Problem behaviors are a commonly associated feature of people diagnosed with autism. Functional assessment has become the norm for evaluating these behaviors. There has been much research concerning functional assessment over the past twenty years, but several important research considerations have yet to be explained. One is the comparison of different types of functional assessment (e.g., experimental functional analysis and behavioral functional assessment). Controversy exists regarding the types of functional assessments and their validity. The current study aims to compare the different methodologies of functional assessment and their effectiveness in ascribing function to a target behavior and in the treatment selection that follows such an assessment. Quantitative synthesis data were used to answer questions regarding behavioral function, assessment type, and treatment effectiveness. Results indicate that assessment type can impact treatment effectiveness and that there is a relationship between type of behavior and ascribed behavioral function.

INDEX WORDS: Autism, Problem behaviors, Functional assessment, Functional analysis
COMPARING FUNCTIONAL ASSESSMENT METHODOLOGIES:

A QUANTITATIVE SYNTHESIS

by

CAITLIN V. HERZINGER

B.S., The University of Florida, 2001

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

MASTER OF ARTS

ATHENS, GEORGIA

2005
COMPARING FUNCTIONAL ASSESSMENT METHODOLOGIES: A QUANTITATIVE SYNTHESIS

by

CAITLIN V. HERZINGER

Major Professor: Jonathan M. Campbell
Committee: David L. Gast
A. Michele Lease

Electronic Version Approved:
Maureen Grasso
Dean of the Graduate School
The University of Georgia
May 2005
TABLE OF CONTENTS

LIST OF TABLES ........................................................................................................................................ vi
LIST OF FIGURES ....................................................................................................................................... vii

CHAPTER

1 Introduction ........................................................................................................................................ 1

   Autism Definition and Prevalence Rates ......................................................................................... 1

   Approaches to Assessing and Treating Problem Behaviors ............................................................. 2

   Purpose of Current Study .................................................................................................................. 13

2 Method .............................................................................................................................................. 15

   Study Identification and Selection .................................................................................................. 15

   Estimating Effects of Behavioral Interventions ............................................................................. 16

   Statistical Analyses ......................................................................................................................... 20

3 Results ............................................................................................................................................. 21

   Characteristics of Participants ........................................................................................................ 21

   Characteristics of Studies .............................................................................................................. 21

   Characteristics of Functional Assessment, Behavioral Intervention, and Experimental Quality .................................................. 22

4 Discussion ....................................................................................................................................... 40

   Summary of Findings ....................................................................................................................... 40

   Implications for Clinicians ............................................................................................................. 41
Limitations of the Current Study.............................................................................42
 Directions for Future Research................................................................................43
 REFERENCES ..............................................................................................................................45
 APPENDICES ...............................................................................................................................53
   A  Coding Form used in Meta-Analysis .................................................................53
   B  Articles used in Meta-Analysis...........................................................................62
   C  Interrater Reliability Data.....................................................................................70
LIST OF TABLES

Table 3.1: Participant Characteristics .................................................................26
Table 3.2: Description of Study Characteristics .....................................................28
Table 3.3: Assessment, Intervention, and Experimental Characteristics.............29
Table 3.4: Descriptive Statistics for Three Effect Sizes .............................................34
Table 3.5: Descriptive Statistics for Assessment Type and Function .....................35
Table 3.6: Treatment Effect Sizes by Function ......................................................36
Table 3.7: $\chi^2$ for Function by Behavior Type .....................................................38
Table 3.8: $\chi^2$ for Behavior Type by Function .....................................................39
LIST OF FIGURES

Page

Figure 1.1: Hypothetical EFA graph results ....................................................................................9
CHAPTER 1: INTRODUCTION

Autism Definition and Prevalence Rates

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 2-20 cases per 10,000 and a male to female ratio of 4:1 (Diagnostic and Statistical Manual IV-TR, American Psychiatric Association [APA], 2000). Autism is often associated with Mental Retardation, which can range from mild to profound. 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.

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 et al., 2002). According to Wicks-Nelson and Israel (1999), 15\% of developmentally disabled children engage in SIB. Emerson et al. (2001) reported that “challenging behaviors are shown by 10-15\% of people with mental retardation who are in contact with educational, health, or social care services” (p. 77). These data, although not comprehensive, highlights the importance of assessing and treating these maladaptive behaviors. Aside from tissue damage and property destruction, problem behaviors can interfere with new learning, prevent acquisition of adaptive behavior and compete with socially acceptable behaviors thus having a negative impact on social relationships, academic performance, and overall personal growth.

**Approaches to Assessing and Treating Problem Behaviors**

*Psychoanalytic Approach*

Once the primary method for treating individuals with autism, the psychoanalytic approach has since fallen out of favor with most researchers and clinicians in the area. This perspective typically focuses on the relationship between parent and child when treating children with autism and their problematic behaviors. Psychodynamic treatments are based on the assumption that autism is caused by emotional trauma or parental mishandling and treatments are focused on procedures that attempt to “reduce the core hypothesized inner conflict the children [were] thought to be exhibiting” (Koegel, Koegel, & McNerney, 2001, p. 19). Bettelheim (1967), perhaps the most well-known proponent of the psychoanalytic model, suggested that children with autism in a supportive environment will gradually develop trust and no longer feel threatened by their parents.
Gentle therapy and holding therapy are two of the more common psychoanalytic treatments that attempt to create a supportive environment for the individual with autism. Holding therapy is described as forcibly holding the child “so as to cause the autistic defense . . . to crumble” (Welch, 1987, p. 48). The theoretical assumptions, treatment procedures, and outcome claims made by Bettelheim and others have been challenged by researchers (Charlop, Schreibman, & Kurtz, 1990). Studies comparing children receiving psychoanalytic treatment to those receiving other forms of treatment and no treatment at all, found the effectiveness of the psychoanalytic treatment to be inadequate in comparison (Levitt, 1963; Bartak & Rutter, 1973). For further reading on the psychoanalytic approach to treating children with autism, see Psychoanalysis with Children: History, Theory, and Practice (Rodriguez, 1999).

**Psychopharmacological Approach**

The psychopharmacological approach to the assessment and treatment of maladaptive behaviors in people with autism 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 have 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
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.

Sensory Integration Approach

Dr. A. Jean Ayres is the founder of Sensory Integration therapy (SI), a sensory-motor
treatment based upon theories developed over the last 30 years. Advocates of this theory
maintain that sensory integration is an innate neurobiological process (Hatch-Rasmussen, 1995). Proponents suggest that individuals with autism and other developmental disabilities experience
dysfunction in which sensory input is not integrated or organized appropriately by the brain. SI
therapy is viewed as a direct intervention that can improve nervous system function. This is
done by providing the child with enhanced levels of sensory information gleaned during physical
activities that are meaningful to the child, and that elicit adaptive behaviors (Koomar & Bundy,

There are three forms of stimulation often used in conjunction with one another during SI
treatment: a) vestibular; b) proprioceptive; and c) tactile (Fisher et al., 1991). Vestibular
treatments refer to engaging the individual in specific movements to provide information about
where the body is in space (e.g., use of suspended swings). Proprioceptive input refers to touch
or pressure applied to the individual to provide general body awareness (e.g., use of weighted
vests, hug machines). Tactile input is used to provide information about objects in the
environment (e.g., use of textured toys). Thus far, little research has supported the anecdotal
claims that SI is successful at reducing problem behaviors in children or adults with autism. In fact, in one study, SI was shown to increase problem behaviors (Mason & Iwata, 1990). Despite the lack of supportive evidence, the SI approach is often used in conjunction with other treatments as a way to treat individuals with autism. The interested reader is referred to Dawson and Wattling (2000) and Baranek (2002) for more in depth reviews of the literature regarding SI treatment for individuals with autism.

**Behavioral Approaches**

The behavioral approach to treating maladaptive behaviors has consistently focused on identifying antecedents and consequences of problem behavior and typically uses the operant conditioning approach during treatment. Prior to the 1980’s, problem behaviors were often treated with punishment, without concern for the contingencies maintaining the behavior. Punishment was a successful treatment strategy (i.e., reduced problem behaviors) because “its effectiveness [did] not depend on its ability to alter a reinforcement contingency” (Iwata, 1982, p. 5). Interventions were often implemented arbitrarily without regard to the contingencies that maintained the problem behavior (Kahng, Iwata, & Lewin, 2001).

With the growth of research defining and describing functional assessment of problem behavior, the behavior analytic perspective is quickly becoming the “gold standard” for the assessment and treatment of problem behaviors. Carr (1977) and Johnson and Baumeister (1978) were some of the first to document the usefulness of identifying the environmental events that maintain problem 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 problem 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 (i.e., self stimulation). 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] (1997) emphasizes the importance of the behavioral approach to assessment and treatment of problematic behaviors (Gresham et al., 2004).

**Functional Assessment**

Functional Assessment (FA) 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 FA are to be used in treatment selection for reducing maladaptive behaviors, rather than selecting treatments arbitrarily. For example, if the FA indicates that an individual is engaging in a problem 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, Iwata, and Lewin (2002) compared the effectiveness of treatments based on FAs to arbitrarily chosen ones and found treatments based on FAs to be more effective.

FA encompasses *indirect assessments* (e.g., interviews, questionnaires, rating scales such as the MAS and QABF); *descriptive assessments* (e.g., A-B-C sheets, direct observation with no variable/environment manipulation); and *experimental functional analyses* (EFA), also known as “true” or traditional FAs (e.g., analogue conditions in which variables are manipulated). “The primary benefit of functional assessments has been to increase the precision
and resulting effectiveness of reinforcement-based interventions” (Kahng et al., 2002, p. 130). Functional assessment has become foundational in the study of aberrant behaviors. For example, Kahng et al. report that there has been a dramatic increase in the number of published data sets investigating functional assessments since the early 1980’s. This increase coincides with and is presumed to be a result of the increase of data sets that included an EFA.

Based on the ability of FAs to ascribe function to aberrant behaviors and thus improve treatment effectiveness of those behaviors through treatment selection (Kahng et al., 2002; Campbell, 2003), FAs have become an integral part of the entire treatment process. In 1997, the amendments to IDEA required the use of functional behavioral assessment 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 does not, however, define what constitutes a valid behavioral functional assessment. Because of this lack of detailed description, school districts and other professionals are using less specific models of assessing function. The current review of the literature aims to delineate the differences in accuracy and effectiveness of different functional assessment methodologies.

*Experimental Functional Analysis (EFA) and Behavioral Functional Assessment (BFA)*

EFA represents simulation of natural environments and is the primary tool for demonstrating causal relationships (Carr, Langdon, & Yarborough, 1999). With EFA, the goal is to isolate the function or purpose of the problem behavior. This allows the therapist to directly and specifically alter the social and physical environment in a way that is most likely to alter the problem behavior. Iwata et al. (1982) defined the first standardized and comprehensive method for conducting functional analyses. As outlined by Iwata et al. there are four different conditions
in an EFA: social disapproval, academic demand, unstructured play, and alone. In the social disapproval condition, the client receives therapist attention contingent on problem behavior. The social disapproval condition is thought to mimic naturalistic occurrences where individuals may respond with attention when the problem behavior is exhibited. This “attention” includes 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 EFA, 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 tangible condition, it is assumed that the purpose of the behavior is to gain access to tangible items. See Figure 1.1 on the following page for a visual representation of this example. This information would be used as an integral part in the treatment selection process. Following the assessment, treatment selection might include some type of differential reinforcement of alternative behavior (DRA) such as having access to items contingent on pointing to a picture of the wanted item.
BFA represents the other two categories mentioned under the umbrella term FA: indirect and descriptive assessments. These assessments 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 (FOAF; O’Neill et al., 1997) is a direct observation method of behavioral functional assessment. Some research has been reported comparing the validity of these methods of assessment to EFA. For example, the MAS ratings
have been correlated with EFA analogue data and were 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 almost all published accounts of comparison data, the EFA represented the gold standard for validity tests of other types of assessment.

Because all BFAs are not administered in the same way (e.g., rating scales, observations, interviews), interpretations of the assessments differ. Some BFAs are simply used to gain qualitative information that is later combined with other sources of information to identify the function of maladaptive behavior. Other BFAs, specifically 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). On 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 problem behavior to continue (Durand & Crimmins). No matter the method of interpretation, the results of BFAs are used in the same manner as those from EFAs. 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 EFA when compared to BFA. Much research has been published regarding the effectiveness of EFA. EFA has provided methods to test the notion that each behavior has a function that can be observed. EFA is able to identify causal and maintaining factors of problem behavior and has led to the development of effective treatments for severe problem behaviors. EFA is a specific and structured assessment method, in which the analog conditions and the contingencies set in place for each condition are defined a priori. The structure and conditional quality of the EFA makes replication possible. EFA allows the clinician/practitioner to base treatment selection on an observable, recorded (not simply reported) function of behavior. The primary benefit of EFA is that this methodology allows for causal inferences to be made regarding the purpose of the targeted maladaptive behavior.

In contrast, BFA 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, BFA methodologies are correlational in nature. BFA is less thorough and is more susceptible to data recorder biases. For example, a teacher completing a functional rating scale about her own observations may be more likely to recall instances in which her attention was drawn from a target activity to a problem behavior that an instance in which a child engaged in problem behaviors in order to get a break from difficult academic tasks.

In 1994, Horner wrote that although different situations may require different functional assessment procedures, EFA “will remain the expectation within published research” (p. 402). Although other procedures for functional assessment 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 BFA when compared to EFA.** According to research by Sturmey (1995), the experimental design of EFA presents a number of procedural and psychometric problems. EFAs are time consuming, require extensive training, and are not cost-effective. Matson et al. (1999) reports that a single EFA can require several hours daily over a two- to three-week-period. Iwata et al. (1994) summarized data from over 150 EFAs and reported that the mean length of assessment was 26 sessions, which represented approximately 6.5 hours of direct observation, not to mention 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 EFA 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 work funding have exposed the monetary constraints on using EFA as an automatic prelude to behavioral treatment (Applegate, Matson, & Cherry, 1999). Martin, Gaffan, and Williams, (1999) also noted problems with poor test-retest reliability of analogue conditions. Martin and colleagues (1999) further identified the potential problems with “imposing a contingency on a response that is likely to occur may result in the response coming under the control of that contingency, even if it was not maintained previously by such a contingency in the person’s natural environment” (p. 126). Other researchers have noted several other limitations of EFA procedures, including non-
naturalistic environment and intentional increases in life threatening problem behaviors (Dawson, Matson, & Cherry, 1998; Axelrod, 1987).

BFAs, in contrast, do not require significant amounts of time and are easier to administer. They are dependent on the non-compulsory training of the raters, observers, and other informants. Also, BFAs do not include environmental manipulations that intentionally increase the rates of problem behaviors. BFAs are naturalistic and do not require careful experimental design.

Purpose of Current Study

The focus of the current study is to compare different types of FA methodologies, including analogue conditions, interviews, observations, rating scales, and combinations of these methods. Similar to Toogood and Timlin (1996), the study aims to compare the different methodologies of FA in a) similarity of ascribing function to a target behavior, b) treatment selection that follows such an assessment, and c) differences in treatment efficacy depending on type of FA methodology. The purpose of the following quantitative synthesis is to answer the following questions regarding the assessment and treatment of problem behaviors for people with autism:

1) Is treatment more effective when following an experimental functional analysis (EFA) or a behavioral functional assessment (BFA)? I hypothesize that treatment based on the more standardized and thorough EFA will be more effective in delineating the specific purpose of aberrant behavior. The EFA is experimental in nature and controls for most observer bias, unlike the BFA.
2) Do different types of FAs lead to different treatment selections? This question was noted by Pelios, Morren, Tesch, & Axelrod (1999) as an important research consideration. This question is exploratory in nature.

3) Is there a predominant observed function based on the type of assessment? This question is exploratory in nature.

4) Does the observed function of the behavior, regardless of FA method used, have an impact on the effectiveness of treatment? I hypothesize that behaviors identified as having an automatic function will be more difficult to treat. This notion is based on the previous research of Vollmer (1994), Piazza, Hanley, and Fisher (1996), and many others who found that treatment development for behaviors maintained by automatic reinforcement is more difficult because the reinforcer may be unknown or not easily manipulated.

5) Is there a relationship between the type of maladaptive behavior observed and the obtained function for that behavior? For example, is self injury more often attributed as having an escape component? I hypothesize that external maladaptive behaviors (e.g., aggression and property destruction) will be often identified as having an attention and/or tangible function. This hypothesis has been suggested by other researchers (Dawson, Matson & Cherry, 1998).
CHAPTER 2: METHOD

Study Identification and Selection

For the years 1998 through 2003, published functional assessments of problem behavior for people with autistic disorder were identified through searches of PsycLit, ERIC, and MedLine databases using appropriate search terms, such as subject descriptions (e.g., autism, autistic disorder), 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 also identified by issue-by-issue hand searches of relevant journals including *Journal of Applied Behavior Analysis*, *Research in Developmental Disabilities*, and *Behavioral Interventions* (see Table 3.2 for complete list). Also, timely references (i.e., citations between 1998 and 2003) 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 1998 and December 2003. Second, single case studies were included only if a participant was diagnosed with autistic disorder, described as meeting the criteria for autistic disorder, or described as autistic or autistic-like. Third, a functional assessment had to be conducted and results reported, with problem behaviors as the target behaviors of treatment. For those articles used in the treatment effectiveness analyses, the 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 that 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). In the current study, three effect sizes based on nonregression approaches were calculated per intervention: MBLR, PND, and PZD. The MBLR was 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 was 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 et al., 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 was calculated by locating the first intervention data point that reached zero and computing the percentage of data points that reached zero subsequent to this first zero (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 (Campbell, 2003).

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

Several rules were established for the coding of assessment type. Functional assessment type was coded as either: a) EFA (strictly adhering to guidelines set forth in Iwata, 1985), b) modified EFA, c) ABC sheet, d) rating scales (e.g., MAS, QABF), e) informal assessment or f) other. Under the modified EFA code, brief EFA sessions, (i.e., shorter sessions than those described by Iwata) such as those described by Northup et al. (1991) and summarized in Derby et al. (1992) were coded. Later, this group (modified EFA) was consolidated with the EFA group. Thus, articles were coded as EFA if they included analogue conditions and manipulated variables in the environment in each condition.

If two different types of FAs (e.g., EFA 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 BFA 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 BFAs yielded different functions for each method. In the current data set, the BFA methods did not identify contradicting functions, thus results never required separate coding.

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. 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 EFA 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 were coded for both EFA and BFA. 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 were reported separately, thus for participant B three effect sizes were calculated for behavior at school and three effect sizes were 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 nontreatment phase and the last treatment phase, per Faith’s et al.’s (1996) recommendations and implemented in Campbell (2003). In designs that compared multiple treatments (A-B-A-C), the final treatment phase was 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 .05. The data conversion procedure has
been used by Allison, Faith, and Franklin (1995) and Campbell (2003) with a high degree of interrater reliability.

The following participant information was coded when available: participant’s age, gender, intellectual functioning, secondary diagnoses, years since autism diagnosis prior to study, and years of prior treatment. The following assessment/pre-intervention data was coded: target behavior, type of FA (e.g., EFA, 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 was coded: type of intervention, type of experimental design, interrater reliability, number of baseline data points, number of final phase treatment points, and attempt to generalize treatments. The types of intervention coded included: reinforcement, punishment, timeout, extinction, sensory extinction, FCT, combined treatments, and other interventions. 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. Refer to Appendix A for copy of the coding form used.

Twelve articles were randomly selected for independent coding by an advanced graduate student in School Psychology, who had experience working with individuals with autism, and interrater agreement was established. The 12 articles (20.69% of all articles) included 17 separate assessments (16.04% of all assessments) and 13 different participants (16.05% of all
participants). Interrater agreement, with a mean of 99.06% and a range of 82.35% to 100% across all coded variables, was determined by the percent agreement method (\( \frac{\text{# of agreements}}{\text{# of agreements} + \text{# of disagreements}} \times 100 \)). See Appendix C for a detailed reliability summary.

Statistical Analyses

*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, level of mental retardation and other variables were summarized. This analysis gave an indication of the representativeness of the sample. Assessment, intervention, and experimental characteristics of the studies were examined as well (e.g., percentage of studies that used EFA methodology; percentage of studies who targeted aggression). One-way ANOVAs were used to examine research question 1 (a comparison of treatment effectiveness for EFA and BFA) and research question 4 (treatment effectiveness as impacted by function). Two variable chi-square tests of non-independence of categorical variables were used to assess research question 2 (FA type’s impact on treatment selection), research question 3 (possible bias in assessment outcomes based on FA type), and research question 5 (relationship between behavior type and behavioral function).
CHAPTER 3: RESULTS

Characteristics of Participants

This review included 58 articles reporting on 81 participants with a total of 106 separate FAs. Similar to prior reviews documenting the higher prevalence of autism in males, the majority of participants included in the review were male. However, the ratio of males to females in this study (2.5 : 1) was somewhat lower than gender ratios reported in the autism literature, typically 4 : 1. Consistent with prior reviews documenting the prevalence of mental retardation in individuals with autism, the majority of the participants (77.7%) included in the present review functioned in the range of mental retardation or were considered “untestable” via standardized, formal intelligence testing. In most cases (84%), the criteria used to diagnose autism were not reported; however, when the criteria were reported, they were most commonly a version of those presented in the DSM (9.8%). 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. Detailed information about the characteristics of participants is presented in Table 3.1.

Characteristics of Studies

Studies included in the research synthesis were collected from a total of six journals, with the Journal of Applied Behavior Analysis contributing the highest percentage of articles (56%). Studies most often included one participant and rarely included more than two. Detailed information about the characteristics of studies included in the meta-analysis is presented in Table 3.2.
Characteristics of Functional Assessment, Behavioral Intervention, and Experimental Quality

Studies employed both experimental (75.47%) and non-experimental (24.53%) methods of FA. Under the EFA umbrella, the majority of assessments were Modified EFA (45%), based on the analogue conditions of Iwata and colleagues, but tailored in terms of time or specific conditions used to the particular needs of the researcher, clinician, and/or participant. The type of BFAs most often reported were described as informal assessment (53.8%).

The studies included in the quantitative synthesis targeted the reduction of more than one problematic behavior more often than any single maladaptive behavior. When assessed individually, aggression (20%) and SIB (18%) were the most commonly targeted behaviors. In combination with other maladaptive behaviors, aggression (28%) and disruption (18%) were the most commonly reported among the participants. Reinforcement only (20%) was the most commonly reported intervention type. Functional communication training (FCT) was the most commonly reported form of treatment included under the “reinforcement” coding. FCT constituted 43% of the treatments coded as reinforcement. A combination of extinction and either reinforcement or punishment (16%) was another common intervention technique. In 16% of the studies that claimed treating the targeted behaviors assessed, no specific intervention information was reported. In a majority of cases (65.6%), behavioral therapist could not be coded due to lack of information reported by the authors. Inpatient hospital setting was the treatment setting that occurred with the highest frequency (23.4%).

The simple AB experimental design was the most commonly used, reported in 37.5% of the studies. Studies in the meta-analysis omitted follow up data collection (67.2%) more often than included follow-up data collection (25%). When follow up data was collected, the intervals
from treatment to follow up data collection ranged from two weeks to two years with an average interval of approximately 8.6 months. The most commonly reported follow up interval was 6 months, reported in 6.3% of studies. In 5 of 64 cases, there was no clear indication of whether follow-up data was collected or not. Generalization data were omitted from the studies more often than reported. When reported, the data indicated that generalization to new situations and settings was the most common form of generalization data reported (50%). Inter-rater reliability for FA sessions was reported in 80.2% of articles. The reliability ranged from 84.0 to 100.0 with a mean of 94.6 and a standard deviation of 3.7. For treatment sessions, inter-rater reliability data was reported in 93.8% of articles. The reliability ranged from 87.0 to 100.0 with a mean of 95.3 and standard deviation of 3.3. Detailed information about functional assessments, behavioral interventions, and experimental quality is presented in Table 3.3.

Treatment Effectiveness Based on FA Type

Three one-way ANOVAs showed that when comparing EFA and BFA, there is a significant difference in treatment effectiveness as measured by the PZD statistic, $F(1, 60) = 7.58, p < .01$). EFA and BFA did not differ when treatment effectiveness was measured with MBLR, $F(1, 60) = 1.35, n.s.$) or PND, $F(1, 60) = .99, n.s.$). Therefore, statistical analyses indicate that treatment is more effective when based on the results of EFA as compared to BFA, when the PZD statistic is used to assess effectiveness. See Table 3.4 for means and standard deviations. In summary of this finding, it appears that when behavior suppression (as measured by the PZD statistic), rather than behavior reduction is the goal of treatment, intervention packages based on the results of EFAs are more effective than those based on results of BFAs.
Is Treatment Selection Impacted by Function?

A two variable $\chi^2$ test of non-independence shows that the type of treatment selected (i.e., reinforcement only, combination of extinction and punishment) is independent from ascribed behavioral function of the behavior, $\chi^2(N = 25, 62) = 32.94, n.s.$ This finding highlights that although the results of FAs are intended to guide intervention selection, there is not a predominant intervention that results from a FA. A one variable $\chi^2$ test was used to assess the distribution of interventions that included FCT, a commonly reported treatment, across ascribed functions. No significant differences exist in the distribution of the FCT intervention packages in comparison to the non-FCT packages. Treatments for problem behavior are tailored to the identified function of behavior, but no specific intervention types (e.g., reinforcement only) or specific treatments (e.g., FCT) were found to have a significant relationship with a particular identified function.

Is Ascribed Function Related to FA Methodology?

The data show that there is not a relationship between the type of methodology (e.g., EFA, BFA) used in the assessment and the result of that assessment, $\chi^2 (N = 6, 106) = 8.01, n.s.$ This finding supports the notion that ascribed behavioral function is not a product of the type of assessment employed. This is positive support for both EFA and BFA as independent assessment methodologies (see Table 3.5 for frequency information).

Does Ascribed Function Impact Treatment Effectiveness?

The results from three one-way ANOVAs indicate that there is no statistical significance for the three calculated effect sizes. Treatment effectiveness, as assessed by the a) MBLR statistic, $F(5, 56) = .90, n.s.$, b) PND statistic, $F(5, 56) = 1.57, n.s.$, and c) PZD statistic, $F(5, 56) = .90, n.s.$ was not significantly affected by the function of the problem behavior. Table 3.6
includes data regarding the ranges, means, and standard deviations for the three calculated effect sizes: MBLR, PND, and PZD.

Is There a Relationship Between Problem Behavior Type and Identified Function?

The two variable $\chi^2$ test of non-independence indicates that there is a statistically significant relationship between the type of problem behavior (i.e., internal versus external) and the behavioral function, $\chi^2 (N = 10,106) = 43.37, p < .001$. The goodness of fit, one variable $\chi^2$ test was then identified for each reported function. Internal behaviors were more likely identified as having automatic functions. External behaviors were more often identified as having escape and tangible functions. Combinations of internal and external behaviors were identified as having combination functions (two functions). The attention, tangible, automatic, and undifferentiated (3 or more functions) functions were significant when tested to determine the specific relationships with types of behavior. This data is reported in Tables 3.7 and 3.8.
<table>
<thead>
<tr>
<th>Characteristic</th>
<th>n</th>
<th>%</th>
<th>r</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>58</td>
<td>71.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>23</td>
<td>28.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (in years)</td>
<td>3.42-49.00</td>
<td>13.53</td>
<td>9.90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race/Ethnicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>African-American</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not reported</td>
<td>80</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level of Mental Retardation (IQ range)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Severe (&lt;39)</td>
<td>38</td>
<td>46.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not reported/Other</td>
<td>17</td>
<td>21.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mild (70 - 55)</td>
<td>12</td>
<td>14.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderate (54 - 40)</td>
<td>12</td>
<td>14.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None (IQ &gt; 70)</td>
<td>1</td>
<td>1.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Untestable</td>
<td>1</td>
<td>1.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diagnostic criteria used</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not reported</td>
<td>68</td>
<td>84.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DSM-IV</td>
<td>7</td>
<td>8.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>“Autistic-like”</td>
<td>5</td>
<td>6.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DSM-III</td>
<td>1</td>
<td>1.2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3.1 continued

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>n</th>
<th>%</th>
<th>r</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Months since diagnosis</td>
<td></td>
<td>100%</td>
<td>Not reported</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Months of prior treatment</td>
<td></td>
<td></td>
<td>(7 - 168)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medication administered</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not reported</td>
<td>70</td>
<td>86.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>6</td>
<td>7.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>5</td>
<td>6.2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note. n = number of participants; % = percentage of participants; r = range; M = mean; SD = standard deviation.*
Table 3.2.

Description of Study Characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Journal of Applied Behavior Analysis</td>
<td>56</td>
<td>52.8</td>
</tr>
<tr>
<td>Research in Developmental Disabilities</td>
<td>24</td>
<td>22.6</td>
</tr>
<tr>
<td>Behavioral Interventions</td>
<td>12</td>
<td>11.3</td>
</tr>
<tr>
<td>Journal of Autism and Developmental Disorders</td>
<td>8</td>
<td>7.5</td>
</tr>
<tr>
<td>Behavior Therapy</td>
<td>4</td>
<td>3.8</td>
</tr>
<tr>
<td>Journal of Clinical Psychopharmacology</td>
<td>2</td>
<td>1.9</td>
</tr>
</tbody>
</table>

Total N 106

Number of participants per study

<table>
<thead>
<tr>
<th>Participants</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>42</td>
<td>72.41</td>
</tr>
<tr>
<td>2</td>
<td>11</td>
<td>18.97</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>6.90</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>1.72</td>
</tr>
</tbody>
</table>

Total N 58

*Note.* n = number of studies; % = percentage of studies.
Table 3.3

Assessment, Intervention, and Experimental Characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggression</td>
<td>20</td>
<td>18.87</td>
</tr>
<tr>
<td>Self-injurious Behavior</td>
<td>18</td>
<td>16.98</td>
</tr>
<tr>
<td>Self-stimulation/stereotypic behavior</td>
<td>16</td>
<td>15.09</td>
</tr>
<tr>
<td>Disruptive Behavior</td>
<td>6</td>
<td>5.66</td>
</tr>
<tr>
<td>Pica</td>
<td>5</td>
<td>4.72</td>
</tr>
<tr>
<td>Property Destruction</td>
<td>3</td>
<td>2.83</td>
</tr>
<tr>
<td>Vocalizations</td>
<td>3</td>
<td>2.83</td>
</tr>
<tr>
<td>Other</td>
<td>1</td>
<td>0.94</td>
</tr>
<tr>
<td>Combinations</td>
<td>34</td>
<td>32.08</td>
</tr>
</tbody>
</table>

Combination Breakdown:

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggression</td>
<td>28</td>
<td>82.40</td>
</tr>
<tr>
<td>Disruption</td>
<td>18</td>
<td>52.90</td>
</tr>
<tr>
<td>Property Destruction</td>
<td>13</td>
<td>38.20</td>
</tr>
<tr>
<td>SIB</td>
<td>7</td>
<td>20.60</td>
</tr>
<tr>
<td>Self stimulatory</td>
<td>4</td>
<td>11.80</td>
</tr>
<tr>
<td>Vocal</td>
<td>3</td>
<td>8.80</td>
</tr>
</tbody>
</table>

Assessment and Intervention characteristics
Table 3.3 continued

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Type of functional assessment</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experimental</td>
<td>80</td>
<td>75.47</td>
</tr>
<tr>
<td>Modified EFA</td>
<td>36</td>
<td></td>
</tr>
<tr>
<td>EFA</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>Brief EFA</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Partial Experimental</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Non-experimental</td>
<td>26</td>
<td>24.53</td>
</tr>
<tr>
<td>Informal Assessment</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>ABC sheet</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Rating Scale</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Descriptive Assessment</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td><strong>Type of intervention</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reinforcement only</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Extinction and Reinforcement or Punishment</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Other/Not reported</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Reinforcement and Punishment</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Punishment only</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Extinction only</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td><strong>Total N</strong></td>
<td>62</td>
<td></td>
</tr>
</tbody>
</table>
Table 3.3 continued

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behavioral Therapist</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not reported or insufficient information</td>
<td>42</td>
<td>65.6</td>
</tr>
<tr>
<td>Other Professional</td>
<td>9</td>
<td>14.1</td>
</tr>
<tr>
<td>Teacher/Special Education teacher</td>
<td>9</td>
<td>14.1</td>
</tr>
<tr>
<td>Student</td>
<td>2</td>
<td>3.2</td>
</tr>
<tr>
<td>Parent</td>
<td>1</td>
<td>1.5</td>
</tr>
<tr>
<td>Paraprofessional</td>
<td>1</td>
<td>1.5</td>
</tr>
<tr>
<td>Treatment Setting</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inpatient Hospital</td>
<td>15</td>
<td>23.4</td>
</tr>
<tr>
<td>Other</td>
<td>13</td>
<td>20.3</td>
</tr>
<tr>
<td>Public School</td>
<td>12</td>
<td>18.7</td>
</tr>
<tr>
<td>Not Reported</td>
<td>8</td>
<td>12.5</td>
</tr>
<tr>
<td>Outpatient Clinic</td>
<td>6</td>
<td>9.4</td>
</tr>
<tr>
<td>Inpatient Classroom</td>
<td>5</td>
<td>7.8</td>
</tr>
<tr>
<td>Home</td>
<td>4</td>
<td>6.3</td>
</tr>
<tr>
<td>Residential Treatment Setting</td>
<td>1</td>
<td>1.6</td>
</tr>
</tbody>
</table>
Table 3.3 continued

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Experimental characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experimental design</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simple AB</td>
<td>24</td>
<td>22.6</td>
</tr>
<tr>
<td>Reversal</td>
<td>21</td>
<td>19.8</td>
</tr>
<tr>
<td>Alternating Treatments</td>
<td>16</td>
<td>15.1</td>
</tr>
<tr>
<td>Multiple Baseline</td>
<td>2</td>
<td>1.9</td>
</tr>
<tr>
<td>Not reported</td>
<td>1</td>
<td>.9</td>
</tr>
<tr>
<td>Follow-up data collected</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>43</td>
<td>67.2</td>
</tr>
<tr>
<td>Yes</td>
<td>16</td>
<td>25.0</td>
</tr>
<tr>
<td>Not Reported</td>
<td>5</td>
<td>7.8</td>
</tr>
<tr>
<td>Follow-up interval (in months) (M = 8.65, SD = 6.61)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attempt to generalize behavior</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>36</td>
<td>56.3</td>
</tr>
<tr>
<td>Yes</td>
<td>28</td>
<td>43.7</td>
</tr>
<tr>
<td>Attempt to generalize to:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>New situation or setting only</td>
<td>14</td>
<td>13.2</td>
</tr>
<tr>
<td>Combination</td>
<td>8</td>
<td>7.5</td>
</tr>
<tr>
<td>New people only</td>
<td>5</td>
<td>4.7</td>
</tr>
<tr>
<td>Other</td>
<td>1</td>
<td>.9</td>
</tr>
</tbody>
</table>
Table 3.3 continued

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Range</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reliability of observations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inter-rater reliability</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FA</td>
<td>84.0-100.0</td>
<td>94.6</td>
<td>3.7</td>
</tr>
<tr>
<td>Treatment</td>
<td>87.0-100</td>
<td>95.3</td>
<td>3.3</td>
</tr>
</tbody>
</table>

*Note.* n = number of studies; % = percentage of studies; M = mean; SD = standard deviation.
Table 3.4

Descriptive Statistics for Three Effect Sizes

<table>
<thead>
<tr>
<th></th>
<th>M</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>EFA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MBLR</td>
<td>83.89</td>
<td>22.43</td>
<td>-.96</td>
<td>100.00</td>
</tr>
<tr>
<td>PND</td>
<td>81.00</td>
<td>31.23</td>
<td>.00</td>
<td>100.00</td>
</tr>
<tr>
<td>PZD</td>
<td>62.55</td>
<td>31.87</td>
<td>.00</td>
<td>100.00</td>
</tr>
<tr>
<td>BFA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MBLR</td>
<td>77.21</td>
<td>16.63</td>
<td>45.33</td>
<td>100.00</td>
</tr>
<tr>
<td>PND</td>
<td>89.02</td>
<td>23.73</td>
<td>10.33</td>
<td>100.00</td>
</tr>
<tr>
<td>PZD</td>
<td>35.61</td>
<td>42.86</td>
<td>.00</td>
<td>100.00</td>
</tr>
</tbody>
</table>

Note. EFA = experimental functional analysis; BFA = behavioral functional 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 62 treatment outcomes.
Table 3.5.

Descriptives for Assessment Type and Function

<table>
<thead>
<tr>
<th>Assessment Type</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ATT</td>
</tr>
<tr>
<td>EFA</td>
<td>9</td>
</tr>
<tr>
<td>BFA</td>
<td>4</td>
</tr>
<tr>
<td>Total</td>
<td>13</td>
</tr>
</tbody>
</table>

Note. EFA = experimental functional analysis; BFA = behavioral functional assessment; ATT = attention; ESC = escape; TAN = tangible; AUT = automatic; UND = undifferentiated; COM = combination.
Table 3.6.

**Treatment Effect Sizes by Function**

<table>
<thead>
<tr>
<th>Function</th>
<th>Effect Size</th>
<th>M</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Attention</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MBLR</td>
<td></td>
<td>88.77</td>
<td>13.31</td>
<td>60.22</td>
<td>98.81</td>
</tr>
<tr>
<td>PND</td>
<td></td>
<td>96.88</td>
<td>8.84</td>
<td>75.00</td>
<td>100.00</td>
</tr>
<tr>
<td>PZD</td>
<td></td>
<td>52.66</td>
<td>43.36</td>
<td>.00</td>
<td>100.00</td>
</tr>
<tr>
<td><strong>Escape</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MBLR</td>
<td></td>
<td>86.27</td>
<td>13.90</td>
<td>46.72</td>
<td>100.00</td>
</tr>
<tr>
<td>PND</td>
<td></td>
<td>86.78</td>
<td>29.06</td>
<td>.00</td>
<td>100.00</td>
</tr>
<tr>
<td>PZD</td>
<td></td>
<td>56.86</td>
<td>39.51</td>
<td>.00</td>
<td>100.00</td>
</tr>
<tr>
<td><strong>Tangible</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MBLR</td>
<td></td>
<td>69.40</td>
<td>33.28</td>
<td>33.97</td>
<td>100.00</td>
</tr>
<tr>
<td>PND</td>
<td></td>
<td>94.05</td>
<td>10.31</td>
<td>82.14</td>
<td>100.00</td>
</tr>
<tr>
<td>PZD</td>
<td></td>
<td>80.30</td>
<td>34.12</td>
<td>40.91</td>
<td>100.00</td>
</tr>
<tr>
<td><strong>Automatic</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MBLR</td>
<td></td>
<td>75.20</td>
<td>28.01</td>
<td>.96</td>
<td>100.00</td>
</tr>
<tr>
<td>PND</td>
<td></td>
<td>87.57</td>
<td>27.56</td>
<td>.00</td>
<td>100.00</td>
</tr>
<tr>
<td>PZD</td>
<td></td>
<td>41.55</td>
<td>36.75</td>
<td>.00</td>
<td>100.00</td>
</tr>
</tbody>
</table>
Table 3.6 continued

Treatment Effect Sizes by Function

<table>
<thead>
<tr>
<th></th>
<th>M</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Undifferentiated</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MBLR</td>
<td>76.27</td>
<td>27.07</td>
<td>39.03</td>
<td>97.34</td>
</tr>
<tr>
<td>PND</td>
<td>72.29</td>
<td>32.31</td>
<td>39.14</td>
<td>100.00</td>
</tr>
<tr>
<td>PZD</td>
<td>40.24</td>
<td>30.18</td>
<td>.00</td>
<td>66.50</td>
</tr>
<tr>
<td><strong>Combination</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MBLR</td>
<td>82.49</td>
<td>20.80</td>
<td>31.47</td>
<td>100.00</td>
</tr>
<tr>
<td>PND</td>
<td>66.67</td>
<td>36.21</td>
<td>.00</td>
<td>100.00</td>
</tr>
<tr>
<td>PZD</td>
<td>63.39</td>
<td>33.56</td>
<td>.00</td>
<td>100.00</td>
</tr>
</tbody>
</table>

*Note. M = mean; SD = standard deviation; Min = minimum value; Max = maximum value; MBLR = mean baseline reduction; PND = percentage on non-overlapping data; PZD = percentage of zero data.*
Table 3.7.

$\chi^2$ for Function by Behavior Type

<table>
<thead>
<tr>
<th>Function</th>
<th>$\chi^2$</th>
<th>df</th>
<th>Asymp. Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attention</td>
<td>8.00</td>
<td>2</td>
<td>.018*</td>
</tr>
<tr>
<td>Escape</td>
<td>5.57</td>
<td>2</td>
<td>.062</td>
</tr>
<tr>
<td>Tangible</td>
<td>7.36</td>
<td>1</td>
<td>.007*</td>
</tr>
<tr>
<td>Automatic</td>
<td>19.36</td>
<td>2</td>
<td>.000*</td>
</tr>
<tr>
<td>Undifferentiated</td>
<td>6.70</td>
<td>2</td>
<td>.035*</td>
</tr>
<tr>
<td>Combination</td>
<td>1.37</td>
<td>2</td>
<td>.504</td>
</tr>
</tbody>
</table>

*Note.* df = degrees of freedom; * = p < .05
Table 3.8.

χ² for Behavior Type by Function

<table>
<thead>
<tr>
<th>Behavior Type</th>
<th>χ²</th>
<th>df</th>
<th>Asymp. Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internal</td>
<td>13.16</td>
<td>4</td>
<td>.011*</td>
</tr>
<tr>
<td>External</td>
<td>10.55</td>
<td>5</td>
<td>.061</td>
</tr>
<tr>
<td>Internal/External</td>
<td>12.90</td>
<td>5</td>
<td>.024*</td>
</tr>
</tbody>
</table>

*Note. df = degrees of freedom; * = p < .05
CHAPTER 4: DISCUSSION

Summary of Findings

In light of the debate regarding the relative benefits of EFA and BFA methodology, the present review focused on a series of questions. First, is treatment effectiveness influenced by the type of FA? Second, do different FAs lead to different treatment selections? Third, does FA type affect the behavioral function identified? Fourth, does observed function impact the effectiveness of treatment? Lastly, is there a relationship between type of behavior (e.g., internal, external) and the identified function? The main findings of the quantitative research synthesis are summarized as follows. Treatment based on the results of both EFAs and BFAs were compared and EFA-based treatments were found to be more effective than BFA-based treatments when the PZD statistic was used to determine effectiveness. This finding indicates that treatments based on EFAs are more successful at suppressing problematic behaviors when compared to those based on BFAs. Behavioral function was not found to be a significant influence in the type of treatment selected. Reinforcement alone and extinction plus reinforcement or extinction were the most commonly reported treatments.

The type of FA used was not related to the outcome of the assessment. The finding indicates that neither EFAs nor BFAs result in predictable outcomes because identified behavioral functions were equally distributed based on type of assessment. The functions of target behavior, in and of themselves, have no significant impact on treatment effectiveness. 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 this study
do not support that notion. A significant relationship was observed between the type of problem behavior and the identified function of that behavior. Internal behaviors were more often identified as having automatic functions while external behaviors were more often identified as having escape and tangible functions. When both internal and external behaviors were targeted, the most common result of FAs was a combination of two functions.

Implications for clinicians

The results of the quantitative review are relevant to practicing clinicians regarding the assessment and treatment of severe problem behaviors exhibited by individuals within the autism population. The data shows that there is a significant relationship between behavioral function and behavioral topography (see Tables 3.7 and 3.8). The relationship may help to identify the function and later possible treatment options for maladaptive behaviors. For example, if an external behavior (e.g., aggression) is targeted it is more likely that the purpose of that behavior is to gain access to tangible items or to gain escape from presented activities. Initial treatment selection could be based on this hypothesis.

When the goal of treatment is suppression of behavior, our findings suggest that treatments based on the outcome of EFAs are more effective than those based on BFAs. For severe problem behaviors that are high in intensity, severity, and frequency or those that lead to tissue damage or other life threatening injury, complete suppression of the targeted behavior is even more important. The treatments based on the functions identified by both EFAs and BFAs were shown to be effective in reducing the rate of problem behavior; however, the results of EFAs were found to be significantly more useful when the treatments were assessed by the PZD statistic.
Limitations of the current study

Conclusions of the review must be considered within the context of its limitations. The main 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. For example, studies that report poor treatment effectiveness may go unpublished and thus the average effect sizes reported within this review represent overestimates. Also, FAs that have undifferentiated results and are not further assessed may not be published and therefore not included in the current dataset. Finally, it is possible that published articles that met inclusion criteria may have been unintentionally excluded and not included in the review.

Another limitation of the current study rests within the primary articles used. Many articles did not include potentially useful information about the characteristics of the participants. Basic demographic data such as race, age, and level of mental retardation 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. 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. 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.
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 EFA category, along with traditional EFAs described by Iwata et al. Categories were combined in order to assess the effectiveness of experimental versus non-experimental assessments rather than specific subtypes of assessment. 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).

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 ABCD 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.

Directions for future research

Future quantitative reviews examining functional assessment and comparing different methodologies should include unpublished studies such as theses and dissertations. This would increase the sample size of the included articles. Also, including studies from a larger time span (i.e., greater than five years) would increase sample size and possibly improve the statistical analyses. Bearing in mind that one commonly reported negative aspect of the EFA is the length of time needed to complete an assessment, an evaluation of “brief” EFAs and BFAs may be a more appropriate comparison. Although the current study included only the assessment and
treatment of individuals with autism, many other clinical populations engage in severe problem behaviors. Articles that report problem behavior assessment and intervention for children diagnosed with mental retardation and individuals without a formal diagnosis could be included in future reviews.
REFERENCES


T.Thompson (Eds.), *Self-injurious Behavior: Gene-brain-behavior Relationships* (p. 119-


theory. In A.G. Fisher, E.A. Murray, & A.C. Bundy (Eds.), *Sensory Integration Theory

Therapy, 1*, 45-51.


challenging behavior: a study of validity and reliability. *Research in Developmental
Disabilities, 20*, 125-146.


Questions about Behavioral Function (QABF) scale: predicting treatment success for
163-176.


children with conduct disorders: A quantitative synthesis of single-subject research.

*Behavioral Disorders, 11,* 260-271.


APPENDIX A

Coding Form used in Meta-analysis

Study ID#: _____

Article ID#: _____Subject ID#: _____

Authors: _____________________________________________________________________________

Title: _______________________________________________________________________________

______________________________________________________________________________

Published? _____ Yes (1) _____ No (2) Publication Year: __________

Journal: ______________________________________________________________________________

Number of subjects in study that meet criteria: _____

Participant Information

Age (at intervention):

_____ Years, _____ Months

Gender:

_____ Male (1) _____ Female (2)

Race:

_____ Caucasian (1) _____ Black/African American (2) _____ Hispanic/Latino (3)

_____ Other (4) _____ Not reported (9)

Enter IQ/Adaptive behavior Score beside test given:

_____ WISC Full Scale (edition:__________) (01)

_____ WISC Verbal (edition:__________) (02)

_____ WISC Performance (edition:__________) (03)

_____ Bayley Mental Developmental Index (04)

_____ Leiter International (05)

_____ Merrill-Palmer Scales (06)

_____ Cattell Infant Intelligence (07)

_____ McCarthy Scales (08)

_____ Stanford-Binet (edition:__________) (09)

_____ Vineland Scales (10)

_____ Other: _________________________________________________________________ (88)

_____ Not reported (99)
Level of Mental Retardation (based on test data above or described in report):

_____ None (1)
_____ Mild (70-55) (2)
_____ Moderate (54-40) (3)
_____ Severe/Profound (39 and below) (4)
_____ Untestable (5)
_____ Other: _____________________________________________ (8)
_____ Not reported (9)

Level of verbal communication ability:

_____ Average language skills (1)
_____ Minimally verbal; some functional language (2)
_____ Nonverbal; mute (3)
_____ Echolalic (4)
_____ Other: _____________________________________________ (8)
_____ Not reported (9)

Secondary diagnoses: ______________________________________

________________________________________________________

Diagnostic criteria used:

_____ National society for Autistic Children (NSAC; 1978) (1)
_____ Rutter’s (1978) Criteria (2)
_____ DSM-III (3)
_____ DSM-IIIR (4)
_____ DSM-IV (5)
_____ Subject described as “autistic”, “autistic-like”, or having “autistic features” (6)
_____ Other: _____________________________________________ (8)
_____ Not reported (9)

Time since autism diagnosis (in months):

____________ _____ Not reported (999)

Time in prior treatment (in months):

____________ _____Not reported (999)
Intervention Information

Target behavior

As defined in report: __________________________________________________________

_____ Aggression (1) _______ Self-injurious Behavior (4)
_____ Property destruction (2) _______ Self-stimulation/stereotypies (5)
_____ Disruptive behavior (3)
_____ Combination: ______________________________________________________(10)
_____ Other: __________________________(8)
_____ Not reported (9)

*If SIB, what type?

_____ Head banging (1) _______ Face slapping (3)
_____ Eye gouging (2) _______ Biting (4)
_____ Combination: ______________________________________________________(10)
_____ Other: __________________________(8)

Functional assessment conducted?:

_____ Yes (1)
_____ No (2)
_____ Other: __________________________(8)

*If yes, complete below:

_____ Experimental Functional Analysis (EFA) (1)
_____ Partial experimental design (2)
_____ Antecedent-Behavior-Consequence sheet (ABC) (3)
_____ Rating Scale (e.g., MAS, QABF) (4)
_____ Informal assessment (5)
_____ Brief EFA (6)
_____ Modified EFA (7)
_____ Other: __________________________(8)

Function of behavior?:

_____ Attention (1)
_____ Escape (2)
_____ Access to tangibles (3)
_____ Communication (4)
_____ Self stimulation /Automatic (5)
_____ Undifferentiated (6)
_____ Combination: ________________________________________________ (10)
_____ Other: ________________________________________________________ (8)
_____ Not reported (9)

Type of intervention used:

_____ Reinforcement only (01)   _____ DRI only (05)
_____ Timeout only (02)   _____ Punishment only (06)
_____ Extinction (03)    _____ Overcorrection (07)
_____ Sensory extinction (04)   _____ Antecedent exercise (08)
_____ Combination: ________________________________________________ (10)
_____ Other: ________________________________________________________ (8)
_____ Not reported (99)

*If reinforcement used in intervention, complete below:

Type:

_____ Food (1)
_____ Token/Money (2)
_____ Sensory stimulation (3)
_____ Social (4)
_____ Toy; play time (5)
_____ Stereotypies (6)
_____ Combination: ________________________________________________ (7)
_____ Other: ________________________________________________________ (8)
_____ Not reported (9)

Schedule:

_____ Continuous (1)
_____ Fixed ratio (every _____ nth response) (2)
_____ Variable ratio (every _____ nth response on average) (3)
_____ Other: ________________________________________________________ (8)
_____ Not reported (9)

*If punishment used in intervention, complete below:

_____ Slap/spank (01)   _____ Social disapproval (e.g. “No!”) (05)
_____ Shock (02)   _____ Water mist (06)
_____ Restraint (03)  _____ Physical exercise (07)
_____ Visual screening (04)  _____ Ammonia (08)
_____ Combination: _____________________________ (10)
_____ Other: _____________________________ (88)
_____ Not reported (99)

Duration (in seconds): __________  _____ Not reported (9999)

Length of exposure to intervention / sessions (in minutes):

_____  _____ Not reported (999)

Medication administered during intervention?:

_____ Yes (1)
_____ No (2)
_____ Not specified; not reported (9)

Treatment setting:

_____ Home (01)  _____ Public School (05)
_____ Laboratory (02)  _____ Inpatient classroom (06)
_____ Inpatient ward (03)  _____ Outpatient clinic (07)
_____ Residential treatment setting (04)  _____ Center for Dev. Disorders (08)
_____ Combination: _____________________________ (10)
_____ Other: _____________________________ (88)
_____ Not reported (99)

Dimensions of treatment setting (specify feet, meters, yards, etc):

__________ by __________ (in__________)
_____ Not reported (99)

Behavioral therapist:

_____ Parent (01)
_____ Paraprofessional (02)
_____ Student research assistants (03)
_____ Teacher (04)
_____ Special Education Teacher (05)
_____ Psychologist (06)
_____ Peers (07)
_____ Self (08)
_____ Other professional (09)
_____ Combination: _____________________________ (10)
Parental involvement in treatment:

- Yes, as therapist (1)
- Yes, in other capacity (2)
- No (3)
- Not reported; unclear (9)

Type of experimental design:

- Simple A-B (1)
- Multiple baseline (Behavior, Setting, Subject) (2)
- Alternating treatments (e.g., A-B-A-C) (3)
- Reversal (e.g., A-B-A) (4)
- Randomization (5)
- Combination: ____________________________ (10)
- Other design: ___________________________ (8)
- Not reported (99)

Number of data points in first baseline phase: ________
Number of data points in final treatment phase: ________

Effort to generalize?:

- Yes
- No
- Other: ____________________________ (8)

*If yes, to:

- New situation; new setting; new context (1)
- New people (2)
- New behavior (3)
- New object (4)
- Combination: ____________________________ (10)
- Other: ____________________________ (8)
- Not reported (99)

Follow up data collected?:

- Yes (1) Interval after treatment, in months: ________
- No (2)
- Unclear (9)
Experimental Quality

Scale used to measure target behavior:

_____ Frequency count (length of interval: __________) (1)
_____ Percentage (length of interval: __________) (2)
_____ Rating on an interval scale (length of interval: __________) (3)
_____ Other: ________________________________________________________________ (8)
_____ Not reported (9)

Units of measurement on X axis:

_____ Days (1)         _____ Sessions (4)
_____ Trials (2)       _____ Weeks (5)
_____ Months (3)       _____ Other: _________________ (8)
_____ Not reported (9)

For functional assessment:

Inter-rater reliability for target behavior:

_________
_____ None reported (999)

Type of interobserver agreement:

_____ Kappa (1)
_____ Phi coefficient (2)
_____ Percent agreement (e.g. agree / agree + disagree x 100 ) (3)
_____ Other: ________________________________________________________________ (8)
_____ Not reported (9)

Interobserver agreement based on:

_____ Partial interval recording system (specify in seconds: __________) (1)
_____ Other: ________________________________________________________________ (8)
_____ Not reported (9)

For treatment:

Inter-rater reliability for target behavior:

_________
_____ None reported (999)

Type of interobserver agreement:

_____ Kappa (1)
_____ Phi coefficient (2)
_____ Percent agreement (e.g. agree / agree + disagree x 100 ) (3)
Interobserver agreement based on:

- Partial interval recording system (specify in seconds: __________) (1)
- Other: ________________________________________________________________ (8)
- Not reported (9)

Source of observations:

- Parent (01)
- Psychologist (05)
- Paraprofessional (02)
- Peers (06)
- Student research assistants (03)
- Self (07)
- Teacher (04)
- Combination: __________________________________________________________ (10)
- Other: ________________________________________________________________ (88)
- Not reported (99)

Effect Size

Mean % reduction from baseline: ______

PND (using Didden et al.’s (1997) modification):

- 1st baseline versus 1st treatment
- 2nd baseline versus 2nd treatment
- 1st baseline versus 2nd treatment
- Total (average of 1st baseline/1st treatment and 2nd baseline/2nd treatment)

PZD:

- 1st treatment phase
- 2nd treatment phase
- Total (average of 1st and 2nd treatment phases)
(NOTE: Data points equal distance from center of data point to X-axis, rounded to closest 0.5 mm if X-axis is labeled zero or if no clearly defined data points exist. If a clear zero data point lies above the X-axis, draw line perpendicular to Y-axis through midpoint of those data points clearly designated as zero in text of manuscript.)

<table>
<thead>
<tr>
<th>Phase</th>
<th>Phase</th>
<th>Phase</th>
<th>Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.____</td>
<td>1.____</td>
<td>1.____</td>
<td>1.____</td>
</tr>
<tr>
<td>2.____</td>
<td>2.____</td>
<td>2.____</td>
<td>2.____</td>
</tr>
<tr>
<td>3.____</td>
<td>3.____</td>
<td>3.____</td>
<td>3.____</td>
</tr>
<tr>
<td>4.____</td>
<td>4.____</td>
<td>4.____</td>
<td>4.____</td>
</tr>
<tr>
<td>5.____</td>
<td>5.____</td>
<td>5.____</td>
<td>5.____</td>
</tr>
<tr>
<td>6.____</td>
<td>6.____</td>
<td>6.____</td>
<td>6.____</td>
</tr>
<tr>
<td>7.____</td>
<td>7.____</td>
<td>7.____</td>
<td>7.____</td>
</tr>
<tr>
<td>8.____</td>
<td>8.____</td>
<td>8.____</td>
<td>8.____</td>
</tr>
<tr>
<td>9.____</td>
<td>9.____</td>
<td>9.____</td>
<td>9.____</td>
</tr>
<tr>
<td>10._____</td>
<td>10._____</td>
<td>10._____</td>
<td>10._____</td>
</tr>
<tr>
<td>11._____</td>
<td>11._____</td>
<td>11._____</td>
<td>11._____</td>
</tr>
<tr>
<td>12._____</td>
<td>12._____</td>
<td>12._____</td>
<td>12._____</td>
</tr>
<tr>
<td>13._____</td>
<td>13._____</td>
<td>13._____</td>
<td>13._____</td>
</tr>
<tr>
<td>14._____</td>
<td>14._____</td>
<td>14._____</td>
<td>14._____</td>
</tr>
<tr>
<td>15._____</td>
<td>15._____</td>
<td>15._____</td>
<td>15._____</td>
</tr>
<tr>
<td>16._____</td>
<td>16._____</td>
<td>16._____</td>
<td>16._____</td>
</tr>
<tr>
<td>17._____</td>
<td>17._____</td>
<td>17._____</td>
<td>17._____</td>
</tr>
<tr>
<td>18._____</td>
<td>18._____</td>
<td>18._____</td>
<td>18._____</td>
</tr>
<tr>
<td>19._____</td>
<td>19._____</td>
<td>19._____</td>
<td>19._____</td>
</tr>
<tr>
<td>20._____</td>
<td>20._____</td>
<td>20._____</td>
<td>20._____</td>
</tr>
</tbody>
</table>
APPENDIX B

Studies used in Meta-analysis


treatment of screaming in a young girl with severe disabilities. *Behavioral Interventions, 14.* 233-239.


APPENDIX C

Interrater Reliability Data

<table>
<thead>
<tr>
<th>Variable</th>
<th>Percent Agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>100.00</td>
</tr>
<tr>
<td>Gender</td>
<td>100.00</td>
</tr>
<tr>
<td>Race</td>
<td>100.00</td>
</tr>
<tr>
<td>IQ/Adaptive behavior score</td>
<td>100.00</td>
</tr>
<tr>
<td>Level of Mental Retardation</td>
<td>88.24</td>
</tr>
<tr>
<td>Level of verbal communication</td>
<td>82.35</td>
</tr>
<tr>
<td>Secondary diagnoses</td>
<td>100</td>
</tr>
<tr>
<td>Diagnostic criteria used</td>
<td>100.00</td>
</tr>
<tr>
<td>Time since autism diagnosis</td>
<td>100.00</td>
</tr>
<tr>
<td>Time in prior treatment</td>
<td>100.00</td>
</tr>
<tr>
<td>Target behavior</td>
<td>100.00</td>
</tr>
<tr>
<td>SIB type</td>
<td>100.00</td>
</tr>
<tr>
<td>FA conducted?</td>
<td>100.00</td>
</tr>
<tr>
<td>FA type</td>
<td>100.00</td>
</tr>
<tr>
<td>Function of behavior</td>
<td>94.12</td>
</tr>
<tr>
<td>Type of intervention</td>
<td>100.00</td>
</tr>
<tr>
<td>Type of reinforcement</td>
<td>100.00</td>
</tr>
<tr>
<td>Reinforcement schedule</td>
<td>100.00</td>
</tr>
<tr>
<td>Type of punishment</td>
<td>100.00</td>
</tr>
<tr>
<td>Punishment duration</td>
<td>100.00</td>
</tr>
<tr>
<td>Length of intervention</td>
<td>100.00</td>
</tr>
<tr>
<td>Medication administered?</td>
<td>94.12</td>
</tr>
<tr>
<td>Treatment setting</td>
<td>100.00</td>
</tr>
<tr>
<td>Dimensions of setting</td>
<td>100.00</td>
</tr>
<tr>
<td>Behavioral therapist</td>
<td>100.00</td>
</tr>
<tr>
<td>Parental involvement</td>
<td>100.00</td>
</tr>
<tr>
<td>Experimental Design</td>
<td>100.00</td>
</tr>
<tr>
<td># of data points in first baseline</td>
<td>100.00</td>
</tr>
<tr>
<td># of data points in final treatment</td>
<td>100.00</td>
</tr>
<tr>
<td>Effort to generalize?</td>
<td>100.00</td>
</tr>
<tr>
<td>Generalization to what?</td>
<td>100.00</td>
</tr>
<tr>
<td>Follow up data</td>
<td>100.00</td>
</tr>
<tr>
<td>Scale used to measure target behavior</td>
<td>100.00</td>
</tr>
<tr>
<td>Units on X axis</td>
<td>100.00</td>
</tr>
<tr>
<td>FA: interrater reliability</td>
<td>100.00</td>
</tr>
<tr>
<td>FA: type of agreement</td>
<td>100.00</td>
</tr>
<tr>
<td>FA: agreement recording system</td>
<td>100.00</td>
</tr>
<tr>
<td>TX: interrater reliability</td>
<td>100.00</td>
</tr>
<tr>
<td>TX: type of agreement</td>
<td>100.00</td>
</tr>
<tr>
<td>TX: agreement recording system</td>
<td>100.00</td>
</tr>
<tr>
<td>Source of Observations</td>
<td>100.00</td>
</tr>
<tr>
<td>MBLR</td>
<td>100.00</td>
</tr>
<tr>
<td>PND</td>
<td>100.00</td>
</tr>
<tr>
<td>PZD</td>
<td>100.00</td>
</tr>
</tbody>
</table>

Mean = 99.06, Range = 82.35 – 100.00