Future modifications to FA methodology may involve collecting data on antecedent events; however, current BA practice is more likely to include information regarding antecedent events. Martin, Gaffan, and Williams (1999) also noted problems with poor test-retest reliability and generalizability of analogue conditions. Researchers have noted other limitations of FA procedures, such as collecting data within non-naturalistic environments and intentional increases in life threatening maladaptive behaviors (Dawson, Matson, & Cherry, 1998; Axelrod, 1987). These are valid reasons that may account for the data that suggests BAs are performed more often than FAs in educational settings (Kern, Hilt, & Gresham, 2004).

BAs, in contrast to FAs, do not require specialized, intensive training for observers and clinicians or significant amounts of time for administration. Also, BAs do not include environmental manipulations that intentionally increase the rates of maladaptive behaviors. BAs are naturalistic and do not require careful experimental design.

META-ANALYSIS AND MEGA-ANALYSIS

A meta-analysis is a type of research synthesis that quantifies the findings of primary studies and results in an integrative, quantitative review. The meta-analysis can be useful because it allows researchers to make comparisons of treatment outcomes within and across related disorders, such as autism and ID. Several meta-analyses exist that cover treatment outcomes for individuals with autism (Horner et al., 2002) and ID (e.g., Didden, Duker, & Korzilius, 1997; Mostert, 2003). The meta-analysis differs from a standard narrative or descriptive review in that direct comparisons can be made between the articles that employ single subject research designs. Meta-analyses of single subject research designs typically utilize different effect size calculations than those used with group design. Instead, more appropriate, non-regression efficacy indicators such as Mean Baseline Reduction (MBLR), Percentage of Zero Data (PZD), and Percentage of Non-overlapping data (PND), are typical (Campbell, 2003, Scotti et al., 1991; Wehmeyer, 1995).

The MBLR is calculated by subtracting the mean of treatment observations from the mean of baseline observations then dividing by the mean of baseline observations and multiplying by 100 (Campbell, 2003; Lundervold & Bourland, 1988; O'Brien & Repp, 1990). The PND statistic is calculated as the percentage of treatment data that did not overlap with baseline data points (Scruggs, Mastropieri, & Casto, 1987). If a baseline phase reports one or more data points of zero, then the same number of data points will be excluded in the treatment

phase prior to calculation of the PND (Didden, Duker & Korzilius, 1997). The PND can range from 0 to 100%. The PZD statistic is calculated by computing the percentage of data points that reached zero after locating the first intervention data point that reached zero. This first zero point is included (Scotti et al., 1991). The PZD score is considered a more stringent efficacy indicator as it requires target behaviors to remain at zero levels throughout treatment to be considered effective. Scotti, Evans, Meyer, and Walker (1991) and Campbell (2003) found PND and PZD to be independent indicators of treatment outcome. Several published studies have used these two statistics to measure the effectiveness of treatments (Didden, Korzilius, van Oorsouw, & Sturmey, 2006; Herzinger & Campbell, 2007). It is important to note that these statistics are not effect sizes as traditionally defined, that is, none report the relative standing of the average treatment point within a distribution of baseline data points. "True" effect sizes (e.g., Cohen's *d*) are difficult to use and interpret within a single-subject meta-analysis (Campbell, 2004).

A mega-analysis is another way to glean information from a large database. The mega-analysis differs from the meta-analysis in that it is a synthesis of existing meta-analyses rather than related primary research studies. The mega-analysis can be used to summarize findings across different, but related research domains (Mostert, 2003). Many mega-analyses have been conducted in the area of special education in general (Mostert, 1996; Mostert, 2001; & Lipsey & Wilson, 1993).

PURPOSE

In light of the growing number of published meta-analyses of treatment outcomes for individuals with autism and ID, the aim of the present literature review is to provide an organized and quantitative review of this literature. With this overall aim in mind, the current review consists of four purposes: (a) review content of meta-analyses in developmental

disabilities with respect to problem behaviors, (b) illustrate themes in research, (c) highlight gaps in current research body and (d) propose a study that will address gaps in the research literature.

METHOD

Predetermined criteria were established for inclusion in the current mega analysis and are reported as follows. Articles were included in this review if they met the following criteria: (a) a meta-analytic, review paper, (b) pertained to the assessment and/or treatment of problem behavior, (c) included participants with developmental disabilities (e.g., autism, ID) and (d) published in a peer-reviewed journal. Meta-analytic articles were found via computer searches of PsycLit, ERIC, etc. Also, reviews of relevant mega-analytic articles (e.g., Mostert, 2003) were searched for crossover meta-analysis articles for inclusion.

Articles that met some criteria (i.e., review of problem behaviors within developmental disability population) but did not meet others (i.e., not a meta-analysis; e.g., Hemmings et al., 2006; Pelios et al., 1999) were not included. Also, book chapters that met most inclusion criteria, such as Marquis et al. (2000) were excluded because they were not published in peer-reviewed journals. Once compiled, the meta-analyses were reviewed for consistent themes and gaps in the available literature.

Twenty-four meta-analytic studies were found that met the inclusion criteria. Table 2.1 highlights the significant variables coded and reported for each article and the major findings.

DESCRIPTIVE ANALYSES

For the purposes of the present review, the findings are organized according to three main themes: (a) the use and outcomes of different theoretical approaches to intervention, (b) the

frequency and impact of FBA inclusion in intervention process, and (c) the impact of participant characteristics as moderating variables in treatment effectiveness.

Approaches to Assessing and Treating Problem Behaviors

In general, the collected meta-analyses supported the use of the behavioral approach to assessment and treatment of severe problem behaviors within the developmental disability literature. Roberts et al. (2003) noted that comprehensive interventions based on social learning theory and applied behavior analysis show significant promise, especially for children with autism. Campbell (2003) found that behavioral treatments were significantly effective in reducing problem behaviors. Other studies indicated that interventions based in psychopharmacology were less effective than other treatments. For example, Didden et al. (1997), Lennox et al. (1988), and Scotti et al. (1991) found that medication based treatments were the least effective treatment types included in their studies. Although the findings overwhelmingly support the use of a behavioral approach, this approach was not flawless. Symons et al. (1999), for example, found that aversive procedures were more likely to be included in behavioral studies than in pharmacological studies. Also, fewer secondary measures, such as “quality of life” were included in behavioral studies (Symons et al., 1999).

For those studies in which indicators of efficacy (i.e., effect sizes) were reported based on treatment approach, Table 2.2 highlights the findings. Generally, higher effect sizes were reported for behavioral approaches to treatment (e.g., differential reinforcement) when compared to non-behavioral approaches (e.g., medication).

Inclusion of Functional Behavioral Assessment in Treatment Process

Several of the meta-analyses included in the mega-analysis reported on functional behavioral assessment and its impact upon the assessment and intervention of maladaptive

behaviors. Shogren et al. (2004) and Gresham et al. (2004) found that the use of FBAs on treatment selection did not have significant therapeutic effect on the reduction of problem behaviors in individuals with developmental disabilities. However, neither found the FBA-based treatments to be less effective than those selected using other criteria. Besides the effectiveness of treatments based on FBAs, the usefulness of FBAs in treatment selection was also addressed. Scotti et al. (1996) found that when FBAs were reported, treatment selection did not always correlate with ascribed function. In fact, in four out of 12 studies that included a functional behavioral assessment and implemented time-out as treatment, the ascribed function of the maladaptive behavior was escape—a clear mismatch between ascribed function and treatment selection. Many researchers such as Barlow, Hayes, and Nelson (1984), Repp, Felce, and Barton (1988) and Durand and Carr (1991) have concluded that using the results of an FBA in treatment selection is a core feature of the assessment process. Clearly, however, the matching of treatment to behavioral function is not always done.

Other researchers found that the use of FBAs in determining treatment selection had a significant, therapeutic impact on the rate of problem behaviors reported. Herzinger and Campbell (2007), Scotti et al. (1991), and Didden et al. (1997) found that conducting an FBA was significantly related to effectiveness, as assessed by two specific effectiveness indicators, the PZD and PND respectively. Campbell (2003) also found that pretreatment FBA was a significant variable moderating the success of treatment. Horner et al. (2002) and Sternberg et al. (1994) both supported the use of FA results as a significant component of any comprehensive behavioral intervention.

Table 2.3 summarizes the reported efficacy indicators (i.e., MBLR, PND, and PZD) for studies that compared the inclusion of FBAs in treatment selection versus those that did not

incorporate the use of FBAs. If specified and reported, the type of FBA included is also reported. If the level of experimental design was not reported (i.e., the authors used the general term functional assessment), the statistics are reported in the BA column. Across the three reported effect size statistics, FAs and BAs appear to produce similar means for PND (e.g., 71 for FA and 69 for BA). For the PZD statistic however, the means differ dramatically (e.g., 59 for FA and 35 for BA). Only one study included used the MBLR statistic to compare FBA methodologies, thus comparisons of FA and BA using MBLR are not appropriate.

Participant Characteristics as Moderators of Treatment Effectiveness

The results of the mega-analysis revealed variable results on the impact of participant characteristics on overall treatment effectiveness. For those who found the variables to be significant, there were contradictory results. For example, some meta-analyses reported that age of participant was a moderating factor. Matson and Gorman-Smith (1986) and Gorman-Smith and Matson (1985) found that older subjects showed greater treatment effects. Sternberg, Taylor and Babkie (1994) and Shogren, Faggella-Luby, Bae and Wehmeyer (2004) reported that interventions conducted with younger subjects were likely to be more effective, conclusions more in line with the overarching notion of the importance of early intervention. Other meta-analyses focused on the diagnosis of participants as it relates to treatment effectiveness and behavioral topography. McClintock, Hall and Oliver (2003), Gorman-Smith and Matson (1985), and Matson and Gorman-Smith (1986) all noted that level of functioning (i.e., range of ID) impacted responsiveness to treatment. McClintock et al. (2003) and Roberts, Mazzucchelli, Taylor and Reid (2003) also found a diagnosis of autism or developmental disability to be a risk marker for certain maladaptive behaviors, such as SIB, aggression, and disruptions.

In contrast, the meta-analytic results of Campbell (2003), Scruggs, Mastropieri, Cook and Escobar (1986), Scotti et al. (1991) and Wehmeyer (1995), found that participant characteristics such as diagnosis, gender, age, level of functioning, and targeted maladaptive behaviors did not moderate the effectiveness of treatment.

CONCLUSIONS

The mega-analysis presented was a collection of over 20 related meta-analyses, comprised of over 2,000 single subject research studies. Three main conclusions resulted from the mega-analysis: (a) most published single-subject research examines the effectiveness of behavioral approaches, (b) inclusion of pre-treatment FBA procedures results in more positive treatment outcomes, and (c) participant characteristics, such as age of participant or psychiatric diagnosis, do not appear to be strongly related to treatment outcomes.

The mega-analysis results indicate that the behavioral approach to assessment and treatment of maladaptive behaviors is more effective than non-behavioral approaches. The superiority of the behavioral approach over other approaches was found across three 'effect size' indicators (i.e., MBLR, PND, and PZD). As evidenced by all three efficacy indicator means, treatments based on the behavioral approach were more successful at reducing rates of maladaptive behaviors than those based on non-behavioral approaches, such as psychopharmacology. Treatment effectiveness has been described by many authors using a three- or four-point scale. For example, the MBLR has been evaluated on a three tier scale (0 = less than or equal to 50% reduction, 1 = between 51% and 74% reduction, 2 = between 75% and 100% reduction; Lundervold & Bourland, 1988). Others have described a range of effectiveness for the PND (i.e., Scruggs et al., 1986). Using this terminology to describe behavioral treatment

outcomes within the current dataset, the PND falls within the “fairly effective” range (Scruggs et al., 1986).

Several meta-analyses examined the impact of FBA on treatment effectiveness with mixed conclusions. Out of 11 meta-analyses that reported specifically about the inclusion of FBAs prior to treatment implementation, nine reported the benefits of including versus excluding pre-treatment FBA procedures. In general, the data suggest that FBAs are associated with greater treatment effectiveness, but the significance of that difference appears to depend on the effect size reported. For example, when comparing treatment outcomes for those that included a FBA and those that did not, the PND means were 70 and 64, respectively. When comparing treatment outcomes for those that include a FBA and those that did not, the PZD means were 47 and 39, respectively. This may suggest a difference in the target of treatment, as PND is a measure of behavior reduction and PZD is a measure of behavioral suppression. More importantly, the results indicate that even when FBAs are conducted, the results of such assessments are not always used in the selection of treatment. This could be one explanation for the contradictory results. Comparisons of FBA methodology (i.e., FA versus BA) were more straightforward. For all three effect size means, interventions that included FAs were more effective than those that included a BA as part of the intervention process. For example, for the MBLR, the means were 83 (FA) and 77 (BA). For the PND, the means were 72 (FA) and 70 (BA). The greatest difference in treatment outcome when comparing FBA methodology was seen when using the PZD statistic with means of 59 (FA) and 35 (BA).

When reviewing the information regarding participant characteristics as moderators of treatment effectiveness, again, the synthesis provided inconclusive results. Some meta-analyses reported participant characteristics, such as age and level of cognitive ability, as moderating

treatment effectiveness. Others reported no significant relationship between participant characteristics and treatment outcome.

Recommendations for Future Research

The mega-analysis findings indicate that there are several gaps within the treatment literature targeting the reduction of maladaptive behavior for individuals with developmental disabilities. One notable aperture within the literature is that there is no direct comparison of FBA methodologies across diagnostic categories (such as autism and ID). With the federal government's support of functional assessment technology and use through legislation (i.e., IDEA and IDEIA), a thorough comparison of the effectiveness of FBA for different diagnostic populations is necessary. Specifically, there is no existing quantitative review comparing FBA methodologies for the ID population. Also, comparisons of different types of experimental FAs would also be useful. For example, in the current study all methods of experimental assessments were subsumed under the FA category. Brief FAs as described by Northup et al. (1991) and later by Derby et al. (1992) were categorized with full-length FAs and were not directly compared with other, less time intensive assessment methodologies. Future research may include a direct comparison between specific subtypes of FAs to BAs. Another line of research that would yield comparable results would be to design a single subject study of original data collection in which comparisons are made for individuals who have been administered a multitude of assessments (e.g., an interview, MAS rating scale, brief FA, and extended FA).

REFERENCES

Allison, D.B., Faith, M.S., & Franklin, R.D. (1995). Antecedent exercise in the treatment of disruptive behavior: A meta-analytic review. *Clinical Psychology: Science and Practice*, 2, 279-303.

Allison, D.B. & Gorman, B.S. (1993). Calculating effect sizes for meta-analysis: The case of the single case. *Behaviour Research and Therapy, 31,* 621-631.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders: Fourth Edition Text Revision.* Washington, DC: American Psychiatric Association.

Applegate, H., Matson, J.L., & Cherry, K.E. (1999). An evaluation of functional variables affecting severe problem behaviors in adults with mental retardation by using the Questions about Behavioral Function scale (QABF). *Research in Developmental Disabilities, 20,* 229-237.

Axelrod, S. (1987). Functional and structural analyses of behavior: Approaches leading to reduced use of punishment procedures. *Research in Developmental Disabilities, 8,* 165-178.

Barlow, D.H., Hayes, S.C., & Nelson, R.O. (1984). *The scientist-practitioner: Research and accountability in clinical and educational settings.* New York: Plenum.

Bryson, S.E., Rogers, S.J., & Fombonne, E. (2003). Autism spectrum disorders: Early detection and intervention, education, and psychopharmacological management. *Canadian Journal of Psychiatry, 48,* 506 –16.

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.

Campbell, J.M. (2004). Statistical comparison of four effect sizes for single-subject designs. *Behavior Modification, 28*, 234-246.

Carr E.G., (1977). The motivation of self-injurious behavior: A review of some hypotheses. *Psychological Bulletin, 84*, 800-816.

Carr, E.G., Langdon, N.A., & Yarborough, S.C., (1999). Hypothesis-based intervention for severe problem behavior. In A.C. Repp and R.H. Horner (Eds), *Functional Analysis of Problem Behavior.* Belmont, CA: Wadsworth Publishing Company, (9-31).

Clarke, D.J. & Boer, H. (1998). Problem behaviors associated with Deletion Prader-Willi, Smith-Magenis, and Cri Du Chat syndromes. *American Journal on MentalRetardation, 103*, 264-271

Cornish, K.M. & and Pigram, J. (1996). Developmental and behavioural characteristics of Cri du Chat syndrome. *Archives of Disease in Childhood, 75*, 448-450.

Dawson, J.E., Matson, J.L., & Cherry, K.E. (1998). An analysis of maladaptive behaviors in persons with autism, PDD-NOS, and mental retardation. *Research in Developmental Disabilities, 19*, 439-448.

Derby, K.M., Wacker, D.P., Sasso, G., Steege, M., Northup, J., Cigrand, K., & Asmus, J. (1992). Brief functional assessment techniques to evaluate aberrant behavior in an outpatient setting: A summary of 79 cases. *Journal of Applied Behavior Analysis, 25*, 713-721.

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

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,* 251-264.

Durand, V.M. & Crimmins, D.B. (1988). Identifying the variables maintaining self-injurious behavior. *Journal of Autism and Developmental Disorders, 18,* 99-117.

Durand, V.M., & Crimmins, D.B. (1992). The Motivation Assessment Scale (MAS) administration guide. Topeka, KS: Monaco and Associates.

Dykens E.M., Hodapp R.M. & Finucane B.M. (2000). *Genetics and Mental Retardation Syndromes: A New Look at Behavior and Interventions*. Paul H. Brookes Publishing, Baltimore, MD.

Emerson, E., Kiernan, C., Alborz, A., Reeves, D., Mason, H., Swarbrick, R., Mason, L., & Hatton, C. (2001). The prevalence of challenging behaviors: A total population study. *Research in Developmental Disabilities*, 22, 77-93.

Faith, M.S., Allison, D.B., & Gorman, B.S. (1996). Meta-analysis of single-case research. In R.D. Franklin, Allison, D.B., & Gorman, B.S. (Eds.), *Design and analysis of single-case research* (pp. 245-277). Hillsdale, NJ: Lawrence Erlbaum.

Fombonne, E. (2005). Epidemiology of autistic disorder and other pervasive developmental disorders. *Journal of Clinical Psychiatry, 66,* 3-8.

Gorman-Smith, D. & Matson, J.L. (1985). A review of treatment for self-injurious and stereotyped responding. *Journal of Mental Deficiency Research, 29,* 295-308.

Greaves, N., Prince, E., Evans, D.W., & Charman, T. (2006). Repetitive and ritualistic behaviour in children with Prader–Willi syndrome and children with autism. *Journal of Intellectual Disability Research, 50,* 92-100.

Gresham, F.M., McIntyre, L.L., Olson-Tinker, H., Dolstra, L., McLaughlin, V., & Van, M. (2004). Relevance of functional behavioral assessment research for school-based interventions and positive behavioral support. *Research in Developmental Disabilities, 25,* 19-37.

Gringas, P. (2000). Practical paediatric psychopharmacological prescribing in autism: The potential and the pitfalls. *Autism,* 4, 229-247.

Hall, S. S. (2005). Comparing descriptive, experimental, and informant-based assessments of problem behaviors. *Research in Developmental Disabilities, 26,* 514-526.

Hemmings, C. P.; Gravestock, S.; Pickard, M.; Bouras, N. (2006). Psychiatric symptoms and problem behaviours in people with intellectual disabilities. *Journal of Intellectual Disability Research, 50,* 269-276.

Herzinger, C.V., & Campbell, J.M. (2007). Comparing functional assessment methodologies: A quantitative synthesis. *Journal of Autism and Developmental Disorders, 37,* 1430-1445.

Horner, R.H. (1994). Functional assessment: Contributions and future directions. *Journal of Applied Behavior Analysis, 27,* 401-404.

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.

Horner, R.H., & Carr, E.G. (1997). Behavioral support for students with severe disabilities: functional assessment and comprehensive intervention. *Journal of Special Education, 31,* 84-104.

Horner, R.H., Day, H.M., & Day, J.R. (1997). Using neutralizing routines to reduce problem behaviors. *Journal of Applied Behavior Analysis, 30,* 601-614.

Hyman, P., Oliver, C., & Hall, S. (2002). Self-injurious behavior, self-restraint, and compulsive behaviors in Cornelia de Lange syndrome. *American Journal on Mental Retardation, 07,* 146-154.

Iwata, B.A., Dorsey, M.F., Slifer, K.J., Bauman, K.E. & Richman, G.S. (1982). Toward a functional analysis of self-injury. *Analysis and Intervention in Developmental Disabilities, 2,* 3-20.

Iwata, B.A., Pace, G.M., Dorsey, M.F., Zarcone, J.R., Vollmer, T.R., Smith, R.G., Rodgers, T.A., Lerman, D.C., Shore, D.A., Mazaleski, J.L., Goh, H., Cowdery, G.E., Kalsher, M.J., McCosh, K.C., & Willis, K.D. (1994). The functional analysis of self-injurious behavior: An experimental-epidemiological analysis. *Journal of Applied Behavior Analysis, 27,* 215-240.

Johnson, W.L., & Baumeister, A. (1978). Self-injurious behavior: A review and analysis of methodological details of published studies. *Behavior Modification, 2,* 465-484.

Kahng, S., Iwata, B.A., & Lewin, A.B. (2002). Behavioral treatment of self-injury, 1964-2000. *American Journal on Mental Retardation, 107,* 212-221.

Kahng, S., Iwata, B.A., & Lewin, A.B. (2002). The impact of functional assessment on the treatment of self-injurious behavior. In S.R. Schroeder, M.L. Oster-Granite, & Thompson

(Eds.), *Self-injurious Behavior: Gene-brain-behavior Relationships* (p. 119-131). Washington, DC: American Psychological Association.

Kern, L., Hilt, A.M. & Gresham, F. (2004) An evaluation of the functional behavioral assessment process used with students with or at risk for emotional and behavioral disorders. *Education and Treatment of Children, 27,* 440-451.

Kincaid D., Knoster T., Harrower J.K., Shannon P., & Bustamante S. (2002). Measuring the impact of positive behavior support. *Journal of Positive Behavior Interventions, 4,* 109-117.

LaGrow, S.J. & and Repp, A.C. (1984). Stereotypic Responding: A review of intervention research. *American Journal on Mental Deficiency, 88.* 595-609.

Lennox, D.B., Miltenberger, R.G., Spengler, P. & Erfanian, N. (1988). Decelerative treatment practices with persons who have mental retardation: A review of five years of the literature. *American Journal on Mental Retardation, 92,* 492-501.

Lundervold, D., & Bourland, G. (1988). Quantitative analysis of treatment of aggression, self-injury, and property destruction. *Behavior Modification, 12,* 590-617.

Marquis, J.G., Horner, R.H., Carr, E.G., Turnbull, A.P., Thompson, M., Behrens, G.A., Magito-McLaughlin, D., McAtee, M.L., Smith, C.E., Ryan, K.A., & Doolabh, A. (2000). A Meta-analysis of Positive Behavior Support. In R. Gersten, E.P. Schiller & S. Vaughn (Eds.), *Contemporary Special Education Research: Syntheses of the Knowledge Base on Critical Instructional Issues* (pp.137-178). Mahwah, NJ: Lawrence Erlbaum Associates.

Martin, N.T., Gaffan, E.A., & Williams, T. (1999). Experimental functional analyses for challenging behavior: a study of validity and reliability. *Research in Developmental Disabilities,* 20, 125-146.

Matson, J.L. & Taras, M.E. (1989). A 20 year review of punishment and alternative methods to treat problem behaviors in developmentally delayed persons. *Research in Developmental Disabilities, 10,* 85-104.

Matson, J.L., Benavidez, D.A., Stabinsky-Compton, L.S., Paclawskyj, T., & Baglio, S. (1996). Behavioral treatment of autistic persons: A review of research from 1980 to the present. *Research in Developmental Disabilities, 17,* 433-465.

Matson, J.L., & Vollmer, T.R. (1995). *User's Guide: Questions About Behavioral Function (QABF).* Baton Rouge, LA: Scientific Publishers, Inc.

Matson, J.L., Bamburg, J.W., Cherry, K.E., & Paclawskyj, T.R. (1999). A validity study on the Questions about Behavioral Function (QABF) scale: predicting treatment success for self-injury, aggression, and stereotypies. *Research in Developmental Disabilities,* 20, 163-176.

Matson, J.L. & Gorman-Smith, D. (1986). A review of treatment research for aggressive and disruptive behavior in the mentally retarded. *Applied Research in Mental Retardation, 7,* 95-103.

McClintock, K., Hall, S., & Oliver, C. (2003). Risk markers associated with challenging behaviours in people with intellectual disabilities: A meta-analytic study. *Journal of Intellectual Disability Research, 47,* 405-416.

Mesibov, G.B., Browder, D.M., & Kirkland, C. (2002). Using individualized schedules as a component of positive behavioral support for students with developmental disabilities. *Journal of Positive Behavior Interventions, 4,* 73-79.

Mostert, M.P. (2003). Meta-analyses in mental retardation. *Education and Training in Developmental Disabilities, 38,* 229-249.

Newcomer, L.L. & Lewis, T.J. (2004). Functional behavioral assessment: An investigation of assessment reliability and effectiveness of function-based interventions. *Journal of Emotional and Behavioral Disorders, 12,* 168-181.

Northup, J., Wacker, D., Sasso, G., Steege, M., Cigrand, K., Cook, J., & Deraad, A. (1991). A brief functional analysis of aggressive and alternative behavior in an outclinic setting. *Journal of Applied Behavior Analysis, 24,* 509-522.

O'Brien, S., & Repp, A.C., (1990). Reinforcement-based reductive procedures: A review of 20 years of their use with persons with severe to profound mental retardation. *Journal of the Association for Persons with Severe Handicaps, 15,*148-159.

O'Neill, R.E., Horner, R.H., Albin, R.W., Storey, K., Sprague, J.R., & Newton, J.S. (1997). *Functional assessment of problem behavior: A practical assessment guide.* Pacific Grove, CA: Brooks/Cole.

Pelios, L., Morren, J., Tesch, D. & Axelrod, S. (1999). The impact of functional analysis methodology on treatment choice for self-injurious and aggressive behavior. *Journal of Applied Behavioral Analysis, 32,* 185-195.

Piazza, C.C., Hanley, G.P., & Fisher, W.W. (1996). Functional analysis of cigarette pica. *Journal of Applied Behavioral Analysis, 29,* 437-439.

Repp, A.C., Felce, D., & Barton, L.E. (1988). Basing the treatment of stereotypic and self-injurious behaviors on hypotheses of their causes. *Journal of Applied Behavior Analysis, 21*, 281-289.

Roberts, C., Mazzucchelli, T., Taylor, K., & Reid, R. (2003). Early intervention for behaviour problems in young children with developmental disabilities. *International Journal of Disability, Development, and Education, 50*, 275-292.

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*, 233-256.

Scotti, J.R., Ujcich, K.J., Weigle, K.L., Holland, C.M., & Kirk, W. (1996). Interventions with challenging behavior of persons with developmental disabilities: A review of current research practices. *Journal of the Association for Persons with Severe Handicaps, 21*, 123-34.

Scruggs, T.E., Mastropieri, M.A., Cook, S.B., & Escobar, C. (1986). Early intervention for children with conduct disorders: A quantitative synthesis of single-subject research. *Behavioral Disorders, 11*, 260-271.

Scruggs, T.E., Mastropieri, M.A., & Casto, G. (1987). The quantitative synthesis of single-subject research: Issues and applications. *Behavior Modification, 22*, 221- 242.

Schultz, R.T., & Anderson, G.M. (2004). The neurobiology of autism and the pervasive developmental disorders. In D.S. Charmey & E.J. Nestler (Eds.) *Neurobiology of Mental Illness 2nd edition.* Oxford University Press.

Shogren, K., Faggella-Luby, M., 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.

Sivagamasundari, U., Fernando H., Jardine P., Rao, J.M., Lunt, P., & Jayewardene, S.L. (1994). The association between Coffin-Lowry syndrome and psychosis: A family study. *Journal of Intellectual Disability Research, 38,* 469-473.

Skinner, B.F. (1953). *Science and Human Behavior.* New York: Free Press.

Sternberg, L., Taylor, R. L., & Babkie, A. (1994). Correlates of interventions with self-injurious behavior. *Journal of Intellectual Disability Research, 38,* 620-630.

Sturmey, P. (1995). Analog baselines: A critical review of the methodology. *Research in Developmental Disabilities, 16,* 269-284.

Sugai, G., Horner, R.H., & Sprague, J.R. (1999). Functional assessment-based behavior support planning: Research to practice to research. *Behavioral Disorders, 24,* 223-227.

Symons, F.M., Koppekin, A., & Wehby, J.H. (1999). Treatment of self-injurious behavior and quality of life for persons with mental retardation. *Mental Retardation, 37,* 297-307.

Toogood, S., & Timlin, K., (1996). The functional assessment of challenging behavior: A comparison of informant-based, experimental and descriptive methods. *Journal of Applied Research in Intellectual Disabilities, 9,* 206-222.

Volkmar, F.R., (2001). Pharmacological Interventions in Autism: Theoretical and Practical Issues. *Journal of Clinical Child Psychology, 30,* 80-87.

Vollmer, T.R. (1994). The concept of automatic reinforcement: implications for behavioral research in developmental disabilities. *Research in Developmental Disabilities, 15,* 187-207.

Wehmeyer, M.L. (1995). Intra-individual factors influencing efficacy of interventions for stereotyped behaviours: A meta-analysis. *Journal of Intellectual Disability Research, 39,* 205-214.

Weiten, W. (1995). *Psychology: Themes and Variations.* Brooks/Cole Publishing Co.

Wieseler, N.A., Hanzel, T.E., Chamberlain, T.P., & Thompson, T. (1985). Functional taxonomy of stereotypic and self-injurious behavior. *Mental Retardation, 23,* 230-234.

Wicks-Nelson, R., & Israel, A.C. (1999). Autism and schizophrenia. In *Behavior Disorders of Early Childhood* (pp. 330-363). Upper Saddle River, N.J.: Prentice Hall.

Table 2.1

Description of all meta-analyses included in the present review

Study	Title	N articles	Important coded variables	Major findings
Campbell (2003)	Efficacy of behavioral interventions for reducing problem behavior in persons with autism: A quantitative synthesis of single-subject research.	117	1. Participant characteristics 2. Treatment characteristics 3. Effectiveness of treatment 4. Experimental design	1. Behavioral treatments are significantly effective in reducing problem behaviors. 2. Pretreatment FBA was a significant variable moderating success of treatment. 3. Participant variables do not significantly influence the general effectiveness of behavioral treatment.
Didden, Duker, & Korzilius (1997)	Treatment effectiveness for problem behaviors among subjects with Mental Retardation	482	1. Topography of problem behavior 2. Treatment type	1. 20.3% of treatment procedures were highly effective, 37.5% fairly effective, 21.9% questionably effective, 20.3% unreliable. 2. External behaviors were less successfully treated than internal behaviors. 3. Response contingent behaviors were the most effective treatment procedure; pharmacological procedures were the least effective. 4. Conducting an EFA was significantly related to PND scores (i.e., effect size).
Didden, Korzilius, van Oorsouw, & Sturmey (2006)	Behavioral treatment of challenging behaviors in individuals with mild mental retardation: Meta-analysis of single-subject research	88	1. Topography of problem behavior 2. Intervention characteristics 3. Participant characteristics 4. BFA methodology 5. Treatment effectiveness	1. Experimental design and assessment methodology were both moderating variables for treatment effectiveness. 2. Experimental, as opposed to descriptive, methods of FBA resulted in significantly higher PND scores. 3. A-B designs had significantly lower PZD scores. 4. There was no statistical difference in treatment effectiveness between behavior type.
Gorman-Smith & Matson (1985)	Treatment for self-injurious and stereotyped responding.	39	1. Participant characteristics (e.g., age, level of mental retardation) 2. Behavioral topography 3. Treatment procedure type	1. Individuals functioning in mild and profound/non-ambulatory range experienced the greatest improvement during treatment. 2. Reinforcement may be better than punishment

				for treating some stereotypy and SIB. 3. Some of most effective treatments: DRO, lemon juice, time out, air splints, and DRO plus overcorrection. 4. Older subjects had greater treatment effect than younger subjects.
Gresham, McIntyre, Olson-Tinker, Dolstra, McLaughlin, & Van (2004)	Relevance of functional behavioral assessment research for school-based interventions and positive behavioral support	150	1. FBA methodology 2. Treatment method 3. Behavioral response class	1. Interventions based on FBAs (BA and FA) were no more effective than those without FA 2. When comparing FBA type, effect size calculations were inconsistent across type of effect size calculation (e.g., MBLR vs. PND). 3. FBAs are not needed in every behavioral intervention plan (BIP) for students served under IDEA.
Herzinger & Campbell (2007)	Comparing functional assessment methodologies: A quantitative synthesis	58	1. Participant characteristics 2. FBA methodology 3. Intervention procedures 4. Effectiveness of treatment	1. FA-based treatments were more effective than those based on BAs, when using the PZD statistic. 2. The type of FBA used did not affect the result of that assessment. 3. Internal behaviors were more often identified as having an automatic function; external behaviors had escape and tangible functions. 4. When multiple behavior response classes were targeted, multiple functions were typical.
Horner, Carr, Strain, Todd, & Reed (2002)	Problem behavior interventions for young children with autism: A research synthesis	9	1. Participant characteristics 2. Problem behavior types 3. Type of FBA 4. Intervention characteristics	1. Disruption/tantrums and aggression were the behaviors most often targeted for reduction. 2. Instruction based intervention procedures were the most commonly reported. 3. Nearly 60% of the studies reported 90% reduction of problem behavior. 4. FBA results should be incorporated into the content of comprehensive behavioral interventions.
Kahng, Iwata, & Lewin (2002)	Behavioral treatment of self-injury, 1964-2000.	396	1. Participant characteristics 2. SIB topographies 3. Treatment setting 4. Experimental design 5. Treatment	1. There has been an increase in the publication of studies regarding SIB treatment since the early 1990's. 2. Head hitting and head

			characteristics	banging were the most commonly reported topographies of SIB. 3. Interobserver agreement (IOA) and experimental design with replication were reported in the majority of studies, highlighting a relatively recent trend toward experimental control. 4. Reinforcement based procedures are the most commonly reported. 5. Most treatments were successful in reducing SIB rates by at least 80%. 6. There has been little increase in the effectiveness of behavioral interventions across the years.
LaGrow & Repp (1984)	Stereotypic Responding: A review of intervention research	60	1. Participant characteristics 2. Stereotypic responses 3. Treatment type 4. Treatment effectiveness	1. Severe and profound MR and severe emotional disturbance were the most commonly reported subject descriptors. 2. Body rocking was the most frequently reported stereotypy. 3. Overcorrection was the most frequently used intervention strategy (38% of studies). 4. Out of the 4 categories of treatment, aversive procedures rated the most effective at reducing stereotypy. 5. Electric shock and 3 combination procedures (physical restraint + over-correction, physical restraint + DRO, and overcorrection + DRI) resulted in highly effective (i.e., responding reduced to 5% of original rates) treatment effects.
Lennox, Miltenberger, Spengler, & Efranian (1988)	Decelerative treatment practices with persons who have mental retardation: A review of five years of the literature	162	1. Participant characteristics 2. Target behaviors 3. Pretreatment functional assessment data 4. Treatment procedures and results	1. Differential reinforcement (e.g., DRO, DRI, DRL, DRA) procedures were reported more often than any other procedure. 2. Medication based treatment regimens had low overall treatment effectiveness across all behavior classes. 3. Aversive procedures, although typically more restrictive, are more commonly reported than positive approaches. 4. The most effective

				treatments for each class of aberrant behavior are: a. SIB-restraint b. stereotypies-overcorrection c. aggression-time out 5. The majority of studies (64%) did not report a pretreatment functional assessment.
Lundervold & Bourland (1988)	Quantitative analysis of treatment of aggression, self-injury, and property destruction	62	1. Subject characteristics 2. Target behaviors 3. Setting 4. FBA methodology 5. Intervention characteristics and effectiveness 6. Interaction effects	1. Pretreatment FBAs were not typically reported; only 36% used some form of FBA 2. Severe and profound MR (60%) and autism (20%) were the most commonly used descriptors for participants. 3. DRO/DRI + punishment was the most commonly reported intervention 4. Response interruption + DRO/DRI and facial screening resulted in the highest efficacy scores.
Matson, Benavidez, Stabinsky Compton, Paclawskyj, & Baglio (1996)	Behavioral treatment of autistic persons: A review of research from 1980 to the present	251	1. Target behaviors 2. Behavioral techniques	1. Stereotypy, aggression, and SIB were the most frequently reported problem behaviors. 2. The majority of behavioral techniques implemented were positive (53%).
Matson & Gorman-Smith (1986)	A review of treatment research for aggressive and disruptive behavior in the mentally retarded	27	1. Participant characteristics 2. Behavior type 3. Treatment characteristics	1. Inappropriate verbal behaviors, aggression, noncompliance were the most frequently reported problem behaviors. 2. Most research in this area has been done with school aged children. 3. More females than males were treated for aggression. 4. Reinforcement, specifically DRI, was the most effective treatment procedure. 5. Children under the age of 7 were the most difficult to treat. 6. Profound/nonambulatory persons were much more effectively treated than others.
Matson & Taras (1989)	A 20 year review of punishment and alternative methods to treat problem	382	1. Problem Behavior topography 2. Participant characteristics	1. Studies that employed painful stimuli in treatment were largely used on children

	behaviors in developmentally delayed persons		3. Data collection (i.e., generalization and follow-up information) 4. Treatment type 5. Side effects of treatment	younger than age 10. 2. Positive treatment methods result in greater long term effects than aversives. 3. Short term effects occur with both positive and aversive procedures.
McClintock, Hall & Oliver (2003)	Risk markers associated with challenging behaviours in people with intellectual disabilities: A meta-analytic study	22	1. Participant characteristics including possible risk markers 2. Challenging behavior topography 3. Study design 4. Data collection methodology 5. Sample characteristics	1. Males were more likely to show aggressive behaviors than females. 2. Level of MR moderated likelihood of challenging behavior (e.g., severe and profound MR more likely to engage in behaviors like SIB and stereotypy). 3. Individuals with autism were more likely to show SIB, aggression, and disruption than those without a diagnosis of autism.
O'Brien & Repp (1990)	Reinforcement-based reductive procedures: A review of 20 years of their use with persons with severe or profound mental retardation	49	1. Reinforcement-based reductive procedure type(i.e., DRI, DRA, DRO, DRL) 2. Demographics 3. FBA Methodology 4. Treatment effectiveness	1. Procedures implemented most often by teachers, staff, and experimenters in schools and residential settings. 2. Social reinforcement was commonly reported, but the least effective. 3. Reinforcer preference assessment data were rarely reported. 4. Reinforcement schedules are not often reported in applicable studies.
Roberts, Mazzucchelli, Taylor & Reid (2003)	Early intervention for behaviour problems in young children with developmental disabilities	48	1. Prevalence rates 2. Intervention types	1. Behavior problems occur more frequently in children with developmental disabilities than in typically developing children. 2. Comprehensive interventions based on social learning theory and applied behavior analysis show significant promise, especially for children with autism.
Scotti, Evans, Meyer, & Walker (1991)	A meta-analysis of intervention research with problem behavior: Treatment validity and standards of practice	318	1. Intervention effectiveness (PND and PZD) 2. Level of intervention intrusiveness 3. Level of behavior severity 4. Experimental design	1. Reinforcement-based procedures were the most common (18%), followed by overcorrection (11%) and contingent aversive stimulation (11%). 2. There is a positive relationship between behavior severity and level of treatment intrusiveness. 3. Using the PND statistic

				of effect size, medication based interventions were significantly lower than all other treatment types. 4. DRO alone is the least effective treatment, but adding DRO to another procedure increases effectiveness. 5. Setting is unrelated to treatment effectiveness. 6. PND scores at follow up were higher for studies that employed a FBA prior to treatment.
Scotti, Ujcich, Weigle, Holland & Kirk (1996)	Interventions with challenging behavior of persons with developmental disabilities: A review of current research practices	150	1. FBA methodology 2. Intervention characteristics including level of intervention intrusiveness 3. Participant characteristics 4. Setting/Location 5. Follow up/Generalization data	1. Fewer intrusive interventions are utilized, however, protective and chemical restraints instead of nonaversive strategies are being used. 2. Almost half of the studies employed some form of data based FBA (e.g., BA or FA). 3. FBA data was more likely reported if target behavior was SIB and least likely when stereotypy was targeted. 4. FBA was rarely used when medication was selected form of treatment. 5. Even when FBAs were used, treatment was not always based on function.
Scruggs, Mastropieri, Cook & Escobar (1986)	Early intervention for children with conduct disorders: A quantitative synthesis of single-subject research	16	1. Subject characteristics 2. Treatment procedures and outcomes 3. Experimental design	1. Tangible reinforcement and punishment/timeout procedures were associated with stronger outcomes. 2. Age, sex, setting, handicapping condition, target behavior, and intervention agent did not demonstrate any consistent overall effects on outcomes.
Shogren, Faggella-Luby, Bae & Wehmeyer (2004)	The effect of choice-making as an intervention for problem behavior: A meta-analysis	13	1. Intervention type 2. Experimental design 3. Participant characteristics 4. Problem behavior topography 5. Treatment characteristics	1. Choice based interventions have clear benefits for individuals with disabilities. 2. Effect sizes (PND and PZD) were higher for individuals aged 4-7 than those who were 8-50. 3. Using FBA results to determine treatment did not have result in noticeable differences in reduction of problem behaviors.
Sternberg, Taylor &	Correlates of	143	1. SIB topography	1. Institutionalization and

Babkie (1994)	interventions with self-injurious behaviour		2. Intervention type	chronicity of SIB have a positive relationship. 2. Intervention type differs regarding age of subject (e.g., medication for >18 years; aversives for < 18 years). 3. There is a significant relationship between intervention type and presence of FBA. 4. Interventions conducted with older subjects are not likely to be as successful as interventions with younger individuals. 5. Type of SIB did not have a significant relationship with treatment efficacy. 6. Intervention should include differential reinforcement in combination with another approach based on FBA results.
Symons, Koppekin & Wehby (1999)	Treatment of self-injurious behavior and quality of life for persons with mental retardation	138	1. Subject characteristics 2. SIB topography 3. Quality of life data 4. Treatment characteristics 5. Intervention outcomes	1. When more aversive interventions were used, quality of life measures were more likely to be reported. 2. Aversive procedures were more likely to be included in behavioral studies than in pharmacological studies.
Wehmeyer (1995)	Intra-individual factors influencing efficacy of interventions for stereotyped behaviours: A meta-analysis	33	1. Treatment efficacy 2. Participant characteristics 3. Topography of stereotypy	1. Older subjects tended to show higher rates of behavior suppression than younger subjects. 2. Gender and level of functioning did not significantly affect treatment outcomes. 3. Interventions were quite successful in reducing rates of stereotyped behavior, but less successful in complete suppression of targeted behavior.

Table 2.2

Efficacy of Intervention by Theoretical Approach for Three Effect Sizes

MBLR

Study	Total	Behavioral	Non-behavioral	Notes
Campbell, (2003)	76.47	76.47	--	
Horner, Carr, Strain, Todd & Reed (2002)	85*	--	--	*Modified MBLR was reported; Specific information about treatment approach outcomes not reported
Kahng, Iwata & Lewin (2002)	83.7*	83.7*	--	*Modified MBLR was reported
Lennox, Miltenberger, Spengler & Erfanian (1988)	63*	65.02*	12*	*Modified MBLR was reported; Data presented was calculated from the original 23 rows of data in Table 2
Mean	**77.04**	**75.06**	**12**	
Range	**(63 - 85)**	**(65.02 - 83.7)**	**(12)**	

PND

Study	Total	Behavioral	Non-behavioral	Notes
Campbell, (2003)	84.40	84.40	--	
Didden, Duker, & Korzilius (1997)	71.06	72.63	49.38	Data presented comes from Table 5 in original article
Gresham, McIntyre, Olson-Tinker, Dolstra, McLaughlin, & Van (2004)	61.94	61.94	--	
Scotti, Evans, Meyer, & Walker (1991)	72	76.55	26	Data calculated from means presented in table 4 and Ns in table 3 in original article
Scruggs, Mastropieri, Cook & Escobar (1986)	79	79	--	Data reported on p. 267 in original article
Shogren, Fagella-Luby, Bae & Wehmeyer (2004)	65.7	--	--	
Wehmeyer (1995)	81	81	--	Although it was not directly stated that all tx were behavioral, a review of the primary articles included indicate this
Mean	**73.59**	**75.92**	**37.69**	
Range	**(61.94 - 84.4)**	**(61.94 - 84.4)**	**(26 - 49.38)**	

Table 2.2 continued

Efficacy of Intervention by Theoretical Approach for Three Effect Sizes

PZD

Study	**Total**	**Behavioral**	**Non-behavioral**	**Notes**
Campbell, (2003)	42.86	42.86	--	
Scotti, Evans, Meyer, & Walker (1991)	49.42	50.34	22	Data calculated from means presented in table 4 and Ns in table 3 in original article
Shogren, Fagella-Luby, Bae & Wehmeyer (2004)	42.3	--	--	
Wehmeyer (1995)	36	36	--	Although it was not directly stated that all tx were behavioral, a review of the primary articles included indicate this
Mean	**42.65**	**43.07**	**22**	
Range	**(36 - 49.42)**	**(36 – 50.34)**	**(22)**	
NOTE: MBLR = mean baseline reduction; PND = percentage of non-overlapping data; PZD = percentage of zero data; * = modified MBLR				

Table 2.3

Efficacy of Intervention by Assessment Type for Three Effect Sizes

MBLR

Study	**Total**	**FA**	**BA**	**No FBA**	**Notes**
Herzinger & Campbell (2007)	80	83.89	77.21	--	
Mean	**80**	**83.89**	**77.21**		
Range	**80**	**83.89**	**77.21**		

PND

Study	**Total**	**FA**	**BA**	**No FBA**	**Notes**
Didden, Duker & Korzilius (1997)	73.58	82.60	66.41	--	Data reported on p. 392 in original article
Didden, Korzilius, van Oorsouw, & Sturmey (2006)	75.00	83.00	62.00	72.00	
Gresham, McIntyre, Olson-Tinker, Dolstra, McLaughlin, & Van (2004)	61.37	51.41	57.89	66.15	Original article excluded all combined FBA procedures
Herzinger & Campbell (2007)	85	81.00	89.02	--	
Shogren, Faggella-Luby, Bae & Wehmeyer (2004)	65.7	--	65.4	62.3	Data reported on p. 231 and p. 232 in original article
Mean	**71.41**	**71.67**	**69.68**	**64.23**	
Range	**(61.37 - 85)**	**(51.41 – 81)**	**(57.89 - 89.02)**	**(62.3 – 66.15)**	

PZD

Study	**Total**	**FA**	**BA**	**No FBA**	**Notes**
Campbell, (2003)	42.86	66.84	36.71	37.02	
Didden, Korzilius, van Oorsouw, & Sturmey (2006)	35.00	49.00	35.00	30.00	
Herzinger & Campbell (2007)	49.00	62.55	35.61	--	
Shogren, Faggella-Luby, Bae & Wehmeyer (2004)	42.30	--	33.2	49.6	
Mean	**42.29**	**59.46**	**35.13**	**38.87**	
Range	**(35.00-49.00)**	**(49.00-66.84)**	**(33.2 - 36.71)**	**30.00 - 49.6)**	
NOTE: MBLR = mean baseline reduction; PND = percentage of non-overlapping data; PZD = percentage of zero data; * = modified MBLR					

CHAPTER 3

A QUANTITATIVE SYNTHESIS OF DEVELOPMENTAL DISABILITY RESEARCH: THE IMPACT OF FUNCTIONAL ASSESSMENT METHODOLOGY ON TREATMENT EFFECTIVENESS

Herzinger, C.V. and J.M. Campbell. To be submitted to Journal of Autism and Developmental Disabilities.

Abstract

Methods and outcomes from functional behavioral assessment have been researched widely over the past twenty years, however, but several important research questions have yet to be examined sufficiently. This quantitative review of developmental disability research aims to make comparisons of different functional behavioral assessment methodologies, both across and within diagnostic categories. Quantitative synthesis data were used to answer questions regarding behavioral function, assessment type, differences based upon diagnostic category, and treatment effectiveness. Results indicate that assessment methodology does not impact treatment effectiveness, but both identified functions and treatment effectiveness are impacted by diagnosis. Implications for clinicians as well as future research directions are also discussed.

KEY WORDS: autism, mental retardation, intellectual disability, developmental disability, maladaptive behaviors, functional assessment, functional analysis

A Quantitative Synthesis of Developmental Disability Research: The Impact of Functional Assessment Methodology on Treatment Effectiveness

INTRODUCTION

Based on the definition given in The Rehabilitation, Comprehensive Services, and Developmental Disabilities Amendment (PL 95-602), a developmental disability is a severe, chronic disability that: (a) is attributed to a mental or physical impairment or combination of mental and physical impairments; (b) is manifested before the person attains the age of 21; (c) is likely to continue indefinitely; and (d) results in substantial functional limitations in some of the following major life activity areas: language and communication, learning, self care skills, and daily living skills. Generally, the term developmental disability refers to any disability that interferes with development and occurs prior to adulthood (Walling, 1995). The most common forms of developmental disabilities are intellectual disability, cerebral palsy, autism, epilepsy, and Down syndrome. Individuals diagnosed with cerebral palsy, or other developmental disabilities that are attributed solely to physical impairments, are not the focus of the current study since cerebral palsy is not typically associated with maladaptive behaviors to the extent of individuals with autism and ID. For the current study, the term developmental disabilities will refer to those individuals diagnosed with autism spectrum disorders and/or intellectual disability.

Autism Spectrum Disorders (ASD) and Intellectual Disability (ID)

ASD is characterized by significant impairment in the following areas: social interaction, communication skills, or the presence of stereotypic patterns of behavior, restricted activities and/or interests. Autistic Disorder, the most common subtype of ASD, is also characterized by onset prior to age 3. For persons with autism, social impairments are often typified by marked impairment in using and understanding nonverbal social cues, such as facial expressions and

body posture. Communication delays noted in persons with autism can range from a total lack of developmentally appropriate verbal or nonverbal language to the inability to initiate or sustain a conversation with others. The restricted patterns of behavior associated with autism include those behaviors that are abnormal in intensity or focus as well as an inflexible adherence to nonfunctional routines and rituals. The prevalence rate of autism is currently estimated at 10-16 cases per 10,000 with a male to female ratio of 1.95 to 5.5 : 1 (Fombonne, 2005).

ID, previously referred to as Mental Retardation, is typified by "significantly subaverage intellectual functioning (an IQ of approximately 70 or below) with onset before age 18 years and concurrent deficits or impairments in adaptive functioning" (DSM-IV-R, p. 39). Adaptive functioning refers to an individual's daily life skills and their personal independence. ID occurs in approximately 1-3% of the population, depending on the criteria used to determine the prevalence rates. Four degrees of severity, based on the level of cognitive ability and adaptive functioning, have been established to better explain the variability of ID: Mild, Moderate, Severe, and Profound.

Individuals with autism and ID are commonly associated with a broad range of maladaptive behaviors including self-injurious behavior [SIB] (e.g., self-pinching, scratching, head banging), property destruction, aggression towards others, severe disruptions (e.g., tantrums), and stereotypic behaviors (e.g., body rocking, hand flapping). These behaviors can impact not only the individual engaging in such behavior but also those around them (i.e., caregivers, teachers, peers). Maladaptive behaviors can lead to serious medical problems, such as tissue damage; destruction or loss of property; poor social relationships; and also have a negative

impact on academic learning. For these reasons, the assessment and treatment of such behaviors in individuals with autism and ID is an important component of any comprehensive approach to rehabilitation.

Behavioral Approach to Assessment and Treatment of Maladaptive Behaviors

A behavioral approach to intervening with maladaptive behaviors has been consistently documented as the most efficacious approach for treating aberrant behaviors (Campbell, Herzinger, & James, 2007). A behavioral approach is focused on identifying the antecedents and consequences of the target behavior during assessment and manipulating either the environment (e.g., establishing operations) or the immediate consequences to target behavior during intervention. Behavior analysts hold the view that all behavior responses have a function or purpose and are observable demonstrations of cause-effect relationships (Skinner, 1953). The key to effective treatment is the identification of this function and the transfer of stimulus control to a more appropriate behavior. The most current taxonomy of behavioral function focuses on three types of reinforcement as the major behavioral mechanisms maintaining behavior: (a) positive reinforcement, (b) negative reinforcement, and (c) automatic reinforcement. The importance of identifying function as part of the intervention process is evinced by recent amendments to the Individuals with Disabilities Education Act [IDEA] (P.L. 105-117, 1997; Gresham et al., 2004).

Reinforcement is defined as a stimulus that increases the likelihood that a behavior will occur in the future. Positive reinforcement refers to the presentation of a stimulus that strengthens behavior. For individuals whose target behaviors are maintained by positive reinforcement, this could include access to leisure items (e.g., videos, toy planes), access to edible items (e.g., cookies, juice) or access to therapist/parent attention. If a child engaged in

aggression and the consequences of that behavior were verbal reprimands (e.g., “Don’t do that. That’s not nice. You are going to hurt someone.”), it may be hypothesized that the aggression is maintained by access to parental attention, a form of positive reinforcement. Lovaas and colleagues (1965) were one of the first to describe the SIB of individuals with autism and ID as maintained by social positive reinforcement in the form of attention.

With negative reinforcement, the target behavior is strengthened following the removal or avoidance of a particular stimulus. Examples of negative reinforcement include escape from academic demands, escape from chores at home, and avoidance of social situations. For example, a teacher instructs a student to complete a math problem. The student responds by throwing a pencil at the teacher and in consequence is sent to the principal’s office. If an increase in disruptive behavior occurs under this contingency, it may be hypothesized that escape from academic demands, a form of negative reinforcement, is the maintaining function of that maladaptive behavior. Similar to Lovaas et al.’s (1965) seminal work documenting social reinforcement as maintaining aggression, Sailor et al. (1968) documented the influence of negative reinforcement in maintaining the maladaptive behavior of an individual with ID. In the Sailor et al. study, the individual’s maladaptive behavior was maintained by access to breaks during academic tasks.

Automatic reinforcement refers to behaviors that are strengthened via access to non-socially mediated stimuli (Skinner, 1953). Rincover (1978) was one of the first to investigate maladaptive behavior maintained by automatic reinforcement. Automatic reinforcement may include proprioceptive, visual or auditory stimulation, self restraint, pain attenuation, or some

combination of mechanisms independent of the social environment. Often, it is suggested that maladaptive behaviors maintained by automatic reinforcement are the most difficult to treat (Vollmer, 1994; Leblanc et al., 2000; Patel et al., 2000).

In the last 20 years, there has been a trend toward developing treatments for maladaptive behaviors following determination of the hypothesized functions of the behaviors. Based on the ascribed function of the target behavior, an appropriate treatment package can be selected. For individuals whose maladaptive behaviors are maintained by positive reinforcement, common treatment packages include altering the establishing operations, extinction, non-contingent reinforcement, and differential reinforcement. For the previous example of target behaviors maintained by positive reinforcement (i.e., aggression maintained by access to parental attention) one possible treatment package might include extinction of attention following aggression and the introduction of differential reinforcement in the form of functional communication training (FCT). FCT might include teaching the child to ask for attention appropriately (e.g., "Talk to me, please,") and reinforcing (i.e., providing attention) only following the appropriate request. Mirenda (1997) conducted a comprehensive review of the literature from 1985 to 1996 regarding the use of FCT in the treatment of maladaptive behaviors. Results indicated that FCT led to significant decreases in target behaviors for the majority of participants in both the short term and at follow-up.

For individuals whose maladaptive behavior is maintained by negative reinforcement, some common treatments include compliance training, extinction, and differential reinforcement with and without the use of signals. For the previous example of target behaviors maintained by negative reinforcement (e.g., the student given escape from work following inappropriate behavior), one treatment option would be extinction. That is, the teacher continues to present

demands following inappropriate behavior and escape from the task is no longer contingent upon disruptive behavior. However, some research has shown that when given a choice, individuals with escape maintained maladaptive behavior choose edible reinforcers more often than functional reinforcers (i.e., a break from work; Lalli et al., 1999). This indicates that at least for escape-maintained maladaptive behavior, other reinforcement based procedures that are not based on the function of the behavior may also be effective interventions.

For individuals whose maladaptive behaviors are hypothesized to be maintained by automatic reinforcement, some treatment approaches include non-contingent reinforcement, differential reinforcement, punishment procedures, and extinction. Some researchers have suggested the use of non-contingent access to preferred stimuli over the other treatment packages based on these advantages: (a) ease of application, (b) does not involve restrictive devices, (c) no deprivation of reinforcement, and (d) appropriate behavior may be strengthened (Roscoe et al., 1998). However, others have successfully decreased the rate of automatically-maintained maladaptive behaviors with other interventions. For example, following a FA that indicated a target behavior was maintained independent of all tested environmental variables, Kennedy and Souza (1995) found response blocking to be an effective treatment for the eye-poking of a 19-year-old male.

Any of the three aforementioned reinforcement types (i.e., positive, negative, and automatic) may be the mechanism maintaining maladaptive behaviors for the individuals who engage in them. However, it is possible that participant characteristics, such as diagnostic category, may influence the function of maladaptive behaviors. In other words, disability type may align with behavioral function. It may be hypothesized that individuals diagnosed with ID, but without any impairment in socialization or communication, are more likely to engage in

behaviors that result in access to social positive reinforcement more frequently than those diagnosed with ASD, a disability characterized by marked impairments in socialization and communication. Alternatively, it may be hypothesized that individuals with ASD may be more likely to engage in behaviors ascribed to automatic functions due to the sensory sensitivities commonly reported in the literature or escape functions to avoid situations in which interactions with others is necessary. In their review of the developmental disability literature, Dawson, Matson, and Cherry (1998) decided to include only individuals who functioned in the severe to profound range of ID and based their decision on the notion that level of cognitive functioning may influence the reasons for the maladaptive behavior (i.e., function). Although Dawson et al. (1998) found no significant differences in ascribed function across diagnostic category; they did find a pattern of mean differences to support their hypothesis that diagnosis mediates function of maladaptive behaviors. Currently, there is little published research comparing the differences in ascribed functions of maladaptive behaviors across diagnostic categories.

Aside from the possible mediating influence of participant characteristics, assessment type must also be assessed for inadvertent influence on assessment outcome. Herzinger and Campbell (2007) compared the treatment outcomes of over 100 individuals with autism who engaged in maladaptive behaviors and found no relationship between FBA type and ascribed function. The results indicated that both experimental (i.e., FA) and non-experimental (i.e., BA) assessment types are independent methodologies and do not falsely ascribe functions based on inherent methodological flaws or biases.

FUNCTIONAL BEHAVIORAL ASSESSMENT METHODOLOGY

The term FBA refers to any methodology used to identify the purpose of behavior and encompasses *indirect assessments*, often referred to as informant assessment, (e.g., interviews,

questionnaires, rating scales), *descriptive assessments* (e.g., A-B-C sheets, direct observation with no variable or environment manipulation); and *functional analyses* (FA), also known as experimental, "true," or traditional functional assessments (e.g., analogue conditions in which antecedent or consequent variables are systematically manipulated within an experimental design). For the purposes of this paper, we are using the term FA to describe all experimental analyses. The term Behavioral Assessment (BA) refers to those assessments which are non-experimental in nature and includes both indirect and descriptive assessments.

Functional Analysis (FA) and Behavioral Assessment (BA) Methodologies

FA represents simulation of natural environments and is the primary tool for demonstrating causal relationships between behavior and reinforcement contingencies (Carr, Langdon, & Yarborough, 1999). With FA, the goal is to isolate the function or purpose of the problem behavior by directly and specifically altering the social and physical environment in a way that is most likely to alter the problem behavior. Although previously theorized and described (Skinner, 1953; Wolfe, Birnbauer, Williams & Lawler, 1965), Iwata and colleagues (1982/1994) defined the first standardized and comprehensive method for conducting FA. As outlined by Iwata et al. there are four conditions in a FA: social disapproval, academic demand, unstructured play, and alone. However, other conditions (e.g., tangible; Day et al., 1988) may be substituted or implemented in addition to the original analogue conditions in order to test all hypotheses regarding behavioral function.

Through each condition of the FA, frequency data are collected. Typically, conditions are presented more than once and in a random order to prevent order bias. The frequency data are graphed and rate of behavior comparisons are made across condition types. If the rate of behavior is significantly higher in one condition, then it is assumed that the problem behavior is

a function of that condition. For example, if the problem behavior occurs at much higher rates in the demand condition, it is assumed that the purpose of the behavior is to escape demands (see Figure 3.1). If target behavior rates are high across conditions, the behavior could be determined to be automatically maintained (i.e., showing little difference across different environmental manipulations), multiply-maintained (i.e., being maintained, at least in part, by a variety of environmental variables), or undetermined, in which case a more thorough analysis of behavior and variable manipulation must take place.

For the purposes of this paper, the term FA refers to any experimental analysis of maladaptive behavior in which variables are manipulated to determine the purpose or function of target behavior. Different analysis models can focus on the different components of the antecedent-behavior-consequence relationship. For example, with structural analysis, antecedent variables such as task difficulty and attention to task are manipulated to alter target behavior rates (Repp & Horner, 1999). Other behavioral assessment and intervention methodologies are consequence driven (i.e., reinforcement or punishment based treatments). The interventions that stem from this type of assessment focus on altering the immediate consequences to aberrant behaviors in order to effect change upon the target behavior. In Hanley, Iwata, and McCord's (2003) review of the FA literature, they described two distinct models of FA: (a) A-B, which includes those that focused on the manipulation of antecedent variables and the effect on target behavior, such as the structural analysis, and (b) A-B-C, those that focused on the manipulation of antecedents and consequences and the effect on target behaviors, such as Iwata's 1982 methodology. Recently, Potoczak, Carr and Michael (2007) directly compared the two FA models and determined that the A-B-C method was more successful in accurately ascribing function to maladaptive behaviors.

BA represents the other two categories mentioned under the umbrella term FBA: indirect and descriptive assessments. Also, BAs can be considered formal (i.e., structured) or informal depending on how systematically the information is collected. Indirect methods can include interviews with parents or teachers in which they are asked questions such as, "Why do you think the problem behavior is occurring?" and a variety of rating scales completed by parents and professionals. The Motivation Assessment Scale (MAS; Durand & Crimmins, 1988), Questions About Behavioral Function (QABF; Matson & Vollmer, 1995) and Motivation Analysis Rating Scale (MARS; Wieseler et al., 1985) are three recognized indirect methods of assessing behavioral function. Some research has been reported regarding the comparability of these methods of assessment to FA. For example, MAS ratings have been correlated with FA analogue data and the relationship between the two sources was found to be highly significant ($r = .99$, $p < .001$) (Durand & Crimmins). Thus, "the teacher ratings on the MAS predicted their student's behavior in the experimental condition" (Durand & Crimmins, p. 112). Descriptive assessments refer to those assessments in which no variables are directly manipulated, but direct observations are made. These include continuous observation methods (Bijou, Petersen, & Ault, 1968), antecedent-behavior-consequence forms (i.e., A-B-C sheets; Groden, 1989), scatter-plot recording (e.g., Touchette, MacDonald, & Langer, 1985), and specified observational forms such as the Functional Assessment Observation Form (FAOF; O'Neill et al., 1997). The Structured Descriptive Assessment (SDA; Anderson & Long, 2002) is another form of descriptive assessment; however, antecedent variables can be slightly manipulated during implementation depending on the natural situations observed.

Due to the variability in BA methods (e.g., rating scales, observations, interviews), assessment outcomes and treatment recommendations may differ. Some BAs are simply used to

gain qualitative information that is later combined with other sources of information to identify the function of maladaptive behavior. Other BAs, such as the rating scales, use cut scores or relative rankings to identify the functions of the targeted behavior. For example, the authors of the MAS, a 16-item questionnaire, note that if one condition (e.g., attention) has clearly received the highest score it is assumed that this is the function of the behavior. The relative rankings of conditions are used to determine the most important influence on the behavior (Durand & Crimmins, 1992). For the MAS, like all other measures of behavioral function, more than one possible function may be identified. The authors provide a guide to interpret most outcome results, including decision rules when two functions are scored within .25 to .50 points of each other. In this scenario, both categories would be considered influences that may be causing the maladaptive behavior to continue (Durand & Crimmins, 1992). Regardless of the method of interpretation, the results of BAs are used in the same manner as those from FAs. The qualitative data from interviews, rating scale results, and direct observations are often combined and, if possible, a single function of behavior is identified. Treatment selection is then based on the outcome of the gathered information.

FUNCTIONAL ANALYSIS AND BEHAVIORAL ASSESSMENT COMPARISONS

Each general FBA approach (i.e., FA and BA) has its strengths and weaknesses. Indirect assessments provide a relatively quick method of gathering information and generating hypotheses regarding behavioral function; however, the validity of the information gathered during these measures is dependent on the ability and willingness of the informant to provide accurate, unbiased information. Descriptive assessments of function are thought to be more valid and accurate than indirect assessments because descriptive assessments are based on direct observations. Descriptive assessments, however, do not allow for systematic manipulation of

environmental variables (antecedents or consequences) and involve a recording of what is seen (or noticed) in the interaction. Therefore, descriptive assessments are correlational in nature and cause-effect relationships between the environment and target behavior cannot be assumed. FA, on the other hand, is a methodology that does allow for the identification of idiosyncratic causal relationships. The structure and experimental design of FAs allow for ease of replication. FAs, however, have been criticized as time-consuming, complicated, and not cost-effective.

Several researchers have made comparisons across FBA methodologies and, in general, the findings support the FA as the "gold standard" for ascribing function and consequently developing function-based treatments. Paclawskyj et al. (2001) evaluated the convergent validity of the QABF with FA data. They studied 13 adults with profound ID who engaged in self-injury, aggression and stereotypic behaviors. Overlap of the two methods producing the same function was 56%; however, when the cases in which no function was identified were excluded, the FBA methodologies overlapped for 75% of the participants. In contrast to Durand and Crimmins (1988) and Paclawskyj et al., (2001) findings regarding indirect assessments, Hall (2005) found that descriptive and experimental methods of FBA agreed only 25% of the time. In almost all published accounts of comparison data, the FA represented the gold standard for validity tests of other types of assessment.

Others have looked past the overlap of ascribed function across the methodologies and to intervention outcome measures. Didden, Korzilius, van Oorsouw, and Sturmey (2006) made comparisons across descriptive and experimental FBAs and found that treatments based on experimental (i.e., FA) methods resulted in significantly higher PND scores. Herzinger and Campbell (2007) conducted a meta-analysis of autism literature on the assessment and treatment of maladaptive behaviors. The authors found that when using the PZD statistic to determine

treatment efficacy, treatments that were based on FA results were more effective that those based on BA results. In a more recent review of meta-analyses of developmental disability intervention literature, Herzinger (2008) reviewed 11 meta-analyses that reported on FBA inclusion in the intervention process. Herzinger found that treatments based on FA results were more effective that those based on BA results led to higher means for all three efficacy indicators used (i.e., MBLR, PND, and PZD).

Researchers assessing maladaptive behaviors agree that identifying the function of the target behavior is integral in the treatment selection process; thus FBAs are a core feature in the development of interventions designed to ameliorate aberrant behaviors (Yarborough & Carr, 2000). Knowing which FBA methodology is associated with more successful treatment outcomes is imperative.

PURPOSE

Based on the brief review of the literature presented, the current study aims to answer the following specific research questions:

1) Is treatment more effective when following an experimental functional analysis (FA) or a non-experimental behavioral assessment (BA) for individuals with autism and ID?
2) Is there a predominant observed function based on the type of assessment (i.e., FA or BA) for either (a) individuals with autism, (b) individuals with ID, or (c) individuals with autism and ID?
3) Does ascribed function differ depending on diagnostic category? For example, are problem behaviors more likely to be reinforced by access to tangible reinforcers for individuals with autism when compared to individuals with ID?

Similarly, do problem behaviors tend to be maintained by attention when compared to children with autism?

4) Does the observed function of the behavior, regardless of FA method used, have an impact on the effectiveness of treatment?

5) Is treatment effectiveness impacted by diagnosis? For example, do individuals with autism who function in the range of ID show poorer response to behavioral treatment than those without a diagnosis of ASD?

METHOD

Study Identification and Selection

For the years 2000 through 2005, published functional assessments of problem behavior for individuals with developmental disabilities were identified through searches of PsycLit, ERIC, and MedLine databases using appropriate search terms, such as subject descriptions (e.g., autism, mental retardation, intellectual disability), target behaviors (e.g., self-injurious behaviors, aggression, problem behaviors), and assessment type (e.g., applied behavior analysis, functional assessment, functional analysis). Published studies were identified by issue-by-issue hand searches of the following journals: *American Journal of Mental Retardation, Behavioral Interventions, Behavior Modification, Education and Training in Mental Retardation and Developmental Disabilities, Journal of Applied Behavior Analysis, Journal of Association of People with Severe Handicaps, Journal of Autism and Developmental Disorders, Journal of Intellectual Disability Research, Mental Retardation,* and *Research in Developmental Disabilities.* Also, timely references (i.e., citations between 2000 and 2005) from each article found through the literature search were reviewed for possible inclusion.

Studies were selected for inclusion if the following criteria were satisfied. First, studies were selected if they were published in peer reviewed journals between January 2000 and December 2005. Second, single case studies were included only if a participant was diagnosed with intellectual disability or autistic disorder. If the participants were described as being "autistic-like," or developmentally delayed, they were also included. Third, a functional behavioral assessment (FBA) had to be conducted and results reported, with maladaptive behaviors as the target behaviors of treatment to be included. For those articles used in the treatment effectiveness analyses, studies were included if: (a) data points, not just mean scores, were reported; (b) baseline data and treatment data were reported; and (c) if the intervention procedures targeted reduction of stereotyped, self-stimulatory, self-injurious, destructive, disruptive, or aggressive behaviors. If an article included multiple participants or studies only partially met inclusionary criteria, only those components that met criteria were included in the review.

Estimating Effects of Behavioral Interventions

Effect Size Calculations

There are several methods for assessing effectiveness data using both regression and non-regression approaches. Frequently reported summary methods have involved the calculation of Mean Baseline Reduction (MBLR), Percentage of Non-overlapping Data (PND), and Percentage of Zero Data (PZD; Campbell, 2003). Other methods, such as the Percentage of data points exceeding the median (PEM; Ma, 2006) and Improvement Rate Difference (IRD; Parker & Hagan-Burke, 2007) could have been selected for use and comparison. Based on the their common usage in related literature reviews, the following three effect sizes based on nonregression approaches were calculated per intervention in the current study: MBLR, PND,

and PZD. The MBLR is calculated by subtracting the mean of treatment observations from the mean of baseline observations then dividing by the mean of baseline observations and multiplying by 100 (Campbell, 2003; Lundervold & Bourland, 1988; O'Brien & Repp, 1990). The PND statistic is calculated as the percentage of treatment data that did not overlap with baseline data points (Scruggs, Mastropieri, & Casto, 1987). If a baseline phase reported one or more data points of zero, then the same number of data points was excluded in the treatment phase prior to calculation of the PND (Didden, Duker & Korzilius, 1997). The PND can range from 0 to 100%. According to Scruggs, Mastropieri, Cook, and Escobar (1986) a PND greater than 90% reflects a highly effective treatment, a PND of 70-90% is considered a fair treatment outcome, and a PND of less than 50% indicates unreliable/ineffective intervention. The PZD statistic is calculated by locating the first intervention data point that reached zero and computing the percentage of data points that reached zero including the first zero point (Scotti et al., 1991). The PZD score is considered a more stringent efficacy indicator as it requires target behaviors to reach and stay at zero levels throughout treatment to be considered effective. Campbell (2004) noted that the PZD score represents a "degree of behavior suppression versus degree of behavior reduction" (p. 235). PND and PZD scores have been found to be independent indicators of treatment outcome (Scotti et al., 1991; Campbell, 2003) and have been used in several studies to measure the effectiveness of treatments (Didden, Korzilius, van Oorsouw, & Sturmey, 2006; Herzinger & Campbell, 2007).

Handling Multiple Outcomes, Participants, Assessment types, and Experimental phases

Several rules have been established for the coding of assessment type. Using Herzinger and Campbell's (2007) coding system, functional assessment type was coded as either: (a) FA (strictly adhering to guidelines set forth in Iwata et al., 1982), (b) modified FA, (c) ABC sheet,

(d) rating scales (e.g., MAS, QABF), (e) informal assessment or (f) other. Under the modified FA code, FAs that included sessions not described by Iwata et al (e.g., tangible session as described by Day et al., 1988) and brief FAs, (i.e., shorter or fewer sessions than those described by Iwata) such as those described by Northup and colleagues (1991) and summarized in Derby and colleagues (1992) were coded. Later, this group (modified FA) was consolidated with the FA group in order to unify the experimental analyses. The studies coded as A-B-C sheets, rating scales, and informal assessments were consolidated to form the BA, or non-experimental category. Thus, articles were coded as FA if they included analogue conditions and manipulated variables in the environment in each condition as opposed to those designated as BA which did not include variable manipulation.

Consistent with the methodology of Herzinger and Campbell (2007), if two different types of FBAs (e.g., FA and MAS) were used with a participant, the methods were coded separately with the possibility of two functions and different treatments identified. If a participant's problem behavior was assessed using multiple BA methods (e.g., MAS, parent interview, and observation) the assessments were coded as a combination. In such a case, the coding resulted in one effect size unless the BAs yielded different functions for each method.

Studies that reported on multiple outcomes or multiple participants required separate effect size calculations for each outcome for each participant. When more than one problem behavior was targeted for a participant and separate data points were reported, individual effect sizes were calculated per problem behavior per participant rather than arbitrary selection of one behavior. This approach was used in order to capture all available data regarding each participant and each problem behavior. For example, a hypothetical study reported data for two different participants using two different assessment methodologies for each and targeting two behaviors

for participant A and one behavior for participant B. For participant A, a teacher interview and FA were conducted both ascribing function to escape from academic tasks. Only one treatment (FCT plus extinction) was implemented. For participant A, 10 baseline data points and 20 treatment data points were reported and all three nonregression effect sizes were computed and reported. Identical effect sizes would have been coded for both FA and BA. For participant B, a functional rating scale assessment was conducted. The data suggested that the behavior was maintained by different functions in different settings (i.e., escape at school; attention at home) and two different treatment packages were implemented. The baseline and treatment data points from each setting specific observation would have been reported separately, thus for participant B three effect sizes would have been calculated for behavior at school and three effect sizes would have been calculated for behavior at home.

Single case designs vary (e.g., A-B; A-B-A-B) and effect sizes can be calculated from varied contrasts (Allison & Gorman, 1993). In the present study, the effect sizes were calculated between the first non-treatment phase and the last treatment phase, per Faith et al.'s (1996) recommendations and implemented in Campbell (2003) and Herzinger and Campbell (2007). In designs that compared multiple treatments (A-B-A-C), the initial baseline and final treatment phase were coded. Although it is not ideal to make comparisons between baseline and subsequent intervention phases that are separated by both time and experience, this was necessary given the limitations of the meta-analysis format used in the current study. In order to accurately and reliably quantify the data presented in the primary articles, a decision rule was made to compare the initial baseline phase and the final treatment phase for each study, regardless of whether the phases were adjacent to one another. In studies using a multi-element

or alternating treatments design, both treatments were coded unless a final "best treatment alone" condition was conducted. In this case, the initial baseline phase and final treatment condition were coded.

Data extraction, Variables coded, and Reliability check

For the necessary analyses in the present study, the graphs provided by the articles were transformed into raw data via a ruler. The distance between each point and the abscissa was calculated in millimeters and rounded to the nearest 0.5. The data conversion procedure has been used by Allison, Faith, and Franklin (1995), Campbell (2003), and Herzinger and Campbell (2007) with a high degree of inter-rater reliability.

The following participant information was coded when available: participant's age, gender, race, level of intellectual functioning, secondary diagnoses, years since diagnosis prior to study, and years of prior treatment. Diagnostic category was grouped into one of three mutually exclusive categories: (a) autism without ID, (b) ID without autism, and (c) autism and ID. A fourth group, developmental delay, was also recorded and later combined with those in the ID only category, due to the lack of autism descriptors. The following assessment/pre-intervention data were coded: target behavior, type of FBA (e.g., experimental FA, informal assessment), function(s) of behavior, type of intervention used (e.g., reinforcement only, combination of extinction and punishment), length of session, treatment setting, and type of therapist. Targeted behaviors were coded as: aggression, property destruction, disruptive behaviors (e.g., spitting), vocalizations, SIB, and stereotyped behaviors. If relevant, specific types of SIB were also recorded. These subtypes included head banging, face slapping, eye gouging, biting, and scratching.

The following intervention data were coded: type of intervention, type of experimental design, inter-rater reliability, number of baseline data points, number of final phase treatment points, and attempt to generalize treatments. The types of intervention coded included: non-contingent reinforcement, differential reinforcement, punishment, timeout, extinction, sensory extinction, FCT, combined treatments, and other interventions. Based on Herzinger and Campbell (2007) these categories were later consolidated into six categories: (a) reinforcement only, (b) punishment only, (c) extinction only, (d) reinforcement and punishment, (e) extinction plus reinforcement or punishment, and (f) other.

Reliability of data extraction and coding decisions

There is some controversy regarding acceptable levels of reliability in the literature, but generally, reliability above 90% is desirable and reliability below 80% indicates significant problems with measurement (Wolery, Ault, & Doyle, 1992). For this study, 80% was chosen as the cut off for acceptable levels of reliability of coding decisions. Eighteen articles were randomly selected for independent coding by advanced graduate students in Psychology, who had experience working with individuals with autism and ID, and interrater agreement was established. The 18 articles (21.69% of all articles) included 30 separate assessments (15.07% of all assessments) and 26 different participants (18.05% of all participants). Interrater agreement, with a mean of 99.71% and a range of 95.98% to 100% across all coded variables, was determined by the percent agreement method (# of agreements / # of agreements + # of disagreements X 100). See Appendix A for a detailed reliability summary.

STATISTICS

Descriptive analysis of participants, studies, interventions, and experimental features

Participants and articles were examined through descriptive statistics. Characteristics of the articles and studies located included: the number of articles located; the number of studies extracted from each article; and the journals contributing to the review. Participant characteristics were examined next. Gender, age, race/ethnicity, primary and secondary diagnoses, level of intellectual functioning, and other variables were summarized. The descriptive analysis provides an indication of how well the sample represents the larger population of individuals with developmental disabilities. Assessment, intervention, and experimental characteristics of the studies were examined as well (e.g., percentage of studies that used FA methodology; percentage of studies who targeted aggression).

Inferential statistical procedures

1. Three one-way ANOVAs were used to examine research question 1 (a comparison of treatment effectiveness for FA and BA for individuals in one of three diagnostic categories).

2. A non-parametric Chi-square test of non-independence was used to assess research question 2 (possible bias in assessment outcomes based on FBA methodology).

3. A non-parametric Chi-square test of non-independence was used to examine research question 3 (assessment of impact of diagnostic category on FBA outcome).

4. Three one-way ANOVAs were used to address research question 4 (treatment effectiveness as impacted by function). Treatment effectiveness means for each effect size statistic were compared for each of seven functional categories.

5. Three one-way ANOVAs were used to examine research question 5 (treatment effectiveness as impacted by diagnostic category). Treatment effectiveness means for each effect size statistic were compared across three diagnostic categories.

RESULTS

Characteristics of Participants

This review included 83 articles reporting on 144 participants with a total of 199 separate studies (i.e., assessments and/or treatments). Based on the previously determined mutually exclusive categories, the majority of participants fell into the ID only category (54.9%), followed by autism and ID (24.3%), followed by autism only (12.5%), and finally unspecified developmental disability (8.3%). Those reported as having unspecified developmental delays were later combined with the ID only category based on the lack of autism characteristics mentioned, following further analysis of the participant descriptions reported in primary articles. The ratio of males to females in this study was 1.5 : 1. For individuals described as having autism or autistic characteristics, the gender ratio was 3.5 : 1 in favor of males, which is similar to prior reviews documenting the higher prevalence of autism in males. Also consistent with prior reviews documenting the prevalence of ID in individuals with autism, the majority of the participants diagnosed with autism functioned in the range of mental retardation (79.9%) or were considered “untestable” via standardized, formal intelligence testing. In most cases (97.9%), the criteria used to diagnose autism were not reported; however, when the criteria were reported, they were based on a version of those presented in the DSM. In the great majority of cases, information regarding a client’s time since diagnosis, time in prior treatment, or current use of prescription medication was excluded. See Table 3.1 for detailed information about the characteristics of participants.

Characteristics of Studies

The 199 studies included in the research synthesis were collected from a total of eight journals, with the *Journal of Applied Behavior Analysis* contributing the highest percentage of articles (48.7%). Studies most often included one (57.8%) or two (18.1%) participants and rarely included more than three. See Table 3.2 for detailed information about the characteristics of the studies included in the meta-analysis.

Characteristics of Functional Assessment, Behavioral Intervention, and Experimental Quality

Studies employed both experimental (77.4%) and non-experimental (22.6%) methods of FBA. Under the FA umbrella, the majority of assessments were modified session FAs (53.9%), which are based on the analogue conditions of Iwata et al., but tailored in terms of the specific conditions used in the analysis. The type of BA most often reported was described as informal assessment (73.3%). See Table 3.3 for detailed information regarding the breakdown of FBA types used in the included studies.

The studies included in the quantitative synthesis targeted the reduction of more than one problematic behavior (33.2%) more often than any single maladaptive behavior, $\chi 2$ (N = 6, 199) 105.17, $p < .001$. When assessed individually, SIB (20.1%) and Aggression (17.1%) were the most commonly targeted behaviors. In combination with other maladaptive behaviors, SIB (44.7%), aggression (46.2%) and disruptive behaviors (32.7%), were the most commonly reported among the participants. Based on the re-categorization of target behaviors as being internal or external, or a combination, the more participants engaged in externally destructive behaviors (38.2%). The most commonly reported intervention type reported was a combination of two or more treatments (37.7%) followed by changes in antecedent events or

settings (29.6%) and reinforcement based interventions (25.6%). In a majority of cases (73.4%), behavioral therapist could not be coded due to lack of information reported by the authors. Centers for individuals with developmental disabilities were the treatment setting that occurred with the highest frequency (21.1%). The use of medications during assessment and interventions was reported in 14.1% of studies; however, no information regarding medication use was given in the majority of studies (59.3%).

The A-B-A-B experimental design (i.e., including both reversal and withdrawal) was the most commonly used, reported in 31.2% of the studies, followed by the multiple baseline design (29.6%). Studies in the meta-analysis omitted follow up data collection (82.4%) more often than they included follow-up data collection (14.1%), $\chi 2$ (N = 2, 199) 219.03, $p < .001$. Generalization data were omitted from the studies (71.4%) more often than reported (28.6%), $\chi 2$ (N = 2, 199) 89.89, $p < .001$. When reported, the data indicated that generalization to new situations and settings was the most common form of generalization data reported (54.4%). Inter-rater reliability for FBA sessions was reported in 88.9% of articles. The reliability ranged from 80.0 to 100.0 with a median of 98.0. For treatment sessions, inter-rater reliability data was reported in 100% of articles. The reliability ranged from 70.2 to 100.0 with a median of 98.0. See Table 3.3 for detailed information about FBA, behavioral interventions, and experimental quality.

Does functional assessment methodology impact treatment effectiveness for individuals with developmental disabilities?

Three independent samples *t*-tests indicated that when comparing FA and BA across diagnostic categories, there were no significant differences in treatment effectiveness as measured by the

three effect sizes calculated. Results from the independent samples t tests and reported means, standard deviations, and ranges of effect size calculations are presented in Table 3.4.

Is ascribed function related to functional assessment methodology?

The data showed that there is a relationship between the type of methodology (e.g., FA, BA) used in the assessment and the result of that assessment, χ^2 (N = 6, 199) 19.81, $p < .01$. This finding indicates that there are significant differences in the function ascribed to target behaviors depending on the type of functional assessment used. The results indicated that FA procedures more likely result in a social positive reinforcement function (i.e., behavior maintained by access to tangible items or social attention) and BA are more likely to result in automatic functions. In addition, BA most often indicated a single function maintaining target behaviors as opposed to target behaviors that are multiply-maintained. See Table 3.5 for frequency information.

Is ascribed function impacted by diagnostic category?

A non-parametric chi-square test of non-independence showed that there are significant differences in the ascribed functions of target behaviors across diagnostic categories, χ^2 (N= 12, 199) 29.22, $p < .01$. The data indicated that individuals diagnosed with autism alone and autism and ID were most often identified as having target behaviors maintained by social negative functions (e.g., escape from tasks). Individuals diagnosed as functioning within the range of ID were more often identified as having maladaptive behaviors maintained by social positive contingencies. See Table 3.6 for frequency information.

Is treatment effectiveness impacted by ascribed function?

The results from three one-way ANOVAs indicated no differences for ascribed function for the three calculated effect sizes. Treatment effectiveness, as assessed by the (a) MBLR statistic, F (6, 198) = .46, *n.s.*, (b) PND statistic, F (6, 198) = 1.10, *n.s.*, and (c) PZD statistic, F (6, 198) = 1.72,

n.s. was not significantly affected by the function of the problem behavior. Table 3.7 includes data regarding the means, and standard deviations for the three calculated effect sizes: MBLR, PND, and PZD.

Is treatment effectiveness impacted by diagnostic category?

The results from three one-way ANOVAs indicated no differences for ascribed function for two of the three calculated effect sizes. Treatment effectiveness, as assessed by the (a) MBLR statistic, F (2, 198) = .45, *n.s.*, and (b) PND statistic, F (2, 198) = .51, *n.s.* was not significantly affected by the diagnosis of the individual. However, treatment effectiveness as assessed by the PZD statistic did indicate significant differences of treatment effectiveness across diagnostic categories, F (2, 198) = 4.36, $p < .01$. When assessing treatment effectiveness with the PZD statistic, the most stringent efficacy indicator, treatment is significantly more effective for individuals with ID than individuals diagnosed with autism spectrum disorders. Table 3.8 includes data regarding the means and standard deviations for the three calculated effect sizes: MBLR, PND, and PZD.

DISCUSSION

Summary of Findings

The current study focused on a series on questions regarding the impact of functional assessment methodology, functional assessment outcome, and diagnostic category on treatment effectiveness. The study focused on the following questions. First, is treatment effectiveness influenced by the type of FBA? Second, is ascribed function influenced by FBA methodology? Third, is ascribed function influenced by participant characteristics such as diagnostic category? Fourth, is the effectiveness of treatment mediated by the function of the maladaptive behavior? Lastly, is there a relationship between diagnostic category and treatment effectiveness outcomes?

The main findings of the quantitative synthesis are summarized as follows. When comparing treatment outcomes for interventions based on both experimental and non-experimental FBA for all three diagnostic groups, no significant differences were found. This finding adds to the mixed results currently found in the literature regarding comparisons of FBA methodologies. Findings also indicate that FBA methodology itself moderates the outcomes of the assessment. For example, FA were more likely to result in maladaptive behaviors being identified as maintained by social positive reinforcement contingencies as opposed to BA which most often identified automatic functions. There were also differences in the ability of FBA methodologies to detect multiply-maintained behaviors. One potential hypothesis to explain these findings is the possibility of rater bias inherent in many forms of BA. The BA is often dependent on the ability of the rater to report accurate information and, like other rating scales, is subject to rater biases. In addition, the goal of many BA is to identify the most significant function of the maladaptive behavior. Potentially useful information may be lost when the assessment methodology assumes maladaptive behaviors are maintained by only one function and are not multiply-maintained. Significant differences in FBA functional outcomes were found depending on the type of assessment conducted. These findings differ from previous research that indicates that FA and BA methodologies themselves do not impact the results of the assessment (e.g., Herzinger & Campbell, 2007).

Diagnostic category was identified as significantly affecting ascribed function, regardless of FBA type used. For individuals with autism, maladaptive behavior was more likely to be identified as being maintained by social negative reinforcement (i.e., escape). These findings support the hypothesis that marked impairment in socialization and communication, as evinced by a diagnosis of ASD, influences the function of maladaptive behavior. It is possible that

individuals with ASD are more likely to avoid situations due to a deficit in communication and socialization rather than a skill deficit. The results indicate that the functions of maladaptive behavior, in and of themselves, have no significant impact on treatment effectiveness. That is, the ascribed function of behavior does not mediate the effectiveness of interventions. Although previous research (e.g., Vollmer, 1994; Piazza, Hanley, & Fisher, 1996), has indicated that behaviors maintained by some functions are more difficult to treat, the results of the current study and others (Herzinger & Campbell, 2007) do not support that notion. A significant relationship was observed when assessing the impact of diagnostic category on treatment effectiveness. Interventions were more successful for individuals with ID, but not diagnosed with autism, than for any other category when assessed using the PZD statistic. Follow-up comparisons were made across diagnostic categories to help explain this finding. Due to the overrepresentation of males in the two autism categories (i.e., autism only and autism/ID), three one-way ANOVAs were conducted. Treatment effectiveness as measured by the (a) MBLR statistic, F (1, 198) = 2.44, *n.s.*, (b) PND statistic, F (2, 198) = 1.20, *n.s*, and (c) PZD statistic, F (1, 198) = .17, *n.s.* was not significantly affected by the gender of the individual. This indicated that the higher PZD scores for individuals with ID (and without characteristics of autism) was not due to unequal gender ratios across diagnostic categories. Comparisons of communication ability were also made across three diagnostic categories to determine if language skills were the moderating variable impacting treatment effectiveness. The results of a one-way ANOVA indicated no significant differences in average level of communication across the diagnostic categories, F (2, 198) = 2.73, *n.s*. The follow-up comparisons across diagnostic categories could not explain the differences in treatment outcome as measured by the PZD statistic.

Implications for clinicians

The results of the quantitative review are relevant to practicing clinicians regarding the assessment and treatment of severe maladaptive behaviors exhibited by individuals with developmental disabilities. The most salient finding, with immediate implications for practitioners, is the effect of diagnostic category on identified function of maladaptive behaviors and treatment outcome effectiveness. If an individual's diagnosis influences the maintaining contingencies of exhibited maladaptive behaviors, knowledge of that diagnosis can inform treatment development from the beginning. Similarly, the impact of diagnostic category on treatment effectiveness, when assessed by the PZD statistic, could inform decisions regarding treatment outcome expectations at the onset of evaluation and treatment development. Currently, some behavioral clinicians and researchers do not use diagnostic categories to describe the participants in their studies. This may be a nod to the behavioral perspective that focuses on observable behavior and making treatment decisions based on that rather than assumed characteristics implied by diagnostic associations. However, the results of the current study suggest that diagnostic category influences the functional relationships for maladaptive behaviors. Procedural concerns regarding the use of FBA methodologies that are prone to particular results (i.e., ascribing function to behavior based on inherent methodological flaws) are also important to consider. If the results of FBAs can be hypothesized by simply knowing the type of FBA methodology used, the results can not be considered valid. This finding has immediate implications for clinicians who interpret the results of any FBA without further assessment and continuous evaluation. The implication that FBA methodologies may impact the

results (i.e., ascribed function of target behaviors) reinforces the notion that interventions should be continuously evaluated for effectiveness, not deemed effective because they are based on the ascribed function of the maladaptive behavior.

Limitations of the literature

Conducting a meta-analysis allows researchers to synthesize the findings of several primary articles that utilize single subject research to determine "general findings". However, these findings are inherently impacted by the quality of the primary articles included. The literature reviewed for the current meta-analysis contained several limitations. One limitation is the possibility that articles that are selected for publication are biased or skewed in some ways. For example, studies that report poor treatment effectiveness may go unpublished and thus the average effect sizes reported within this review represent overestimates. Also, FBAs that have undifferentiated results and are not further assessed may not be published and therefore not included in the current dataset.

In addition, many articles did not include potentially useful information about the characteristics of the participants. Basic demographic data such as race, age, and even diagnosis were often not reported in the primary articles. Location of assessment and treatment sessions was also excluded from most articles. The lack of information provided and the possible effects of this exclusion have been reported by others. However encouraged, many researchers are still excluding important information about participants and methodological design from their studies.

Though not initially coded and recorded, follow-up reviews of a sample of the included literature indicated that less than 25% of the studies reported procedural fidelity inter-rater reliability. However, when it was reported it was typically reported only for intervention phases

and never for BAs. IT is difficult to make comparisons across FBA methodologies if there is no guarantee that the methods were implemented as intended. Also, data common to multiple investigations may have unintentionally been coded more than once in a quantitative synthesis such as this if not noted by the primary article author. That is, some investigators may have presented treatment outcomes on the same participants in separate published articles without acknowledging these circumstances. In some cases, articles did not meet inclusion criteria because a diagnosis of autism or claim that participant was "autistic-like" was not explicitly stated. The lack of information presented could affect not only the results of the analyses but also attempts to generalize the findings.

In contrast, there were several exemplary examples within the quantitative synthesis. For example, Beare, Severson, and Brandt (2004), provided a thorough description of the participant included in the study. They described the participant's intellectual functioning (including scores on standardized measures of cognitive ability), expressive and receptive language ability, adaptive functioning, as well as the maladaptive behaviors he engaged in and the direct impact of those behaviors. Adelinis, Piazza, and Goh (2001) acknowledged another article that reported data on the same participant. Hanley, Piazza, Fisher, and Maglieri (2005) described the analog conditions included in the FA, as well as the reasons for their inclusion. These authors demonstrated how practitioners can use information from a variety of sources (i.e., BA methods) when developing thorough assessment and intervention strategies.

Limitations of the current study

Conclusions of the review must be considered within the context of its limitations. One limitation of this research synthesis is the exclusion of unpublished studies, including unpublished theses and dissertations. It is possible that the studies included represent a skewed

portion of the population and are not representative of the whole. It is also possible that published articles that met inclusion criteria may have been unintentionally excluded from the review.

Another limitation of the current study is that subgroups of both assessment and treatment types were combined throughout the analyses. For example, FAs reported as "modified" and "brief" analogue sessions were included under the FA category, along with traditional FAs. Categories were combined in order to assess the effectiveness of experimental versus non-experimental assessments rather than specific subtypes of assessment. Existing research indicates possible differences in outcomes for subtypes of experimental analyses (e.g., Hanley, Iwata, & McCord, 2003) as well as subtypes of non-experimental analyses (e.g., Arndorfer et al., 1994; Cunningham & O'Neill, 2000). Combining all types of experimental analyses under the FA category may have influenced the validity of the results for all subtypes. The same is true for the BA category. Similarly, coding intervention groups into six categories, including three groups comprised of multiple components, may not capture the differences between specific types of treatment (e.g., verbal reinforcement, tangible reinforcement).

The procedure used to calculate effect sizes used for comparison could be considered another limitation of the current study. Treatment effectiveness was summarized by examining the first baseline and last treatment phase reported in the primary studies. The choice to use these phases was necessary for legitimate comparison of non-regression-based effect sizes. For example, in some studies, several different treatments were assessed and reported in an A-B-C-D design. In this case, the rate of behavior reported in phase D was compared to the baseline data reported in phase A to determine effectiveness of treatment. However, this choice resulted in a

loss of information available in published reports that may have altered effect sizes in unknown ways.

Recommendations for Future Research

The results of the current study indicate that diagnostic category impacts the assessment and treatment of individuals who engage in maladaptive behaviors. Systematic assessment of diagnostic categories and possible influencing characteristics (i.e., level of intellectual disability; level of communication ability) may help further guide treatment development for these individuals. As indicated as a limitation of the current study, direct comparisons of different types of experimental FAs and non-experimental BAs might be useful. For example, in the current study all methods of experimental assessments were subsumed under the FA category. Brief FAs as described by Northup et al. (1991) and later by Derby et al. (1992) were categorized with full-length FAs and were not directly compared with other, less time intensive assessment methodologies. Future research may include a direct comparison between specific subtypes of FAs to BAs. Another line of research that would yield comparable results would be to design a single subject study of original data collection in which comparisons are made for individuals who have been administered a multitude of assessments (e.g., an interview, MAS rating scale, brief FA, and extended FA), comparing both assessment outcome and treatment effectiveness.

REFERENCES

Allison, D.B., Faith, M.S., & Franklin, R.D. (1995). Antecedent exercise in the treatment of disruptive behavior: A meta-analytic review. *Clinical Psychology: Science and Practice,* 2, 279-303.

Allison, D.B., & Gorman, B.S. (1993). Calculating effect sizes for meta-analysis: The case of the single case. *Behaviour Research and Therapy, 31,* 621-631.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders: Fourth Edition Text Revision.* Washington, DC: American Psychiatric Association.

Anderson, C.M. & Long, E.S. (2002). Use of a structured descriptive assessment methodology to identify variables affecting problem behavior. *Journal of Applied Behavior Analysis, 35,* 137-154.

Arndorfer, R.E., Miltenberger, R.G., Woster, S.H., Rortvedt, A.K., & Gaffaney, T. (1994). Homebased descriptive and experimental analysis of problem behaviors in children. *Topics in Early Childhood Special Education, 14,* 64-87.

Bijou, S.W., Peterson, R.F., & Ault, M.H. (1968). A method to integrate descriptive and experimental field studies at the level of data and empirical concepts. *Journal of Applied Behavior Analysis, 1,*175–191.

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.

Campbell, J.M. (2004). Statistical comparison of four effect sizes for single-subject designs. *Behavior Modification, 28,* 234-246.

Campbell, J.M., Herzinger, C.H., & James, C.L. (2007). Evidenced based therapies for autistic disorder and pervasive developmental disorders. In R.G. Steele, T.D. Elkin and M.C. Roberts, (Eds.) *Handbook of evidence-based therapies for children and adolescents* (pp.371-388). Springer.

Carr, E.G., Langdon, N.A., & Yarborough, S.C. (1999). Hypothesis-based intervention for severe problem behavior. In A. C. Repp, & R. H. Horner, (Eds.), *Functional analysis of problem behavior* (pp. 9–31). Belmont, CA: Wadsworth Publishing Company.

Cunningham, E., & O'Neill, R.E. (2000). Comparison of results of functional assessment and analysis methods with young children with autism. *Education and Training in Mental Retardation and Developmental Disabilities, 35,* 406-414.

Dawson, J.E., Matson, J.L., & Cherry, K.E. (1998). An analysis of maladaptive behaviors in persons with autism, PDD-NOS, and mental retardation. *Research in Developmental Disabilities, 19,* 439–448.

Day, R.M., Rea, J.A., Schussler, N.G., Larsen, S.E., & Johnson, W.L. (1988). A functionally based approach to the treatment of self-injurious behavior. *Behavior Modification, 12,* 565–589.

Derby, K.M., Wacker, D.P., Sasso, G., Steege, M., Northup, J., Cigrand, K., & Asmus, J. (1992). Brief functional assessment techniques to evaluate aberrant behavior in an outpatient setting: A summary of 79 cases. *Journal of Applied Behavior Analysis, 25,* 713-721.

Didden, R., Duker, P.C., & Korzilius, H. (1997). Meta-analytic study on treatment effectiveness for problem behaviors with individuals that 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 on Mental Retardation, 111,* 290-298.

Durand, V.M., & Crimmins, D.B. (1988). Identifying the variables maintaining self-injurious behavior. *Journal of Autism and Developmental Disorders, 18*, 99–117.

Durand, V.M., & Crimmins, D.B. (1992). *The motivation assessment scale* (*MAS*) *administration guide*. Topeka, KS: Monaco and Associates.

Faith, M.S., Allison, D.B., & Gorman, B.S. (1996). Meta-analysis of single-case research. In R.D. Franklin, D.B. Allison, & B.S. Gorman, (Eds.), *Design and analysis of single-case research* (pp. 245-277). Hillsdale, NJ: Lawrence Erlbaum.

Fombonne, E. (2005). Epidemiology of autistic disorder and other pervasive developmental disorders. *Journal of Clinical Psychiatry, 66,* 3-8.

Gresham, F.M., McIntyre, L.L., Olson-Tinker, H., Dolstra, L., McLaughlin, V., & Van, M. (2004). Relevance of functional behavioral assessment research for school-based interventions and positive behavioral support. *Research in Developmental Disabilities, 25*, 19-37.

Groden, G. (1989). A guide for conducting a comprehensive behavioral analysis of a target behavior. *Journal of Behavior Therapy and Experimental Psychiatry, 20,* 163–169.

Hall, S.S. (2005). Comparing descriptive, experimental, and informant-based assessments of problem behaviors. *Research in Developmental Disabilities, 26,* 514-526.

Hanley, G.P., Iwata, B.A., & McCord, B.E. (2003). Functional analysis of problem behavior: A review. *Journal of Applied Behavior Analysis, 36,* 147-185.

Herzinger, C.V., & Campbell, J.M. (2007). Comparing functional assessment methodologies: A quantitative synthesis. *Journal of Autism and Developmental Disorders, 37,* 1430-1445.

Herzinger, C.V. (2008). Assessment and treatment of maladaptive behaviors within the developmental disabilities population: A mega-analysis. Unpublished manuscript.

Individuals with Disabilities Education Act Amendments (1997). Pub.L. No. 105–117, USC 1400 et seq.

Iwata, B.A., Dorsey, M.F., Slifer, K.J., Bauman, K.E. & Richman, G.S. (1982). Toward a functional analysis of self-injury. *Analysis and Intervention in Developmental Disabilities, 2,* 3-20.

Kennedy, C.H., & Souza, G. (1995). Functional analysis and treatment of eye poking. *Journal of Applied Behavior Analysis, 28,* 27-37.

Lalli, J.S., Vollmer, T.R., Progar, P.R., Wright, C., Borrero, J., Daniel, D., Barthold, C.H., Tocco, K., & May, W., (1999). Competition between positive and negative reinforcement in the treatment of escape behavior. *Journal of Applied Behavior Analysis, 32,* 285-296.

Leblanc, L., Patel, M.R., & Carr, J.E. (2000). Recent advances in the assessment of aberrant behavior maintained by automatic reinforcement in individuals with developmental disabilities. *Journal of Behavior Therapy and Experimental Psychiatry, 31,* 137-154.

Lovaas, O.I., Freitag, G., Gold, V.J., & Kassorla, I.C., (1965). Experimental studies in childhood schizophrenia: Analysis of self-destructive behavior. *Journal of Experimental Child Psychology, 2,* 67-84.

Lundervold, D., & Bourland, G. (1988). Quantitative analysis of treatment of aggression, self-injury, and property destruction. *Behavior Modification, 12,* 590-617.

Ma, H. (2006). An alternative method for quantitative synthesis of single-subject researches. *Behavior Modification, 30,* 598-617.

Matson, J.L., & Vollmer, T.R. (1995). *User's guide: Questions about behavioral function* (*QABF*). Baton Rouge, LA: Scientific Publishers Inc.

Mirenda, P. (1997). Supporting individuals with challenging behavior through functional communication training and AAC: Research review. *Augmentative and Alternative Communication, 13,* 207-225.

Northup, J., Wacker, D., Sasso, G., Steege, M., Cigrand, K., Cook, J., & DeRaad, A. (1991). A brief functional analysis of aggressive and alternative behavior in an outclinic setting. *Journal of Applied Behavior Analysis, 24,* 509-522.

O'Brien, S., & Repp, A.C., (1990). Reinforcement-based reductive procedures: A review of 20 years of their use with persons with severe to profound mental retardation. *Journal of the Association for Persons with Severe Handicaps, 15,*148-159.

O'Neill, R.E., Horner, R.H., Albin, R.W., Storey, K., Sprague, J.R., & Newton, J.S. (1997). *Functional assessment of problem behavior: A practical assessment guide*. Pacific Grove, CA: Brooks/Cole.)

Parker, R.I., & Hagan-Burke, S., (2007). Median-based overlap analysis for single case data. *Behavior Modification, 31,* 919-936.

Patel, M.R., Carr, J.E., Kim, C., Robleson, A., & Eastridge, D., (2000). Functional analysis of aberrant behavior maintained by automatic reinforcement: Assessments of specific sensory reinforcers. *Research in Developmental Disabilities, 21,* 393-407.

Paclawskyj, T.R., Matson, J.L., Rush, K.S., Smalls, Y., & Vollmer, T.R. (2001). Assessment of the convergent validity of the Questions About Behavioral Function scale with analogue

functional analysis and the Motivation Assessment Scale, *Journal of Intellectual Disability Research, 45,* 484-494.

Piazza, C.C., Hanley, G.P., & Fisher, W.W. (1996). Functional analysis of cigarette pica. *Journal of Applied Behavioral Analysis, 29,* 437–439.

Potoczak, K., Carr, J.E., & Michael, J. (2007). The effects of consequence manipulation during functional analysis of problem behavior maintained by negative reinforcement. *Journal of Applied Behavior Analysis, 40,* 719-724.

Repp, A.C., Felce, D., & Barton, L.E. (1988). Basing the treatment of stereotypic and self-injurious behaviors on hypotheses of their causes. *Journal of Applied Behavior Analysis, 21,* 281-289.

Repp, A., & Horner, R.H. (Eds.). (1999). *Functional analysis of problem behavior: From effective assessment to effective support.* Belmont, CA: Wadsworth Publishing

Rincover, A. (1978). Sensory extinction: A procedure for eliminating self-stimulatory behavior in developmentally disabled children. *Journal of Abnormal Child Psychology, 6,* 299-310.

Roscoe, E.M., Iwata, B.A., & Goh, H. (1998). A comparison of noncontingent reinforcement and sensory extinction as treatments for self-injurious behavior. *Journal of Applied Behavior Analysis, 31,* 635-646.

Sailor, W., Guess, D., Rutherford, G., & Baer, D.M. (1968). Control of tantrum behavior by operant techniques during extinction of self-injurious escape behavior. *Journal of Applied Behavior Analysis, 1,* 237-243.

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,* 233-256.

Scruggs, T.E., Mastropieri, M.A., Cook, S.B., & Escobar, C. (1986). Early intervention for children with conduct disorders: A quantitative synthesis of single-subject research. *Behavioral Disorders, 11,* 260-271.

Scruggs, T.E., Mastropieri, M.A., & Casto, G. (1987). The quantitative synthesis of single-subject research: Issues and applications. *Behavior Modification, 22*, 221- 242.

Skinner, B.F. (1953). *Science and Human Behavior.* New York: Free Press.

Touchette, P.E., MacDonald, R.F., & Langer, S.N. (1985). A scatter plot for identifying stimulus control of problem behavior. *Journal of Applied Behavior Analysis, 18,* 343–351.

Vollmer, T.R. (1994). The concept of automatic reinforcement: implications for behavioral research in developmental disabilities. *Research in Developmental Disabilities, 15,* 187–207.

Walling, L.L. & Irwin, M.M. (Eds.). (1995). *Information services for people with developmental disabilities: The library manager's handbook.* Westport, CN: Greenwood Press.

Weiten, W. (1995). *Psychology: Themes and Variations.* Brooks/Cole Publishing Co.

Wieseler, N.A., Hanzel, T.E., Chamberlain, T.P., & Thompson, T. (1985). Functional taxonomy of stereotypic and self-injurious behavior. *Mental Retardation, 23,* 230–234.

Wolery, M., Ault, M., & Doyle, P.M. (1992). *Teaching students with moderate to severe disabilities: Use of response prompting strategies.* White Plains, NY: Longman.

Wolf, M.M., Birnbrauer, J., Williams, T., Lawler, J. (1965). A note on apparent extinction of the vomiting behavior of a retarded child. In L.P. Ullmann, L. Krasner, (Eds.), *Case studies in behavior modification* (pp. 364–366). New York: Holt, Rinehart & Winston.

Yarborough, S.C. & Carr, E.G. (2000). Some relationships between informant assessment and functional analysis of problem behavior. *American Journal on Mental Retardation, 105,* 130-151.

Table 3.1

Participant Characteristics

Characteristic	n	%	r	M	SD
Gender					
Male	87	60.4			
Female	57	39.6			
Age (in years)			1.30-51.00	15.95	11.97
Race/Ethnicity					
Not reported	142				
African-American	1				
Caucasian	1				
Main Diagnostic Category					
ID	79	54.9			
Autism/ID	35	24.3			
Autism	18	12.5			
Dev. Disability	12	8.3			
Level of Intellectual Disability (IQ range)					
Severe (<39)	79	54.9			
Not reported	28	19.4			
Moderate (54-40)	26	18.1			
Mild (70 - 55)	9	6.3			
Untestable/Other	2	1.4			

Table 3.1 continued

Participant Characteristics

Characteristic	n	%	r	M	SD
Autism diagnostic criteria used					
Not reported	141	97.9			
DSM-IV	3	2.1			
ID diagnostic criteria used					
Not reported	134	93.1			
Other	4	2.8			
Stanford Binet	2	1.4			
Vineland	2	1.4			
WISC	1	.7			
WPPSI	1	.7			
Months since diagnosis	100% Not reported				
Language Ability					
Nonverbal/Mute	52	36.1			
Not reported	48	33.3			
Some functional language	39	27.1			
Average language	4	2.8			
Echolalic	1	.7			

Table 3.1 continued

Participant Characteristics

Characteristic	n	%	r	M	SD
Medication administered					
Not reported	118	59.3			
No	53	26.6			
Yes	28	14.4			

Note. n = number of participants; % = percentage of participants; r = range; M = mean; SD = standard deviation.

Table 3.2

Description of Study Characteristics

Characteristic	n	%
Journal		
Journal of Applied Behavior Analysis	97	48.7
Behavioral Interventions	39	19.6
Research in Developmental Disabilities	27	13.6
American Journal on Mental Retardation	18	9.0
Journal of Autism and Developmental Disorders	9	4.5
Behavior Modification	7	3.5
Other	2	1.0
Total N	199	
Number of participants per article		
1	48	57.8
2	15	18.1
3	14	16.9
4	4	4.8
5	2	2.4
Total N	83	

Note. n = number of studies; % = percentage of studies.

Table 3.3

Assessment, Intervention, and Experimental Characteristics

Characteristic	n	%
Assessment and Intervention characteristics		
Behavior targeted for reduction		
Self-injurious behavior	40	20.1
Aggression	34	17.1
Self-stimulation/Stereotypy	30	15.1
Disruptive behavior	25	12.6
Other	3	0.02
Property Destruction	1	0.01
Combinations	66	33.2
Combination Breakdown:		
Aggression	58	87.8
SIB	49	74.2
Disruptions	45	68.2
Property destruction	30	45.5
Other	10	15.2
Self-stimulation	1	1.5

Table 3.3 continued

Assessment, Intervention, and Experimental Characteristics

Characteristic	n	%
Type of functional behavioral assessment		
Experimental	154	72.9
Modified session EFA	83	
EFA (Iwata et al.)	49	
Brief EFA	16	
Antecedent Analysis	4	
Other	2	
Non-experimental	45	22.6
Informal Assessment	33	
Combination of BA types	4	
Descriptive Assessment	4	
ABC sheet	3	
Not reported	1	
Type of intervention		
Reinforcement only	106	53.3
Extinction and Reinforcement or Punishment	53	26.6
Other/Not reported	20	10.1
Reinforcement and Punishment	10	5.0
Extinction only	6	3.0
Punishment only	4	2.0

Table 3.3 continued

Assessment, Intervention, and Experimental Characteristics

Characteristic	n	%
Behavioral Therapist		
Not reported or insufficient information	146	73.4
Parent	13	6.5
Combination of therapists	11	5.5
Teacher	9	4.5
Undergraduate/Graduate student	7	3.5
Psychologist	6	3.0
Other	6	3.0
Self	1	0.5
Treatment Setting		
Center for Developmental Disabilities	42	21.1
Inpatient ward/Hospital	38	19.1
Public School	33	16.6
Home	22	11.1
Residential treatment setting	22	11.1
Combination of settings	13	6.5
Inpatient classroom	8	4.0
Not reported	8	4.0
Outpatient clinic	7	3.5
Other	6	3.0

Table 3.3 continued

Assessment, Intervention, and Experimental Characteristics

Characteristic	n	%
Experimental characteristics		
Experimental design		
Reversal/Withdrawal	62	31.2
Multiple Baseline	59	29.6
Multiple Treatment Comparison	27	13.6
Alternating Treatments	24	12.1
Combination	15	7.5
Simple A-B	11	5.5
Other	1	0.5
Follow-up data collected		
No	164	82.4
Yes	28	14.1
Not Reported	7	3.5
Attempt to generalize behavior		
No/Not reported	142	71.4
Yes	57	28.6
Attempt to generalize to:		
New situation, setting, context	37	64.9
Combination	14	24.6

Table 3.3 continued

Assessment, Intervention, and Experimental Characteristics

Characteristic	n	%
New behavior	3	5.3
Other	2	3.5
New people	1	1.8

Characteristic	Range	M	SD
Reliability of observations			
Inter-rater reliability			
FBA	80.0-100.0	75.6	36.8
Treatment	84.0-100.0	94.3	9.5

Note. n = number of studies; % percentage of studies; M = mean; SD = standard deviation.

Table 3.4

Descriptive Statistics for Three Effect Sizes by FBA Methodology

	M	SD	Min	Max
FA				
MBLR	81.29	25.93	-49.03	100.00
PND	81.06	30.53	0.00	100.00
PZD	58.77	35.36	0.00	100.00
BA				
MBLR	75.86	29.65	-14.14	100.00
PND	77.09	32.67	0.00	100.00
PZD	51.12	35.66	0.00	100.00

Results from One-way ANOVAs

MBLR: F (1, 198) = 1.43, *n.s*

PND: F (1, 198) = .57, *n.s*

PZD: F (1, 198) = 1.62, *n.s*

Note. FA = functional analysis; BA = behavioral assessment; MBLR = mean baseline reduction; PND = percentage of non-overlapping data; PZD = percentage of zero data; M = mean; SD = standard deviation; Min = minimum value; Max = maximum value. Descriptive statistics are presented for 199 treatment outcomes.

Table 3.5

Frequencies for Assessment Type and Function

	Function					
Assessment Type	POS RF	NEG RF	AUT	COM	UND	OTH
FA	49	36	31	27	9	2
BA	12	12	16	1	0	4
Total	61	48	47	28	9	6

Note. FA = functional analysis; BA = behavioral assessment; POS RF = social positive reinforcement; NEG RF = social negative reinforcement; AUT = automatic; COM = combination; UND = undifferentiated; OTH = other

Table 3.6

Frequencies for Diagnostic Category and Function

	Function					
	POS RF	NEG RF	AUT	COM	UND	OTH
Diagnostic Category						
AUT	8	14	4	1	3	0
ID	36	15	26	16	1	6
AUT/ID	17	19	17	11	5	0
Total	61	48	47	28	9	6

Note. AUT = autism diagnosis; ID = intellectual disability diagnosis; AUT/ID = autism and ID diagnoses; POS RF = social positive reinforcement; NEG RF = social negative reinforcement; AUT = automatic; COM = combination; UND = undifferentiated; OTH = other

Table 3.7

Means and Standard Deviations for Three Effect Sizes by Ascribed Function

Function	MBLR	PND	PZD
POS RF	82.29 (24.29)	83.81 (24.91)	61.57 (29.20)
NEG RF	77.74 (27.54)	75.46 (32.30)	54.16 (38.85)
AUT	79.22 (30.35)	82.60 (34.37)	50.10 (39.24)
COM	77.87 (28.24)	71.43 (36.72)	58.61 (34.57)
UND	81.62 (26.11)	82.82 (29.10)	61.37 (36.13)
OTH	95.02 (5.79)	98.07 (3.24)	89.55 (15.18)

Note. POS RF = social positive reinforcement; NEG RF = social negative reinforcement; AUT = automatic; COM = combination; UND = undifferentiated; OTH = other; MBLR = mean baseline reduction; PND = percentage of non-overlapping data; PZD = percentage of zero data

Table 3.8

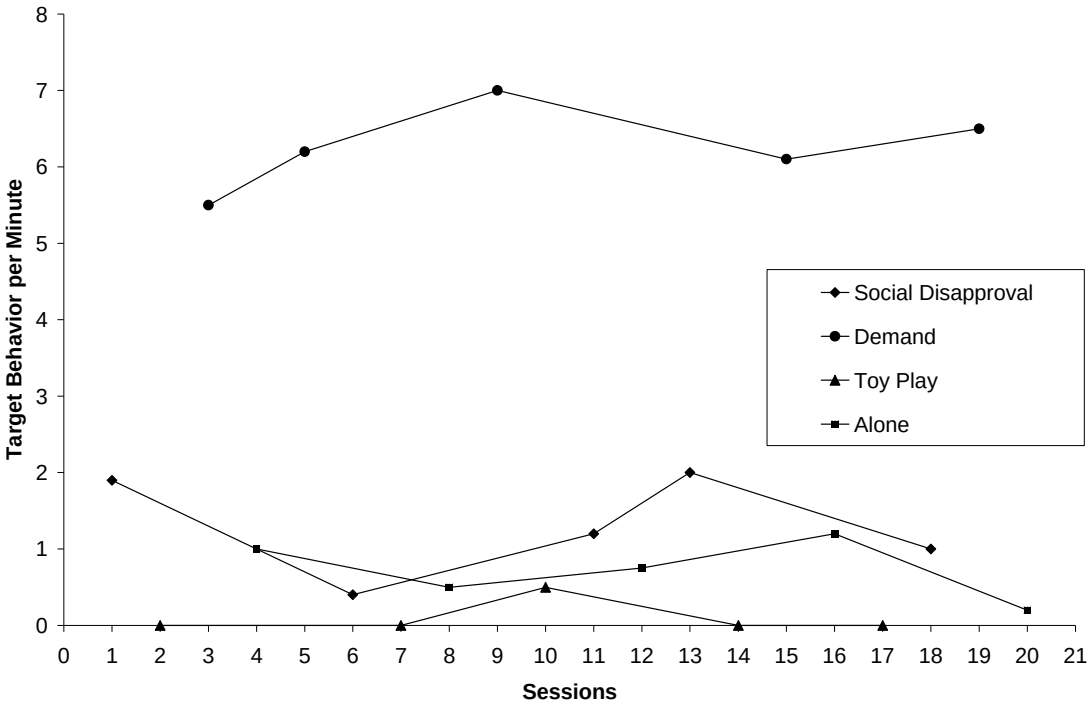
Means and Standard Deviations for Three Effect Sizes by Diagnostic Category

Diagnostic Category	MBLR	PND	PZD
AUT	80.84 (13.47)	75.81 (29.39)	45.84 (32.04)
ID	81.27 (28.62)	81.89 (30.30)	63.23 (34.52)
AUT/ID	77.13 (28.57)	78.95 (33.44)	50.32 (37.07)

Note. AUT = autism diagnosis; ID = intellectual disability diagnosis; AUT/ID = autism and ID diagnoses; MBLR = mean baseline reduction; PND = percentage of non-overlapping data; PZD = percentage of zero data

Figure Caption

Figure 3.1. Results of a Hypothetical Functional Analysis.



CHAPTER 4

GENERAL DISCUSSSION

Given the prevalence of autism and intellectual disability and the propensity for individuals diagnosed with these disabilities to engage in maladaptive behaviors, the systematic assessment and treatment of maladaptive behavior are important areas of research. The behavioral approach to assessment and intervention for maladaptive behavior, which has been heralded as the most efficacious (Campbell, Herzinger, & James, 2007), includes identification of the purpose or function of the target behavior to aid in treatment development. This approach has been supported by the federal government through amendments to IDEA, which require a functional behavioral assessment (FBA) prior to placement changes for individuals with individualized education programs who engage in maladaptive behaviors and violate the school code of conduct (P.L. 105-117, 1997; Gresham, 2004). The goal of various FBA methodologies is to establish the purpose or function of the behavior in order to generate hypotheses about how to approach treatment. The current research identified general themes within the developmental disability research that focused on FBA methodology and highlighted gaps within the literature. Those findings guided the meta-analysis, in which existing single subject literature was synthesized to assess potential moderating variables on treatment outcome.

The findings of the mega-analysis (i.e., review of meta-analyses) indicate that the behavioral approach to assessment and treatment of maladaptive behavior is more effective than non-behavioral approaches. The superiority of the behavioral approach over other approaches was found across three ‘effect size’ indicators (i.e., MBLR, PND, and PZD). As evidenced by all three efficacy indicators, treatments based on the behavioral approach were more successful at

reducing rates of maladaptive behavior than those based on non-behavioral approaches, such as psychopharmacology. Several meta-analyses examined the impact of FBA on treatment effectiveness with mixed conclusions. In general, the data suggest that FBAs are associated with greater treatment effectiveness, but the significance of that difference appears to depend on the effect size reported. For example, when comparing treatment outcomes for studies that included a FBA and those that did not, the PND means were 70 and 64, respectively. When comparing treatment outcomes for studies that include a FBA and those that did not, the PZD means were 47 and 39, respectively. The findings suggest a difference in the goal of intervention, as PND is a measure of behavior reduction and PZD is a measure of behavioral suppression (Campbell, 2004). More importantly, the results indicate that even when FBAs are conducted, the results of such assessments are not always used in the selection of treatment. This could be one explanation for the contradictory results. The current meta-analysis only included those studies which included some form of FBA, thus were predominantly behavioral in nature. No comparisons were made across theoretical approaches or regarding FBA inclusion in the current meta-analysis.

In the mega-analysis, comparisons of FBA methodology (i.e., FA versus BA) were conclusive. For all three effect size means, interventions that included FAs were more effective than those that included a BA as part of the intervention process. Contradictorily, in the current meta-analysis, the findings indicate that there were no clinically significant differences in treatment outcome based on FBA methodology across diagnostic category. However, meta-analysis findings did indicate that FBA methodology itself moderates assessment outcome. For example, FAs were more likely to result in maladaptive behavior being identified as maintained by social positive reinforcement contingencies as opposed to BA which most often identified

automatic functions. There were also differences in the ability of FBA methodologies to detect multiply-maintained behavior. FAs were more likely to identify more than one function of target behavior and BA methodologies were more likely to identify only one. These findings illustrate the impact of FBA methodology on assessment outcome.

When reviewing the information regarding participant characteristics as moderators of treatment effectiveness, again, the mega-analysis provided inconclusive results. Some meta-analyses reported participant characteristics, such as age and level of cognitive ability, as moderating treatment effectiveness. Others reported no significant relationship between participant characteristics and treatment outcome. In the current meta-analysis, participant characteristics such as diagnostic category were found to significantly influence functional assessment outcome and treatment effectiveness. Regarding FBAs, diagnosis moderated outcomes regardless of the methodology employed. For individuals with autism, maladaptive behavior was more likely to be identified as being maintained by social negative reinforcement (i.e., escape). The findings support the hypothesis that marked impairment in socialization and communication, as evinced by a diagnosis of ASD, influences the function of maladaptive behavior. It is possible that individuals with ASD are more likely to avoid situations due to a deficit in communication and socialization rather than a skill deficit. Regarding intervention, outcomes were more successful for individuals with ID, but not diagnosed with autism, than for any other category when assessed using the PZD statistic. Follow-up comparisons across diagnostic categories on gender and communication ability could not account for this finding.

The dissertation attempts to synthesize the growing literature base regarding assessment and treatment of maladaptive behavior within the developmental disability population. Together, the review and meta-analysis highlight the need for additional evaluation of assessment

methodologies and their effects on the outcome of interventions aimed at reducing maladaptive behavior. Future research in this area should include comparisons of specific FBA methodologies that are matched on time and effort and thorough evaluation of potential moderating variables such as diagnosis, level of ID, communication ability, and prior history of treatments. Another line of research that may prove useful would be the inclusion of additional efficacy indicators (e.g., PEM; Ma, 2006) to assess treatment outcome.

REFERENCES

Campbell, J.M. (2004). Statistical comparison of four effect sizes for single-subject designs. *Behavior Modification, 28,* 234-246.

Campbell, J.M., Herzinger, C.H., & James, C.L. (2007). Evidenced based therapies for autistic disorder and pervasive developmental disorders. In R.G. Steele, T.D. Elkin and M.C. Roberts, (Eds.) *Handbook of evidence-based therapies for children and adolescents* (pp.371-388). Springer.

Gresham, F.M., McIntyre, L.L., Olson-Tinker, H., Dolstra, L., McLaughlin, V., & Van, M. (2004). Relevance of functional behavioral assessment research for school-based interventions and positive behavioral support. *Research in Developmental Disabilities, 25,* 19-37.

Individuals with Disabilities Education Act Amendments (1997). Pub.L. No. 105–117, USC 1400 et seq.

Ma, H. (2006). An alternative method for quantitative synthesis of single-subject researches. *Behavior Modification, 30,* 598-617.

APPENDIX A

Interrater Reliability Data

Variable	Percent Agreement
Age	100.00
Gender	100.00
Race	100.00
Diagnostic Category	100.00
Number of autism characteristics	98.99
Autism diagnostic criteria used	100.00
Level of intellectual disability	100.00
IQ/Adaptive behavior scores	98.99
Level of communication	95.98
Secondary diagnoses	100.00
Time in prior treatment	100.00
Target behavior	100.00
SIB type	100.00
FBA conducted?	100.00
FA type	100.00
Function of behavior	100.00
Type of intervention	97.99
Type of extinction	98.99
Type of differential reinforcement	98.99
Type of punishment	100.00
Intervention schedule	100.00
Length of sessions	100.00
Medication administered?	100.00
Treatment setting	100.00
Dimensions of setting	100.00
Behavioral therapist	100.00
Parental involvement	99.49
Experimental Design	97.49
Effort to generalize?	100.00
Generalization to what?	100.00
Follow up data collected?	100.00
Scale used to measure target behavior	100.00
Units on X axis	100.00
FA: interrater reliability	100.00
FA: type of agreement	100.00
FA: agreement recording system	100.00
TX: interrater reliability	100.00
TX: type of agreement	100.00
TX: agreement recording system	100.00
Source of observations	100.00
# of data points in first baseline	100.00
# of data points in final treatment	100.00
MBLR	100.00
PND	100.00
PZD	100.00

Mean = 99.71, Range = 95.98 – 100.00