PREDICTORS OF OUTCOME FOR CHILDREN WITH AUTISM RECEIVING A BEHAVIORAL INTERVENTION

A Dissertation
Submitted to
The Temple University Graduate Board

In Partial Fulfillment
of the Requirements for the Degree
DOCTOR OF PHILOSOPHY

By
Melanie Pellecchia
August, 2013

Examining Committee Members:
James E. Connell, Ph.D., Advisory Chair, University of Pennsylvania
Joseph Ducette, Ph.D., Temple University Educational Psychology
Erin Rotheram-Fuller, Ph.D., Temple University School Psychology
Catherine Fiorello, Ph.D., Temple University School Psychology
Philip N. Hineline, Ph.D., External Member, Temple University Psychology

ABSTRACT

Autism spectrum disorders (ASD) are associated with many different levels of language and social impairment, differences in levels of cognitive impairment, varying levels of aberrant behavior, and discrepancies in the presence or amount of restrictive and repetitive behavior. The heterogeneity found within the ASD population is coupled with significant heterogeneity in outcome for these individuals. Although interventions based on the principles of applied behavior analysis have been repeatedly cited as evidence-based for individuals with autism, significant discrepancies in outcome are evident within the literature. The heterogeneity in treatment outcome has been hypothesized as related to the heterogeneity of children included within the autism spectrum as well as differences related to treatment variables. The current study evaluated individual child characteristics and differences related to intervention intensity and fidelity as predictors of outcome after one year of exposure to a behaviorally based intervention. The primary goal of this study was to identify individual and treatment level characteristics that were predictive of differences in outcome for children with ASD. Information was gathered from a total of 368 students with autism spectrum disorders in kindergarten through second grade classrooms in the School District of Philadelphia. Correlational analyses and multiple regression analyses indicated that increased levels of expressive language skills at the start of the intervention year were correlated with and predictive of improved outcome. However, no other child level variables were related to differential outcomes. Additionally, higher levels of treatment intensity and treatment fidelity were associated with improved outcome.

ACKNOWLEDGEMENTS

I am extremely grateful for the guidance and support provided by James E. Connell throughout this journey. His direction and encouragement have been instrumental in the completion of this project. He has been an invaluable advisor, mentor, and friend during the completion of this research and throughout my doctoral training. He has provided me with countless opportunities to expand and improve upon my professional skills, and has helped to make me a better school psychologist. Words cannot adequately express my gratitude for his support and guidance.

I am also thankful to David Mandell for allowing me to access and utilize data from the Philly AIMS project in this dissertation research. This dissertation would not exist without his sponsorship. I am also grateful for the opportunity to further develop my consultation and leadership skills while serving as a coach for his research project.

I am very appreciative for the assistance and guidance of Joseph Ducette and Erin Rotheram-Fuller while serving on this advisory committee and for helping to make this dissertation what it is today.

I'd like to thank the faculty of Temple University's School Psychology program for encouraging me to critically evaluate my philosophical orientation throughout my doctoral training. The training and clinical experiences provided throughout this program have helped me to gain a deep understanding and appreciation of differing views within school psychology, while strengthening my behavioral orientation. I am a more well-rounded and knowledgeable clinician because of this program.

I am especially thankful for Jean Boyer's support throughout my doctoral training. She has been an excellent mentor and has enthusiastically provided me with beneficial opportunities to improve my leadership and professional skills.

I am also thankful to Philip Hineline for his advising during the completion of my master's degree and for encouraging me to pursue a doctoral degree, as well as for pointing me in the direction of Temple's School Psychology program and suggesting I work with Jim Connell. This was the start of the path to my doctoral degree and I am very grateful for this suggestion.

I'd like to thank my friends and colleagues who have taken this journey with me: Emily, Celeste, Amber, Sarah, and Greg. It has been a pleasure to get to know you and work closely with you over the past four years. Thank you for sharing in the good times and bad times with me along the way.

I am forever grateful to my parents, Marta and Jose, for their love, support, and encouragement throughout this journey, and along every path I have ever taken. Thank you for your unwavering belief in me, for always pushing me to do the best I can, and for instilling an appreciation for education in me for as long as I can remember. You are amazing parents who have helped to make me the person I am today. Thank you for enthusiastically celebrating every accomplishment and for taking such pride in me.

I'd also like to thank my parents, my grandmother, Lola, and my sister Deliah, for their unconditional love and support along this road and throughout my life. Thank you for your patience and your encouragement. Thank you for believing in me. Most importantly, thank you for your countless hours of childcare. I could not have done this

without your support. I take enormous pride in being a member of this family, and I share this doctorate and all it stands for with you.

DEDICATION

I dedicate this work in its entirety to my amazing husband, Joe, our beautiful son Jake, and our soon-to-be-born baby girl. Everything I do is for this family. Joe, you are my rock and my strength. Thank you for taking this journey with me and for supporting me along the way. There is absolutely no way I could have done this without you. Thank you for the countless times you placed my needs above yours, for your emotional support, for encouraging me to keep going, and for always believing in me. I have constantly felt your love and support, and I am forever grateful. I dedicate this dissertation, and my entire doctorate, to you and our precious family. I can't wait to begin the next chapter of our lives together and to continue sharing every moment with you.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

CHAPTER 1

INTRODUCTION

Leo Kanner first identified autism in 1943 as a rare childhood condition observed and chronicled in eleven children. In his original description of these children he identified a range of abilities and symptoms associated with the disorder, including delayed, absent, or peculiar speech, aloneness, and an insistence on sameness (Blacher & Christensen, 2011). One year after Kanner's description of autism, Hans Asperger (1944) published an account of four children with typical intellectual ability and speech, but with significant impairments in social interactions and repetitive interests and behaviors. As chronicled in the detailed observations of Kanner (1943) and Asperger (1944), children with autism displayed substantial heterogeneity in presentation, with varying degrees of impairments in social interactions, language, and repetitive behaviors. These forerunners of the present day Diagnostic and Statistical Manual of Mental Disorders, fourth edition, Text Revision (DSM-IV-TR; APA, 2000) classification of Autism Spectrum Disorders (ASD), laid the foundation for decades of research and clinical practice regarding etiology, intervention effectiveness and intervention dissemination in community settings.

Currently autism spectrum disorders are listed under the diagnostic category of Pervasive Developmental Disorders in the DSM-IV-TR and include Autistic Disorder, Asperger's Disorder, Pervasive Developmental Disorder-not otherwise specified, and two less commonly diagnosed conditions, Rett's Disorder and Childhood Disintegrative Disorder (APA, 2000). However, increased consideration of the range of the spectrum

disorder has prompted the use of one name for the disorder (Autism Spectrum Disorder), in the fifth revision of the DSM (APA, 2012). Under the new DSM label, ASD continues to reflect the significant heterogeneity in presentation including a wide range in language and social impairment, differences in levels of cognitive impairment, varying levels of aberrant behavior, and discrepancies in the presence or amount of restrictive and repetitive behavior.

In addition to symptom heterogeneity, significant variation in intervention outcome has been identified for these individuals. That is, in most studies evaluating intervention outcomes for students with ASD, some students make significant progress, while others make minimal to no progress (Eikeseth, Smith, Jahr, & Eldevik, 2002; Lovaas, 1987; & Smith, Groen, & Wynn, 2000). Given that some children who receive high quality evidence-based interventions make significant gains while others do not, it is possible that outcome differences are a function of individual characteristics (Shcreibman, 2000).

Research evaluating the relationship between child characteristics and outcomes has produced mixed findings. Some researchers identified differences in cognitive level, as indicated through an IQ score, as a predictor of later outcomes, where higher cognitive scores are correlated with greater positive change scores on outcome measures (Gillberg & Steffenburg, 1987; Volkmar, Cohen, Bergman, Hooks, & Stevenson, 1989; Waterhouse et al., 1996). However, other researchers found no relationship between cognitive scores and later outcomes for children with autism (Ben-Itzchak, Lahat, Burgin, &Zachor, 2008). Researchers have also identified differences in communication and language abilities as predictors of outcomes, with greater communication and language

abilities at an early age as predictive of improved outcomes in later cognitive scores, language and communication skills, social skills, and less restrictive school placements (Goldstein, 2002; Luyster, Qiu, Lopez, & Lord, 2007; Mawhood, Howlin, & Rutter, 2000). Higher levels of early play and social skills (e.g., joint attention, imitation, and social interaction) have also been correlated with improved outcomes (Charman et al., 2003; Sigman & Ruskin, 1999; Schreibman & Stahmer, 2001; Toth, Munson, Melzoff, & Dawson, 2006), and a relationship between age and differences in outcomes has also been found, with younger ages at the start of intervention predictive of improved outcomes (Baker-Ericzen, Stahmer, & Burns, 2007). Additionally, researchers have identified differences in initial diagnosis among classifications within the autism spectrum (Lord et al., 2006) and levels of restrictive and repetitive behavior (Helt et al., 2008) as predictive of outcome. Overall, the research evaluating child characteristics associated with differences in outcome has identified a number of variables that may predict differential outcomes for children with ASD. Collectively, the extant research suggests that child-level characteristics may be predictors of intervention outcome success for children with ASD.

In addition to child-level characteristics, researchers have identified treatment variables, such as type of intervention and treatment intensity to be associated with differential outcomes. Overwhelmingly, interventions based on applied behavior analysis are considered best practice and have the most empirical support for individuals with ASD (National Research Council, 2001). Additionally, the research indicates that higher levels of treatment intensity of behavioral interventions are associated with improved outcomes for children with ASD (Eikeseth, Smith, Jahr, & Eldevik, 2002; Granpeesheh,

Kenzer & Tarbox, 2011; Lovaas, 1987; Smith, Groen, & Wynn, 2000). That is, children who received a greater number of weekly hours of behaviorally based interventions were more likely to have significantly better outcomes than children who received fewer hours of intervention. The research evaluating the combined effects of child characteristics and treatment intensity indicates that child-level characteristics such as cognitive ability, language and social skills, play skills, and age of diagnosis, may be more predictive of outcome than treatment intensity (Sallows and Graupner, 2005; Turner & Stone, 2007). Therefore, it appears that treatment intensity and child characteristics are related to differential outcomes for children with ASD, however the extent to which each predicts outcome is unclear. Additionally, which child characteristics are most predictive of outcome remain indistinguishable.

There are numerous behavior analytic approaches available, adding to the confusion about which interventions are most suited for this population. These interventions range from highly structured instructional approaches, such as Discrete Trial Training (DTT), to less structured naturalistic approaches, such as Pivotal Response Training (PRT; Prizant & Wetherby, 1998). There is a substantial body of research supporting both structured and naturalistic behavioral intervention approaches as effective for children with ASD (Birnbauer & Leach, 1993; Eikeseth, Smith, Jahr, & Eldevik, 2002; Koegel, Koegel, & Carter, 1999; Koegel, Koegel, Harrower, & Carter, 1999; Lovaas, 1987; Sheinkopf & Seigel, 1998). Research evaluating the comparative effectiveness of structured and naturalistic behavioral approaches yield mixed results, with some research indicating that naturalistic approaches are more effective for students with autism (Delprato, 2001), and other research indicating that naturalistic interventions

are more effective for skill acquisition and maintenance while structured approaches resulted in greater generalization of skills (Kane, Connell, & Pellecchia, 2010). That is, it appears that some students have improved outcomes with structured approaches while others have improved outcomes with naturalistic approaches (Elliott, Hall, & Soper, 1991; Kok, Kong, & Bernard-Opitz, 2002). However, the active mechanisms (i.e., the components of the intervention associated with change in outcome) of each approach associated with the differential outcomes have yet to be convincingly identified.

An additional variable associated with differences in outcome is treatment fidelity, or the extent to which an intervention is implemented accurately (Hagermoser, Sanetti, & Kratochwill, 2009; Pellecchia et al., 2011). Overall, research evaluating the relationship between treatment fidelity and outcomes indicates that higher levels of treatment fidelity result in better outcomes (Durlack & Dupree, 2008; O'Donnell, 2008). However, very limited research has specifically evaluated the relationship between treatment fidelity and outcomes for children with ASD. In the only study of its kind, Strain and Bovey (2011) found significantly improved outcomes for children with ASD related to increased levels of treatment fidelity. This study provides strong preliminary support for the need to monitor and address treatment fidelity in order to maximize outcomes for children with ASD. Further research is needed to support this finding.

The purpose of the current study is to extend and clarify the research evaluating predictors of outcome for children with ASD. A current research project funded by the National Institute of Mental Health (NIMH) and the Institute of Education Sciences (IES), the Philadelphia Autism Instructional Methods Study (Philly AIMS), provides a fortuitous opportunity to extend this line of research. Philly AIMS is a large-scale project

conducted by the Center for Autism Research within the University of Pennsylvania and The Children's Hospital of Philadelphia evaluating the implementation and sustainability of an evidence-based program for children with autism within the School District of Philadelphia. The project provides staff training and support for the implementation of a program based on the principles of applied behavior analysis, the Strategies for Teaching based on Autism Research (STAR) program in kindergarten through second grade autism support classrooms. The STAR program includes both structured (Discrete Trial Training) and naturalistic (Pivotal Response Training) behavioral approaches as primary intervention components. Additionally, research evaluating the STAR program indicates the program is effective for children with ASD (Arick et al., 2003).

The primary goal of the current study was to use the Philly AIMS data to identify child-level characteristics and treatment level variables that were predictive of change in IQ score as the outcome measure for children with ASD who participated in the study. Child characteristics and intervention variables were evaluated after one year of intervention using the STAR program by asking the following questions:

Question 1. Do the individual student characteristics measured predict outcome within a behavioral intervention for students with autism? Specifically, researchers evaluated various child level characteristics using measures of language ability, adaptive skills, autism severity, challenging behavior, and social skills. Those measures were then investigated as possible predictors of differential outcome for children with ASD.

Hypothesis 1. It was hypothesized that child level characteristics such as language skills, aberrant behavior, and social skills would be correlated with and predictive of outcome.

Question 2. Is there a relationship between intervention intensity and outcome? In the Philly AIMS trial, researchers collected treatment intensity data (e.g., number of intervention hours per week) for each intervention component in the STAR program (e.g., DTT, PRT, and FR) and evaluated the relationship with changes in IQ score.

Hypothesis 2. It was hypothesized that increased levels of treatment intensity (a greater number of weekly intervention hours) would be correlated with and predictive of outcome.

Question 3. Is there a relationship between intervention fidelity and outcome? Lastly, the relationship between levels of intervention fidelity for each intervention and student outcomes were evaluated.

Hypothesis 3. It was hypothesized that higher levels of intervention fidelity would be correlated with and predictive of outcome.

CHAPTER 2

LITERATURE REVIEW

Autism is a complex neurobiological disorder usually diagnosed in early childhood and lasting throughout a person's life (NICHD, 2005). It has become one of the most recognized disabilities in the United States and it is one of the most common developmental disabilities diagnosed. Currently, it is estimated that 1 out of every 88 children in the United States is diagnosed with an Autism Spectrum Disorder (CDC, 2012). Autism is characterized by impairments across three main areas: social reciprocity, communication, and restrictive and repetitive behavior (APA, 2000). However, individuals diagnosed with autism display varying levels of impairments in these domains, ranging from mild to severe (Jensen & Spannagel, 2011). Two related diagnoses, Asperger's disorder and Pervasive Developmental Disorder – not otherwise specified (PDD-NOS), represent milder forms of the disorder (APA, 2000).

Autism Heterogeneity

The various terms used to describe autism reflect the significant presentation of heterogeneity in the disorder. Autism is associated with many different levels of language and social impairment, variant mental ages, diversity in levels of cognitive impairment, varying levels of aberrant behavior, and discrepancies in the presence of restrictive and repetitive behavior (Waterhouse et al., 1996). The diagnosis of autism includes individuals found on a wide-range spectrum, from those exhibiting severe impairments across multiple areas of functioning to those displaying mild deficits in language and social areas of functioning. Formal diagnostic systems have assumed the presence of a core disorder (e.g., Autistic Disorder) and non-core groups (e.g., Asperger's

disorder and PDD-NOS; see APA, 1987; & APA, 2000). However, other researchers have suggested an autistic continuum, rather than a sharp boundary between core and non-core groups (Cohen, Paul, & Volkmar, 1986; Wing, 1988). Yet other areas of research have proposed two types of autism, high functioning and low functioning (Cohen, Paul, & Volkmar, 1987; Tsai, 1992). As the disorder takes shape in its many presentations, the term "Autism Spectrum Disorders" (APA, 2012) will be the classification used to represent all presentations of autism, including Asperger's disorder, and PDD-NOS.

The symptom heterogeneity found within the ASD population is parallel only to the outcome heterogeneity after intervention. Interventions based on the principles and techniques of applied behavior analysis have been repeatedly cited as the most effective evidence-based practices (EBP) for individuals with autism (National Autism Center, 2009); however, significant outcome differences are evident within the behavioral literature (Schreibman & Anderson, 2001). That is, in most studies evaluating the effectiveness of behavioral treatments for students with autism, some students make significant progress, while others make minimal to no progress (Eikeseth, Smith, Jahr, & Eldevik, 2002; Lovaas, 1987; Smith, Groen, & Wynn, 2000).

These inconsistent intervention outcomes lead researchers to question which variables influence treatment outcomes. The heterogeneity in treatment outcome has been hypothesized as related to the symptom heterogeneity found on the spectrum (Shcreibman, 2000). Given that some children who receive high quality EBP fail to make significant gains, it is hypothesized that differences in outcome for children with ASD are related to differences in levels of impairment. That is, a combination of

treatment and child characteristics are most likely related to differences in outcome. The heterogeneous presentation of symptomology within the ASD population has served to complicate the scientific and clinical detection of predictors of intervention outcomes.

Variables Associated With Outcome

Evaluations of Child Characteristics on Differential Student Outcomes

Limited research exists evaluating differences in outcome related to child characteristics, with mixed findings evident in the literature. Several studies have evaluated the relationship between initial IQ scores and intervention outcome, in which cognitive level (as indicated by the initial IQ score) was repeatedly identified as a predictor of outcome for children with autism (e.g., Gillberg & Steffenburg, 1987; Volkmar, Cohen, Bergman, Hooks, & Stevenson, 1989; Waterhouse et al., 1996). Harris and Handleman (2000) examined the relationship between age and IQ with educational placement for 27 children with autism following four-to-six years of behavioral intervention. They found that younger children with higher IQ scores were more likely to be placed in inclusive classrooms following intervention, whereas children who had lower IQ scores and were older at intake were more likely to be placed in special education classrooms. In a related study, Szatmari, Bryson, Boyle, Streiner, and Duku (2003) examined the extent to which early cognition and language skills predicted outcome after two and six years of intervention. They found that higher IQ scores and language skills strongly predicted improved outcomes, and that these results were stable over time. Furthermore, Ben-Itzchak and Zachor (2007) examined the relationship between cognition, communication, and socialization skills with outcome after one year of behavioral intervention. Results of this study indicated that children with higher IQ

scores and social skills on standardized measures were more likely to have improved outcomes. Similarly, Remington and colleagues (2007) evaluated differential outcomes for 23 preschool-aged children who received intensive behavioral intervention. They found that children who had higher initial IQ scores, as well as communication and social scores and more behavioral problems were more likely to have better outcomes following two years of intervention.

There is also research demonstrating no relationship between IQ and outcome for children with autism. Ben-Itzchak, Lahat, Burgin, and Zachor (2008) examined the relationship between cognitive ability and intervention outcome for children with ASD after one year of intervention with results indicating that baseline IQ score did not predict improved outcomes. However, as a whole, the research examining the relationship between cognitive ability and intervention outcome indicates that higher IQ scores, along with other variables, are prognostic indicators of improved outcome for children with autism, with some limited research finding opposing results.

Researchers have also examined the predictive relationship of communication and language abilities with intervention outcome for children with autism, indicating that early communication and language abilities are likely indicators of later improved outcomes. Luyster, Qiu, Lopez, and Lord (2007) evaluated the predictive relationship of measures of early language ability and later outcome for 63 children with autism. The authors evaluated the relationship between communication scores at ages 2 and 3 years with outcomes at 9 years of age. They found that children who had higher receptive and expressive language scores at age 3, as well as those who used symbolic gestures at age 2, were likely to have improved IQ scores, expressive language, and adaptive skills at

nine years of age. Similarly, Goldstein (2002) evaluated language skills as predictors of outcome and found that verbal imitation skills, combined with higher IQ scores at an early age strongly predicted language outcome following communication intervention for young children with autism. Furthermore, Mawhood, Howlin, and Rutter (2000) conducted a longitudinal study of males with autism or language disorders, matched on non-verbal IQ in childhood, and evaluated differences in outcomes for the two groups. They also found that language ability in childhood strongly predicted improved outcomes in early adulthood for individuals with autism. Overall, the research evaluating early language skills as predictors of outcome indicates that the presence or absence of early language is a strong predictor of later outcome for children with autism.

In addition to cognitive and language skills, researchers have evaluated the predictive relationship of other child-level characteristics and outcome. Charman and colleagues (2003) evaluated the association between diagnosis, play, joint attention, and imitation skills with language outcomes for 18 children with ASD. They found that the presence of joint attention and imitation skills at 20 months of age were strongly associated with improved language outcomes at 42 months of age. Additionally, children who were diagnosed with PDD-NOS at an early age were more likely to have better language outcomes than those diagnosed with autism. Early play skills were not associated with outcome in this study. Similarly, other researchers also evaluated the relationship between early joint attention and play skills with later outcome. Toth, Munson, Melzoff, and Dawson (2006) investigated the relationship between joint attention, imitation, and play skills with language outcome for 60 preschool-aged children with ASD. They found that these skills were strongly associated with later

language ability. Sigman and Ruskin (1999) also evaluated joint attention and social skills as predictors of outcome. They conducted a longitudinal investigation of changes in social competency for 155 children with autism, Down syndrome, and developmental delays. Children in this study were followed for an average of eight years during which they received a variety of special education and therapy services. For children diagnosed with autism, early joint attention skills were found to predict later improvements in expressive language and play skills. Additionally, children with greater nonverbal communication skills were more likely to have improved social skills in later childhood.

Investigations into the predictive relationship between social skills and later outcomes have also been conducted. Ingersoll, Schreibman, and Stahmer (2001) evaluated levels of peer social avoidance exhibited by students with autism and its effect on treatment outcome within an inclusive preschool classroom. Six children with autism participated in this study, and were divided into two groups matched on mental age and language level with differences in levels of peer social avoidance. The results of this study indicated that students with higher levels of peer social avoidance showed less improvement in language skills and social behavior; thus, level of peer social avoidance may be predictive of outcome within an inclusive educational environment.

Expanding upon this line of research, Sherer and Schreibman (2005) included a wider range of predictive variables in order to identify specific behavioral profiles associated with outcome. The results of this study found that students who exhibited low rates of toy play, low rates of approach behavior, low rates of verbal self-stimulatory behavior, and moderate to high rates of non-verbal self-stimulatory behavior had little or no gains after intervention. However, students who exhibited moderate to high interest in

toys, were tolerant of others within close proximity, exhibited moderate to high rates of verbal self-stimulatory behavior, and low to moderate rates of non-verbal self-stimulatory behavior made significant gains in the areas of language, play, and social skills following intervention. Both groups were similar regarding measures of intelligence, language, and autism severity. Thus, the primary difference between the two groups was their performance in the behavioral domains mentioned above, indicating that these variables may significantly affect treatment outcome.

Additionally, limited research has evaluated the relationship between demographic variables and outcome. Baker-Ericzen, Stahmer, and Burns (2007) evaluated whether gender, age, or ethnicity were associated with intervention outcomes. The results of this analysis revealed no significant differences between gender and ethnicity related to outcome. However, younger children (under three years of age) showed the most improvement following intervention.

In addition to research evaluating child characteristics predictive of improved outcome, other researchers have evaluated child characteristics that predict recovery (loss of the ASD diagnosis) or optimal outcome for children with ASD. Sutera and colleagues (2007) evaluated predictors of optimal outcome for toddlers diagnosed with ASD. They assessed differences between groups of children at age two who lost the diagnosis at age four and those who retained the diagnosis at this age. Children diagnosed with PDD-NOS were more likely to lose the diagnosis than those diagnosed with autism. Additionally, children who lost the diagnosis had significantly higher motor skills at age two than those who retained the diagnosis. No differences were found in symptom severity, socialization skills, or communication skills between the two groups of children.

Lord et al. (2006) also evaluated differences between children who lost the diagnosis of ASD and those who retained the diagnosis. They examined changes in diagnosis from ages two to nine. Children who received a non-spectrum diagnosis at age nine were more likely to have had an original diagnosis of PDD-NOS and not the core disorder, autism. More than 10% of children with a PDD-NOS diagnosis at age two moved off of the spectrum by age nine, while only 1 child out of 84 with a diagnosis of autism at age two was given a non-spectrum diagnosis at age nine. Additionally, Kelley, Naigles, and Fein (2010) compared differences in individual characteristics for 13 children who lost the ASD diagnosis and 14 high functioning children with ASD who retained the diagnosis, matched on age, gender, and non-verbal IQ. They found that children who lost the diagnosis were more likely to have average language skills and adaptive behavior, as well as exhibit average levels of problem behaviors. They were also more likely to have been diagnosed early, and receive intensive early intervention. In addition to these direct analyses, Helt and colleagues (2008) conducted a comprehensive review of existing research evaluating early predictive variables associated with later autism recovery. The review indicated that cognitive and motor development as well as an initial diagnosis of PDD-NOS instead of autism, were the strongest predictors of recovery. Additionally, they found that several physiological factors may be predictive of later outcome. Specifically, accelerated head circumference growth, the presence of seizures, and high levels of restrictive and repetitive behaviors were indicated as prognostic indicators of poorer outcomes.

In summary, research evaluating individual child characteristics associated with later outcomes has produced mixed findings. For example, researchers identified early

cognitive ability, as indicated by an IQ score, as a strong predictor of later improved outcomes. Others have identified the presence of language and communication skills as the strongest predictors of improved outcome. Some studies indicate that initial diagnosis was the strongest predictor of improved outcome, yet there was no relationship found between autism severity and outcome. Yet others have identified play, social, and motor skills as predictors of improved outcome. Finally, researchers have identified some physiological factors that may predict differences in outcome, such as head circumference growth and the presence of seizures. Collectively, this research suggests many child-level characteristics that may serve as predictors of outcome for children with ASD. It should be noted that none of this research evaluated differences in treatment variables (e.g., intensity, time, fidelity) associated with outcome.

Evaluations of Intervention Variables on Differential Student Outcomes

Differences In Intervention Intensity Related to Outcomes

In addition to research evaluating the relationship between child characteristics and outcomes, researchers have also evaluated the effects of intervention variables associated with differences in outcome. Several studies have evaluated the effects of differences in treatment intensity, or amount of time in intervention, on differential outcomes. Lovaas (1987) was one of the first to demonstrate significant differences in outcome based on treatment intensity. Children who participated in this study were assigned to one of two groups, an experimental group, in which 19 children received more than 40 hours of one-to-one DTT each week, or a control group in which 19 children received ten hours or less of one-to-one DTT each week. The results of this study indicated that the experimental group scored significantly higher on measures of IQ

as compared to the control group, and that 47% of the children in the experimental group achieved "normal functioning." Lovaas defined normal functioning as those students who had successfully completed first grade within a regular education classroom and had obtained an IQ score within the average range on the Weschler Intelligence Scale for Children-Revised. The outcomes for children who received high intensity intervention were in striking contrast to those of the students in the control group, with only 2% of the students in this group achieving normal functioning. On average, the students in the experimental group gained 30 more IQ points than those in the control group. Although this study did not evaluate differences in outcome based in child-level characteristics, it clearly indicated that differences in treatment intensity were strongly associated with differences in outcome.

However, some critics questioned the validity of the results and pointed to design limitations found in the study, such as a lack of random assignment and selection criteria that limited participation to children with relatively high IQs (Gresham & MacMillan, 1997; Schopler, Short, & Mesibov, 1989). Smith, Groen, and Wynn (2000) replicated and extended the Lovaas (1987) study and addressed some of the major criticisms by conducting a randomized clinical trial that included children from a wider range of functioning. Participants in this study were randomly assigned to either an intensive treatment group, which received 30 hours per week of intensive intervention, or a control group, which received five hours per week of intervention for two years. Results of this study revealed that students in the intensive treatment group had significantly better outcomes on measures of intelligence, visual-spatial ability, language, and academic achievement following intervention and had less restrictive school placements. However,

students in the intensive treatment group did not differ from those in the control group on measures of adaptive functioning or behavior problems. The Smith et al study did not evaluate the relationship between child level-characteristics and outcome; however, it supported Lovaas' finding that differences in treatment intensity were associated with differences in outcome.

More recently, Granpeesheh, Kenzer and Tarbox (2011) evaluated differences in outcomes in the areas of diagnostic classification, intellectual functioning, executive functions, challenging behavior, language, socialization, and independent living skills after two years of high or low intensity behavioral treatment for 60 children with autism. The high intensity group received 25 to 35 hours of weekly intervention, while the low intensity group received 8 to 15 hours per week of intervention. The high intensity group significantly outperformed the low intensity group on all measures following intervention. These studies indicate that treatment intensity, which is defined as hours of intervention per week in the studies above, is positively related to changes in IQ, language ability, social skills, diagnostic classification, and educational placement for children with autism.

A limited number of studies have evaluated differences in outcome based on a combination of child-level factors as well as treatment intensity for children with ASD. Granpeesheh, Dixon, Tarbox, Kaplan, and Wilke (2009) evaluated the effects of age and treatment intensity on behavioral intervention outcomes for 245 children with ASD and found both age and treatment intensity to be strong predictors of outcome. Specifically, an increase in treatment hours and a decrease in child age strongly predicted an increase in the number of mastered program objectives. Sallows and Graupner (2005) evaluated

differential outcomes for 24 children with autism after four years of intensive behavioral intervention. Results of this study indicated differential outcomes for participants, with 48% of the children demonstrating significant gains in IQ, as indicated by an average increase in 25 points and IQ scores in the average range following intervention. The authors evaluated the relationship between outcome and individual child characteristics as well as the number of weekly hours of intervention. The number of hours of intervention was found to be less related to outcome than child characteristics. Instead, several child characteristics were found to be strong predictors of outcome. Specifically, children with higher levels of imitation, daily living, language, and socialization skills were more likely to achieve better outcomes. Higher IQ scores were also found to be strong predictors of outcome. The results of this study clearly demonstrate differential outcomes for children with autism related to individual child characteristics, rather than treatment intensity. Turner and Stone (2007) found similar results when they evaluated child and environmental factors that contribute to variability in outcome for 48 children diagnosed with autism at age two. These authors found that children who were diagnosed before 30 months of age, had milder social impairments, and higher intelligence scores were more likely to lose their ASD diagnosis by age 4. No differences were found in outcome related to the amount of intervention services the children received. Collectively, these studies evaluating treatment intensity and outcome indicate that child-level characteristics such as cognitive ability, language and social skills, play skills, and age of diagnosis, may be more predictive of outcome than treatment intensity.

Differences In Type of Intervention Related to Outcomes

It is likely there is no intervention "silver bullet" for children with ASD. Indeed, the symptom heterogeneity suggests that intervention variables may be just as important when considering how best to improve the outcomes for children on the spectrum (Schreibman, 2000). Interventions based on applied behavior analysis are considered EBP and have the strongest empirical support regarding effectiveness for individuals with ASD (Green, 1996; Howard, Sparkman, Cohen, Green, & Stanislaw, 2005; National Research Council, 2001; Smith, Groen, & Wynn, 2000). In 2009, the National Autism Center published a comprehensive review of the EBP for this population (The National Standards Report, 2009). The authors reviewed more than six thousand studies based on specific search criteria and agreed that 775 met the rigid criteria needed for a scientific review of the extant literature. Those studies were then placed in four treatment categories (established treatments, emerging treatments, unestablished treatments, and ineffective/harmful treatments). Of those in the established category, over two thirds were from the behavior analytic community and 75% of the remaining one-third used behavioral techniques. Overall, behaviorally based interventions are described as the most effective for individuals with ASD. Although behavior analytic interventions are described as most effective for this population, a wide range of behavior analytic interventions was developed over the past four decades. These interventions fall on a continuum from highly structured instructional approaches (eg., Discrete Trial Training) utilizing massed discrete trials of isolated skills, to less structured naturalistic approaches utilizing training in the natural environment (Prizant & Wetherby, 1998).

Discrete Trial Training. Highly structured, contrived approaches for children with autism are most commonly referred to as Discrete Trial Training (Lovaas, 1987). The early application of DTT resulted in the first significant behavioral changes in children with autism (Lovaas, Berberich, Perloff, & Schaeffer, 1966; Lovaas & Simmons, 1969). Discrete Trial Training is implemented using an intensive one-to-one teaching session in a setting free from distractions. Discrete Trial Training generally involves mass trials, or the repeated practice of the same response for several successive teaching episodes and the use of reinforcers that are functionally unrelated to the response (e.g., providing access to a small amount of food for correctly identifying a car). Instruction within DTT involves breaking down complex skills into small component parts, and teaching each component part individually. For example, in order to teach a student with autism to play appropriately with toys, an instructor may first teach them to imitate actions with objects, such as pushing a car when provided with an imitative cue. This method of breaking down complex skills into smaller instructional targets coupled with the repeated practice of mass trials has been identified as a highly effective strategy for teaching individuals with autism (Ghezzi, 2007).

Several comprehensive reviews of the efficacy of DTT have demonstrated significant improvements in overall functioning, as demonstrated by increases in IQ score and inclusive educational placements. The most widely cited article demonstrating the effectiveness of DTT for students with autism is the seminal Lovaas (1987) study described above. Additionally, several subsequent studies conducted partial replications of the original Lovaas (1987) study (Birnbauer & Leach, 1993; Eikeseth, Smith, Jahr, & Eldevik, 2002; Sheinkopf & Seigel, 1998; Smith, Groen, & Wynn, 2000), with results for

all of these replication studies demonstrating substantial average increases in nonverbal IQ (22 to 29 points) as a result of discrete trial behavioral intervention. This research supports DTT as effective and demonstrates that DTT can lead to significant improvements in overall functioning for students with autism, with those gains maintained over long periods of time. Although DTT has been proven to be extremely effective in teaching a variety of skills to individuals with autism, several limitations have been identified. Researchers have described numerous disadvantages of DTT including limited generalization of learned skills, a lack of spontaneous responding, prompt dependency, robotic or stereotyped responding, the presence of escape and avoidance behaviors said to be caused by the nonfunctional nature of the training, and the primarily teacher-directed nature of the training (Schreibman & Anderson, 2001; Sundberg & Partington, 1998).

Naturalistic behavioral approaches. Alternative behavioral approaches to the treatment for individuals with autism have also been established and are purported to address the problems identified with DTT. These approaches, described as naturalistic behavioral approaches, typically involve more naturalistic and less artificially structured training protocols and assume many titles, including incidental teaching (Hart & Risley, 1975), natural language paradigm (Koegel et al., 1987), mileau training (Kaiser, Yoder, & Keetz, 1992), and pivotal response training (Koegel, Koegel, & Carter, 1999). Naturalistic approaches typically consist of loosely structured sessions that are initiated and paced by the child, take place in a variety of locations, and employ a variety of stimuli. Additionally, naturalistic procedures often require that the teaching materials be selected by the child and vary often during the teaching episode (Cowan & Allen, 2007).

Lastly, the reinforcers used within naturalistic approaches are functionally related to the target responses (e.g., providing a yellow piece of candy for correctly identifying the yellow candy). Although naturalistic approaches are less formally structured, they are no less behavioral than the highly structured approaches and have been repeatedly found to be effective for individuals with ASD (Schreibman, & Anderson, 2001).

Pivotal response training (PRT), originally referred to as the natural language paradigm (Koegel, O'Dell, & Koegel, 1987), is one of the most commonly used naturalistic behavioral approaches to instruction used with students with ASD. Pivotal Response Training has been used to effectively teach children with autism a wide range of skills including: a variety of language and communication skills such as increasing spontaneous vocalizations (Koegel, O'Dell, & Koegel, 1987), improving speech intelligibility (Koegel, Camarata, Koegel, Ben-Tall, & Smith, 1998), and improving conversational skills (Koegel, Camarata, Valdez-Menchaca, & Koegel, 1998). Pivotal Response Training has also been used as an effective method for reducing disruptive behavior (Koegel, Koegel, & Surratt, 1992), improving play skills (Stahmer, 1995; Thorp, Stahmer, & Schreibman, 1995), and improving joint attention initiation skills (Jones, Carr, & Feeley, 2006; Vismara & Lyons, 2007). Additionally, researchers have demonstrated the effectiveness of PRT in improving the social skills of students with autism (Pierce & Schreibman, 1995), including improving the number of social initiations and turn-taking behavior (Harper, Symon, & Frea, 2008). Furthermore, while trained professionals primarily implement most structured behavioral approaches for individuals with autism, parents have been taught to use PRT procedures resulting in improvements in their children's expressive language skills as well as parental affect

(Koegel, Symon, & Koegel, 2002), and PRT procedures have been effectively implemented by typically developing peers, with improvements in spoken language and toy play (Pierce & Schreibman, 1997a), as well as social behavior in children with autism (Pierce & Schreibman, 1997b).

Comparisons of Discrete Trial Training and naturalistic approaches. Due to the research supporting the effectiveness of both structured approaches, (e.g., DTT), and naturalistic approaches, (e.g., PRT), to intervention for children with ASD, there has been some debate regarding the comparative effectiveness of these different approaches, with DTT and PRT being the benchmark for each approach, respectively. Arguments and statements regarding the superiority of one treatment approach over the other have flourished within the behavioral literature (Schreibman & Anderson, 2001). In an attempt to provide evidence regarding the efficacy of these approaches, several researchers have conducted comparative analyses of structured and naturalistic approaches to intervention for autism. Research evaluating the comparative effectiveness of DTT and naturalistic approaches has produced conflicting results. Some research revealed no differences in acquisition rates, retention over time, or efficiency of instruction when comparing the interventions but improved generalization of skills taught using naturalistic intervention (McGee, Krantz, & McClannahan, 1985). Other research indicated faster acquisition of skills using discrete trial instruction and improved generalization of skills using naturalistic intervention (Miranda-Linne & Melin, 1992). Finally some researchers have found results favoring naturalistic behavioral approaches pertaining to skill acquisition (Koegel, Camarata, Koegel, Ben-Tall, & Smith, 1998; Koegel, O'Dell, & Koegel, 1987). Additionally, two broad scale reviews have been

conducted comparing the effectiveness of these approaches. Delprato (2001) reviewed studies in which DTT and naturalistic behavioral approaches to language intervention were implemented with students with autism. The main conclusion of the review was that naturalistic approaches to language instruction were more effective for students with autism, producing more successful acquisition and generalization of responses than structured approaches. In a more quantitative analysis, Kane, Connell, and Pellecchia (2010) conducted a meta-analysis comparing the effects of DTT and naturalistic approaches to language instruction for students with autism. The results of the meta-analysis indicated that naturalistic interventions were more effective than DTT for teaching language skills and for maintaining the skills learned. However, the data also indicated that DTT resulted in greater generalization than naturalistic approaches, a contradictory finding to the previous research discussed. This collective body of research reveals differing results regarding the comparative effectiveness of one behavioral strategy over the other. Given the heterogeneity found within the autism population, it is likely that different individuals may benefit from different intervention approaches (Schreibman, 2000).

Comparisons of behavioral interventions based on child characteristics. Limited research has evaluated the effectiveness of different intervention approaches for individuals with autism based on certain individual characteristics. Elliott, Hall, and Soper (1991) compared the acquisition rates and generalization of receptive and expressive labels using DTT and naturalistic approaches to intervention for twenty-three adults with autism. Differences in the acquisition and generalization of language skills were found based on cognitive level and order of instruction. The participants with

higher levels of cognitive functioning acquired more nouns when taught using DTT first. In contrast, participants functioning within the profound range of mental retardation (now called Intellectual Disability) performed better when the naturalistic approach was presented first. These results support the possibility of an interaction effect between individual level of cognitive functioning and type of instruction provided.

Differences in performance based on level of functioning and intervention approach have also been found in the acquisition of play skills. Kok, Kong, and Bernard-Opitz (2002) compared the effects of DTT and naturalistic approaches to instruction for eight young children with autism. Gains in communication and play behaviors were observed in both intervention conditions across participants. However, the children with higher scores on assessments of mental age, speech and language, and social skills exhibited more frequent communicative responses during the DTT conditions; and they initiated play more often and exhibited more appropriate play behaviors during the naturalistic condition. Thus, the results of this study indicate that there may be no difference in terms of the acquisition of play and communicative behaviors for children with autism demonstrating lower levels of cognition and language. However, there may be differences in the acquisition of these behaviors based on the type of intervention provided for higher functioning children with autism. Overall, the research evaluating differences in outcome based on type of behavioral intervention and child characteristics indicates that differences in outcome are related to a combination of type of intervention and child characteristics. However, the exact combination of variables contributing to these differences in outcome remains unclear.

Differences In Intervention Fidelity Related to Outcomes

A final area that is likely related to differences in intervention outcome is treatment fidelity. Broadly speaking, treatment fidelity refers to the extent to which an intervention was implemented with accuracy (Hagermoser Sanetti & Kratochwill, 2009). Dane and Schneider (1998) identified several distinct components of treatment fidelity: interventionist adherence, the extent to which intervention components are implemented as planned; and quality of implementation, qualitative aspects of intervention delivery, such as interventionist enthusiasm. Overall, research evaluating the relationship between treatment fidelity and outcomes indicates that higher levels of treatment fidelity result in better outcomes (Durlack & Dupree, 2008; O'Donnell, 2008). Wolery (2011) provides four reasons for measuring fidelity in intervention studies: 1) measuring fidelity allows investigators to document differences in outcomes between two conditions as being more likely related to the differences in conditions instead of some other factor; 2) fidelity measurements provide information regarding the transportability of the intervention to other contexts; 3) fidelity data provides information for replication; and 4) fidelity measures provide information about children's experiences within the intervention. Given these advantages to measuring treatment fidelity within intervention research, it is clear why increasingly more reports of intervention effectiveness include descriptions of intervention fidelity (Gresham, 2009). However, very limited research has documented the effects of treatment fidelity when evaluating intervention effectiveness for children with ASD. To date, only one comprehensive review of the effectiveness of different intervention models for children with autism reported treatment fidelity data. Strain and Bovey (2011) conducted a randomized control trial evaluating the effects of differing

intervention models for classrooms of preschool aged children with autism. The authors closely monitored implementation fidelity within both intervention groups throughout the study. Results indicated significant correlations between higher levels of treatment fidelity and improved outcomes for all classrooms. This study provides support for the need to monitor and provide training related to treatment fidelity regarding comprehensive interventions developed and implemented in community settings for children with ASD. Additional research is needed to replicate and strengthen these findings.

The STAR Curriculum

As previously mentioned, a wide range of behavior analytic intervention approaches are now available and widely disseminated for individuals with ASD. Several comprehensive interventions (sometimes called a curriculum or program) have recently been developed that include components from both structured (DTT) and naturalistic (PRT) behavioral approaches. The Strategies for Teaching based on Autism Research (STAR) program is one such example. The STAR program is a fully developed, manualized program for children with autism. It consists of three methods of instruction: Discrete Trial Training, Pivotal Response Training, and teaching within Functional Routines (FR). The curriculum is divided into six major areas: expressive language, receptive language, spontaneous language, functional routines, pre-academic concepts, and play and social interaction skills. Table 1, adapted from Arick, Loos, Falco, and Krug (2004), shows the interaction between each instructional strategy and the curriculum areas for which it is primarily used within the STAR curriculum.

The STAR program materials include detailed lesson plans, teaching materials, and forms and instructions for developing learning profiles (i.e., initial assessment and progress monitoring tool) and identifying appropriate curriculum content (i.e., scope and sequence) for each student. The program has three levels of instruction to meet the needs of children at different developmental stages, up to 8 years of age, and provides lesson plans for highly specific activities designed to improve skills in each curriculum area. Lessons provide a detailed plan to teach a specific skill or routine. Additionally, the lesson plans prescribe the instructional method (DTT/ PRT/FR) to be used for each

Table 1. Overview of the STAR Curriculum Areas and Intervention Approach Used Within the STAR Program.

Curriculum Area	Intervention Approach		
	Pivotal Response Training	Discrete Trial Training	Functional Routines Instruction
Expressive Language	*All expressive language lessons*	Specific imitative sounds and words Specific labels Most Level II and III lessons	Develop generalization of expressive language
Receptive Language	Taught incidentally within context of other PRT lessons	*All receptive language lessons*	Generalize use of receptive language within routines
Spontaneous Language	*All spontaneous language lessons*	Reinforce spontaneous language when it occurs	Set up situations in which the student needs to use spontaneous language
Functional Routines	Expand expressive language using PRT strategies within routine	Expand receptive language using DTT strategies within routine	*All activities comprise a predictable chain of behaviors*
Pre-academic Skills	Expand and generalize use of pre-academic skills	*All pre-academic lessons*	Generalize use of pre-academic skills
Play and social interaction concepts	*Play skills are taught within PRT play lessons and incidentally during PRT language lessons*	Social interaction and play are taught incidentally during one-to-one DT lessons	Develop appropriate play and social interaction during all appropriate routines

specific lesson. The STAR program also consists of specific instructional materials to be used during instruction, including data collection sheets. Each instructional strategy has particular data sheets that are unique to each strategy and track progress on each lesson.

Preliminary research regarding the efficacy of the STAR program shows promising results. Arick and colleagues (2003) described how teachers throughout the state of Oregon implemented the STAR curriculum and monitored the progress of 67 children with autism. The study showed that significant progress in social interaction, expressive speech, and use of language concepts was observed in the majority of students receiving instruction based on the STAR program. On average, 40% of the students in the study gained more than one month of language age for each calendar month, with most other students making smaller but significant gains. In addition, the student's functional communication and social skills improved, and decreases in behaviors associated with autism spectrum disorders were observed. The STAR program was the intervention used in the Philly AIMS randomized controlled trial and the dataset used in this investigation.

Philly AIMS

Although the preliminary research evaluating the effectiveness of comprehensive behavioral programs comprising both structured and naturalistic approaches such as the STAR program is promising, further research is needed to replicate and generalize these findings. To expand upon this research, the STAR program recently received funding from the National Institute of Mental Health (NIMH) and the Institute of Education Sciences (IES), for the Philadelphia Autism Instructional Methods Study (Philly AIMS). Philly AIMS is the largest evaluation of an autism intervention to date and was conducted

by the Center for Autism Research within the University of Pennsylvania and The Children's Hospital of Philadelphia. The Philly AIMS project was a comparative analysis of two behavioral programs STAR and Structured Teaching (another behavioral program) and included training and support for both programs for teachers and staff in kindergarten through second grade autism support classrooms. The primary purpose of this comparative evaluation was to examine student outcomes in the areas of school readiness skills, communication skills, socialization, adaptive behaviors, and challenging behaviors after one year of intervention. The results of the comparative evaluation are in review (Mandell, Stahmer, Shinn, & Marcus, 2012).

The Philly AIMS trial provided a fortuitous opportunity to evaluate a range of child characteristics captured by a battery of assessments including: cognitive ability, adaptive skills, social skills, language and communication skills, challenging behavior, and restrictive and repetitive behavior, with intervention variables including: time in instruction and intervention fidelity. The principal investigator of Philly AIMS has agreed to allow the researcher to access and evaluate existing data collected as part of Philly AIMS for use in this dissertation research.

Current Study

The purpose of this study is to extend the research evaluating the variables related to the effectiveness of intervention approaches for children with ASD by evaluating child-level and environmental factors related to differences in outcome, in an attempt to identify predictors of outcome for children with ASD. The study evaluated individual child characteristics and differences related to intervention for children after one year of

intervention within the STAR program. Three separate but related questions were asked in this study:

1. *Do the individual student characteristics measured in the Philly AIMS trial predict differences in outcome?*

As discussed above, previous research evaluating the predictive relationship between individual child characteristics and outcome has yielded mixed findings. Many different factors have been identified as predictive of outcome. However, when evaluated collectively, the research related to child characteristics that predict outcome identifies conflicting variables, with some research identifying cognitive and language skills as most predictive of improved outcome, and other research identifying play, social, and motor skills as most predictive of improved outcome. Additionally, most research evaluating predictive variables and outcome included a limited number of predictor variables and relatively small sample sizes, possibly contributing to the discrepant findings. This study attempted to clarify the research regarding individual child factors associated with outcome by identifying specific variables that predict differences in outcome following one year of intervention with the STAR program. This study also extended the current literature on this topic by including a large number of predictor variables related to a wide range of skill areas with a larger sample of students.

2. *Is there a relationship between intervention intensity and outcome?*

Research evaluating the relationship between intervention intensity and outcome for children with ASD has produced mixed results. Some research has clearly indicated that higher levels of intervention intensity are strongly related to improved outcomes. However, other researchers have found that there is little to no relationship between

intervention intensity and outcome for this population. Additionally, research evaluating the combined effects of child characteristics and intervention intensity on outcome indicates that child characteristics are stronger predictors of outcome than treatment intensity. This study evaluated whether differences in intervention intensity for DTT, PRT, and FR interventions provided within the STAR program were related to differences in outcome. Additionally, the relationship between individual child characteristics and treatment intensity were investigated.

3. *Is there a relationship between intervention fidelity and outcome?*

Overall, research evaluating the relationship between treatment fidelity and outcomes indicates that higher levels of treatment fidelity are associated with improved outcomes. However, very limited research exists specifically examining the relationship between treatment fidelity and outcomes for children with ASD. This study extends the research on this topic by evaluating the relationship between intervention fidelity and outcome after one year of intervention within the STAR program. Specifically, the study investigated the relationship between outcome and intervention fidelity for DTT, PRT, and FR. Additionally, the relationship between individual child characteristics and treatment fidelity were investigated.

CHAPTER 3

METHODOLOGY

Participants

Participants were selected from an existing research project, the Philadelphia Autism Instructional Methods Study (Philly AIMS), conducted through the Center for Autism Research at Children's Hospital of Philadelphia and the University of Pennsylvania, and funded by the National Institute of Mental Health (NIMH) and the Institute of Education Sciences (IES). All participants were active participants in the Philly AIMS study and were enrolled in kindergarten-through-second grade autism support classrooms in the School District of Philadelphia. Mean participant age was 6 years 8 months with an overall age range of 5 years 0 months to 9 years 8 months. Additionally, 87% of the participants were male. The participants represented a diverse sample and consisted of the following ethnic distribution: 53% African-American, 30% Caucasian, 10% Hispanic, 5% Asian, and 2% other ethnicities. Approximately 75% of the participants were eligible for free or reduced-price lunches. Finally, all participants had a diagnosis of autism through the School District of Philadelphia.

Information gathered from a total of 368 students who participated in Philly AIMS for the 2010 – 2011 academic year was used in the study. However, missing data from various assessment scales necessitated the use of different sample sizes for various analyses within the study. Missing data resulted from various uncontrollable events, such as student absence, incomplete surveys, or missing surveys. Missing data qualified as missing completely at random (MCAR: Little & Rubin, 2002), thus bias was not introduced as a result of the missing data. Additionally, sample sizes remained high

enough to preserve power within the analyses (Davy & Savla, 2009). A summary of the sample sizes used for each variable can be found in Table 3.

Setting

Participants were enrolled in the autism support classrooms in the School District of Philadelphia, a large urban school district. The School District of Philadelphia is the eighth largest school district in the country by enrollment, with more than 163,000 students enrolled. More than 1,000 of these students are identified with an autism spectrum disorder. The School District of Philadelphia students are racially and ethnically diverse, 61.2% are African American, 17.6% are Hispanic, 13.3% are Caucasian, 6.2% are Asian, and 1.8% are of other races or ethnicities. The district is also economically diverse, with 71% of its students eligible for free or reduced-price lunches.

Measures

Primary Measures

Multiple measures were used to monitor student outcomes within Philly AIMS. Children's functioning across a variety of domains was assessed using direct observation measures, parent report measures, and teacher report measures. These comprehensive assessments were administered once or twice each academic year, by assessment teams trained and considered "research reliable" to administer these measures. The results of several of these assessment scales were used for analyses within the study. The following scales were used as indicators of individual student characteristics:

Aberrant Behaviors Checklist (ABC)

Challenging behavior was measured using the ABC via parent survey. The ABC evaluates the presence of problem behaviors in individuals with developmental delays

and takes 10-15 minutes to complete. It can be used with individuals from 5 to 54 years of age. The ABC includes 58 items and produces scores across five subscales, including irritability, agitation, lethargy/social withdrawal, stereotypic behavior, hyperactivity/noncompliance and inappropriate speech. Psychometric studies of the ABC have shown high internal consistency (subscale range of 0.86 - 0.95), moderate inter-rater reliability (0.63), and very high test-retest reliability (0.96 - 0.99). Criterion validity is also high (0.87) (Aman, Singh, & Turbott, 1987). Results of the ABC were used to identify levels of inappropriate behaviors in students for this study.

Adaptive Behavior Assessment System (ABAS-II)

The parent survey version of the ABAS-II was administered to all participants in Philly AIMS. The ABAS-II uses a behavior-rating format to assess adaptive behavior and related skills for individuals, birth through 89 years of age. The scores derived from the ABAS-II describe an individual's general adaptive behavior as well as his or her functioning in ten related adaptive skill areas: communication, community use, functional academics, school/home living, health and safety, leisure, self-care, self-direction, social, and work (for older adolescents and adults). These skill areas encompass the practical, everyday skills required to function and meet environmental demands, including those needed to effectively and independently care for oneself and to interact with others. Reliability estimates for the ABAS-II are high; estimates of internal consistency and test-retest reliability are above .90. Average reliability coefficients for the 10 skill areas across all ages range from .85 to .97. Inter-rater reliability and cross-form consistency also are high (.91 to .99) (Harrison & Oakland, 2003). Factor analytic, concurrent validity and clinical studies provide strong support for its validity. The correlation between the

school-age form and the Vineland Adaptive Behavior Scales–Classroom Edition (VABS) Adaptive Behavior Composite is .82. Furthermore, all scores from nine ABAS skill areas and the three VABS subdomains correlate significantly ($p < .01$) (Harrison & Oakland, 2003). Results of the ABAS-II were used to evaluate levels of independence with adaptive skills for participants.

Autism Diagnostic Observation Schedule (ADOS)

Autism symptoms were measured using the ADOS by the Philly AIMS research team. The ADOS is a semi-structured standardized observational measure of social interaction, communication skills, and play or imaginative use of materials for the assessment of autism spectrum disorders. The ADOS consists of standard activities that allow the examiner to observe behaviors that have been identified as important to the diagnosis of autism spectrum disorders at different developmental levels and chronological ages (Lord, Rutter, DiLavore, & Risi, 1999). It takes approximately 45 to 60 minutes to complete. The ADOS is widely considered the gold standard to confirm an autism spectrum diagnosis. One of four possible modules is administered to each child, determined by the child's expressive language ability and chronological age. The measure can be validly used with children who have a minimum cognitive developmental level of 15 months. The ADOS has been reported to have strong psychometric properties, with high sensitivity and specificity reported for each of the four modules. Sensitivity and specificity in identifying autism spectrum disorders for Module 1 were both reported as 94%. Sensitivity for Module 2 was reported as 89%, and specificity for this module was 88%. For Module 3 sensitivity was reported as 80%, while specificity was reported as 94%. Lastly, sensitivity for Module 4 was reported as 86% and

specificity for this module was reported as 93% (Lord et al., 1999). For the purposes of this study, the ADOS was used to provide information regarding each student's overall autism severity, as indicated by the autism severity score produced by the ADOS.

Child Symptom Inventory (CSI)

Psychiatric symptoms were measured using the Child Symptom Inventory by the Philly AIMS team via parent survey. The CSI is based on diagnostic criteria from the Diagnostic and Statistical Manual of Mental Disorders, shows high predictive and concurrent validity and is used across a wide variety of settings. The CSI can be completed in less than 20 minutes. The test-retest reliability ranged from 0.47 to 0.88 and internal consistency for the CSI ranges from 0.70 to 0.96 (Sprafkin, Gadow, Salisbury, Schneider, & Loney, 2002). It has been used and validated in a number of studies of co-occurring psychiatric symptoms among children with ASD. Children are categorized on the CSI as meeting or not meeting the clinical cut-offs for three disorders whose symptoms commonly occur in children with ASD: attention deficit/hyperactivity disorder, anxiety disorder, and conduct/oppositional defiant disorder. Additionally, the CSI provides information regarding the presence of other psychiatric disorders including: major depressive disorder, dysthymic disorder, and social phobia. The results of the CSI were used to evaluate the presence of each of these disorders in the participant sample.

PDD Behavior Inventory (PDD-BI)

The PDD-BI is administered via teacher survey at the beginning and end of each academic year by the Philly AIMS team. The PDD-BI is a rating scale that assesses problem behaviors, social skills, language skills, and learning or memory

skills in children who have been diagnosed with Pervasive Developmental Disorders. It can be used with children between the ages of 1.6 and 12.5 years. The teacher rating form used in this study consists of 124 items and is administered in approximately 20 to 30 minutes. The PDD-BI has high reliability and validity, test-retest reliability for the teacher ratings ranged from .65 to .99 over an average 2-week interval and inter-rater reliability among teachers ranged from .85 to .92. Construct validity was also high and ranged from .63 to .73 for Factor 1 and .58 to .89 for Factor 2 (Cohen, 2003). The PDD-BI is highly correlated with the Childhood Autism Rating Scale (CARS) and the Vineland Adaptive Behavior Scales. Results of the PDD-BI were used to identify levels of problem behavior and social skills for the participants.

Social Responsiveness Scale (SRS)

The Social Responsiveness Scale (SRS) was administered by teacher report to measure children's social skills by the Philly AIMS research team. The SRS provides an overall evaluation of the social impairments associated with autism as they occur within natural settings. The measure provides an overall assessment of social functioning through the assessment of social awareness, social information processing, capacity for reciprocal social communication, social anxiety/avoidance, and autistic preoccupations and traits. The SRS is appropriate for children between 4 and 18 years of age and consists of a 65-item rating scale that is completed by a parent or teacher in approximately 15 to 20 minutes. It has been demonstrated as sensitive and reliable across a wide range of symptom severity. Measures of reliability for the SRS indicate strong test-retest reliability for both males (.85) and

females (.77), as well as high inter-rater reliability ranging from .75 to .91. Internal consistency measures for the sub-scales were also high and ranged from .77 to .92 (Constantino, Gruber, & Christian, 2005). Measures of validation for the SRS against other diagnostic measures for autism revealed a strong association between the SRS and the Autism Diagnostic Interview-Revised, which is largely considered the gold standard for autism diagnostic interviews. The SRS was particularly useful for this study because it measures social impairment on a quantitative scale across a wide range of severity, thus providing information regarding the social functioning of children with varying degrees of autism severity. Results of the SRS were used to identify children with varying degrees of social impairment.

Differential Ability Scales – II (DAS-II)

In addition to the child characteristic measures described above, intervention outcome (i.e., change in IQ score) was measured by using the Differential Ability Scale, 2nd edition (DAS –II). That is, cognitive ability was measured using the DAS-II at the beginning and end of the academic year by the Philly AIMS research team. Specifically, the change score derived from the difference in overall cognitive score from the DAS-II assessment administered at the start of the academic year (in September) and the DAS-II assessment administered at the end of the academic year (in June) was used as the outcome measure for the study. The DAS–II is a comprehensive, individually administered, clinical instrument for assessing the cognitive abilities that are important to learning. The test may be administered to children ages 2 years 6 months through 17 years 11 months across a broad range of developmental levels. The diagnostic subtests measure a variety of

cognitive abilities including verbal and visual working memory, immediate and delayed recall, visual recognition and matching, processing and naming speed, phonological processing, and understanding of basic number concepts. The DAS-II may be considered a more appropriate assessment tool for evaluating the cognitive abilities of children with autism and other disabilities because it relies less on expressive language ability than other cognitive assessments and its primary purpose is as a tool for identifying and understanding the strengths and weaknesses in individuals rather than establishing a general IQ score. Although specific information regarding the utility of the DAS-II for children with autism is not available, the measure has been reported to have strong psychometric properties with the general population. Internal reliability for the sub-tests within the school age battery was high and ranged from .74 to .96. Similarly, test-retest reliability was also high and ranged from .78 to .83. Lastly, the DAS-II has been shown to be highly correlated with other measures for cognitive ability, with correlations with the WISC-IV generally found to be in the .70's and correlations with the Bayley-III ranging from .56 to .67 (Elliott, 1990). The DAS-II was used to evaluate differences in treatment outcome based on changes in overall cognitive score.

Secondary Measures

Additional variables were measured to evaluate possible effects on student outcome. Specifically, the study evaluated whether levels of intervention intensity and intervention fidelity affected outcome.

Intervention intensity

The Philly AIMS research team monitored levels of intervention intensity for each classroom by direct observation and teacher report for each of the EBP (DTT, PRT & FR) included within the STAR program. Intervention intensity was identified as the frequency of delivery of each EBP. Classroom teachers varied in their frequency of implementation for each component of the STAR program, largely based on individual teacher preference. Intervention intensity was coded using a Likert scale ranging from 0 to 4 with the following criteria for each score: 0: less than one time per week, 1: one time per week, 2: two to four times per week, 3: one time per day, and 4: two times per day. Information regarding intervention intensity was gathered for each child and used to evaluate the relationship between intervention intensity and student outcome.

Intervention fidelity

Fidelity of implementation of the STAR program was measured via monthly observations for each classroom by the Philly AIMS team. During these observations, the research team evaluated levels of program fidelity for each EBP included within the STAR program using implementation fidelity checklists. The research team coded fidelity of implementation using a Likert scale with the following scores: 0 (does not implement), 1 (poor use), 2 (somewhat accurate), 3 (mostly accurate), and 4 (highly accurate). The results of fidelity assessments were gathered for each classroom team and used to evaluate the relationship between intervention fidelity and student outcome.

Procedure

Question 1: Do the individual student characteristics measured predict outcome within a behavioral intervention for children with autism?

Several analyses were used to evaluate the relationship between individual student characteristics and outcome. First, intervention outcome was calculated by computing the DAS-II change score from Time 1 (September) to Time 2 (June). The resulting change score was used as the variable representing student outcome in all subsequent analyses. Next, a factor analysis using principal component analysis and Varimax rotation was conducted to reduce the number of variables used to classify individual child characteristics. Information gathered from the six primary measures described above, consisting of 37 total sub-scales, was included in the factor analysis.

The factors identified through the factor analysis were then used as variables representing individual student characteristics in all subsequent data analyses. The relationship between individual student characteristics and student outcome was then evaluated. Pearson correlation coefficients were calculated to evaluate the relationship between individual child characteristics and outcome. Multiple regression analyses were also conducted to examine the predictive relationship between individual child characteristics, as indicated by the factors established through the factor analysis, and outcome, as indicated by change score.

Additional analyses were conducted to further evaluate differences in individual child characteristics across the range of outcomes. To conduct these analyses the sample was divided into three groups based on the magnitude of change score. Group 1 included all students with a negative change score indicating a lower cognitive score after

intervention from Time 1 of the DAS-II administration to Time 2 (9 months later). Group 2 consisted of students who made no or minimal change with change scores ranging from 0 to a +5 increase from Time 1 to Time 2. Group 3 consisted of students who had a substantial improvement in change score (+6 to +43 points).

Analyses were then conducted to evaluate differences between groups regarding individual child characteristics. Specifically, a one-way ANOVA was used to evaluate differences between the three groups for each factor representing individual child characteristics.

Question 2: Is there a relationship between intervention intensity and outcome?

Students who participated in the Philly AIMS study received three different types of intervention: DTT, PRT, and FR, as part of the comprehensive curriculum provided within the study. Intervention intensity varied substantially across students. Several analyses were conducted to evaluate the relationship between intervention intensity and outcome. First, Pearson correlation coefficients were calculated to evaluate the relationship between intervention intensity for each intervention and outcome. Next, multiple linear regression analyses were conducted to evaluate whether intervention intensity for each intervention predicted outcome. That is, regression analyses were used to evaluate whether higher levels of intervention intensity predicted better student outcome. Additional analyses were conducted to evaluate differences in intervention intensity between change score groups. Specifically, analyses were conducted to evaluate whether students with changes scores demonstrating a loss of points, no change, or a gain in points received differing levels of intervention intensity for each intervention.

One-way ANOVA analyses were used to compare differences in intervention intensity for the three groups.

Question 3: Is there a relationship between intervention fidelity and outcome?

Several analyses were conducted to evaluate the relationship between intervention fidelity and outcome. Pearson correlation coefficients were conducted to evaluate the relationship between intervention fidelity for each EBP (DTT, PRT & FR) and student outcome. Additionally, a multiple regression was used to evaluate the predictive relationship between intervention fidelity and outcome. That is, multiple regression analyses were conducted to evaluate whether higher levels of intervention fidelity for each type of EBP within the STAR program predicted better outcome. Lastly, one-way ANOVA analyses were also used to compare differences in intervention fidelity for the three change score groups.

CHAPTER 4

RESULTS

Information gathered from a total of 368 students who participated in Philly AIMS for the 2010 – 2011 academic year was used in the study. Additionally, six primary assessment measures were used to identify individual child characteristics, and measures of intervention intensity and fidelity were used to identify intervention variables. These variables were used to investigate whether there was a relationship between the independent variables and the outcome variable used in the study.

Question 1: Do the individual student characteristics measured predict outcome within a behavioral intervention for children with autism?

Several analyses were conducted to evaluate the relationship between individual child characteristics and outcome. Six assessment scales, comprising of 37 total subscales, were used to identify individual characteristics. These 37 sub-scales were included in the factor analysis using principal component analysis and Varimax rotation. The analyses yielded eight factors explaining a total of 71.65% of the variance (see Table 2). Factor 1 was labeled Adaptive Behavior due to the high loadings of the following items from the Adaptive Behavior Assessment System (ABAS-II): social skills, leisure skills, home living, self-care, self-direction, functional academics, health and safety, community use, communication skills. This factor also included social-approach behaviors from the PDD Behavior Inventory (PDDBI). Factor 1 accounted for 26.13% of the total variance. Factor 2 was labeled Social Skills and Hyperactivity and accounted for 17.97% of the total variance. This factor included the following items from the Social

Table 2. Summary of Factor Analysis for Individual Child Characteristics

Factor	Eigen Value	% of Variance
1: Adaptive Behavior (AB)	9.669	26.133
2: Social Skills & Hyperactivity (SH)	6.650	17.974
3: Challenging Behavior (CB)	3.552	9.600
4: Depression & Anxiety (DA)	1.728	4.671
5: Stereotypic Behavior (SB)	1.410	3.811
6: Expressive Language (EL)	1.292	3.491
7: Conduct (CD)	1.200	3.243
8: Autism Severity (AS)	1.011	2.733
Total		71.655

Responsiveness Scale (SRS): social communication, social cognition, social awareness, autistic preoccupations, and social motivation; and the following items from the Child Symptom Inventory (CSI): hyperactivity-impulsivity, inattentiveness, and oppositional defiant behavior. Factor 3 was labeled Challenging Behavior due to the high loadings of the following items from the Aberrant Behavior Checklist (ABC): inappropriate speech, hyperactivity and noncompliance, lethargy and withdrawal, stereotypic behavior, irritability, agitation, and crying. Factor 3 accounted for 9.6% of the total variance. Factor 4 was labeled Depression and Anxiety and accounted for 4.67% of the total variance. This factor included the following items from the CSI: major depressive disorder, dysthymic disorder, social phobia, separation anxiety, and generalized anxiety. Factor 5 was labeled Stereotypic Behavior and accounted for 3.81% of the variance.

Items within this factor were all from the PDDBI and included: ritualistic behavior and resistance to change, social pragmatic problems, sensory-perceptual behaviors, semantic pragmatic problems, and autistic behavior. Factor 6 was labeled Expressive language, accounted for 3.49% of the variance, and included two items from the PDDBI: the expressive language scale and the social approach/withdrawal Problems scale. Factor 7 was labeled Conduct and included 1 item from the CSI: conduct disorder scale. This factor accounted for 3.24% of the variance. Lastly, Factor 8 was labeled Autism Severity and accounted for 2.73% of the variance. This factor included the autism severity score from the Autism Diagnostic Observation Scale (ADOS). The individual factor loadings for each variable are presented in Table 3. The factors identified through the factor analysis were used as the predictor variables in subsequent analyses.

Student outcome was assessed using the IQ change score for each student derived from pre- and post-intervention cognitive ability scores as measured by the Differential Ability Scales, 2nd Edition (DAS-II). The mean overall change score was 3.95 (see Table 4). However, change scores were distributed across a wide range, from -43 to 43, indicating substantial variability in outcome across participants. Figure 1 shows the distribution of change scores across participants, as well as the frequency of each change score.

Pearson correlation coefficients were calculated to evaluate the relationship between individual child characteristics and outcome (see Table 5). Results of the Pearson correlations indicate that expressive language skills and outcome had a weak positive correlation that was significant, $r = .18$, $p < .05$. However, none of the other individual child characteristics were significantly correlated with outcome. Multiple

Table 3. Factor Loadings and Sample Sizes for Individual Child Characteristics

	Factor Loadings								
Item	1	2	3	4	5	6	7	8	Sample
ABAS: Social skills	**.894**	-.214	-.035	.027	.019	-.026	.122	-.085	159
ABAS: Leisure Skills	**.868**	-.130	-.033	.004	.006	.057	-.134	.110	159
ABAS: Home Living	**.837**	-.092	-.005	.020	-.005	-.071	.126	-.156	159
ABAS: Self-care	**.799**	-.201	.052	.010	-.096	.096	-.102	.021	159
PDDBI: Social-approach	**.798**	-.270	-.086	.028	.037	.181	-.161	.104	159
ABAS: Self-direction	**.793**	.002	-.001	-.066	-.060	-.097	.176	-.013	159
ABAS: Functional Academics	**.792**	.128	-.039	-.175	.022	.037	.139	-.031	159
ABAS: Health & Safety	**.788**	-.136	-.118	.084	-.027	.085	.016	-.003	159
ABAS: Community use	.**725**	-.043	-.018	.021	-.020	-.157	.303	-.121	159
ABAS: Communication Skills	**.701**	-.118	-.139	-.092	-.066	.201	-.059	.031	159
SRS: Communication	-.252	**.783**	.302	.190	-.014	-.032	.046	-.095	149
SRS: Cognition	-.212	**.777**	.236	.118	-.027	-.005	.136	-.085	153
SRS: Awareness	-.238	**.756**	.136	.026	.005	-.045	.002	-.003	158

Table 3. (continued)

	Factor Loadings								
Item	1	2	3	4	5	6	7	8	Sample
SRS: Autistic Preoccupations	-.049	**.740**	.406	.246	.018	-.001	-.157	.099	154
SRS: Motivation	-.239	**.624**	.144	.343	.033	-.139	.100	-.092	155
CSI: Inattentive	-.266	**.495**	.267	.279	-.002	-.033	.226	.347	151
CSI: Oppositional Defiant	.081	**.416**	.386	.292	-.101	.327	.388	-.044	157
ABC: Inappropriate Speech	.117	.214	**.825**	-.017	-.071	.068	-.081	.014	154
ABC: Hyperactivity & Noncompliance	-.163	.335	**.730**	.275	.025	.053	.223	.041	154
ABC: Lethargy	-.135	.208	**.712**	.293	-.008	-.164	.180	-.101	156
ABC: Stereotypy	-.112	.268	**.709**	.218	.004	-.145	-.077	.079	156
ABC: Irritability	-.114	.227	**.685**	.344	.025	.154	.293	-.123	159
CSI: Major Depression	-.022	.120	.236	**.838**	.002	-.041	.061	-.138	157
CSI: Dysthymia	.071	.125	.159	**.819**	-.069	-.018	.049	-.209	157
CSI: Social Phobia	.030	.281	.016	**.618**	.033	-.005	-.196	.260	152
CSI: Separation Anxiety	-.051	.151	.225	**.601**	-.066	.073	.119	.302	154
CSI: Generalized Anxiety	-.056	.209	.363	**.562**	.061	.069	.213	-.100	157
PDDBI: Autism	-.078	.014	-.027	.004	**.887**	-.413	-.022	.069	205

Table 3. (continued)

	Factor Loadings								
Item	1	2	3	4	5	6	7	8	Sample
PDDBI: Ritualistic	.037	-.070	.016	.013	**.882**	.248	.044	-.028	219
PDDBI: Social-pragmatic	.016	.037	-.082	-.035	**.737**	-.314	-.033	.032	215
PDDBI: Sensory	-.178	.025	-.037	.018	**.672**	-.124	.136	.221	217
PDDBI: Semantic-Pragmatic	.025	-.029	.066	-.055	**.654**	-.130	-.322	-.115	193
PDDBI: Expressive	.070	-.019	.004	-.022	-.201	**.807**	.066	-.026	191
PDDBI: Approach Withdrawal Probs.	.157	-.080	-.021	.034	-.296	**.765**	-.041	-.178	193
CSI: Conduct disorder	.267	.140	.254	.123	-.072	.020	**.755**	.022	158
ADOS: Severity	.018	-.063	-.036	-.083	.115	-.160	-.018	**.709**	251

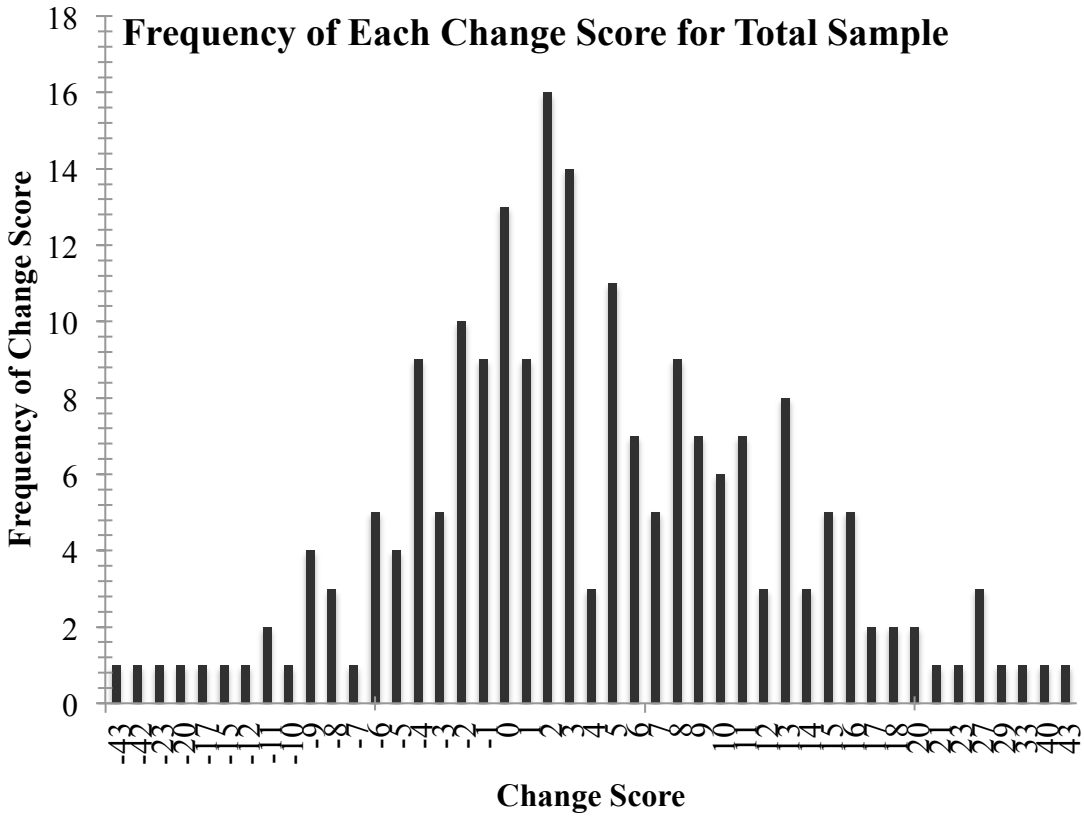

Figure 1. Distribution and frequency of change scores across participants.

Table 4. Change in Overall Cognitive Ability as Measured by the DAS-II for the Total Sample

Mean change score	Median change score	Range
3.95	3.00	-43 to 43

regression analyses were also conducted to examine the predictive relationship between individual child characteristics, as indicated by the factors established through the factor analysis, and outcome, as indicated by change score. A significant regression equation

Table 5. Correlations Between Child Characteristics and Change Score

Variable	Pearson correlation	N	Significance
1: Adaptive behavior	-.051	206	.464
2: Social Skills & Hyperactivity	.054	206	.440
3: Challenging behavior	.018	206	.799
4: Depression & Anxiety	.016	206	.817
5: Stereotypic behavior	-.079	206	.261
6: Expressive language	.178*	206	.010
7: Conduct	.030	206	.669
8: Autism severity	-.022	206	.749
Age	-.097	206	.164

* $p < .05$.

was found ($F(12, 355) = 2.244$, $p < .05$), with an R^2 of .039, indicating that individual child characteristics significantly predicted outcome. However, analyses of individual factors indicated that only one factor, Factor 6: Expressive Language, significantly predicted outcome ($\beta = .168$, $p < .05$). These results indicate that expressive language skills significantly predicted outcome, specifically, students with higher expressive language skills had a better outcome. All remaining factors did not significantly predict outcome (see Table 6). These results support those of the Pearson correlation analysis. In both analyses, expressive language skills were the sole factor found to have a significant relationship with outcome.

Overall, the results of the initial analyses for Question 1 indicate that expressive language skills are significantly correlated with outcome, however the correlation is weak. Additionally, expressive language skills were found to significantly predict student outcome. However, adaptive behavior, social skills and hyperactivity, challenging behavior, depression and anxiety, stereotypic behavior, conduct, autism severity, and age did not predict outcome and were not significantly correlated with outcome.

Given the range of outcome found across the sample, additional analyses were conducted to further evaluate differences in individual child characteristics across the range of outcomes. To conduct these analyses the sample was divided into three groups based on the magnitude of change scores. Group 1 was comprised of students who had a negative change score, indicating a loss in overall cognitive score after intervention. The mean change score for this group was -6.68, ranging from -1 to -43, and consisted of 29% of the total sample. Group 2 was comprised of students who demonstrated no or minimal change in score after intervention. The mean change score for Group 2 was 2.27, with a

Table 6. Summary of Linear Regression Analysis for Variables Predicting Change Score

Variable	B	β	*p*
Factor 1:			
Adaptive behavior	-.645	-.071	.181
Factor 2:			
Social Skills & Hyperactivity	.439	.048	.349
Factor 3:			
Challenging behavior	.250	.028	.592
Factor 4:			
Depression & Anxiety	.180	.020	.699
Factor 5:			
Stereotypic behavior	-.609	-.067	.193
Factor 6:			
Expressive Language	1.521	.168*	.001*
Factor 7:			
Conduct	.088	.010	.852
Factor 8			
Autism severity	-.238	-.0326	.611
Age	-.052	-.070	.244

*$p < .05$.

range of 0 to 5. Group 2 consisted of 32% of the total sample. Lastly, students who demonstrated a substantial gain in overall cognitive score were included in Group 3. The mean change score for Group 3 was 13.32, ranging from 6 to 43, and included 38% of the total sample (see Table 7).

Table 7. Change Score for Total Sample Divided into Groups Based on Magnitude of Change

Group	Description	Mean	Range	Group Size
1	Negative Change	-6.68	-1 to -43	60
2	No Change	2.27	0 to 5	66
3	Positive Change	13.32	6 to 43	72

Additional analyses were then conducted to evaluate whether or not differences in individual child characteristics existed between these groups. For example, analyses were conducted to evaluate whether students with a substantial gain in change score differed significantly in individual child characteristics than students with no change or a loss in points. To accomplish this, each factor depicting individual child characteristics was compared for the three groups using a one-way ANOVA. No significant differences were found between groups for any of the factors (see Table 8). Results of the one-way ANOVA for Factor 1 indicated that the groups did not significantly differ in adaptive behavior skills ($F(2, 190) = .206$, $p > .05$). The groups also did not significantly differ according to Factor 2, social skills and hyperactivity ($F(2, 190) = 1.632$, $p > .05$), Factor

3, challenging behavior ($F(2, 190) = 1.303$, $p > .05$), or Factor 4, depression and anxiety ($F(2, 290) = .191$, $p > .05$.). Similarly, no significant difference was found between groups for Factor 5, stereotypic behavior ($F(2, 190) = .369$, $p > .05$), Factor 6, expressive language ($F(2, 190) = 2.710$, $p > .05$), or Factor 7, conduct ($F(2, 190) = 2.145$, $p > .05$). Lastly, no significant differences were found between groups according to Factor 8, autism severity, ($F(2,190) = .291$, $p > .05$). Overall these analyses indicate that no significant differences in individual child characteristics were found between groups whose IQ scores decreased, remained relatively the same or made substantial gains following intervention.

In summary, the results of the analyses conducted to evaluate Question 1 revealed that Factor 6, expressive language skills, significantly predicted outcome. Students who demonstrated higher expressive language skills at the start of the intervention year were more likely to have a better outcome. All other individual child characteristics evaluated did not significantly predict outcome. Additionally, no significant differences were found when comparisons of individual child characteristics were made among students who demonstrated differences in outcome. Overall, these results indicate that expressive language skills are significantly related to improved outcome, students with higher levels of expressive language skills were more likely to have improved outcome. However, no other child characteristics were found to predict differences in student outcome.

Table 8. Summary of Analysis of Variance for Change Score Group Characteristics

Source	*df*	*F*	Mean Square	*p*
Factor 1				
Between groups	2	.206	.218	.814
Within groups	190		1.059	
Factor 2				
Between groups	2	1.632	1.768	.198
Within groups	190		1.084	
Factor 3				
Between groups	2	1.303	1.515	.274
Within groups	190		1.163	
Factor 4				
Between groups	2	.191	.217	.826
Within groups	190		1.137	
Factor 5				
Between groups	2	.369	.350	.692
Within groups	190		.949	
Factor 6				
Between groups	2	2.710	2.858	.069
Within groups	190		1.055	

Table 8. (continued)

Source	*df*	*F*	Mean Square	*p*
Factor 7				
Between groups	2	2.145	2.445	.120
Within groups	190		1.140	
Factor 8				
Between groups	2	.291	.296	.748
Within groups	190		1.016	

Question 2: Is there a relationship between intervention intensity and outcome?

Students who participated in the Philly AIMS study received three different types of intervention: DTT, PRT, and FR, as part of the comprehensive curriculum provided within the study. Intervention intensity varied substantially across students, largely depending upon the classroom teacher's intervention implementation. Intensity for each intervention ranged from less than one time each week to as much as two times per day. Given the differences present in intervention intensity, several analyses were conducted to evaluate the relationship between intensity and outcome.

First, Pearson correlation coefficients were calculated to evaluate the relationship between intervention intensity and outcome (see Table 9). Weak but significant positive relationships were found for DTT intensity ($r(196) = .147$, $p < .05$) and PRT intensity ($r(196) = .207$, $p < .01$), indicating that more of each of these interventions was related to better student outcome. Intensity of FR was not significantly related to outcome.

Table 9. Correlations Between Intervention Intensity and Change Score

Intervention Intensity	Pearson Correlation	N	Significance
Discrete Trial Training	.147*	196	.040
Pivotal Response Training	.207**	196	.004
Functional Routines	.118	196	.098

* $p < .05$ (2 –tailed)

** $p < .01$ (2 –tailed)

Additionally, multiple linear regression analyses were conducted to evaluate whether intervention intensity predicted outcome (see Table 10). A significant equation was found ($F(3, 330) = 2.948$, $p < .05$) with an R^2 of .026, indicating that intervention intensity predicted outcome. Analyses of the predictive relationship for each type of intervention indicated that PRT was the only intervention found to significantly predict outcome ($\beta = .293$, $p < .05$). Discrete Trial Training intensity ($\beta = -.113$, $p > .05$) and FR intensity ($\beta = .019$, $p > .05$) were not significant predictors of outcome. These results indicate that higher levels of intervention intensity for PRT predicted better student outcome. However, intervention intensity levels of DTT and functional routines did not predict differences in outcome.

Table 10. Summary of Linear Regression Analysis for Intervention Intensity as Predictor of Outcome

Intervention	B	β	*p*
Discrete Trial	-1.186	-.113	.407
Pivotal Response Training	3.379	.293*	.036
Functional Routines	.181	.019	.822

* $p < .05$.

Further analyses were conducted to evaluate differences in intervention intensity between change score groups. Specifically, analyses were conducted to evaluate whether differences in intervention intensity existed between groups with changes scores demonstrating a loss of points, no change, or a gain in points. One-way ANOVA analyses were used to compare differences in intervention intensity for the three groups.

Significant differences were found between groups for all three types of intervention (see Table 11). Groups differed significantly in intensity of DTT ($F(2, 180) = 3.769$, $p < .05$),

Table 11. Analysis of Variance for Differences Between Change Score Groups in Intervention Intensity

Intervention	*df*	F	Mean Square	*p*
Discrete Trial				
Between groups	2	3.769*	3.695	.025
Within groups	180		.981	
Pivotal Response Training				
Between groups	2	5.248*	4.158	.006
Within groups	180		.792	
Functional Routines				
Between groups	2	3.478*	4.237	.033
Within groups	180		1.218	

* $p < .05$.

PRT ($F(2, 180) = 5.248$, $p < .05$), and FR ($F(2, 180) = 3.478$, $p < .05$). Post-hoc analyses using Tukey's HSD were then conducted to determine the nature of the differences between groups (see Table 12). These analyses revealed that students who lost change score points had significantly less intense intervention for all three types of intervention than the students who made no gains or improved. Intensity of DTT for

Table 12. Tukey HSD Comparison for Differences Between Change Score Groups in Intervention Intensity

Group (A)	Group (B)	Mean Difference (A – B)	*p*
		Discrete Trial Intensity	
Loss	No change	-.47122*	.037
Loss	Gain	-.39999	.059
No change	Gain	.07123	.917
		Pivotal Response Training Intensity	
Loss	No change	-.49708*	.011
Loss	Gain	-.42751*	.019
No change	Gain	-.42751	.902
		Functional Routines Intensity	
Loss	No change	-.38238	.170
Loss	Gain	-.50817*	.029
No change	Gain	-.11949	.821

* $p < .05$.

students who lost points was significantly lower (m = 1.14, sd = 1.01) than DTT intensity for students who had no change (m = 1.610, sd = .90). No significant differences in discrete trial intensity were found between students who had a substantial gain in points (m = 1.54, sd = 1.03) and the other groups. Similarly, students who lost change score points received significantly less PRT (m = 1.23, sd = .96) than students who made no change (m = 1.73, sd = .85) and students who made substantial gains (m = 1.54, sd = .86). No significant differences in PRT intensity were found between students who had a substantial gain in points and students who made no change. Differences in intensity of FR were also found between students with a negative change score (m = 2.48, sd = .126) and students who made substantial gains (m = 2.98, sd = 1.04). However, no significant differences were found for intensity of FR between students who made no change (m = 2.86, sd = .99) and the other groups.

Overall, the analyses conducted to evaluate Question 2 indicate a significant relationship between intervention intensity and outcome. Intensity of PRT and DTT were significantly related to outcome, indicating that increased levels of intensity for both interventions were related to improved outcome. Additionally, intensity of PRT significantly predicted outcome. That is, students who received more PRT were more likely to have a better outcome. Lastly, comparisons of levels of intervention intensity between groups with differing change scores, revealed significant differences between groups. Students who had a negative change in IQ score significantly differed in levels of intervention intensity for all three types of intervention. Specifically, intervention intensity levels were lower for students who lost IQ points than for students who made no change or those who made substantial gains. These results indicate that overall students

who received more intense levels of intervention were more likely to have improved outcome.

Question 3: Is there a relationship between intervention fidelity and outcome?

Several analyses were also conducted to evaluate the relationship between intervention fidelity and outcome. First, Pearson correlation coefficients evaluating the relationship between intervention fidelity and outcome were conducted (see Table 13). Results of the correlational analyses indicated that PRT fidelity was significantly correlated with outcome ($r(196) = .177$, $p < .05$), however the correlation was weak.

Table 13. Correlations Between Intervention Fidelity and Change Score

Intervention Fidelity	Pearson Correlation	N	Significance
Discrete Trial Training	.124	196	.083
Pivotal Response Training	.177*	196	.013
Functional Routines	.052	196	.473

* $p < .05$ (2 –tailed)

Significant correlations were not found between fidelity of DTT ($r(196) = .124$, $p > .05$) or fidelity of FR ($r(196) = .052$, $p > .05$) and outcome. Overall, these results indicate that greater PRT fidelity was significantly correlated with improved outcome, however, no relationship was found between fidelity of DTT or FR with student outcome. A multiple regression was then used to evaluate the predictive relationship between the fidelity of each intervention and outcome (see Table 14). The results of the regression indicated that intervention fidelity significantly predicted outcome (F (3, 192) = 3.067, $p < .05$).

Table 14. Summary of Multiple Regression Analysis for Intervention Fidelity as Predictor of Outcome

Intervention	B	β	*p*
Discrete Trial	1.361	.168	.244
Pivotal Response Training	2.044	.208*	.045
Functional Routines	-2.220	-.228	.089

* $p < .05$.

However, individual analyses of the predictive relationship between each type of intervention and outcome revealed that fidelity of PRT significantly predicted outcome ($\beta = .208$, $p < .05$). However, fidelity of DTT ($\beta = .168$, $p > .05$) and FR ($\beta = -.228$, $p > .05$) did not predict outcome. These results are in agreement with the correlational analyses, again indicating that fidelity of pivotal response training was the only intervention related to student outcome.

Additional analyses were conducted to evaluate the presence of differences in intervention fidelity between change score groups. Specifically, analyses were conducted to evaluate whether differences in levels of intervention fidelity for each type of intervention existed between groups with differing change scores. One-way ANOVA analyses were used to compare differences in intervention fidelity for the three change score groups. No significant differences were found between groups related to intervention fidelity (see Table 15). Results of the one-way ANOVA for fidelity of DTT indicated no significant difference between groups ($F(2, 180) = 2.07$, $p > .05$). The groups also did not significantly differ according to fidelity of PRT ($F(2, 180) = 1.55$, $p >$

Table 15. Analysis of Variance for Differences Between Change Score Groups in Intervention Fidelity

Intervention	*df*	F	Mean Square	*p*
Discrete Trial				
Between groups	2	2.069	3.369	.129
Within groups	180		1.628	
Pivotal Response Training				
Between groups	2	1.551	1.744	.215
Within groups	180		1.124	
Functional Routines				
Between groups	2	1.298	1.512	.276
Within groups	180		1.165	

.05) or FR ($F(2, 180) = 1.29$, $p > .05$). These results indicate that all three groups received intervention with the same levels of fidelity, thus differences in outcome are not related to differences in intervention fidelity.

Overall, results of the analyses evaluating Question 3 indicated that fidelity of PRT was significantly related to outcome and significantly predicted outcome. Students who received PRT implemented with higher levels of fidelity were more likely to have a better outcome. However, fidelity of DTT and FR were not related to outcome. Additionally, no significant differences were found between change score groups and intervention fidelity.

In summary, most individual child characteristics evaluated were not found to be related to, or predictive of student outcome. Expressive language ability was the only individual characteristic found to have a significant relationship with outcome. Similarly, expressive language ability was the only individual characteristic found to significantly predict outcome. Students with higher levels of expressive language ability at the start of the intervention year were more likely to have a better outcome. Additional analyses comparing differences in individual child characteristics among groups of students with varying outcome indicated that no significant differences existed between groups. That is, the groups did not significantly differ according to the individual characteristics measured. Overall, expressive language skills were significantly related to outcome. However, all other individual characteristics were not related to outcome and there were no differences between groups of varying outcome related to any of these characteristics.

Results of the analyses evaluating the relationship between intervention intensity and outcome indicated that PRT intensity and DTT intensity were significantly correlated with outcome. However, intensity of FR was not related to outcome. Additionally, intervention intensity was found to significantly predict outcome, overall. However, individual analyses revealed that PRT was the only intervention found to significantly predict outcome. Students who received PRT at higher levels of intensity were more likely to have a better outcome. Intensity of DTT and FR did not predict outcome. Lastly, significant differences were found between change score groups and intervention intensity. Students who lost IQ points, as indicated by a negative change score, received significantly less intervention in all three EBP than students who made no change or those who made substantial progress.

Finally, analyses evaluating the relationship between intervention fidelity and outcome revealed that the fidelity of the PR was the only intervention fidelity found to have a significant relationship with outcome. Analyses of the predictive relationship between intervention fidelity and outcome indicated that overall intervention fidelity significantly predicted outcome. However, individual analyses indicated that PRT fidelity significantly predicted outcome, but neither DTT nor FR predicted intervention outcome. Students who received PRT with higher levels of fidelity were more likely to a have better outcome. However, no significant differences were found between change score groups related to intervention fidelity for any of the EBP.

CHAPTER 5

DISCUSSION

Children with ASD comprise a wide spectrum of ability, with substantial heterogeneity in presentation. Differences in outcome for children with ASD have been identified based on a number of individual child characteristics. However, previous research has produced conflicting results, with differing characteristics identified in the research as most predictive of outcome. Additionally, differences in the relationship between intervention intensity and outcome have been reported. This study evaluated differences in outcome for a large sample of children with ASD after one year of intervention with the STAR program. Differences in outcome were evaluated based on a number of individual child characteristics in an attempt to identify specific characteristics that were most predictive of improved outcomes. Furthermore, differences in outcome were evaluated based on levels of treatment intensity and treatment fidelity.

Evaluation of Child Characteristics on Differential Student Outcome

The results of the analyses evaluating the relationship between individual child characteristics and outcome indicated that only one characteristic was significantly related to outcome. Expressive language skills had a weak, but positive significant relationship with student outcome. Similarly, expressive language skills were the only characteristic found to significantly predict outcome. Children with higher levels of expressive language skills at the start of the intervention year were most likely to have a better outcome at the end of the year. These results are consistent with previous research evaluating language skills as predictors of outcome (Goldstein, 2002; Luyster, Qiu, Lopez, & Lord, 2007; Mawhood, Howlin, & Rutter, 2000). Overall, this research

indicates that children with ASD who display higher levels of expressive language and communication skills are more likely to have improved outcomes following intervention. This is important information for parents, educators, and clinicians of children with ASD. These results may imply that an earlier focus on improving language skills in young children with ASD could lead to improved outcomes for these students. Given that children with higher levels of expressive language skills are more likely to have better outcomes following intervention, one might assume that improving expressive language skills during the preliminary phases of intervention would lead to improved outcomes overall. This is an empirical question, worthy of further investigation.

The results of the analyses evaluating the relationship between individual child characteristics and outcome were surprising. It was unexpected to find that the majority of the characteristics evaluated were not significantly related to outcome. These findings are somewhat inconsistent with previous research. For example, this research found no significant relationship between adaptive skills and outcome. Although very little research has specifically investigated the relationship between adaptive behavior and outcome for children with ASD, the limited research that did evaluate this relationship has indicated a positive relationship between higher levels of adaptive skills and improved outcomes. For example, Sallows and Graupner (2005) found that children with higher levels of daily living skills, as well as higher levels of cognitive ability, language, social, and imitation skills, were most likely to have better outcomes. Similarly, Kelley, Naigles, and Fein (2010) found that children who displayed adaptive behavior skills within the average range of functioning were more likely to display optimal outcomes than children with deficits in adaptive behavior. Although these studies did not evaluate

adaptive skills in isolation as a predictor of outcome, they did find these skills to be predictive of improved outcomes. This is in conflict with the results of the current study, which did not find adaptive skills to be related to outcome. However, differences in findings may be due to several reasons, including differences in sample sizes and duration of intervention time. The Sallows and Graupner (2005) study comprised a much smaller sample of only 24 children, and the children were evaluated after four years of behavioral intervention, while the results of the current study report results after only one year of intervention. Similarly, Kelley, Naigles, and Fein (2010) included 27 children with ASD and reported results after eight years. It is possible that these differences in sample sizes and length of intervention could have resulted in discrepant findings. Further research evaluating predictors of outcome should include adaptive behavior as a possible predictor, given the discrepant findings.

Similarly surprising results were found regarding the lack of relationship between social skills and outcome. The results of this study indicated that social skills were not significantly related to outcome and did not predict differences in outcome. Although this is consistent with the findings of Sutera and colleagues (2007), who found no differences in social skills between children who had optimal outcomes following intervention and those who did not, the majority of the research evaluating the relationship between social skills and outcome for children with ASD has indicated that higher levels of social skills predict improvements in later outcome (Ingersoll, Schreibman, and Stahmer, 2001; Sallows & Graupner, 2005; & Turner & Stone, 2007). Possible explanations for the inconsistent findings may again be related to differences in sample sizes, with most of the previous research including much smaller sample sizes of

6 to 48 children. Additionally, differences in the operational definition and measurement of social skills could be related to different results. For example, Ingersoll, Schreibman and Stahmer (2001) defined social skills as levels of peer social avoidance and reported results based on direct observations of peer social avoidance. Other researchers have defined early social skills as joint attention skills, and have reported differences in outcome based on increased levels of joint attention initiation at an early age (Charman et al., 2003; Toth, Munson, Meltzoff, & Dawson, 2006; Sigman & Ruskin, 1999). In contrast, the present study identified social skills from the results of standardized assessments of social behavior as measured through rating scales such as the Social Responsiveness Scale. It is likely that the social behaviors included in these standardized assessments are more comprehensive than a single measure of social behavior, such as peer avoidance or joint attention, and would thus explain differences in findings. It may be that social skills, in general, are not related to differences in outcome, but individual social behaviors are related to differences in outcome. Further research evaluating social skills at a macro and micro level and their relationship to outcome for children with ASD is needed to clarify these discrepant findings.

Evaluations of the relationship between challenging behavior and outcome also failed to indicate a significant relationship between the two variables. Challenging behavior was not significantly related to outcome and was not found to predict differences in outcome. Very few previous studies have evaluated the relationship between challenging behavior and outcome for children with ASD. However, those that have evaluated this relationship have found significant relationships between the presence of challenging behavior and later outcomes. Surprisingly, Remington and colleagues

(2007) found that children who had higher levels of challenging behavior were more likely to have better outcomes following two years of intervention. In contrast, Kelley, Naigles, and Fein (2010) found that children who exhibited average levels of problem behavior, as indicated by the Behavior Assessment System for Children (BASC; Reynolds & Kamphaus, 1992), were more likely to lose their ASD diagnosis than high functioning ASD children with higher levels of problem behavior. Overall, the limited research evaluating the relationship between challenging behavior and outcomes yield mixed findings. The current study did not find a significant relationship between challenging behavior and outcome. Yet other research has identified a relationship between challenging behavior and outcomes, with some research indicating that increased levels of challenging behavior are related to improved outcome and other research indicating that less challenging behavior is related to improved outcomes. This is an area that requires further research to clarify these findings.

This study also evaluated the relationship between internalizing behaviors, such as depression and anxiety, and outcome for children with ASD. Depression and anxiety have been described as being among the most common co-morbid conditions for individuals with ASD (Simonoff et al., 2008), with a growing body of research documenting the presence of depression and anxiety symptoms in children with ASD (Bellini, 2004; Lecavalier, 2006; & Strang et al., 2012). The current study found no significant relationship between depression and anxiety and outcome for children with ASD. Additionally, these variables did not significantly predict differences in outcome. Despite the growing body of research indicating high comorbidity rates for depression and anxiety among children with ASD, no other research evaluating the relationship

between these internalizing behaviors and outcomes for children with ASD was found. As such, it appears that the current study is the first to evaluate the presence of these behaviors and their effects on outcome for this population. However, depression and anxiety have been repeatedly shown to have a negative correlation with outcomes for non-ASD children (e.g., Ansary & Luthar, 2009; Jones, 2009; & Vaughn et al., 2011). Therefore, it might be assumed that similar negative correlations between depression and anxiety and outcomes would be evident for children with ASD. It is unclear why a similar relationship was not found between internalizing behaviors and children with ASD. However, it may be that deficits in other areas, such as language and communication, overshadow depression and anxiety symptoms in children with ASD; thus minimizing their relationship with outcomes. Additionally, much of the research documenting the comorbidity of depression and anxiety has indicated a higher prevalence of these internalizing behaviors within high functioning children with ASD, such as those with Asperger's disorder. This study did not differentiate inclusion based on level of functioning, and thus included children at all levels of functioning within the autism spectrum. The inclusion of lower functioning children within the analyses may have negated the relationship between these comorbid conditions and outcome. Future research should more closely examine the relationship between depression and anxiety with outcomes for higher functioning children with ASD.

Another unexpected finding was in the lack of a significant relationship between levels of restrictive and repetitive behavior (stereotypy) and outcome. Results of the current study indicated that levels of restrictive and repetitive behavior were not significantly related to outcome and did not predict differences in outcome. This is again

in contrast to previous research evaluating the relationship between levels of restrictive and repetitive behavior and improved outcomes, which has indicated that higher levels of restrictive and repetitive behavior were associated with poorer outcomes for children with ASD. For example, Sherer and Schreibman (2005) found that children who exhibited higher rates of non-verbal self-stimulatory behavior had poorer outcomes than children who did not engage in this behavior. Similarly, Helt and colleagues (2008) found that children who had higher rates of stereotypy were more likely to have poor outcomes. Differences in findings related to the relationship between levels of restrictive and repetitive behavior and outcome may again be related to the differences in definition and measurement methods used to report this behavior. Sherer and Schreibman (2005) used direct observation to record the rates of restrictive and repetitive behavior, while Helt et al., (2008) describe a review of existing research with no direct manipulations or observations conducted. Both of these methods differ from the measurement system used in the current study, which involved identifying the presence and rate of restrictive and repetitive behaviors through behavior rating scales completed by a parent, such as the PDD-BI. It is possible that differences in the definition and measurement systems in each of the studies led to differing results. Further research is needed to clarify this discrepancy.

Autism severity was also found to be unrelated to differences in outcome in the current study. This was another surprising result, as it was hypothesized that autism severity would predict differences in outcome. Previous research regarding the relationship between autism severity and outcome has also indicated that autism severity is not predictive of differences in outcome. Similar to the current findings, Sutera and

colleagues (2007) found no differences in outcome based on autism severity. However, they did find differences in outcome based on initial diagnosis of PDD-NOS or autism, with children diagnosed as PDD-NOS having better outcomes. Other researchers have also found that an initial diagnosis of PDD-NOS rather than autism is associated with improved outcomes (Charman et al., 2006; Helt et al., 2008; & Lord et al., 2006). It seems contradictory that levels of autism severity would have no relationship with differences in outcome, while differing diagnoses on the autism spectrum are in fact related to outcome, given that a common assumption is that differing diagnoses are based on levels of autism severity. Helt and colleagues (2008) offer two possible explanations for this. They propose that the difference may be due to differences in the presence of restrictive, repetitive behaviors more commonly found in autism rather than PDD-NOS, or it may be that children with autism tend to be of lower intellectual functioning, thus accounting for the differences in outcome. The reason for this finding is unclear, however the current study supports previous research indicating that autism severity is not related to outcome. Differences in outcome based on initial diagnoses were not evaluated in the current study, as specific diagnostic classification information within the autism spectrum was not available.

The relationship between age and outcome was also evaluated in the current study. Previous research evaluating differential outcomes based on age has indicated that children of younger ages are more likely to have improved outcomes (Baker-Ericzen, Stahmer, &Burns, 2007; Granpeesheh, Dixon, Tarbox, Kaplan, & Wilke, 2009; Harris & Handleman, 2000). This has led to a common recommendation for intervening early while children are still young in order to maximize the potential for improved outcomes

(NRC, 2001). However, the current study found no relationship between age and outcome, and age was not found to predict differences in outcome. It is possible that failure to find differences in outcome may have been due to the restricted age range of the current sample. All participants were enrolled in kindergarten through second grade classrooms, with an age range of 5 years 0 months to 9 years 8 months. Most of the research reporting differences in outcome based on age has indicated that children who were under three years of age were more likely to have improved outcomes than older children (e.g., Baker-Ericzen, Stahmer, &Burns, 2007). The current sample did not include children younger than five years of age. Therefore, it may be possible that as children with ASD age, differences in outcome based on age become negligible. That is, as children with ASD become older, age no longer predicts differences in outcome. This is an interesting finding, worthy of further research.

Evaluation of Intervention Variables and Differential Outcome

In addition to the child-level characteristics evaluated in this study, several variables related to the intervention were evaluated to identify potential relationships with outcome. The relationship between treatment intensity for all three intervention components within the STAR program and outcome were evaluated. Overall, the findings of the current study indicated that higher levels of treatment intensity predict improvements in outcome. Specifically, higher levels of treatment intensity with PRT significantly predicted improved outcome. Additionally, weak but significant correlations were found between higher levels of treatment intensity for DTT and PRT with improved outcome. However, more instruction using FR were not related to outcome. Overall, the results of the current study are consistent with previous research

evaluating the relationship between treatment intensity and outcome, which has repeatedly indicated that increased levels of treatment intensity are strongly related to improved outcomes for children with ASD (Eikeseth, Smith, Jahr, & Eldevik, 2002; Granpeesheh, Kenzer and Tarbox, 2011; Lovaas, 1987; Smith, Groen, & Wynn, 2000). This data supports the recommendation that children with ASD should be provided with more hours of weekly intervention in order to improve outcomes.

However, it should also be noted that despite the weak but significant relationship between increased levels of treatment intensity and DTT found in the current study, DTT intensity did not significantly predict improved outcome. That is, children who received more hours of DTT, specifically, did not have better outcomes than those who received fewer hours of DTT. This was a surprising finding, given that the research evaluating treatment intensity for DTT overwhelmingly indicates that increased hours of intervention produce better outcomes for children with ASD (Eikeseth, Smith, Jahr, & Eldevik, 2002; Lovaas, 1987; Smith, Groen, & Wynn, 2000). Differences in findings may be due to the levels of treatment intensity provided within the current study and in previous research. High intensity intervention in the current study consisted of at least eight to ten hours of instruction each week with the given intervention. This is a much lower level of treatment intensity than has been reported as high intensity in previous research. For example, Lovass (1987) reported high intensity as 40 hours of weekly DTT. Similarly, high intensity in Smith, Groen, and Wynn (2000) consisted of as least 30 hours of intervention each week and Granpeesheh, Kenzer and Tarbox (2011) described high intensity as 25 to 35 hours of weekly intervention. It is possible that the levels of intervention in this study considered high intensity were not high enough to

produce significant differences in outcome with DTT. This may indicate that in order for DTT to be most effective, it should be provided at intensity levels of 25 to 30 hours each week, as indicated in the previous research. Surprisingly, PRT provided at high intensity levels of eight to ten hours each week were sufficient to predict significant improvements in outcome in the current study. This may indicate that PRT is a more efficient EBP and produces significant improvements in outcome when provided at lower levels of intensity. This is valuable information for parents and educators, especially those with children who are educated in classroom settings where the staffing and resources available prevent the application of very high intensity individualized intervention. Providing eight to ten hours of pivotal response training may be a more feasible option for these educational settings. This is an interesting finding worthy of further research.

Lastly, this study evaluated the relationship between treatment fidelity and outcome for each intervention provided within the STAR program. Results indicated that higher levels of PRT fidelity were slightly correlated with improved outcome. Similarly, PRT fidelity was also found to predict improved outcome. However no relationship was found between DTT or FR fidelity and student outcome. This finding is consistent with the results of the analyses evaluating levels of treatment intensity and outcome, which indicated that only higher levels of PRT intensity predicted improved performance on an IQ measure. Consequently, it seems logical that only increased PRT fidelity would predict improved outcome given that only PRT intensity was predictive of improved outcome. Previous research evaluating the relationship between treatment fidelity and outcomes indicates that higher levels of treatment fidelity result in better outcomes (Durlack & Dupree, 2008; O'Donnell, 2008; Strain & Bovey, 2011). However, it should

be noted that in order for this relationship to exist, the intervention itself must also be predictive of improved outcomes. That is, intervention fidelity will only lead to improved outcomes if the intervention itself leads to improved outcomes. Adhering to an intervention that is not related to improved outcomes should not produce improved outcomes. The current study supports this notion, as it indicated that increased PRT intensity predicted improved outcome and increased PRT fidelity also predicted improved outcome. However, increased treatment intensity for DTT and FR did not predict improved outcome, thus DTT and FR fidelity also did not predict outcome. That is, only fidelity of the intervention found to predict improved outcomes when delivered at a higher intensity was predictive of outcome. This is a logical finding, indicating that treatment effectiveness is directly related to treatment fidelity. However, it has not been specifically proven in the treatment fidelity literature, thus, this is the first study to empirically demonstrate this finding.

Study Limitations

Several limitations are worthy of mention in the current study. As previously mentioned the findings regarding the relationship between individual child characteristics and outcome were surprising, with only expressive language skills found as predictive of improved outcome. It was expected that several other variables would also be predictive of differences in outcome. Further analyses into the reasons for this surprising finding pointed to the mean change score for the overall sample. The mean change score for the sample was 3.95, a considerably small or even negligible difference. This indicates that as a whole, this student sample made very little change following one year of intervention with the STAR program. It is very likely that most of the individual child variables were

not found to predict outcome because overall there was very little change obtained. It is possible that improved outcomes for the group would have produced a greater number of individual child characteristics as predictive of differences in outcome. However, further analyses were conducted in an attempt to reconcile this possibility by dividing the sample into groups based on the magnitude of change score. As previously mentioned, these analyses indicated no significant differences between change score groups according to the predictive variables. That is, children who had substantial improvements in outcome did not differ from those who did not make progress along dimensions of the child level characteristics measured in this study. However, it may still be plausible that different outcomes regarding the predictive relationship of the characteristics measured would have been found with more substantial improvements in outcome for the group as a whole.

A second limitation is related to the outcome measure used in this study. Outcome was reported as a change in overall cognitive ability as measured by a standardized cognitive assessment, the Differential Ability Scales, 2nd Edition (DAS-II). Changes in IQ score are commonly reported as the outcome measure in research evaluating intervention outcomes for children with ASD (e.g., Eikeseth, Smith, Jahr, & Eldevik, 2002; Lovass, 1987; Smith, Groen, & Wynn, 2000). However, researchers have reported that the interpretation of intellectual assessment results for children with autism should be interpreted with caution, as a result of the possible interfering effects of disruptive behaviors that may underestimate the child's ability (Ben-Itzchak, Lahat, Burgin, & Zachor, 2008; Koegel, Koegel, & Smith, 1997; MacMullan, Manfredi, & Connell, 2012). Researchers have reported that several interfering behaviors common in

children with autism, such as limited attention, communication deficits, disruptive behavior, and lack of motivation may invalidate the results of standardized intellectual assessments for this population (Dietz, Swinkels, Buitelaar, Daalen, & Engeland, 2007). Furthermore, robust increases in IQ scores for children with autism following intervention have been suggested as an indication that the baseline IQ score reflects a lack of interest in children with autism to cooperate or participate in the assessment. A reduction in interfering behaviors commonly found in children with autism following intervention may enable the children to better participate in the assessment, allowing a more accurate reflection of their cognitive ability (Ben-Itzchak, Lahat, Burgin, & Zachor, 2008; MacMullan, Manfredi, & Connell, 2012). Although the overall mean change score in the current study was minimal, there was substantial variability in outcome, encompassing a wide range of differences in change scores, with some students gaining or losing upwards of 25 points. It is difficult to explain how students gained over 25 (or as much as 43) points on a cognitive assessment following one year of intervention. It is likely that this robust change in score is a result of a reduction of the interfering behaviors described above, which facilitated improved compliance and participation with the assessment and allowed a more accurate assessment of the child's ability. Similar explanations can be offered for those students who exhibited a drastic loss in points following intervention. It is unlikely that the students lost over 25 IQ points following one year of intervention. It is more likely that these children exhibited higher rates of interfering and noncompliant behaviors during the post-intervention cognitive assessment, producing scores that underestimated their true cognitive ability. Information regarding rates of problem behavior during the assessment would provide

evidence for this hypothesis. Future research should include information regarding the rates of interfering and challenging behavior during cognitive assessments for children with ASD.

A final limitation to the current study relates to the amount of missing data involved with the study. Much of the information used to identify child level characteristics was obtained through the administration of standardized parent and teacher survey reports. There were many instances of missing data in which all or portions of specific rating scales were not completed or completed inaccurately, causing them to be invalidated. Missing data resulted from various uncontrollable events, such as student absence, incomplete surveys, or missing surveys. As a result, missing data from various assessment scales necessitated the use of different sample sizes for various analyses within the study. However, missing data qualified as missing completely at random (MCAR: Little & Rubin, 2002), thus bias was not introduced as a result of the missing data. Additionally, sample sizes remained high enough to preserve power within the analyses (Davy & Savla, 2009).

Additional Recommendations for Future Research

Future research evaluating the relationship between individual child characteristics and outcomes for children with ASD should include a different outcome measure. Given the concerns discussed regarding the use of standardized intellectual assessments for individuals with ASD, and the results of the current study, it is recommended that future research include an outcome measure that is more sensitive to change as a direct result of intervention and can withstand the effects of interfering behavior. Curriculum-based measurements or other assessments that capture incremental

changes in performance would be better suited for this population. Future research should evaluate the utility of curriculum based measurement probes as an outcome measure for children with ASD. Additionally, future research should report outcomes following longer exposure to intervention. It is possible that more robust changes in outcome would be evident following longer exposure to the intervention. School-based research reporting outcomes following two to three academic years of intervention may be more likely to identify stable characteristics that predict differences in outcome. Lastly, it would be interesting to include younger students, below the age of five, in future research evaluating predictors of outcome. Although the age range of the current sample is appropriate for school-based research, valuable information regarding characteristics predictive of outcome may also be gathered from research involving younger children with ASD.

Summary

In summary, the current study provided surprising, yet valuable information for parents and educators of children with autism spectrum disorders. The results of the study are directly related and applicable to the education of young children with autism. Specifically, the results indicate that an early focus on expressive language skills may produce improved outcomes for children with ASD. Additionally, higher levels of treatment intensity, especially with pivotal response training intervention, can produce improved outcomes for this population. Furthermore, the results indicate that pivotal response training may be an effective intervention option for children in classroom settings with fewer staffing resources, given that significant changes in outcome were found with as much as eight to ten hours of weekly intervention. The current study

provides a solid foundation for further research evaluating variables that predict differences in outcome for children with autism spectrum disorders.

REFERENCES CITED

Aman, M. G., Singh, N. N., & Turbott, S. H. (1987). Reliability of the Aberrant Behavior Checklist and the effects of variations in instructions. *American Journal of Mental Deficiency, 92*, 237-240.

American Psychiatric Association (2012). American Psychiatric Association, DSM5 development. Retrieved March 10, 2012 from http://www.dsm5.org/ProposedRevisions/Pages/proposedrevision.aspx?rid=94.

American Psychiatric Association (2000). Diagnostic and statistical manual of mental disorders (4th ed., text revision). Washington, DC: American Psychiatric Association.

American Psychiatric Association (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., rev.). Washington, DC: American Psychiatric Association.

Ansary, N. S., & Luthar, S. S. (2009). Distress and academic achievement among adolescents of affluence: A study of externalizing and internalizing problem behaviors and school performance. *Development And Psychopathology, 21,* 319-341.

Arick, J., Loos, L., Falco, R., & Krug, D. (2004). *The STAR Program: Strategies for Teaching based on Autism Research*. Austin, TX: PRO-ED.

Arick, J.R., Young, H.E, Falco, R. A., Loos, L.A., Krug, D.A., Gense, M.A., et al (2003). Designing an outcome study to monitor the progress of students with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities,18*, 75-88.

Asperger, H. (1991/1944). ‘‘Autistic psychopathology’’ in childhood. (translated and annotated by U. Frith). In U. Frith (Ed.), Autism and Asperger

syndrome (pp. 37–92). New York: Cambridge University Press.

Baker-Ericzen, M. J., Stahmer, A. C., & Burns, A. (2007). Child demographics associated with outcomes in a community-based pivotal response training program. *Journal of Positive Behavior Interventions, 9(1),* 52-60.

Bellini, S. (2004). Social skill deficits and anxiety in high-functioning adolescents with autism spectrum disorder. *Focus on Autism and Other Developmental Disabilities, 19*, 78–86.

Ben-Itzchak, E., & Zachor, D. A. (2007). The effects of intellectual functioning and autism severity on outcome of early behavioral intervention for children with autism. *Research In Developmental Disabilities, 28,* 287-303.

Ben-Itzchak, E., Lahat, E., Burgin, R., & Zachor, A.D. (2008). Cognitive, behavior, and intervention outcome in young children with autism. *Research in Developmental Disabilities, 29,* 447-458.

Blacher, J., & Christensen, L. (2011). Sowing the seeds of the autism field: Leo Kanner (1943). *Intellectual and Developmental Disabilities, 49,* 172-191.

Birnbrauer, J. S., & Leach, D. J. (1993). The Murdoch Early Intervention Program after two years. *Behaviour Change, 10*, 63-74.

Center for Disease Control and Prevention. (2012). Prevalence of Autism Spectrum Disorders–Autism and Developmental Disabilities Monitoring Network, United States, 2008. *Morbidity and Mortal Weekly Report, 61(3),* 1-19.

Charman, T., Baron-Cohen, S., Sweeteneham, J., Baird, G., Drew, A., & Cox, A. (2003). Predicting language outcome in infants with autism and pervasive developmental

disorder. *International Journal of Language and Communication Disorders, 38,* 265-285.

Cohen, D.J., Paul, R., & Volkmar, F.R. (1987). Issues in the classification of pervasive developmental disorders and associated conditions. In D.J. Cohen, A.M. Donnellan, & R. Paul (Eds.). *Handbook of autism and pervasive developmental disorders* (pp. 20-40). New York: Wiley.

Cohen, I.L. (2003). Criterion-related validity of the PDD Behavior Inventory. *Journal of Autism and Developmental Disorders,* 33(1), 47-53.

Constantino, J. M., Gruber, C. & Christian, P. (2005). A Review of the Social Responsiveness Scale. *Mental Measurements Yearbook, 17.*

Cowan, R.J., & Allen, K.D. (2007). Using naturalistic procedures to enhance learning in individuals with autism: Focused on generalized teaching within the school setting. *Psychology in the Schools, 44(7),* 701-715.

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

Davy, A., & Savla, J. (2009). Estimating Statistical Power with Incomplete Data. *Organizational Research Methods, 12,* 320 – 346.

Dietz, C., Swinkels, S., Buitelaar, J., Van Daalen, E., & Van Engeland, H. (2007). Stability and change of IQ scores in preschool children diagnosed with autistic spectrum disorder. *European Child & Adolescent Psychiatry*, *16*(6), 405-410.

Delprato, D. (2001). Comparisons of discrete trial and normalized behavioral language intervention for young children with autism. *Journal of Autism and*

Developmental Disabilities, 31(3), 315- 325.

Durlack, J. A., & Dupree, E.P. (2008). Implementation matters: A review on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology, 41,* 327-350.

Eikeseth, S., Smith, T., Jahr, E., & Eldevik, S. (2002). Intensive behavioral treatment at school for children with autism: A 1-year Comparison controlled study. *Behavior Modification, 31,* 264-278.

Elliott, C. D. (1990). Differential Ability Scales- 2nd Edition. Pearson: San Antonio.

Elliott, R.O., Hall, K., & Soper, H.V. (1991). Analog language teaching versus natural language teaching: Generalization and retention of language learning for adults with autism and mental retardation. *Journal of Autism and Developmental Disorders, 21(4),* 433-447.

Ghezzi, P.M. (2007). Discrete trials teaching. *Psychology in the Schools, 44(7),* 667-679.

Gillberg, C. & Steffenberg, S. (1987). Outcome and prognostic factors in infantile autism and similar conditions: A population-based study of 46 cases followed through puberty. *Journal of Autism and Developmental Disorders, 17,* 272-287.

Goldstein, H. (2002). Communication intervention for children with autism: A review of treatment efficiency. *Journal of Autism and Developmental Disorders, 32,* 373-396.

Green, G. (1996). Early behavioral intervention for autism: What does the research tell us? In C. Maurice, G. Green, & S. Luce (Eds.), Behavioral intervention for young children with autism: A manual for parents and professionals (pp. 29-44). Austin,

TX: Pro-Ed.

Granpeesheh, D., Dixon, D.R., Tarbox, J., Kaplan, A.M., & Wilke, A.E. (2009). The effects of age and treatment intensity on behavioral intervention outcomes for children with autism spectrum disorders. *Research in Autism Spectrum Disorders, 3,* 1014-1022.

Granpeesheh, D., Kenzer, A., & Tarbox, A. (2011). Comparison of two-year outcomes for children with autism receiving high and low intensity behavioral intervention. *European Psychiatry, 26,* 1839.

Gresham, F.M. (2009). Evolution of the treatment integrity concept: Current status and future directions. *School Psychology Review, 38,* 533-540.

Gresham, F. M., & MacMillan, D. L. (1997). Autistic recovery? An analysis and critique of the empirical evidence on the Early Intervention Project. Behavioral Disorders, 22, 185-201.

Hagermoser Sanetti, L.M., & Kratochwill, T.R. (2009). Toward developing a science of treatment integrity: Introduction to the special series. *School Psychology Review, 38,* 445-459.

Harris, S. L., & Handleman, J. S. (2000). Age and IQ at intake as predictors of placement for young children with autism: A four-to-six year follow up. *Journal of Autism and Developmental Disorders, 30,* 137-142.

Harrison, P., & Oakland, T. (2003). Adaptive Behavior Assessment System, Second Edition. Psychological Corporation: New York.

Harper, C., Symon, J., & Frea, W. (2008). Recess is time-in: Using peers to improve social skills of children with autism. *Journal of Autism and Developmental*

Disorders, *38*(5), 815-826.

Hart, B., & Risley, T.R. (1975). Incidental teaching of language in the preschool. *Journal of Applied Behavior Analysis, 8,* 411-420.

Helt, M., Kelley, E., Kinsbourne, M., Pandey, J., Boorstein, H., Herbert, M., et al. (2008). Can children with autism recover? If so, how? *Neuropsychological Review, 18,* 339-366.

Howard, J.S., Sparkman, C.R., Cohen, H.G., Green, G., & Stanislaw, H. (2005). A comparison of intensive behavior analytic and eclectic treatments for young children with autism. *Research in Developmental Disabilities, 26,* 359-383.

Ingersoll, B., Schreibman, L., & Stahmer, A. (2001). Brief Report: Differential treatment outcomes for children with autistic spectrum disorder based on level of peer social avoidance.

Jensen, V. K., & Spannangel, S.C. (2011). The spectrum of autism spectrum disorder: A spectrum of needs, services, and challenges. *Journal of Contemporary Psychotherapy, 41,* 1-9.

Jones, R. (2009). The effects of depressed mood on academic outcomes in adolescents and young adults. *Dissertation Abstracts International Section A, 70*, 630.

Jones, E., Carr, E., & Feeley, K. (2006). Multiple Effects of Joint Attention Intervention for Children With Autism. *Behavior Modification*, *30*(6), 782-834.

Kaiser, A.P., Yoder, P.J., & Keetz, A. (1992). Evaluating milieu training. In S.F.Warren & J. Reichle (Eds.), *Causes and effects in communication and language intervention* (pp. 9-47). Baltimore: P.H. Brookes Publishing.

Kane, M., Connell, J.E., & Pellecchia, M. (2010). A quantitative analysis of language

interventions for children with autism. *The Behavior Analyst Today, (in press)*.

Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child, 2,* 217-250.

Kelley, E., Naigles, L., & Fein, D. (2010). An in-depth examination of optimal outcome children with a history of autism spectrum disorders. *Research in Autism Spectrum Disorders, 4,* 526-538.

Koegel, R.L., Camarata, S., Koegel, L.K., Ben-Tall, A., & Smith, A.E. (1998). Increasing speech intelligibility in children with autism. *Journal of Autism and Developmental Disorders, 28(3),* 241-251.

Koegel, L.K., Camarata, S.M., Valdez-Menchaca, M., & Koegel, R.L. (1998). Setting generalization of question-asking by children with autism. *American Journal on Mental Retardation, 102,* 346-357.

Koegel, L.K., Koegel, R.L., Harrower, J.K., & Carter, C.M. (1999). Pivotal response intervention I: Overview of approach. *Journal of the Association for Persons with Severe Handicaps, 24(3),* 174-185.

Koegel , R.L., Koegel, L.K., & Carter, C.M. (1998). Pivotal responses and the natural language paradigm. *Seminars in Speech and Language, 19(4),* 355-372.

Koegel , R.L., Koegel, L.K., & Carter, C.M (1999). Pivotal teaching interactions for children with autism. School Psychology Review, 28(4), 576-594.

Koegel, L., Koegel, R., Smith A. (1997). Variables related to differences in standardized test outcomes for children with autism. *Journal of Autism & Developmental Disorders*, *27*(3), 233-243.

Koegel, R.L., Koegel, L.K., & Surrat, A. (1992). Language intervention and disruptive behavior in preschool children with autism. *Journal of Autism and*

Developmental Disorders, 22(2), 141-153.

Koegel, R.L., O'Dell, M.C., & Koegel, L.K. (1987). A natural language paradigm for nonverbal autistic children. *Journal of Autism and Developmental Disorders, 17(2),* 187-200.

Koegel, R.L., Symon, J., & Koegel, L.K. (2002). Parent education for families of children with autism living in geographically distant areas. *Journal of Positive Behavior Interventions*, *4*(2), 88-103.

Kok, A.J., Kong, T.Y., & Bernard-Opitz, V. (2002). A comparison of the effects of structured play and facilitated play approaches on preschoolers with autism. *Autism, 6(2),* 181-196.

Lecavalier, L. (2006). Behavioral and emotional problems in young people with pervasive developmental disorders: Relative prevalence, effects of subject characteristics, and empirical classification. *Journal of Autism and Developmental Disorders, 36*, 1101–1114.

Little, R. J., & Rubin, D. B. (2002). Statistical analysis with missing data (2nd ed.). New York: John Wiley.

Lord, C., Rissi, S., DiLavore, P. S., Shulman, C., Thurm, A., & Pickles, A. (2006). Autism from 2 to 9 years of age. *Archives of General Psychiatry, 63,* 694-701.

Lord, C., Rutter, M., DiLavore, P.C., & Risi, S. (1999). Autism Diagnostic Observation Schedule Manual. Los Angeles, CA: Western Psychological Services.

Lovaas, O.I. (1987). Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal of Consulting and Clinical*

Psychology, 55(1), 3-9.

Lovaas, O.I., Berberich, J.P., Perloff, B.F., & Schaeffer, B. (1966). Acquisition of imitative speech in schizophrenic children. *Science, 151,* 705-707.

Lovaas, O.I., & Simmons, J.Q. (1969). Manipulation of self-destruction in three retarded children. *Journal of Applied Behavior Analysis, 2,* 143-157.

Luyster, R., Qiu, S., Lopez, K., & Lord, C. (2007). Predicting outcomes of children referred for autism using the MacArthur-Bates communication inventory. *Journal of Speech, Language, and Hearing Research, 50,* 667-681.

MacMullen, L., Manfredi, R., & Connell, J.E. (2012). *Cognitive Assessment in children with an Autism Spectrum Disorder: What are we measuring?* Manuscript submitted for publication.

Mandell, D., Stahmer, A., Shinn, S., & Marcus (2012). Results of a comparative effectiveness trial of two classroom-based intervention. Manuscript submitted for publication.

Mawhood, L., Howlin, P., & Rutter, M. (2000). Autism and developmental language disorder – A comparative follow-up in early adult life: Cognitive and language outcomes. *Journal of Child Psychology and Psychiatry, 41,* 547-559.

McEachin, J. J., Smith, T., Lovaas, O. I. (1993). Long-term outcome for children with autism who received early intensive behavioral treatment. *American Journal of Mental Retardation, 97(4), 359-372.*

McGee, G. G., Krantz, P. J., & McClannahan, L. E. (1985). The facilitative effects of incidental teaching on preposition use by autistic children. *Journal of Applied Behavior Analysis, 18,* 17-31.

Miranda-Linne, F., & Melin, L. (1992). Acquisition, generalization, and spontaneous use of color adjectives: A c comparison of incidental teaching and traditional discrete trial procedures for children with autism. *Research in Developmental Disabilities, 13,* 191-210.

National Autism Center. (2009). *The National Standards Report.* Retrieved July 7, 2010 from http://www.nationalautismcenter.org/pdf/NAC%20NSP%20Report_FIN.pdf.

National Institute of Child Health and Human Development. (2005). *Autism Overview.* Retrieved March 20, 2011 from http://www.nichd.nih.gov/publications/pubs/upload/introduction_autism.pdf.

National Research Council. (2001). *Educating Children with Autism.* Committee on Educational Interventions for Children with Autism. Catherine Lord and James P. McGee, eds. Division of Behavioral and Social Sciences and Education. Washington DC: National Academy Press.

O'Donnell, C.L. (2008). Defining, measuring, and conceptualizing fidelity of implementation and its relationship to outcomes in K-12 curriculum intervention

Pellecchia, M., Connell, J. E., Eisenhart, D., Kane, M., Schoener, C., Turkel, K., et al. (2011). We're all in the together now: Group performance feedback to increase classroom team data collection. *Journal of School Psychology, 49,* 411-431.

Pierce, K., & Schreibman, L. (1995). Increasing complex social behaviors in children with autism: Effects of peer-implemented pivotal response training. *Journal of Applied Behavior Analysis, 28,* 285-295.

Pierce, K., & Schreibman, L. (1997a). Using peer trainers to promote social behavior in

autism: Are they effective at enhancing multiple social modalities? *Focus on Autism and Other Developmental Disabilities*, *12*(4), 207-218.

Pierce, K. & Schreibman, L. (1997b). Multiple peer use of pivotal response training to increase social behaviors of classmates with autism: Results from trained and untrained peers. *Journal of Applied Behavior Analysis, 30,* 157-160.

Prizant, B.M., & Wetherby, A.M. (1998). Understanding the continuum of discrete-trial traditional behavioral to social-pragmatic developmental approaches in communication enhancement for young children with autism/PDD. *Seminars in Speech and Language, 19(4),* 329-352.

Remington, B., Hastings, R. P., Kovshoff, H., degli Espinosa, F., Jahr, E., Brown, T., et al., (2007). Early intensive behavioral intervention: Outcomes for children with autism and their parents after two years. *American Journal On Mental Retardation, 112,* 418-438.

Reynolds, C. R., & Kamphaus, R.W. (1992). *Behavior assessment system for children.* Circle Pines, MN: Academic Guidance Services.

Sallows, G.O., & Graupner, T.D. (2005). Intensive behavioral treatment for children with autism: Four-year outcome and predictors. *American Journal on Mental Retardation, 110,* 417-438.

Schopler, E., Short, A., & Mesibov, G. B. (1989). Relation of behavioral treatment to "normal functioning": Comment on Lovaas. Journal of Consulting and Clinical Psychology, 57, 162-164.

Schreibman, L. (2000). Intensive behavioral /psychoeducational treatments for autism: Research needs and future directions. *Journal of Autism and Developmental*

Disorders, 30(5), 373-378.

Schreibman, L., & Anderson, A. (2001). Focus on integration: The future of the behavioral treatment of autism. *Behavior Therapy, 32,* 619-632.

Sheinkopf, S., & Siegal, B. (1998). Home-based behavioral treatment for young autistic children. Journal of Autism and Developmental Disorders, 28, 15-23.

Sherer, M.R., & Schreibman, L. (2005). Individual behavioral profiles and predictors of treatment effectiveness for children with autism. *Journal of Consulting and Clinical Psychology, 73(3),* 525-538.

Sigman, M., & Ruskin, E. (1999). Continuity and change in the social competence of children with autism, down syndrome, and developmental delays. *Monographs of the Society for Research in Child Development, 64* (1, Serial No. 256).

Simonoff, E. M., Pickles, A. P., Charman, T. P., Chandler, S. P., Loucas, T., & Baird, G. F. (2008). Psychiatric disorders in children with autism spectrum disorders: Prevalence, comorbidity, and associated factors in a population-derived sample. *Journal of the American Academy of Child & Adolescent Psychiatry, 47*, 921–929.

Smith, T. (1996). Are other treatments effective? In C. Maurice, G. Green, & S. Luce (Eds.), Behavioral interventions for young children with autism: A manual for parents and professionals (pp.45-59). Austin, TX: Pro-Ed.

Smith, T. (2001). Discrete trial training in the treatment of autism. *Focus on Autism and Other Developmental Disabilities, 16(2),* 86-92.

Smith, T., Groen, A.D., & Wynn, J.W. (2000). Randomized trial of intensive early intervention for children with pervasive developmental disorder. *American*

Journal on Mental Retardation, 105(4), 269-285.

Sprafkin, J., Gadow, K. G., Salisbury, H., Schneider, J., & Loney, J. (2002). Further evidence of reliability and validity of the Child Symptom Inventory-4: Parent checklist in clinically referred boys. *Journal of Clinical Child and Adolescent Psychiatry, 31,* 513-524.

Stahmer, A.C. (1995). Teaching symbolic play skills to children with autism using pivotal response training. *Journal of Autism and Developmental Disorders, 25(2),* 123-141.

Strain, P.S., & Bovey II, E.H. (2011). Randomized controlled trial of the LEAP model of early intervention for young children with autism spectrum disorders. *Topics in Early Childhood Special Education, 31,* 133-154.

Strang, J. F., Kenworthy, L., Daniolos, P., Case, L., Wills, M. C., Martin, A., et al., (2012). Depression and anxiety symptoms in children and adolescents with autism spectrum disorders without intellectual disability. *Research In Autism Spectrum Disorders, 6*, 406- 412.

Sundberg, M.L., & Partington, J.W. (1998). Teaching language to children with autism and other disabilities. Pleasant Hill, CA: Behavior Analysts, Inc.

Sutera, S., Pandey, J., Esser, E.L., Rosentahl, M. A., Wilson, L.B., Barton, M., et al., (2007). Predictors of optimal outcome for toddlers diagnosed with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 37,* 98-107.

Szatmari, P., Bryson, S. E., Boyle, M.H., Streiner, D.L., & Duku E. (2003). Predictors of outcome among high functioning children with autism and Asperger syndrome.

Journal of Child Psychology and Psychiatry, 22, 520-528.

Thorp, D., Stahmer, A., & Schreibman, L. (1995). Effects of sociodramatic play training on children with autism. *Journal of Autism and Developmental Disorders*, *25*(3), 265-282.

Toth, K, Munson, J., Melzoff, A. N., & Dawson, G. (2006). Early predictors of communication development in young children with autism spectrum disorder: Joint attention, imitation, and toy play. *Journal of Autism and Developmental Disorders, 36,* 993-1005.

Tsai, L. (1992). Diagnostic issues in high functioning autism. In E. Schopler & G. Mesibov (Eds.), *High functioning individuals with autism* (pp. 11-40). New York: Plenum Press.

Turner, L. M. & Stone, W. L. (2007). Variability in outcome for children with an ASD diagnosis at age 2. *Journal of Child Psychology and Psychiatry, 48,* 793-802.

Vaughn, M. G., Wexler, J., Beaver, K. M., Perron, B. E., Roberts, G., & Fu, Q. (2011). Psychiatric correlates of behavioral indicators of school disengagement in the United States. *Psychiatric Quarterly, 82,* 191-206.

Vismara, L.A., & Lyons, G.L. (2007). Using perseverative interests to elicit joint attention behaviors in young children with autism: Theoretical and clinical implications for understanding motivation. *Journal of Positive Behavior Interventions, 9(4),* 214-228.

Volkmar, F., Cohen, D., Bergman, J., Hooks, M., & Stevenson, J. (1989). An examination of social typologies in autism. *Journal of the American Academy of Child and Adolescent Psychiatry, 28,* 82-86.

Waterhouse, L., Morris, R., Allen, D., Dunn, M., Fein, D., Feinstein, C, et al. (1996). Diagnosis and classification in autism. *Journal of Autism and Developmental Disorders, 26(3),* 59-86.

Wolery, M. (2011). Intervention research: The importance of fidelity measurement. *Topics in Early Childhood Special Education, 31,* 155-157.