SCHOOL CHARACTERISTICS AND THEIR RELATIONSHIP TO INTERVENTION FIDELITY AND STUDENT OUTCOMES IN AUTISM SUPPORT CLASSROOMS

A Dissertation
Submitted to
the Temple University Graduate Board

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

By
Jessica Dodge
August, 2014

Examining Committee Members:
Frank Farley, Advisory Chair, School Psychology
Catherine Fiorello, School Psychology
Joseph DuCette, Educational Psychology
Jerry Stahler, Geography and Urban Studies
David Mandell, External Member, University of Pennsylvania

ABSTRACT

Given the rising prevalence of children diagnosed with autism and the emergence of evidence-based autism interventions, schools are now faced with the challenge of delivering high quality instruction to this unique population. Comprehensive packaged curricula have been developed to address this growing need and to allow educators to transport research-based instruction into their classroom settings. However, there is a dearth of research examining the factors associated with intervention effectiveness with children with autism in public schools. The purpose of this study is to examine the associations between school level factors, the fidelity of interventions, and student outcomes within autism support classrooms. Data were gathered from 171 students with autism spectrum disorders in kindergarten-through-second grade classrooms across 40 schools in the Philadelphia School District. Correlational analyses and linear regression with random effects analyses indicated that school level factors were not associated with and were not moderators of intervention fidelity and student outcomes. The findings suggest that autism support classrooms are like islands within the school building, such that the practices and outcomes within these classrooms were unrelated to the school context. This study indicates that when transporting an evidence-based practice into a public school classroom, it may be more necessary to focus on the classroom context rather than the school building. Future research is needed to fully delineate these relationships between school building level factors and the dissemination and implementation of evidence-based practices for children with autism within a classroom setting.

TABLE OF CONTENTS

Page

LIST OF TABLES

CHAPTER 1

INTRODUCTION

The Centers for Disease Control and Prevention (CDC, 2014) estimates that 1 in 68 children have a diagnosis of autism, an increase of 30% over the last two years. Autism is a pervasive developmental disability that significantly impairs social communication and social interaction, and is accompanied by restrictive, repetitive patterns of behavior, interests, or activities (American Psychiatric Association, 2013). Autism affects all areas of functioning and without intervention can significantly limit an individual's ability to live a successful and independent life (CDC, 2014). With early intervention and instructional programs based on Applied Behavior Analysis (ABA), children with autism have demonstrated significant improvement in many areas of functioning (Foxx, 2008; Remington et al., 2007). A large body of literature published over the last three decades has refined evidence-based practices for children with autism (see for reviews: Fitzer & Sturmey, 2007; Iovannone, Dunlap, Huber & Kincaid, 2003; National Autism Center, 2009; National Research Council, 2001; Rogers & Vismara, 2008; Wong et al., 2013). Recently, Odom, Collet-Klingenberg, Rogers, and Hatton (2010) identified 24 evidence-based practices for children with autism, including discrete trial teaching (DTT), pivotal response training (PRT), and functional routines.

Discrete trial teaching (DTT) consists of specific instructional cues, prompting and error correction procedures, positive reinforcement contingencies, and targeted responses (Delprato, 2001; Ghezzi, 2007; Smith, 2001). DTT is teacher-directed and intensive and typically occurs at a small table in a one-on-one setting. Instructional cues are delivered by the teacher repeatedly and quickly to ensure skill acquisition and fluency in

responding. Correct responses are followed by contingent reinforcement through social praise, edibles, and tangible items. For example, a child receives a piece of a cookie for matching identical cups. Advantages to using DTT include its effectiveness when teaching basic learning skills (e.g., attending, imitating, following directions), the large literature base supporting its use with children with autism, and the controlled and orderly learning environment that supports more efficient skill acquisition (Ghezzi, 2007).

In contrast to DTT procedures, pivotal response training (PRT) involves a more naturalistic, less structured approach to teaching children with autism (Cowan & Allen, 2007). For example, to teach a child to identify the color red during a DTT session, a child may be presented with a red color card and is asked, “What is it?” When the child responds “red,” the child receives a piece of candy. In a PRT session, a child who is playing with toy cars may be presented with red, yellow, and blue cars and is taught to request the red car. In both teaching procedures, the child is learning to identify the color red, but in PRT the instruction occurs in a more natural setting during an activity that is selected by the child. During PRT, tasks change fluidly during teaching, direct reinforcers are utilized, and attempts are reinforced and shaped (Koegel, O’Dell, & Koegel, 1987). Advantages to using PRT include increased motivation, emphasis on the student’s interests, and greater alignment with the naturally occurring teaching environment (Sundberg & Partington, 1998).

Another commonly used teaching procedure for children with autism is functional routines. Functional routines are predictable and consistent events that occur within the natural environment. A functional routine is a chain of behaviors that is broken down into discrete units. One example of a functional routine is the morning routine. The morning

routine includes the individual tasks of getting dressed, eating breakfast, and brushing teeth. Each of these steps is functional, as they each are necessary in successfully getting ready in the morning. Functional routines are integral in teaching children with autism within a classroom setting and subsequently have been included within a variety of comprehensive autism intervention models (Arick et al., 2003; Dawson et al., 2010; Landa, Holman, O'Neill, & Stuart, 2011).

DTT, PRT, and functional routines each offer advantages as instructional practices for children with autism. Even though there are distinct differences in how a skill is targeted across these procedures, when used together they can offer a wide-ranging program for teaching children with autism. For example, a child may acquire matching skills in a contrived and tightly controlled DTT session, learn to generalize these skills in the natural environment during a PRT session, and then apply these skills during a transition routine when the child is required to match a schedule picture card to its corresponding location. As a result, comprehensive treatment packages for children with autism have been developed that include both structured (i.e., DTT) and naturalistic (i.e., PRT and FR) instructional procedures (Arick et al., 2003; Dawson et al., 2010; Landa, Holman, O'Neill, & Stuart, 2011).

Given the large body of literature supporting a variety of evidence-based practices for children with autism and the development of comprehensive treatment packages, schools are now faced with the task of providing evidence-based instruction to this growing population of students. However, the implementation of these evidence-based practices for students with autism in public schools often is complicated by school-based barriers including lack of administrative support, scarce resources, dearth of trainings in

evidence-based practices, and poor collaboration among professionals and families (Tincani, 2007). As a result, evidence-based practices for students with autism have been rarely or inconsistently implemented in schools (Hess, Morrier, Heflin, & Ivey, 2008; Morrier, Hess, & Heflin, 2011). Given the apparent discrepancy between what research suggests and what currently occurs in classrooms, it is important to study the factors that may be related to the use, delivery, and effectiveness of evidence-based practice for students with autism in public school settings.

There have been a limited number of studies examining the factors that predict intervention effectiveness with children with autism. Current studies have focused primarily on child characteristics and teacher characteristics as predictors of child outcomes. Research has shown that children with higher scores on cognitive assessments, who are younger, and who have greater language abilities have better outcomes following intensive interventions (Baker-Ericzen, Stahmer, & Burns, 2007; Ben-Itzchak, Lahat, Burgin, & Zachor, 2008; Ben-Itzchak & Zachor, 2007; Helt et al., 2008; Luyster, Qiu, Lopez, & Lord, 2007; Sherer & Schreibman, 2005; Turner & Stone, 2007). Teacher quality and attitudes have also been shown as predictors of intervention effectiveness with children with autism (Ghaith & Yaghi, 1997; Johnson & Hastings, 2002; Peters-Scheffer, Didden, Korzilius, & Sturmey, 2013).

Another factor that has been implicated in the effectiveness of evidence-based practices is intervention fidelity. Fidelity is defined as how well an intervention is delivered in relation to its intended design (Dane & Schneider, 1998; Gresham, Gansle, Noell, & Cohen, 1993; O'Donnell, 2008; Sanetti, Gritter, & Dobey, 2011; Schulte, Easton, & Parker, 2009). Two components of fidelity are adherence (i.e., number of steps

followed in an intervention) and dosage (i.e., time spent implementing the intervention) (Sanetti, Gritter, & Dobey, 2011; Schulte, Easton, & Parker, 2009). While there have been few studies linking fidelity and child outcomes, research has generally shown that greater fidelity predicts better child outcomes (Benner, Nelson, Stage, & Ralston, 2010; O'Donnell, 2008). Few studies have examined the association between intervention fidelity and outcomes for children with autism (Granpeesheh, Dixon, Tarbox, Kaplan, & Wilke, 2009; Pellecchia, 2013; Strain & Bovey, 2011).

In addition to teacher and child characteristics, factors within schools may be related to student outcomes. While little research has examined the association between contextual factors and autism interventions, education researchers have examined the associations between a school's disciplinary practices, academic performance, and student composition and student outcomes (Hogrebe & Tate, 2010; Klecker & Pollock, 2005; Perry & McConney, 2013). One way to capture a school's disciplinary practice is through the total number of suspensions reported per year. Suspensions have been used to assess the effectiveness of positive behavior support programs (Howard, Mann, & Lebrun, 2008) and have been associated with lower academic achievement (Hogrebe & Tate, 2010). Scores on state standardized reading and mathematics tests can be a proxy for a school's overall academic performance. Limited research has been conducted regarding state standardized test scores; however, one study suggested that state standardized test scores are associated with the type of instructional practices used in classrooms (Klecker & Pollock, 2005). More frequently, state standardized test scores are used in studies as an outcome measure to evaluate academic performance (Hogrebe & Tate, 2010; Southworth, 2010). Another commonly studied school factor is student

composition (i.e., percentage of students from low-income families). Generally, in schools with higher percentages of students from low-income families, student outcomes are poorer (Perry & McConney, 2013). No research exists, however, on the association of these school factors with instructional programming or student outcomes in classrooms for children with autism.

Little attention has been paid to the school factors that may predict outcomes for students with autism. This issue is important because barriers to implementing evidence-based practices for students with autism may be at the building level (Tincani, 2007). To bridge the research-to-practice gap, researchers have suggested a shift from studying the efficacy of autism interventions to first examining the implementation of autism interventions within their community and school setting (Dingfelder & Mandell, 2011; Kasari & Smith, 2013). This study offers a first look at the school-level factors that may be associated with the implementation and effectiveness of autism interventions in schools.

Two separate datasets were combined in this research study to examine the association between interventions in autism support classrooms and school characteristics. One of the datasets was gathered from the Philadelphia Autism Instructional Methods Study (Philly AIMS), a recently completed research project funded by the National Institute of Mental Health and the Institute of Education Sciences. Philly AIMS examined the implementation and sustainability of evidence-based practices for students with autism in a large urban school district by collecting data on teacher, parent, and student characteristics as well as student outcomes and intervention fidelity. As part of the Philly AIMS study, kindergarten-through-second grade autism support teachers

received training and coaching on the implementation of Strategies for Teaching Based on Autism Research (STAR). STAR is a manualized program that incorporates three evidence-based instructional strategies for children with autism: discrete trial teaching, pivotal response training, and functional routines. Preliminary research on the effectiveness of STAR in public school settings revealed student gains in social interaction, expressive speech, and cognitive ability (Arick et al., 2004; Mandell, Stahmer, Shin, Xie, Reisinger, & Marcus, 2013).

The second dataset was compiled from publicly available data on school characteristics provided by the Philadelphia School District. Data from the participating schools in Philly AIMS were collected on total number of suspensions per year, percentage of students receiving free or reduced price lunch, and state standardized test scores. These data were then combined with the Philly AIMS data to allow comparison between school characteristics, intervention effectiveness, and student outcomes in the autism support classrooms.

In light of current gaps in the autism literature and the call for more research focusing on the implementation of evidence-based practices, the purpose of this study is to examine the relationships between school level factors, the fidelity of interventions, and student outcomes within autism support classrooms using the Philly AIMS data set and publicly reported school data. This study will address three primary research questions:

1) Do school factors such as academic performance, disciplinary practices, and school composition, predict teacher intervention fidelity to evidence-based practices for students with autism?

2) Do school factors such as academic performance, disciplinary practices, and school composition predict gains on a cognitive assessment for students with autism?
3) Do school factors such as academic performance, disciplinary practices, and school composition moderate the association between intervention fidelity and student outcomes, such that schools that are high performing (i.e., strong academics and fewer behavioral problems) show a stronger association between intervention fidelity and student outcomes than low-performing schools?

CHAPTER 2

LITERATURE REVIEW

Autism spectrum disorder (ASD) is a pervasive developmental disability characterized by qualitative impairments in social interaction, communication skills, and restricted, repetitive and stereotyped patterns of behavior (American Psychiatric Association, 2013). It is estimated that 1 in 68 children have a diagnosis of autism (CDC, 2014). While the specific cause of autism is unknown, research has suggested that brain mechanisms and genetic factors likely contribute to the occurrence of autism (Volkmar & Pauls, 2003). Presently, there is no cure for autism and as such research has focused on developing treatments to improve the outcomes of individuals diagnosed with autism.

Autism Interventions

Across the past three decades, evidence-based practices for children with autism have been developed and refined (see for reviews: Fitzer & Sturmey, 2007; Iovannone, Dunlap, Huber & Kincaid, 2003; National Autism Center, 2009; National Research Council, 2001; Rogers & Vismara, 2008; Wong et al., 2013). Recently, Odom, Collet-Klingenberg, Rogers, and Hatton (2010) identified 24 evidence-based practices for children with autism including discrete trial teaching (DTT), pivotal response training (PRT), and functional routines (i.e., task analysis and visual supports). As a result, researchers have developed comprehensive treatment packages that incorporate these three practices (Arick et al., 2003; Dawson et al., 2010; Landa, Holman, O'Neill, & Stuart, 2011).

Strategies for Teaching based on Autism Research (STAR)

One treatment package for children with autism that has been developed for use in schools is the STAR program. The STAR program is a manualized curriculum based on the principles of Applied Behavior Analysis (ABA) that incorporates DTT, PRT, and functional routines. Each of these instructional methods will be described in detail in the following sections. The STAR program contains three kits, a curriculum based assessment, lesson plan folders, datasheets, and program materials. The 169 lesson plans cover six instructional domains: receptive language, expressive language and spontaneous language, functional routines, pre-academic skills, and play, and social concepts. There have been two published studies that have examined the effectiveness of the STAR program (Arick et al., 2003; Mandell et al., 2013). Arick et al. (2003) conducted a study that provided training, coaching, and consultation to preschool teachers on the implementation of the STAR program. Following 16 months of STAR implementation, the authors reported student gains in social interaction, expressive language, and adaptive language concepts.

In a recently completed study funded by the National Institute of Mental Health and the Institute of Education Sciences, the Philadelphia Autism Instructional Methods Study (Philly AIMS), examined the effects of implementing the STAR program in a large urban school district. Philly AIMS was a three year-long study conducted at the University of Pennsylvania. In its first year, the study examined the effects of two programs for students with autism in a large randomized field trial (Mandell et al., 2013). Thirty-three kindergarten-through-second grade autism support classrooms and 119 students participated in the study. The two programs compared were STAR and Structured

Teaching, a program based on the use of visual supports, direct reinforcement, classroom organization, and the development of consistent routines (Mesibov et al., 2004).

Overall, both the STAR and Structured Teaching classroom teachers showed moderate fidelity following trainings and ongoing coaching, 48% and 57%, respectively. Mandell et al. suggested that even given the intensity of trainings, high fidelity may not have been achieved given the complex nature of these programs, the under-resourced and understaffed classrooms, and possibly the adaptations teachers made in programming based on student needs. Nonetheless, Mandell et al. found that greater IQ gains were associated with students in the STAR classrooms that demonstrated higher intervention fidelity than in high performing Structured Teaching classrooms, suggesting that teachers who implemented the STAR curriculum as it was designed had better student outcomes.

However, in classrooms with moderate levels of fidelity, students in the Structured Teaching group demonstrated larger gains in IQ. While this finding was surprising, Mandell et al. posited that one possible explanation was that teachers who had moderate fidelity to STAR were struggling to manage and implement the program and therefore had difficulty delivering effective instruction to the students. However, in the Structured Teaching classrooms, moderate fidelity was the result of partial implementation of the program but teachers were still able to provide other forms of effective instruction to their students. Following the study, the Philadelphia School District elected to adopt STAR across all of the kindergarten-through-second grade autism support classrooms. Therefore in subsequent years, Philly AIMS examined more closely the implementation and sustainability of the STAR program within these autism support classrooms. The present study seeks to extend this initial research on the implementation of STAR in the

Philadelphia School District by examining the possible moderating effects of school level factors on the association between intervention fidelity and student outcomes.

First, a brief review of the three instructional components of STAR (DTT, PRT, and functional routines) will be provided below. The research supporting the effectiveness and efficacy of DTT, PRT, and functional routines will be discussed. Second, the factors related to intervention effectiveness will be examined including intervention fidelity, student characteristics, contextual factors, and school factors. A more extensive review of three school factors: disciplinary practices, academic performance, and school composition, will also be included. Lastly, the current aims of this study will be reviewed.

Discrete Trial Teaching (DTT)

One methodology utilized in the STAR program is discrete trial teaching (DTT). DTT consists of specific instructional cues, prompting and error correction procedures, positive reinforcement contingencies, and targeted responses (Delprato, 2001; Ghezzi, 2007; Smith, 2001). DTT can be used with children across ages, from early childhood to adolescence (Smith, 2001). DTT is teacher-directed and intensive. Professionals have recommended between 20-40 hours per week of DTT with children with autism (Smith, 2001). Typically DTT sessions occur at a small table in a one-on-one setting. Distractions are minimized by using room dividers and by keeping the area clean and orderly. DTT is a highly effective teaching procedure that involves breaking down complex skills into small teachable tasks (Ghezzi, 2007). For example, to teach a generalized matching-to-sample repertoire, the instructor first teaches matching identical objects, then matching identical pictures, and then matching pictures to objects and objects to pictures.

Instructional cues are delivered by the teacher repeatedly and quickly to ensure skill acquisition and fluency in responding. Correct responses are followed by contingent reinforcement through social praise, edibles, and tangible items. Given the fast-paced and rigid procedures utilized in DTT, professionals require significant training and practice to become fluent in the delivery of DTT.

Across the past two decades, studies have consistently demonstrated the effectiveness of using DTT to teach children with autism a variety of skills (Eikeseth, Smith, Jahr, & Eldevik, 2007; Lovaas, 1987; McEachin, Smith, & Lovaas, 1993; Smith, Groen, & Wynn, 2000; Smith, 1999; Young, Krantz, McClannahan, & Poulson, 1994). Rogers and Vismara (2008) reported that DTT meets the criteria for a "well-established" treatment for children with autism. Additionally, in a meta-analysis of DTT research, Eldevik et al. (2009) reported that across 34 studies, changes in full-scale intelligence and adaptive behavior composite scores had effect sizes of 1.10 and .66, respectively.

Lovaas (1987) conducted what is considered the seminal study of DTT. In his multi-year study, 19 children with autism (mean age=34.6 months) who received intensive 40 hours per week DTT instruction were compared with a control group (n=19, mean age=40.9 months) who received less than 10 hours of one-on-one treatment per week. As the children gained skills in the DTT group, intensive one-on-one instruction was decreased and education was provided in a more natural setting with typically developing peers. Prior to treatment, Lovaas assessed the children using a variety of cognitive and behavioral scales and found that the groups were similar across functioning levels. However, by the end of first grade, 47% of the children in the DTT group, compared with 2% of children in the control group, had IQ scores within the normal range and were

receiving education in a mainstream classroom. Conversely, 2% of children in the DTT group compared to 53% of children in the control group had IQ scores in the "severely retarded" range. Children in the experimental group were reported as having gained more than 30 IQ points over the control group students. In fact, a follow-up study was conducted six years later and researchers found that the original 19 participants in the DTT group maintained their gains in cognitive ability, behavioral skills, and educational placement (McEachin Smith, & Lovaas, 1993).

While critics of this study suggested that the use of different assessments batteries pre- and post-treatment render the comparison less valid, the Lovaas study had a large impact on the field of autism and sparked subsequent research of DTT effectiveness. Sallows and Graupner (2005) implemented a similar 38-hour per week DTT program for 24 children with autism. Following four years of DTT treatment, 48% of children were able to attend regular education classrooms and showed large gains in cognitive and language skills. Similarly, Smith, Groen, and Wynn (2000) conducted a randomized control trial study that examined the effects of DTT with seven children with autism and eight children with Pervasive Developmental Disorder, Not Otherwise Specified (PDD-NOS). They included a control group of seven children with autism and six children with PDD-NOS who received three-to-nine months of parent training. Prior to treatment the groups were similar across cognitive, language, and behavior measures. However, following two-to-three years of 30 hours of intensive treatment (primarily DTT) per week, the experimental group showed greater gains in intelligence, visual spatial skills, language, and academics. Adaptive and behavioral problems were not different across groups. The results were less substantial than the original Lovaas (1987) study with

differences in IQ of 16 points compared to 31 points, and percentage of children in regular education classes at 27% compared to 47%. The difference in findings could be accounted for by the lower dosage of DTT provided (30 hours compared with 40 hours per week of instruction).

Eikeseth, Smith, Jahr, and Eldevik (2002) in a quasi-experimental design compared one year of DTT with an intensive eclectic intervention approach (e.g., TEACCH, sensory motor therapies, ABA therapies, one-on-one teaching with therapists as determined by a multidisciplinary team). They found that children in the DTT group displayed gains of 17 points in IQ and showed significant progress in language comprehension, expressive language, and adaptive behavior. In comparison, the eclectic group showed only a four-point gain in IQ, regression in language comprehension and expressive language, and no change in adaptive behavior. In a follow up study, Eikeseth, Smith, Jahr, and Eldevik (2007) reported that after three years of treatment, the children in the DTT group continued to show greater gains than the eclectic group. Similarly, Howard, Sparkman, Cohen, Green, and Stanislaw (2005) found that intensive behavioral therapy (i.e., DTT) was more effective than eclectic interventions.

Adaptations to Lovaas's (1987) original study have also been conducted. Research has shown that 15-20 hours per week of DTT resulted in smaller, but comparable gains in language, IQ, and behavior (Anderson, DiPietro, Edwards, & Christian, 1987; Birnbrauer & Leach, 1993). Cohen, Amerine-Dickens, and Smith (2006) further extended the DTT literature by implementing the program in a community-based agency setting rather than in a clinic or home setting. Similar to previous studies the authors reported that children in the DTT group who received 35-40 hours of intensive therapy had higher IQs and

superior adaptive behavior than the control group who received only public special education services. By the third year of instruction, six of the 21 children who received DTT were included in regular education. Eleven of the remaining 15 children were included in regular education with some support still provided. On the other hand, only one child from the control group was placed in a regular education setting. While critics of DTT have often cited the intensity and upfront cost of the program as a significant barrier to implementation, researchers have suggested that the long-term benefits from a cost perspective reveal a significant savings (Jacobson, Mulick, & Green, 1998). The authors suggested that up to one million dollars per child is saved if the initial investment in intensive therapy is made.

Researchers have also suggested that DTT used in isolation may not meet the multitude of needs displayed by children with autism (Smith, 2001; Steege, Mace, Perry, & Longenecker, 2007). In fact, of the studies reviewed, as children progressed in DTT, instruction was systematically faded to a more natural environment setting. Steege et al. (2007) noted some of the disadvantages of DTT including the DTT setting may not mirror what occurs in the natural environment (e.g., DTT is teacher directed, reinforcers during DTT are unique, interaction between speaker and listener is different), DTT may result in rote responding, and DTT instruction may create an escape-maintained learning environment (i.e., the student may engage in problem behavior to avoid DTT). The authors suggested that a more comprehensive approach to programming based on Applied Behavior Analysis (ABA) should be utilized by also including naturalistic teaching methods.

Pivotal Response Training (PRT)

One commonly used naturalistic teaching method is Pivotal Response Training (PRT). PRT, another component of the STAR program, involves a less structured approach to teaching children with autism (Cowan & Allen, 2007). Koegel, Koegel, and Carter (1999) described the importance of identifying and prioritizing pivotal behaviors in teaching children with autism. Pivotal behaviors are repertoires that when acquired, affect many areas of functioning. There are four primary pivotal areas: motivation, responsiveness to multiple cues (e.g., being able to find a blue crayon among a field of other colors of crayons), self-initiated learning (e.g., asking for information), and self-management. PRT generally focuses on teaching these four pivotal areas to children with autism to increase a wide range of skills. Koegel, Koegel, and Carter (1998) described the core components of PRT as utilizing child choice, delivering natural reinforcers, reinforcing communicative attempts, and interspersing maintenance tasks. During PRT sessions, tasks change fluidly during teaching, direct reinforcers are utilized, and attempts are reinforced and shaped (Koegel, O'Dell, & Koegel, 1987). Advantages to using PRT include increased motivation, less frustration, emphasis on student's interests, increased engagement, less need for aversive control, embedded generalization procedures, and alignment with the naturally occurring teaching environment (Sundberg & Partington, 1998).

PRT has strong empirical support in improving social and language outcomes for children with autism (Koegel, Carter, & Koegel, 2003; Rogers, 2000; Stahmer, Suhrheinrich, Reed, Bolduc, & Schreibman, 2011). Baker-Ericzen, Stahmer and Burns (2007) delivered a 12-week parent education program on PRT with 158 families in an

outpatient clinic in California. Following the PRT program, all parents reported that their children showed improvement in adaptive behavior. Similarly, Lydon, Healy, and Leader (2011) conducted a study with five children with autism ages three to six years. They compared video modeling and PRT in teaching play and language skills. The authors found that play actions increased significantly more in the PRT group than in the video-modeling group and these play actions generalized to novel environments.

Stahmer (1999) reviewed the literature on using PRT to teach play skills. She described the components of PRT to teach play as clear instructions and questions, interspersal of maintenance tasks, child choice, direct reinforcement, reinforcement of goal-directed attempts, and turn-taking. She reported that across studies, PRT improved play skills including socio-dramatic play and symbolic play. She also found that PRT was an effective strategy for parents to implement as well. Finally, she reiterated the importance of targeting generalization of play skills during PRT to ensure that children engaged in play activities with novel peers (Stahmer, 1995).

In a recent review by Stahmer, Suhrheinrich, Reed, Bolduc, and Schreibman (2011), the authors reported that PRT has been successfully applied to a variety of behaviors including joint attention, communication, play skills, peer interactions, and homework skills. The authors also suggested that PRT can be applied to classroom settings, but presented a few barriers to the implementation of PRT. For example, they indicated that teachers may not have the prerequisite knowledge of applied behavior analysis (ABA) necessary to implement PRT, teachers may have difficulty with data collection, may struggle to incorporate PRT in an Individualized Education Plan (IEP), may disagree with

using child selected activities, and may have difficulty training paraprofessional staff on PRT.

To better understand how PRT can be applied to the school environment, Suhrheinrich, Stahmer, Reed, Schreibman, Reisinger, and Mandell (2013) conducted a study comparing PRT implementation with 41 research trained and clinically trained teachers. Overall, the authors found that the teachers frequently demonstrated clear opportunities/instruction and utilized child choice but had difficulty with including turn-taking and teaching responsitivity to multiple cues. Overall, higher fidelity to PRT was found in the research trained versus clinically trained group. The authors posited that differences in training and coaching likely accounted for this discrepancy. More specifically, the clinically trained teachers also received training in two other instructional procedures, as part of a packaged curriculum, whereas, the research trained teachers received only PRT training. This suggests that teachers who are required to learn a greater number of procedures may have greater difficulty with fidelity and the implementation of the program.

PRT has also been demonstrated as an effective peer implemented program for children with autism. Kuhn, Bodkin, Devlin, and Doggett (2008) found that a peer implemented PRT program increased social interactions in two children with autism. Similarly, Pierce and Schreibman (1995) conducted a study with two 10-year-old children with autism who received peer implemented PRT sessions within a classroom setting. Social behavior (e.g., conversations, play with peer, initiations, joint attention) and language skills increased following the intervention. Harper, Symon, and Frea (2008) also found that a peer implemented PRT sessions resulted in an increase in social skills

and initiations during recess. Overall, research has indicated that PRT can be effectively implemented by peers, parents, and teachers to increase social skills, language, and behavior in children with autism.

Functional Routines (FR)

The third component of STAR is functional routines (FR). Functional routines are predictable events that occur within the natural environment. A functional routine is a chain of behaviors that is broken down into discrete units. One example of a functional routine is the morning routine, which includes the individual tasks of getting dressed, eating breakfast, and brushing teeth. Each of these steps is functional, as they each are necessary in successfully getting ready in the morning. Functional routines are integral in teaching children with autism within a classroom setting and subsequently have been included within a variety of comprehensive autism intervention models (Arick et al., 2003; Dawson et al., 2010; Landa, Holman, O'Neill, & Stuart, 2011).

One of the primary methods for teaching functional routines is through visual schedules. Visual schedules consist of written words or pictures representing each activity or task for the child to follow. Schedules help structure an activity for children with autism, as well as, ensure predictability and consistency (Dettmer, Simpson, Myles, & Ganz, 2000; Massey & Wheeler, 2000). Banda and Grimmett (2008) reviewed 13 studies using visual schedules with children with autism. Across all 13 studies, improvements were noted for every child with autism. More specifically, gains were reported in the areas of social interaction, transitions and on-task behavior. Disruptive behavior also decreased. The authors concluded that possible reasons for improvements following the use of visual schedules were the predictability and systematic display of

complex behaviors into their component parts and that visual schedules acted as prompts for the children to follow. The authors described the benefits of using activity schedules as useful for teaching a variety of routines, can be applied across settings, and decreases problem behavior during transitions.

Morrison, Benchaaban, and Endo (2002) taught children with autism to use an activity schedule to increase their play skills. After teaching the children to follow a play schedule (i.e., pictures of play activities) the children were able to independently follow the schedule and subsequently spent greater time on-task, engaged in appropriate playing. Dettmer, Simpson, Myles, and Ganz (2000) also conducted a study using visual activity schedules for two children with autism. The authors used pictures representing each task in a day to help improve transitions between activities. In addition to the visual schedule, a timer was also used to show the children when each activity ended. In this study, Dettmer et al. reported that transition time decreased and independent transitions increased.

On the other hand, when functional routines are not established, challenges may frequently arise. Marquenie, Rodger, Mangohig, and Cronin (2011) interviewed mothers of children with autism and found that home routines were often chaotic and disorganized. In the absence of functional routines, there were fewer interactions between parent and child. The authors suggested that family cohesion suffered and parental stress subsequently heightened. Overall, functional routines are an important instructional strategy for children with autism that can help increase independence skills and decrease problem behavior.

Factors Related to Intervention Effectiveness

Fidelity

Intervention or treatment fidelity is defined as how well an intervention is delivered in relation to its intended design (Dane & Schneider, 1998; Gresham, Gansle, Noell, & Cohen, 1993; O'Donnell, 2008; Sanetti, Gritter, & Dobey, 2011; Schulte, Easton, & Parker, 2009). Fidelity has been divided into five categories: adherence, duration (dosage), quality, participant responsiveness (acceptability), and program differentiation (which parts of an intervention are essential) (Dane & Schneider, 1998). While studies in the education literature have generally focused on adherence and dosage (Benner, Nelson, Stage, & Ralston, 2010; Swanson, Wanzek, Haring, Ciullo, & McCully, 2011), there has been an overall lack of fidelity measures included in education research. Swanson, Wanzek, Haring, Ciullo, and McCully (2011) reviewed the education and special education literature and found that only 67% of studies reported fidelity measures, 47% provided quantitative measures of fidelity, and 88% used observation to code fidelity. Similarly, across research in psychology and mental health there is considerable agreement that fidelity measures are reported far too infrequently (Domitrovich & Greenberg, 2000; O'Donnell, 2008; Sanetti, Gritter, & Dobey 2011).

Wolery (2011) identified four reasons why it is important to assess treatment fidelity: it ensures that the results of the intervention were not due to a lack of fidelity, it allows researchers to determine how "transportable" the intervention is to applied settings, it offers necessary information for replication studies, and it provides a picture of the child's experience in the intervention. In addition to the methodological benefits of assessing intervention fidelity, researchers have shown that higher intervention fidelity is

generally associated with better outcomes (Benner, Nelson, Stage, & Ralston, 2010; Durlak & Dupre, 2008). For example, Durlak and DuPre (2008) reviewed five meta-analyses, which included 500 studies total. They found that the effect sizes were two to three times higher if interventions were implemented with fidelity. Furthermore, higher intervention fidelity was associated with better outcomes especially when adherence and dosage were examined, suggesting the importance of measuring these variables when implementing programs. In a recent study, Benner, Nelson, Stage, and Ralston (2010) examined the association between fidelity (including adherence and dosage) to a reading program and student outcomes. Consistent with Durlak and Dupre's review, they found a statistically significant association such that higher fidelity was associated with better student outcomes. Given the evident link between fidelity and outcomes, the authors concluded that teachers need ongoing support and coaching to ensure that interventions continue to be implemented as they were designed.

Studies have also examined the effects of low intervention fidelity and the factors that may be associated with implementation. Dusenbery, Brannigan, Falco, and Hansen (2003) reviewed the drug abuse prevention literature and found that poor implementation was associated with reduced program effectiveness. They identified several factors that predicted fidelity: amount of teacher training, program complexity, teacher attitudes, and organizational factors (support, school culture). Barriers that were identified included lack of time, money, and resources.

Few studies have assessed fidelity in autism intervention delivery. Studies that have included measures of fidelity have generally examined the factors associated with fidelity and the association between fidelity and outcomes. Peters-Scheffer, Didden, Korzilius,

and Sturmey (2013) conducted a study with 22 teachers implementing DTT in a preschool for children with intellectual disabilities. The authors found that teachers who rated their relationship with the child as poorer had lower fidelity. Conversely, teachers who reported having a positive attitude about disability had higher fidelity. As a result, the authors suggested that personality traits may be related to DTT fidelity. Granpeesheh, Dixon, Tarbox, Kaplan, and Wilke (2009) found that age and treatment intensity (dosage) of behavioral interventions was also a strong predictor of outcome (number of programs mastered).

Some studies have examined PRT fidelity. Coolican, Smith, and Bryson (2010) conducted a brief PRT training for eight parents of young children with autism using a non-concurrent multiple baseline design across parents. Three, two-hour trainings were conducted for the parents in the PRT program. At the two to four month follow up, child gains in communication skills were found. Following the training, the average fidelity increased from 27% to 62%. In addition, the authors found that fidelity was associated with student outcomes. Similarly, Minjarez, Williams, Mercier, and Hardan (2011) examined the effectiveness of a group parent training in PRT and associated child gains. Seventeen parents participated in the ten-week PRT training program with weekly 90-minute group sessions and one 50-minute individual session. Results showed an increase in both parent PRT fidelity and frequency of child communication attempts.

A few studies have attempted to increase PRT fidelity by changing staff training procedures. Robinson (2011) found that fidelity improved greatly following video feedback and generalized to new students and activities. An average of 48 minutes of feedback was required before the paraprofessional staff members were able to meet

fidelity criteria. Suhrheinrich (2011) trained 20 teachers on PRT using a six-hour didactic workshop. Following the training, only 15% of the teachers reached mastery criteria for PRT. However, with in-vivo coaching (average of seven hours), the majority of teachers were able to meet mastery criteria.

There is a scarcity of peer-reviewed studies that have focused on the association between fidelity and outcomes with children with autism. In one study, Strain and Bovey (2011) found that teacher fidelity predicted student outcomes in the implementation of LEAP, an inclusive preschool program for children with autism. Training and ongoing coaching resulted in fidelity of 90% after two years. However, in the training alone condition, fidelity was only 38%, indicating the need for continued support to sustain interventions.

In a recent dissertation study conducted using the same dataset as the present study, Pellecchia (2013) evaluated the association between intervention fidelity (dosage and adherence) and student outcomes. Pellecchia found that only PRT fidelity, measured as a function of dosage and adherence, predicted student outcomes. DTT and FR, however, did not predict student outcomes. This finding was surprising given the large literature base supporting DTT and FR in improving child outcomes. Pellecchia suggested that differences in how fidelity was assessed in the study may have accounted for the lack of significant associations in her findings. In addition, teachers who implemented PRT with fidelity were also more likely to implement DTT with fidelity and have strong functional routines, whereas teachers who implemented FR or DTT with fidelity were not necessarily following the other instructional programs, further explaining why PRT was the only intervention associated with student outcome. Given the importance of assessing

fidelity when implementing interventions, there is a clear need for more research delineating the associations among intervention fidelity, training procedures, and student outcomes.

Teacher and Student Characteristics

Studies have also focused on teacher and student characteristics as predictors of student outcomes and intervention effectiveness. Preliminary investigations have found that teacher quality facilitates successful implementation of DTT in a home setting (Johnson & Hastings, 2002). In a teacher training on cooperative learning, Ghaith and Yaghi (1997) concluded that if the intervention is congruent with current practices, teachers are more likely to implement the program and also report that it will be easier to implement. Similarly other studies have found that the teacher's personality traits (e.g., attitude toward disability, perceived relationship with the child, and self-efficacy) were associated with their ability and willingness to implement an intervention (Peters-Scheffer, Didden, Korzilius, & Sturmey, 2013; Wenz-Gross & Upshur, 2012).

Studies evaluating the association between student characteristics and outcomes have generally found that younger children with higher cognitive ability, more expressive language skills, and fewer nonverbal self-stimulatory behaviors (e.g., hand flapping) respond more favorably to intervention (Baker-Ericzen, Stahmer & Burns, 2007; Ben-Itzchak & Zachor, 2007; Luyster, Qiu, Lopez, & Lord, 2007; Sherer & Schreibman, 2005). Helt et al. (2008) also reviewed the autism literature and concluded that predictors of "recovery" included cognitive ability, receptive language, verbal and motor imitation, and motor development, where children who had more skills at intake showed greater

gains following intervention. Age of diagnosis and a diagnosis of PDD-NOS versus autism spectrum disorder were also indicators of symptom improvement.

On the other hand, a recent analysis of the current study's Philly AIMS dataset was conducted as part of a dissertation study and found only a weak positive association between one student characteristic (expressive language skills) and student outcome (Pellecchia, 2013). No associations were found between student outcomes and the following student characteristics at baseline: adaptive skills, social skills, challenging behavior, internalizing behaviors (e.g., depression, anxiety), self-stimulatory behavior, autism severity, and age. These findings directly contrast the current literature, which suggests that these variables are predictors of student outcomes. Pellecchia posited that differences in sample sizes, measurement procedures, and length of intervention likely accounted for the differences in findings.

Contextual Factors

The environmental context also may be associated with intervention effectiveness. A contextual factor is one that describes the surrounding environment in which the intervention is applied. Detrich (1999) suggested that to ensure higher rates of treatment fidelity, it is important to match the intervention to this environment. By matching the contextual variables with the necessary components of the intervention it is more likely that the intervention will be implemented as designed. More specifically within schools, Detrich suggested that the variables of classroom resources, classroom climate, and similarity of the intervention to current practices were critical factors in designing and studying interventions. Lochman (2001) also posited the importance of examining the contextual factors, as they allow researchers insight into which individual interventions

work best in specific settings. Beidas et al. (2013) examined the influence of contextual factors (race and disadvantaged neighborhoods) on internalizing symptoms of anxiety referred youth. They found that diverse children from disadvantaged neighborhoods revealed the most severe symptoms of anxiety, suggesting an association between the environment and anxiety.

With regards to autism intervention research, Dingfelder and Mandell (2011) also proposed that researchers should focus on the implementation of interventions within their community setting from the onset rather than as a subsequent step following an efficacy trial. The authors explained that these implementation outcomes should be assessed throughout the intervention development process to bridge the research-to-practice gap especially in designing and studying interventions for children with autism. Kasari and Smith (2013) further explained that very few autism interventions have been studied within the school context and therefore it is necessary to consider the match of the intervention and environment when selecting evidence-based practices.

School Factors

Schools in particular provide a unique context in which to study intervention effectiveness. Students are nested within classrooms, which are nested within schools. As a result, school level factors may also play a role in how well an intervention can be implemented with a student and how well a student may respond to the given intervention. Educational researchers have suggested the importance of considering the school context when evaluating student outcomes and intervention fidelity. Thrupp and Lupton (2006) posited that focusing on the school context allows for a more fair and valid evaluation of school performance and allocation of resources. Thrupp and Lupton

recommended that schools examine their context to inform interventions and to ensure a goodness of fit between the procedures and school setting. They also explained that adaptations must be made when implementing research-based interventions given the unique context of each school. Wrigley (2006) also suggested that school context (specifically poverty) should be considered when conducting research on school effectiveness and school improvement. Previous paradigms have largely ignored this factor and studies have used primarily homogenous schools in which resources are abundant. It is therefore unclear how these theories would apply to schools with high poverty.

Gottfredson, Jones, and Gore (2002) examined the implementation of a Cognitive Behavioral Therapy (CBT) social skills program in a "disorganized" school (inner-city middle school) with high rates of absenteeism, low staff morale, and low academic achievement. Even though CBT is considered an efficacious treatment, effectiveness as measured by its implementation in a school setting was much lower. The authors hypothesized that the school factors were barriers to the implementation of CBT and subsequently student outcomes were therefore lower.

Research has also suggested that school factors such as principal support and racial climate may also predict intervention fidelity and student outcomes. Kam, Greenberg, and Walls (2003) examined the association between fidelity to Providing Alternative Thinking Strategies (PATHS) program and student outcomes as moderated by principal support. Student outcomes (as measured by behavior ratings from the teacher) were higher in schools with support from principals and when the intervention was implemented as designed. In schools that lacked principal support, student outcomes

were lower regardless of intervention fidelity. Mattison and Aber (2007) found that racial climate also predicted student outcomes (discipline and academics). Positive perceptions of the school's racial climate were associated with greater student outcomes. The authors stated that African American students reported more negative perceptions of racial climate, had lower grades, and more detentions and suspensions than Caucasian students.

Kallestad and Olweus (2003) examined the factors that were related to the implementation of the Olweus Bullying Program in schools from 1983-1985 in Norway. They assessed a variety of school factors and found that school climate (openness in communication and school attention to bullying problems) was a predictor of implementation. At the teacher level, perceived staff importance and perceived level of bullying were important predictors of fidelity as well. Pas and Bradshaw (2012) also evaluated the implementation of a behavioral program in schools. They found that in their Maryland sample of schools implementing School Wide Positive Behavior Supports, teachers who were recently trained and had their teacher certifications displayed better implementation. Better implementation also was associated with higher math and reading test scores and lower levels of truancy.

Overall, the small literature in this area suggests that school factors are related to the implementation and effectiveness of many interventions. Researchers have studied a variety of factors ranging from principal attitudes to school climate. Below is a more comprehensive review of the literature on three commonly reported school factors: suspensions, percentage of students from low-income families, and state standardized test scores. These three factors were selected, as they are considered representative of a school's disciplinary practice, school composition, and academic performance,

respectively. In addition, following the No Child Left Behind Act (2001), publically funded schools are required to submit an annual report card that includes these variables and therefore these data are readily available across the nation. Even though these school characteristics are easily accessible, there have been a limited number of peer-reviewed studies that have analyzed these data. The preliminary research that has been conducted using these school characteristics is presented below.

Disciplinary Practices

One way to assess a school's overall behavioral functioning and disciplinary practice is through the number of suspensions. Out-of-school suspensions are used in schools as a punishment and result in the removal of the student from the school setting for up to ten days. Research on suspensions reveals that the majority of suspensions are generally related to disobedience and/or general disruption rather than more serious offenses (e.g., drugs, weapons) (Morrison & Skiba, 2001). Morrison and Skiba (2001) reviewed the literature on suspensions and reported that often income and race are associated with the use of suspensions in schools. In schools with high percentages of students from low-income families and diversity, suspensions were generally higher. In addition, student-teacher ratios, academic quality, school climate, teacher attitudes, and teacher perceptions of student achievement were implicated in rates of suspensions in schools.

Similarly, Christle et al. (2004) reviewed the research on school suspensions and described how schools with high rates of suspension may represent differences in schools policies and procedures and principal's beliefs, philosophies, and attitudes. Student gender (male), race (African-American) and teacher beliefs and attitudes also played a role in the rate of school suspensions. Schools with low rates of suspensions may have

more strategies, and alternative ways of discipline. Christle et al. (2004) compared Kentucky public middle schools that had high versus low rates of suspension. They found that high rates of suspension were correlated with greater proportions of students from low-income families. Academic achievement and attendance rates were both negatively correlated with suspension rates. Observations of the classrooms revealed that in high suspension rate schools teachers yelled more frequently at students and in low suspension rate schools teacher and student relationships were more caring and respectful. Furthermore the authors found that in high suspension rate schools there was more of an "institutional feel," behavioral and academic programs were not as effective, and the schools' strategies were more reactive than proactive. Even though the sample was limited to schools in Kentucky, the authors provided a snapshot of the differences between high- and low-rate suspension schools that point to classroom and teacher dynamics as likely predictors.

Hogrebe and Tate (2010) reviewed the literature on school suspensions and found that increased use was associated with higher dropout rates and lower academic achievement. In their study of science proficiency scores and school factors, they found that increased student suspensions were associated with lower scores on academic tests. In diverse schools with students from low-income families, lower science scores predicted dropout rates. Raffaele et al. (2002) also found a moderate association between the rate of teacher turnover and the frequency of suspensions, suggesting that teacher burnout was related to high rates of student problem behavior. Raffaele et al. (2002) compared schools with high and low rates of student suspension and found that socioeconomic status was also correlated to rate of suspensions

Researchers have also found a strong correlation between ethnicity and suspension rates. Wallace, Goodkind, Wallace, and Bachman (2008), using a 10th grade sample of students across the US (1999-2005), found that African Americans and Hispanics reported 3.3 times more suspensions than Caucasian students. African-American girls were five times more likely to be suspended or expelled. In this study, the authors controlled for family structure, education, and urban residence. Welch and Payne (2010) similarly concluded that suspension rates were related to the racial composition of the student body. More frequent and harsher disciplinary measures were used in schools with more African American students. The authors suggested that racial threat may be a contributing factor to disciplinary actions. Bradshaw et al. (2010) examined the variables associated with office disciplinary referrals. They also found that African-American males were more likely to receive office referrals. Given the evident relationship between ethnicity, school composition, and suspensions, Gregory and Weinstein (2008) coined the term "discipline gap" to describe how suspensions are used more frequently in schools with more African Americans, more students from low-income families, and in schools with poorer academics.

State Standardized Tests

Since the No Child Left Behind Act (2001), greater emphasis has been placed on state standardized tests. Schools are required to produce an annual report card, detailing their school's performance on these tests. As a result, these scores allow for comparison of how well a school is performing academically in comparison to other schools in its district.

In a study conducted in the Florida public school system, Peabody (2011) categorized schools as high- and low-performing based on the state reading test scores. High-performing was defined as 65% or more of students passing the test and low-performing was defined as 40% of students or less passing the test. In high-performing schools, instructional activities were more student-centered and student-led than in low-performing schools. In addition, teachers reported higher levels of self-efficacy and greater flexibility in making instructional decisions in high-performing schools. Reprimands were more frequent in low-performing schools. However, this study lacked rigor and generalizability given that only four schools were included in the sample.

Klecker and Pollock (2005) also compared classroom practices in schools in which state standardized test scores were high to those in which they were low. The authors examined 39 public schools in Kentucky and found in high-performing schools, teachers reported including significantly more written assignments, greater modification of the vocabulary used in the textbook, and more varied student groupings (i.e., small group, large group, dyads). In low-performing schools, teachers reported delivering significantly more specific task-demands to students during an activity. While research on state standardized test scores is limited, these preliminary findings suggest that schools can be categorized into high- and low-performing and instructional practices may differ across each group.

School Composition

School composition refers to the economic make-up of the student population. One way of capturing this characteristic is by calculating the percentage of students who receive free or reduced price lunches. Research on school composition has generally

focused on the adverse effects seen in schools with high percentages of students from low-income families. Palardy (2008) conducted a study using the National Longitudinal Study data and found differences in student outcomes across schools based on their overall composition. More specifically, students in schools with high percentages of children from low-income families had slower learning rates even when controlling for other school factors and individual student characteristics. Palardy suggested a "synergetic interpretation" such that schools with high proportions of students from poverty also have teachers with less experience and have more frequent disciplinary problems. Taken together, these factors may significantly affect the learning environment. The author suggested that these variables alone may not impact learning but their interaction can be more significant. In schools with low socioeconomic composition, school factors (e.g., disciplinary practices, teacher quality) were more strongly associated with student outcomes.

Similarly, Lytton and Pyryt (1998) showed that in 142 schools in Canada, family income was the main predictor of student achievement scores in reading and math, explaining 45% of the variance. Likewise, Ma and Klinger (2000) examined the data from sixth graders and concluded that gender, school composition, and native ethnicity were strong predictors of academic achievement. In a related study, Hogrebe and Tate (2010) examined tenth grade science proficiency scores in Missouri high schools and their relationship to school and family characteristics. The authors determined that student composition variables (i.e., percentage of free or reduced price lunches) were predictors of science proficiency scores and moderated the relationship between student dropout rates and science proficiency scores. For example, lower science scores were

associated with higher dropout rates in schools with high percentages of students from low-income families.

Research also has found that school composition predicts student outcomes irrespective of the individual student's background (Hogrebe &Tate, 2010; Lytton & Pyryt, 1998; Palardy, 2008; Perry & McConney, 2010; Rumberger & Palardy, 2005; Southworth, 2010). Research also has been conducted across cultures with similar results. Perry and McConney (2013) conducted a study comparing schools in Canada and Australia. Overall, they found a strong association between school composition and academic achievement with a more significant association found in Australia than in Canada. Willms (2010) also focused on cross-cultural research by utilizing the 2006 Programme for International Student Assessment (PISA) data from the United States, Japan, France, Spain, and Finland. He found that teacher variables (i.e., experience and advanced degrees) moderately predicted student performance on tests of math and reading, but overwhelmingly school composition strongly predicted academic performance.

On the other hand, Fantuzzo, LeBoeuf, and Rouse (2014) suggested that poverty may not be an accurate measure to account for the association between student outcomes and school composition. Rather, the authors suggested that factors such as maternal education, birth risks, and maltreatment are more precise variables for assessment. They examined birth weight, age of the mother, maternal education, high lead exposure, homeless shelter stay, and child maltreatment and their association with attendance rates and academic achievement. Maternal education was the primary predictor with child maltreatment and homelessness as secondary predictors of educational well-being. More

specifically, after controlling for student risk factors and demographics, attendance at a school with high concentrations of mothers who did not complete high school, was significantly correlated with lower academic performance and attendance rates. However, future research is necessary in fully delineating the relationships between school composition, poverty, and student outcomes.

The Current Study

The present study extends the autism intervention literature by evaluating the contextual factors that may predict intervention effectiveness and student outcomes. More specifically, the current study examines the association between school factors, intervention fidelity, and student outcomes in autism support classrooms. It is unclear whether building level variables predict the delivery and effectiveness of evidence-based practice in autism support classrooms. This study is a preliminary investigation of these associations.

The first goal of this study is to determine if school factors, including school discipline practices, academic performance, and school composition predict intervention fidelity in autism support classrooms. Autism support classrooms are self-contained rooms that represent one of the most restrictive settings for children with autism given the lack of typically developing peers. School factors may not be associated with an autism support teacher's ability to implement an evidence-based practice in her classroom because of the isolated nature of these classrooms. While school factors may have a large impact on regular education classrooms, a self-contained special education classroom may function separately from and independently of the rest of the school and therefore be less impacted by the surrounding school environment.

Determining predictors of intervention fidelity in classrooms allows researchers to identify malleable factors to address when developing strategies to increase the use of evidence-based practices in schools. If school factors are strong predictors of intervention fidelity, building level interventions may be necessary in helping improve the overall school's climate rather than spending more time coaching the teacher on the evidence-based practice. On the other hand, if school factors are not strong predictors of fidelity, this indicates that evidence-based practices may be effectively transported to classrooms within both high- and low-performing schools with less of a pressing need to address the contextual factors. In this case, greater focus should be placed on coaching the classroom team on the implementation of the intervention.

The second goal of this study is to determine if school factors predict student outcomes in autism support classrooms. Similar to the first goal, it is possible that the school may matter less than the classroom characteristics. Again, by identifying predictors of student outcomes, this study offers greater understanding of the effectiveness and implementation of evidence-based practices in public school settings. While there are a few studies that have reported predictors of student outcomes (Arick et al., 2003; Mandell et al., 2013; Pellecchia, 2013), no study has examined school factors as predictors of student outcomes.

The third goal of this study is to determine if school factors moderate the association between fidelity and student outcomes in autism support classrooms. In high-performing schools, does fidelity predict student outcomes more than in low-performing schools? This study will serve as an initial investigation of possible moderating variables in the implementation of evidence-based practices for students with autism. It is hypothesized

that school factors moderate the relationship between fidelity and outcomes. In schools that have a high percentage of students from low income families, poor academic performance, and high rates of suspension, teachers who consistently adhere to the intervention may not see subsequent gains in student outcomes. In contrast, in schools with lower percentages of students from low income families, strong academic performance, and low rates of suspension, teachers who consistently adhere to the intervention will see greater gains in student outcomes.

CHAPTER 3

METHODS

Participants

Participants were drawn from the Philadelphia Autism Instructional Methods Study (Philly AIMS). Philly AIMS was a research project funded by the National Institute of Mental Health and The Institute of Education Sciences that examined the effects of implementing evidence-based practices for children with autism in a large urban school district. In its first year, Philly AIMS included 33, K-2 autism support classrooms across the school district representing 79% of the total number of classrooms. The school district sent letters to each autism support staff member for participation in the study and 90% of classrooms agreed to participate. No students consented to participate in three classrooms. In two classrooms, data collection at baseline on participating students was incomplete, leaving the 33 classrooms. There were no statistically significant differences in demographics or years teaching between staff in the 33 classrooms and the classrooms that did not participate.

There were 121 consented students, representing 52% of all eligible students in the participating classrooms. Baseline testing of two students was later deemed invalid and subsequently these students were excluded from the analysis, leaving a total sample of 119. Students were recruited through a consent form and flyer describing the program that teachers sent home with the student. Parents received $50 for the first wave of data collection, $100 for the second wave of data collection, and a summary report of their child's assessment that was designed for use in IEP planning.

Inclusion criteria for the Philly AIMS study were an educational classification of autism through the Philadelphia School District and enrollment at least half time within a kindergarten-through-second grade autism support classroom. During the subsequent years, additional incoming kindergarten students were recruited. As a result, in the present study, there were 171 participants included who were enrolled across 47 kindergarten-through-second grade autism support classrooms in the Philadelphia School District. These students were five to nine years old (mean age of 6.5 years) and had a primary diagnosis of autism through the Philadelphia School District. The majority of participants were male (84.8%). Students were from diverse backgrounds: 39.2% African American, 9.4% Hispanic, 13.5% Caucasian, 5.3% Asian, and 32.8% other.

Table 1. Philadelphia Autism Instructional Methods Study Student Characteristics (N = 171)

Student Characteristics	Frequency	Percentage
Gender		
Male	145	84.8%
Female	25	14.62%
Ethnicity		
African American	67	39.2%
Hispanic	16	9.4%
Asian	9	5.3%
Caucasian	23	13.5%
Other	56	32.8%
Age		
5 years	42	24.6%
6 years	64	37.4%
7 years	58	33.9%
8 years	7	4.1%

Setting

Participants in the present study were enrolled in autism support classrooms in the Philadelphia School District. The Philadelphia School District is a large, diverse, and urban school district. During the 2010-2011 academic year, there were more than 165,000 enrolled students and more than 1,000 students with an educational classification of autism spectrum disorder. The students in the Philadelphia School District reflected diverse ethnic backgrounds: 58.3% of students were African American, 18.2% Hispanic, 13.4% Caucasian, 6.9% Asian, and 3.3% were of other ethnicities. The Philadelphia School District reported a high proportion of students from disadvantaged backgrounds; 77% of students were eligible for free or reduced-price lunches.

Measures

To assess the relationships between school factors, intervention fidelity, and student outcomes, a variety of measures were included (see Table 2 for a summary of the measures and their corresponding data sets).

Table 2. Study Measures

Domain	Measures	Data Set
Student Outcomes	Differential Ability Scales, Second Edition	Philly AIMS
Intervention Fidelity Data	1. Functional Routines Fidelity 2. Pivotal Response Training Fidelity 3. Discrete Trial Teaching Fidelity	Philly AIMS
School Factors	1. Academic Performance 2. School Composition 3. Disciplinary Practices	Philadelphia School District

Student Outcomes

Differential Ability Scales, Second Edition (DAS-II). The DAS-II (Elliott, 1990) is a comprehensive, individually administered, clinical instrument for assessing cognitive abilities in children ages 2 years 6 months through 17 years 11 months. The subtests measure 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 has excellent reliability and validity. The mean internal consistency reliability and test-retest reliability for the global conceptual ability composite (GCA) were .95 to .96 and .86 to .92, respectively. In addition, the DAS-II is adequately correlated with other measures of cognitive ability. More specifically, correlations ranged from .59 to .87 on other cognitive assessments including, the Wechsler Intelligence Scale for Children, Fourth Edition (WISC-IV), the Wechsler Preschool and Primary Scale of Intelligence, Third Edition (WPPSI-III), and the Bayley Scales of Infant and Toddler Development, Third Edition (Bayley-III) (Elliott, 1990).

The DAS-II was administered at the beginning of the year (September, 2010) and at the end of the year (June, 2011) by the Philly AIMS research team to each of the participants. The research team, composed of licensed psychologists and school psychology graduate students, administered the test. The assessors received training on the DAS-II prior to administering the test and were required to demonstrate 100% adherence to the testing procedures. The following subtests were administered to children under the age of seven years to obtain the global conceptual ability composite (GCA): verbal comprehension, naming vocabulary, picture similarities, pattern construction, and

copying. For children ages seven to eight years, word definitions, verbal similarities, matrices, sequential and quantitative reasoning, recall of designs, and pattern construction were administered to obtain the GCA composite score. Change in scores between the beginning of the school year and the end of school year were computed to assess student outcomes in relation to intervention fidelity and school factors by subtracting the end of year score from the beginning of the year score.

Fidelity Data

Intervention fidelity adherence and intensity measures were created based on the Strategies for Teaching based on Autism Research (STAR) manual. During monthly observations of each classroom, the Philly AIMS research team rated adherence to and intensity of the STAR instructional programs (DTT, PRT and FR) using implementation fidelity checklists and rating forms. The intervention fidelity scores utilized in the current study reflected the ratings from the last observation (May, 2011). These scores represented the teacher's implementation of the instructional programs at the end of the year following a full school year of professional development training and coaching on the STAR program. The Philly AIMS research team included six Master's level school-based consultants who had received training in DTT, PRT, and FR from the STAR program developers. The consultants provided bi-weekly classroom coaching and consultation on the STAR program to the autism support teachers and support staff. The research team also received training on the fidelity coding procedures. Bi-weekly meetings occurred following the training for the research team to discuss the classroom ratings to ensure reliability among coders.

The research team coded intervention fidelity adherence 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). Highly accurate DTT included gaining the child's attention, providing clear and appropriate instructions, using appropriate prompting strategies, providing clear and correct consequences, and utilizing error correction procedures. Highly accurate PRT implementation included gaining child's attention, providing clear and appropriate instructions, providing the child a choice of stimuli, interspersing maintenance and acquisition tasks, and using natural reinforcers. Highly accurate FR implementation included using individual schedules and providing predictable and consistent routines.

In addition to intervention adherence, a measure of intensity was also included. Through direct observation and teacher report, intervention intensity was assessed as the frequency of the implementation of DTT, PRT, and FR. Intervention intensity was also rated using a Likert scale with the following scores: 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).

Fidelity of each instructional procedure (DTT, PRT, FR) was calculated as a combination of adherence and intensity. More specifically, intervention adherence was multiplied by intensity to obtain an overall fidelity code ranging from 0 to 16. Teachers who showed high adherence to the instructional program (i.e., adherence score of 4) but did not consistently implement the instructional program (i.e., intensity score of 1), received an overall fidelity score of 4. Since fidelity is a measure of how accurately and how consistently an intervention is implemented, this ensures that a teacher who

demonstrates high adherence in an observation did not receive inflated fidelity scores if she did not implement the strategy consistently.

School Factors

The Philadelphia School District published a public use dataset that details demographic variables, state standardized test scores, and disciplinary information for each school in the school district. Data from the 2010-2011 school year were obtained from this dataset and compiled with the Philly AIMS data to assess the associations between school factors, intervention fidelity, and student outcomes within autism support classrooms.

An overall academic performance score for each school was computed by averaging the percentage of third grade students scoring advanced or proficient on math state standardized tests and on reading state standardized tests. An overall disciplinary practices score for each school was determined by totaling the number of school suspensions across the school year. School composition was reported as the percentage of students who received free or reduced price lunch.

Procedure

Question 1: Do school factors such as academic performance, disciplinary practices, and school composition predict teacher intervention fidelity to evidence-based practices for students with autism?

Participants in the Philly AIMS study received DTT, PRT, and FR interventions by their classroom teacher as part of the Strategies for Teaching Based in Autism Research (STAR) program. Intervention fidelity to each instructional strategy varied substantially across teachers (range: 0-16). Several analyses were used to evaluate the association

between school factors and intervention fidelity. First, to create an overall fidelity to STAR variable, the variables representing fidelity to DTT, PRT, and FR (the three components of STAR) were summed. This variable was used as the variable representing overall intervention fidelity to the STAR program in all subsequent analyses. A factor analysis using principal component analysis and varimax rotation was conducted to confirm the use of this unitary STAR fidelity variable.

The association between school factors and intervention fidelity was then evaluated. Pearson correlation coefficients were calculated to evaluate the association between individual school factors (disciplinary practices, school composition, and academic performance) and intervention fidelity to STAR. Linear regression analysis was used to examine the predictive relationship between school factors and intervention fidelity. The number of years the teacher had taught in an autism support classroom and the average baseline DAS-II score across the classroom were added as covariates in the model to account for differences in teaching experience and the overall functioning level of the students in the classroom.

Question 2: Do school factors such as academic performance, disciplinary practices, and school composition predict student gains on a cognitive assessment for students with autism?

Two analyses were conducted to evaluate the correlation and predictive relationship between school factors and student outcomes. First the correlation between school factors and student outcomes was evaluated. Pearson correlation coefficients were calculated to evaluate the association between individual school factors (disciplinary practices, school composition, and academic performance) and student outcomes. Linear regression with

random effects for schools was then used to examine the predictive relationship between school factors and student outcomes. Regression with random effects analysis was selected as the analytic strategy because of the hierarchal nature of the data. Students in this study were nested within classrooms, which in turn are nested within schools. Students in a given classroom may be more similar to each other than to students from a different school. Regression with random effects accounts and adjusts for the potential homogeneity of students within a school and heterogeneity of students across schools. Regression with random effects for schools was selected given that there was generally one classroom per school building included in the Philly AIMS study (47 classrooms across 40 schools). In addition, the number of years the teacher had taught in an autism support classroom, the teacher's STAR fidelity score, and the student's baseline DAS-II score were added as covariates in the model to account for differences in teaching experience, intervention implementation, and baseline cognitive ability.

Question 3: Do school factors such as academic performance, disciplinary practices, and school composition moderate the association between intervention fidelity and student outcomes, such that schools that are high performing (i.e., strong academics and fewer behavioral problems) show a stronger association between intervention fidelity and student outcomes than low-performing schools?

To address this question, the same analysis as for Question 2 was conducted, but now including the interaction term for school factors and intervention fidelity. Baseline DAS-II scores were added as a covariate in the model to account for baseline cognitive ability.

CHAPTER 4

RESULTS

This study combined data from Philly AIMS for the 2010-2011 academic year, and the Philadelphia School District data on school characteristics during the same year. There was one measure of student outcome (change in DAS-II score), three measures of intervention fidelity (DTT, PRT, and FR), and three measures of school characteristics (disciplinary practices, school composition, academic performance) used to investigate the associations between school factors, fidelity, and student outcomes. A summary of the descriptive statistics for each of these measures is presented in Table 3.

Table 3. Summary of Descriptive Statistics

Variable	Mean	SD	Range
School Factors (n=40)			
Academic Performance Average % scoring proficient or advanced on reading and math state tests	63.4%	16.6	29.2-95
School Composition % Receiving free or reduced price lunch	81.3%	14.2	50.3-95.3
Disciplinary Practices Total number of suspensions	52.5	47	3-229
Intervention Fidelity (n=47)			
DTT Fidelity	3.7	3.3	0-12
PRT Fidelity	2.8	2.3	0-9.5
FR Fidelity	7.5	4.5	0-16
Student Outcome (n=171)			
DAS-II Change Score	3.4	10.5	-43-43
Teacher Characteristics (n=47)			
Number of years teaching autism support	11.4	11.3	1-41

Question 1: Do school factors such as academic performance, disciplinary practices, and school composition predict teacher intervention fidelity to evidence-based practices for children with autism?

Several analyses were conducted to evaluate the association between school factors and intervention fidelity. DTT, PRT, and FR fidelity measures were used to assess intervention fidelity. First, to create an overall fidelity to STAR variable, the variables representing fidelity to DTT, PRT, and FR (the three components of STAR) were summed. To confirm the use of one unitary variable to encompass overall fidelity to STAR, the three fidelity variables were included in a factor analysis using principal component analysis and varimax rotation. The analysis yielded one factor that explained 79% of the variance (see Table 4). The one factor included all three fidelity variables, verifying the use of one STAR intervention fidelity variable.

Table 4. Factor Loadings for Intervention Fidelity

Item	Factor 1
DTT Fidelity	.61
PRT Fidelity	.58
FR Fidelity	.54

Pearson correlation coefficients were calculated to evaluate the correlation between school factors and intervention fidelity (see Table 5). Results of the Pearson correlations indicate that there were no significant associations between school factors and intervention fidelity (see Appendix A for the complete correlation matrix).

Table 5. Correlations between School Factors and STAR Intervention Fidelity

School Factor	Pearson correlation	N	Significance
Disciplinary Practices	-.087	47	.559
Academic Performance	.034	47	.818
School Composition	-.167	47	.262

A linear regression analysis was also conducted to examine the predictive relationship between school factors and intervention fidelity, as indicated by the STAR fidelity factor established by summing the three intervention fidelity variables (see Table 6). The number of years the teacher had taught in an autism support classroom and the average baseline DAS-II score across the classroom were added as covariates in the model to account for differences in teaching experience and the overall cognitive functioning level of the students in the classroom. Results of the regression analysis were consistent with the Pearson correlation, such that disciplinary practices, school composition, and academic performance did not predict STAR fidelity.

In summary, the results of the analyses conducted to evaluate Question 1 found that none of the school factors were correlated with or predicted intervention fidelity.

Table 6. Summary of Regression Analysis for School Factors Predicting STAR Intervention Fidelity

	Unadjusted		Adjusted	
	Estimate	P value	Estimate	P value
Academic Performance	.021	.806	-.104	.392
School Composition	-.134	.191	-.182	.149
Disciplinary Practices	-.018	.552	-.024	.500
Number of Years Teaching	---	---	.025	.837
Average Baseline DAS-II Score	---	---	-.098	.275

Question 2: Do school factors such as academic performance, disciplinary practices, and school composition predict student gains on a cognitive assessment?

Two analyses were conducted to evaluate the associations between school factors and student outcomes. Student outcome was measured by computing the change in IQ score for each student derived from pre- and post-intervention cognitive ability scores as measured by the general conceptual ability (GCA) on the Differential Ability Scales, Second Edition (DAS-II). The mean overall change score was 3.36 (See Table 7). Change scores were distributed across a wide range, from -43 to 43, indicating variability in outcome across participants.

Table 7. Change in Overall Cognitive Ability as Measured by the DAS-II for the Total Sample

Mean change score	SD	Range
3.36	10.5	-43 to 43

Pearson correlation coefficients were calculated to evaluate the correlation between school factors and student outcome (see Table 8). Results of the Pearson correlations indicate that there were no significant associations between school factors and student outcome.

Table 8. Correlations between School Factors and Student Outcome

Variable	Pearson correlation	N	Significance
Disciplinary Practices	-.018	171	.811
Academic Performance	-.030	171	.699
School Composition	.047	171	.543

A linear regression with random effects for schools analysis was also conducted to examine the predictive relationship between school factors and student outcomes, as indicated by the DAS-II change score (see Table 9). The number of years the teacher had taught in an autism support classroom, STAR intervention fidelity score, and student's baseline DAS-II score were added as covariates in the model to account for differences in teaching experience, intervention implementation, and baseline cognitive ability.

Results of the regression analysis were consistent with the lack of findings in the Pearson correlations, such that academic performance, school composition, and school disciplinary practices did not predict student outcome.

Table 9. Summary of Regression Analysis for School Factors Predicting Student Outcome

	Unadjusted		Adjusted	
	Estimate	P value	Estimate	P value
Academic Performance	-.019	.714	-.048	.539
School Composition	.018	.778	.002	.974
Disciplinary Practices	-.008	.699	-.018	.474
Fidelity to STAR	---	---	.148	.129
Number of Years Teaching	---	---	-.037	.635
Baseline DAS-II Score	---	---	-.157	<.001

In summary, the results of the analyses conducted to evaluate Question 2 found that none of the school factors were correlated with or predicted student outcome.

Question 3: Do school factors, such as academic performance, disciplinary practices, and school composition, moderate the association between intervention fidelity and student outcomes, such that schools that are high performing (i.e., strong academics and fewer behavioral problems) show a stronger association between intervention fidelity and student outcomes than low-performing schools?

To determine if school factors moderate the association between fidelity and student outcomes, a regression with random effects for schools was conducted examining the

predictive relationship between the interaction of school factors and intervention fidelity and student outcomes. Baseline DAS-II scores were included as a covariate in the model. The results of the analysis indicated that school factors were not moderators, given the overall lack of significance (see Table 10).

In summary, across all analyses, there was an overall lack of significant findings indicating that school factors were not correlated with or predictors of intervention fidelity or student outcome. In addition, the interaction between fidelity and school factors did not predict student outcomes, indicating that school factors did not moderate the association between fidelity and outcomes.

Table 10. Summary of Regression Analysis for School Factors Moderating the Relationship Between Fidelity and Student Outcome

	Estimate	P value
Disciplinary Practices	-.024	.585
STAR Fidelity	.077	.626
STAR Fidelity x Disciplinary Practices	.001	.656
Baseline DAS-II Score	-.158	<.0001

	Estimate	P value
Academic Performance	-.007	.953
STAR Fidelity	.221	.594
STAR Fidelity x Academic Performance	-.001	.840
Baseline DAS-II Score	-.156	<.0001

	Estimate	P value
Student Composition	-.093	.444
STAR Fidelity	-.417	.407
STAR Fidelity x Student Composition	.007	.256
Baseline DAS-II Score	-.156	<.0001

CHAPTER 5

DISCUSSION

In the present study, a preliminary investigation regarding the relationships between school factors, intervention fidelity, and student outcomes was conducted. The data were evaluated to determine if school factors predicted intervention fidelity or student outcomes, and if school factors moderated the association between fidelity and outcomes.

Association of School Factors and Intervention Fidelity

The results of the analyses evaluating the associations between school factors and intervention fidelity were not significant. In addition, none of the school factors predicted intervention fidelity, indicating that the school context was not associated with a teacher's willingness and ability to implement evidence-based practices. Interestingly, educational research has previously suggested that school composition, disciplinary practices, and academic performance predict intervention fidelity and teacher practices and attitudes within a classroom (Christle et al., 204; Morrison & Skiba, 2001; Klecker & Pollock, 2005; Palardy, 2008; Peabody, 2011). However, in the current study, these relationships were not found. There may be a few reasons for this finding. First, autism support classrooms represent a self-contained, highly restrictive environment within a school. Anecdotally, consultants within the school have noted that general education students were often unable to locate the autism support classroom and were frequently unaware of its existence within a school. As a result, given the isolated nature of the autism support classroom, it is not surprising that a teacher's ability or willingness to implement an intervention within an autism support classroom is unrelated to what may

be occurring in the surrounding school building. Numerous autism support teachers have commented that while their schools were chaotic, their classrooms felt like undisturbed islands within the school building.

Another possible reason for this finding is that the measures of school factors used in this study may not render valid representations of academic performance, school disciplinary practices, and school composition. For example, researchers have suggested that large constructs such as poverty may not allow for precise assessment of the relationship between the school context and student outcomes (Fantuzzo, LeBoeuf, & Rouse, 2014). In the future it may be necessary to include additional measures of these variables to supplement these data (e.g., surveys). Nonetheless given that these data are publicly available, it is important for researchers to continue to take advantage of these readily available datasets by including them in analyses.

Lastly, it is possible that there are other school characteristics that may be more salient to autism support teachers. For example measures of school climate regarding special education and principal's involvement in special education may be more related to an autism support teacher's ability to implement evidence-based practices in her classroom. For example, a teacher who has a principal who is involved and supportive within special education classrooms may be more motivated to implement the evidence-based practice. In addition, if the school as a whole shows greater sensitivity and awareness of differences, and special education is more integrated within the school environment the autism support teacher may feel greater pressure to deliver high quality evidence-based instruction to her students.

Association of School Factors and Student Outcomes

The results of the analyses evaluating the associations between school factors and student outcomes similarly revealed non-significant results. School disciplinary practices, school composition, and academic performance were not correlated with student outcomes. Again, this finding is inconsistent with previous research in general education classrooms. Research has generally found that in schools with lower ratios of students from low-income families, higher academic performance on state standardized tests, and fewer suspensions, student outcomes are greater (Gottfredson, Jones, & Gore, 2002; Hogrebe & Tate, 2010; Perry & McConney, 2010; Southworth, 2010). However, analogous to the first question, these results were not surprising given the isolated nature of the autism support classroom. It was therefore anticipated that school factors would be less related to student outcomes in autism support classrooms.

School Factors as Moderators

Not surprisingly, given the lack of significant results in the previous analyses, when testing the interactions between school factors and intervention fidelity, there were no predictive relationships with student outcomes. Given that this was the first evaluation of school factors as moderators, this study begs replication.

Study Limitations

Several limitations to this study may account for the absence of significant findings. The lack of a robust and sensitive student outcome measure may have affected the results. While the DAS-II (Locke, Rotheram-Fuller, Xie, Harker, & Mandell, 2013; Mandell et al., 2013; Pellecchia, 2013) and other cognitive assessments (Ben-Itzchak, Lahat, Burgin, & Zachor, 2008; Howlin, Goode, Hutton, & Rutter, 2004; Lovaas, 1987;

Sallows & Graupner, 2005) have been utilized across research studies as an outcome variable, the significant variability in change scores across nine months (-43 to 43) may suggest that there were other factors associated with the change in scores (e.g., interfering behavior). For example, the score on the DAS-II merely represents a snapshot of the student's functioning on that day. If the student engaged in frequent problem behavior, the examiner may not have been able to obtain a valid estimate of the student's cognitive ability. In addition, improvement on the DAS-II may not capture a comprehensive picture of a student's progress. More specifically, a student's ability to utilize functional communication (an area focused on in PRT) may reflect a significant gain but may not be captured in the DAS-II score. As a result, future research should include additional measures of student progress that may be sensitive to greater areas of improvement.

One student outcome measure that was collected but not included in the present study was the Autism Diagnostic Observation Schedule (ADOS). ADOS scores were also collected at the beginning and end of the year, however little change was observed across the school year. Given that the ADOS was designed as a diagnostic tool rather than as a way to measure progress, it was expected that little change would be recorded across the school year. In the future, curriculum based assessments should be included to provide a criterion-referenced assessment (e.g., Verbal Behavior Milestones Assessment and Placement Program) of skill acquisition across a year as a complementary measure.

A second limitation was the selection of school factors included in the study. The use of total number of suspensions, scores on state tests, and percentage of students receiving free or reduced price lunches may not be valid representations of the larger constructs of disciplinary practices, academic performance, and school composition, respectively. It is

evident that there are other factors that may impact these data. For example, total number of suspensions may represent a lack of other disciplinary options, a high rate of student problem behavior, or both. Much criticism has arisen regarding state standardized tests, their lack of alignment to classroom curricula, and the subsequent result of teachers teaching primarily to the test. However, these data are publically available across the nation and offer researchers a rich dataset to utilize in their studies. Future studies should include additional measures of discipline, academic performance, and school composition to provide a more precise assessment of these factors.

There were other contextual variables that may also merit further study. The current study did not address the differences in staffing ratios across classrooms, or the relationships between staff within the classrooms. In the Philly AIMS study, data were not collected regarding the number of adults in the classroom. In addition, given the staffing changes that occurred frequently throughout the district and movement of students across schools, the ratios at the beginning and end of year often differed considerably. Future research should address this gap in the literature, especially in reference to recommendations for staffing ratios in schools.

A third limitation was that fidelity data were indirect measures that lacked reliability data. Intensity was ranked using a self-reported Likert scale rather than as a direct observation of the exact amount of instruction that was provided. In addition, adherence was measured through a rating scale administered by the research team rather than through a more precise metric (i.e., percentage of steps followed in an intervention). While bi-weekly meetings occurred for the research team to discuss coding, reliability data were not collected among the coders. In the future, interobserver reliability data

should be collected to ensure reliable fidelity ratings. In addition, it is notable that overall fidelity data to DTT and PRT were very low, indicating that these instructional practices may be too complex for the teachers to learn in a school year. Even with professional development training and ongoing coaching support, teachers struggled to implement the program as it was designed. Given that none of the teachers achieved high fidelity across DTT, PRT, and FR it was not surprising that there was an overall lack of significant findings. If there was greater intervention fidelity across practices, more significant associations may have emerged.

Lastly, given the homogenous sample utilized in this study, it is unclear whether these findings would generalize across other settings. For example, school factors may play a larger role in suburban school districts, where there are schools with greater variability in resources. Teachers who elected to participate in this study may also represent a unique population. More specifically, given that these teachers consented to the study and the presence of university-based consultants, trainers and observers, teachers may have been less aware of the school factors surrounding them given the high rates of attention and support provided from the outside university. It is uncertain if school factors may have a greater impact on autism support classrooms if the interventions were recommended and supported by the School District rather than through an outside university. Future studies should include both schools from low-resource school districts and affluent school districts, as well as, assess sustainability of an evidence-based practice when support is shifted from an external university to internal personnel from the school district.

Future Directions

It is evident that future research is needed in parsing out the associations between intervention fidelity, student outcomes, and contextual factors within schools. While this study provides initial information regarding these relationships, the use of additional measurements for fidelity, student outcomes, and school factors is merited to gather a more comprehensive picture of these associations. Future studies should continue to focus on fidelity variables but incorporate additional metrics including more precise fidelity measurements (e.g., number of hours spent implementing the program and percentage of intervention steps followed). Curriculum based assessments for children with autism such as the Verbal Behavior Milestones Assessment and Placement Program (Sundberg, 2008) may offer a more comprehensive picture of a student's progress across a school year that may be better matched to the instructional practices within the classroom and thus more sensitive to change. Additional surveys and measures of school factors are also needed to offer a more complete picture of the school context. For example teacher and student surveys on their perceptions of the disciplinary practices within the school may offer greater insight regarding the number of suspensions per year. Again, it is possible that frequent suspensions may represent high rates of problem behavior, lack of disciplinary options, or both. Lastly, the measures of classroom context factors (e.g., staff to student ratio, availability of resources within the classroom, staff attitudes and perceptions) should be incorporated to provide information regarding the classroom functioning in addition to the student outcomes and teacher fidelity.

Future research should also contain greater variability in the sample. Schools from both affluent and low-resource districts should be included to ensure greater

generalizability. It would be interesting to study the different levels of influence school factors may have on classrooms within high versus low resource school districts. It would also be important to study the implementation of these practices when delivered internally by the school district rather than through an external university. Studies should also include a sustainability plan to ensure that the evidence-based practice can continue once the university no longer provides the support. Internal capacity needs to be developed and later studied to determine its long-term effectiveness.

Lastly, given the overall poor intervention fidelity across teachers, it is recommended that future research focus more specifically on the consultation and coaching strategies that can be used to more effectively teach teachers how to implement evidence-based practices in their classrooms. Greater focus should be placed on classrooms that consistently fail to implement an intervention to determine if building level or classroom level barriers are present that preclude the use of the program. To ensure that evidence-based practices are effectively transported from a research setting to a school environment, researchers should shift their attention to developing efficient and effective training programs that foster greater intervention fidelity and improved student outcomes.

Summary

The rising prevalence of children diagnosed with autism and the emergence of evidence-based autism interventions has added urgency to the challenge schools face in delivering high quality instruction to this unique population. The reauthorization of IDEIA (2004) emphasizes the use of evidence-based practices in special education. Comprehensive packaged curricula have been developed to address this growing need

and to allow educators to transport research-based instruction into their classroom settings.

However, researchers have largely ignored the contextual factors that may be related to the successful implementation of autism interventions in public school settings. Existing studies have focused primarily on intervention fidelity, student, and teacher characteristics as predictors of student outcomes. While researchers have reiterated the importance of evaluating the contextual factors to inform curriculum design and to address potential barriers to implementation (Dingfelder & Mandell, 2011; Tincani, 2007), there have been no published studies to date that have assessed these associations. The present study serves as the first preliminary investigation of the associations between the school context and the implementation of evidence-based practices for children with autism.

The results of this study revealed that school composition, academic performance, and disciplinary practices did not predict intervention fidelity or student outcomes. The findings suggest that autism support classrooms are like islands within the school building, such that the practices and outcomes within these classrooms were unrelated to the school context. This study indicates that when transporting an evidence-based practice into a public school classroom, it may be more necessary to focus on the classroom context rather than the school building. However, given that overall fidelity to the STAR program was low across classrooms irrespective of their school's composition, disciplinary practices, and academic performance, it is not surprising that significant relationships were not found between school factors, intervention fidelity, and student outcomes. Given the limitations of this study, future research is needed to fully delineate

the relationships between school building level factors and the dissemination and implementation of evidence-based practices within a classroom setting.

REFERENCES CITED

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Anderson, S. R., DiPietro, E. K., Edwards, G. L., & Christian, W. P. (1987). Intensive home-based early intervention with autistic children. *Education and Treatment of Children, 10,* 352-366.

Arick, J. R., Falco, L. L. R., & Krug, D. A. (2004). *The STAR Program: Strategies for teaching based on autism research, levels I, II & III.* Austin: PRO-ED.

Arick, J. R., Young, H. E., Falco, R. A., Loos, L. M., Krug, D. A., Gense, M. H., & Johnson, S. B. (2003). Designing an outcome study to monitor the progress of students with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities, 18*(2), 74-86.

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.

Banda, D. R. & Grimmett, E. (2008). Enhancing social and transition behaviors of persons with autism through activity schedules: a review. *Education and Training in Developmental Disabilities, 43*(3), 324–333.

Beidas, R. S., Suarez, L., Simpson, D., Read, K., Wei, C., Connolly, S. & Kendall, P. (2013). Contextual factors and anxiety in minority and European American youth presenting for treatment across two urban university clinics. *Journal of Anxiety Disorders, 26*(4), 544-554.

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.

Benner, G. J., Nelson, J. R., Stage, S. A., & Ralston, N. C. (2010). The influence of fidelity of implementation on the reading outcomes of middle school students experiencing reading difficulties. *Remedial and Special Education, 32*(1), 79-88.

Birnbrauer, J. S., & Leach, D. J. (1993). The Murdoch early intervention program after 2 years. *Behavior Change, 10*, 63-74.

Bradshaw, C. P., Mitchell, M. M., O'Brennan, L. M., & Leaf, P. J. (2010). Multilevel exploration of factors contributing to the overrepresentation of black Students in office disciplinary referrals. *Journal of Educational Psychology, 102*(2), 508-520.

Centers for Disease Control and Prevention. (2014). *Autism spectrum disorders*. Retrieved from http://www.cdc.gov/ncbddd/autism/data.html

Christle, C., Nelson, M. C., & Jolivette, K.(2004). School characteristics related to the use of suspension. *Education and Treatment of Children, 27(*4), 509-526.

Cohen, H., Amerine-Dickens, M., & Smith, T. (2006). Early intensive behavioral treatment: replication of the UCLA model in a community setting. *Developmental and Behavioral Pediatrics, 27*(2), 145-155.

Coolican, J., Smith, I. M., & Bryson, S. E. (2010). Brief parent training in pivotal response treatment for preschoolers with autism. *Journal of Child Psychology and Psychiatry, 51*(12), 1321-1330.

Cowan, R. J., & Allen, K. D. (2007). Using naturalistic procedures to enhance learning in individuals with autism: A focus 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.

Dawson, G. Rogers, S., Munson, J., Smith, M., Winter, J., Greenson, J., … Varley, J. (2010). Randomized controlled trial of an intervention for toddlers with autism: the early start denver model. *Pediatrics, 125*(1), 17-23.

Delprato, D. J. (2001). Comparisons of Discrete-Trial and Normalized Behavioral Language Intervention for Young Children with Autism.*Journal Of Autism & Developmental Disorders, 31*(3), 315-325.

Detrich, R. (1999). Increasing treatment fidelity by matching interventions to contextual variables within the educational setting. *School Psychology Review, 28*(4), 608-620.

Dettmer, S., Simpson, R. L., Myles, B., & Ganz, J. B. (2000). The Use of Visual Supports to Facilitate Transitions of Students with Autism. *Focus On Autism & Other Developmental Disabilities, 15*(3), 163.

Dingfelder, H. E. & Mandell, D. S. (2011). Bridging the research-to-practice gap in autism intervention: an application of diffusion of innovation theory. *Journal of Autism and Developmental Disorders, 41,* 597-609.

Domitrovich, C. E., & Greenberg, M. T. (2000). The study of im- plementation: Current findings from effective programs that prevent mental disorders in school-aged children. *Journal of Educational and Psychological Consultation, 11*(2), 193–221.

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

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

Eikeseth, S., Smith, T., Jahr, E., & Eldevik, S. (2002). Intensive behavioral treatment at school for 4- to 7-year old children with autism: a 1-year comparison controlled study. *Behavior Modification, 26*(1), 49-68.

Eikeseth, S., Smith, T., Jahr, E., & Eldevik, S. (2007). Outcome for children with autism who began intensive behavioral treatment between ages 4 and 7: A comparison controlled study. *Behavior Modification, 31*(3), 264-278.

Eldevik, S., Hastings, R. P., Hughes, J. C., Jahr, E., Eikeseth, S., & Cross, S. (2009). Meta-analysis for early intensive behavioral intervention for children with autism. *Journal of Clinical Child and Adolescent Psychology, 38*(3), 439-450.

Elliott, C. D. (2007a). *Differential Ability Scales – Second Edition: Administration and scoring manual.* San Antonio: Harcourt Assessment.

Elliott, C. D. (1990). *Differential Ability Scales – Second Edition: Introductory and technical handbook.* San Antonio: Harcourt Assessment.

Fantuzzo, J. W., LeBoeuf, W. A., & Rouse, H.L. An investigation of the relations between school concentrations of student risk factors and student educational well-being. *Educational Researcher, 43*(1), 25-36.

Fitzer, A., & Sturmey, P. (2007). *Language and autism: applied behavior analysis, evidence, and practice.* Austin, TX: PRO-ED.

Foxx, R. M. (2008). Applied behavior analysis treatment of autism: The state of art. *Child and Adolescent Psychiatric Clinics of North America, 17(4),* 821-834.

Ghaith, G. & Yaghi, H. (1997). Relationships among experience, teacher efficacy, and attitudes toward the implementation of instructional innovation. *Teaching and Teacher Education, 13*(4), 451-458.

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

Gottfredson, G. D., Jones, E. M., & Gore, T. W. (2002). Implementation and evaluation of a cognitive–behavioral intervention to prevent problem behavior. *Prevention Science, 3*(1), 43-56.

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.

Gregory, A. & Weinstein, R. S. (2008). The discipline gap and African Americans: Defiance or cooperation in the high school classroom. *Journal of School Psychology, 46,* 455-475.

Gresham, F. M., Gansle, K. A., Noell, G. H., & Cohen, S. (1993). Treatment integrity of school-based behavioral intervention studies: 1980–1990. *School Psychology Review, 22,* 254–272.

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.

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.

Hess, K. L., Morrier, M. J., Heflin, L. J., & Ivey, M. L. (2008). Autism treatment survey: services received by children with autism spectrum disorders in public school classrooms. *Journal of Autism and Developmental Disorders, 38,* 961-971.

Hogrebe, M. C. & Tate, W. F. (2010). School composition and context factors that moderate and predict 10th grade science proficiency. *Teachers College Record, 112*(4), 1096-1136.

Howard, M. S., Mann, E. L., & LeBrun, M. R. (2008). Positive behavioral interventions and supports in New Hampshire. *Journal of Positive Behavior Interventions, 10*(3), 190-205.

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,* 369-383.

Howlin, P., Goode, S., Hutton, J., & Rutter, M. Adult outcome for children with autism. *Journal of Child Psychology and Psychiatry, 45*(2), 2004.

Iovannone, R., Dunlap, G., Huber, H., & Kincaid, D. (2003). Effective educational practices for students with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities*, *18*(3), 150-165.

Jacobson, J. W., Mulick, J. A., & Green, G. (1998). Cost-benefit estimates for early intensive behavioral intervention for young children with autism – general model and single state case. *Behavioral Interventions, 13*(4), 201-226.

Johnson, E. & Hastings, R. P. (2002). Facilitating factors and barriers to the implementation of intensive home-based behavioural intervention for young children with autism. *Child: Care, Health, and Development, 28*(2), 123-129.

Kallestad, J. H. & Olweus, D. (2003). Predicting teachers' and schools' implementation of the olweus bullying prevention program: a multilevel study. *Prevention and Treatment*, *6*(21), 1-31.

Kam, C. M., Greenberg, M. T., & Walls, C. T. (2003). Examining the role of implementation quality in school-based prevention using the PATHS curriculum. *Prevention Science, 4*(1), 55-63.

Kasari, C. & Smith, T. (2013). Interventions in schools for children with autism spectrum disorder: Methods and recommendations. *Autism, 17*(3), 254-267.

Klecker, B. M. & Pollock, M. A. (2005). Congruency of research-based literacy instruction in high and low performing schools. *Reading Improvement, 42*(3), 149-157.

Koegel, L. K., Carter, C. M., & Koegel, R. L. (2003). Teaching children with autism self-initiations as a pivotal response. *Topics in Language Disorders, 23*(2), 134-145.

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

Koegel, L. K., Koegel, R. L., & Carter, C. M. (1998). Pivotal responses and the natural language teaching 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.

Kuhn, L. R., Bodkin, A. E., Devlin, S. D., & Doggett, R. A. (2008). Using pivotal response training with peers in special education to facilitate play in two children with autism. *Education and Training in Developmental Disabilities, 43*(1), 37-45.

Landa, R. J., Holman, K. C., O'Neill, A. H., & Stuart, E. A. (2011). Intervention targeting development of socially synchronous engagement in toddlers with autism spectrum disorder: a randomized controlled trial. *The Journal of Child Psychology and Psychiatry, 52*(1), 13-21.

Lochman, J. E. (2001). Issues in prevention with school-aged children: Ongoing intervention refinement, developmental theory, prediction and moderation, and implementation and dissemination. *Prevention and Treatment, 4*(4), 1-7.

Locke, J., Rotheram-Fuller, E., Xie, M., Harker, C., & Mandell, D. Correlation of cognitive and social outcomes among children with autism spectrum disorder in a randomized trial of behavioral intervention. *Autism, 18*(4), 370-5.

Lovaas, O. (1987). Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal Of Consulting And Clinical Psychology, 55*(1), 3-9.

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.

Lydon, H., Healy, O., & Leader, G. (2011). A comparison of Video Modeling and Pivotal Response Training to teach pretend play skills to children with Autism Spectrum Disorder. *Research in Autism Spectrum Disorders, 5*, 872-884.

Lytton, H. & Pyryt, M. (1998). Predictors of achievement in basic skills: a Canadian effective schools study. *Canadian Journal of Education, 23*(3), 281-301.

Ma, X. & Klinger, D. A. (2000). Hierarchical linear modelling of student and school effects on academic achievement. *Canadian Journal of Education*, *25*(1), 41-55.

Mandell, D. S., Stahmer, A. C., Shin, S., Xie, M., Reisinger, E., & Marcus, S. C. (2013). The role of treatment fidelity on outcomes during a randomized field trial of an autism intervention. *Autism, 17*(3), 281-295.

Marquenie, K., Rodger, S., Mangohig, K., & Cronin, A. (2011). Dinnertime and bedtime routines and rituals in families with a young child with an autism spectrum disorder. *Australian Occupational Therapy Journal, 58*, 145–154.

Massey, N., & Wheeler, J. J. (2000). Acquisition and generalization of activity schedules and their effects on task engagement in a young child with autism in an inclusive pre-school classroom. *Education & Training In Mental Retardation & Developmental Disabilities*,*35*(3), 326-335.

Mattison, E. & Aber, M. S. (2007). Closing the achievement gap: the association of racial climate with achievement and behavioral outcomes. *American Journal of Community Psychology, 40*, 1-12.

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

Mesibov, G. B., Shea, V., Schopler, E., Adams, L., Merkler, E., Burgess, S., … Van Bourgondien, M. E. (2004). *The teach approach to autism spectrum disorders*. New York: Springer.

Minjarez, M. B., Williams, S. E., Mercier, E. M., & Hardan, A. Y. (2011). Pivotal response group treatment program for parents of children with autism. *Journal of Autism and Developmental Disorders, 41*, 92-101.

Morrier, M. J., Hess, K. L., & Heflin, L. J. (2011). Teacher training for implementation of teaching strategies for students with autism spectrum disorders. *Teacher Education and Special Education, 34*(2), 119-132.

Morrison, R. S., Sainato, D. M., Benchaaban, D., & Endo, S. (2002). Increasing play skills of children with autism using activity schedules and correspondence training. *Journal of Early Intervention, 25*(1), 58-72.

Morrison, G. M. & Skiba, R. (2001). Predicting violence from school misbehavior: promises and perils. *Psychology in the Schools, 38*(2), 173-184.

National Autism Center (2009). *National standards report.* Randolph, MA: National Autism Center.

National Research Council. (2001). *Educating children with autism*. Washington, DC: National Academy Press.

Odom, S. L., Collet-Klingenberg, L., Rogers, S. J., & Hatton, D. D. (2010). Evidence-based practices in interventions for children and youth with autism spectrum disorders. *Preventing school failure: Alternative education for children and youth, 54*(4), 275–282

O'Donnell, C. L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K-12 curriculum intervention research. *Review of Educational Research, 78*(1), 33-84.

Palardy, G. J. (2008). Differential school effects among low, middle, and high social class composition schools: a multiple group, multilevel latent growth curve analysis. *School Effectiveness and School Improvement, 19*(1), 21-49.

Pas, E. T. & Bradshaw, C. P. (2012). Examining the association between implementation and outcomes: State-wide scale-up of school-wide positive behavior intervention and supports. *Journal of Behavioral Health Services and Research, 39*(4), 417-433.

Peabody, D. (2011). Beliefs and instructional practices among secondary teachers within selected high- and low-performing high schools. *Florida Journal of Educational Administration and Policy, 4*(2), 181-192.

Pellecchia, M. (2013). *Predictors of outcome for children with autism receiving behavioral intervention.* Temple University, Philadelphia, PA.

Perry, L. B. & McConney, A. (2010). Does the SES of the school matter? An examination of socioeconomic status and student achievement using the PISA 2003. *Teachers College Record, 112*(4), 1137-1162.

Perry, L. B. & McConney, A. (2013). School socioeconomic status and student outcomes in reading and mathematics: a comparison of Australia and Canada. *Australian Journal of Education, 57*(2), 124-140.

Peters-Scheffer, N., Didden, R., Korzilius, H., & Sturmey, P. (2013). Therapist characteristics predict discrete trial teaching procedural fidelity. *Intellectual and Developmental Disabilities, 51*(4), 263-272.

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* (3), 285-295.

Raffaele Mendez, L. M. Knoff, H. M., & Ferron, M. (2002). School demographic variables and out-of-school suspension rates: a quantitative and qualitative analysis of a large ethnically diverse school district. *Psychology in the Schools, 39*(3), 259-277.

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.

Robinson, S. E. (2011). Teaching paraprofessionals of students with autism to implement pivotal response treatment in inclusive school settings using a brief video feedback training package. *Focus on Autism and Other Developmental Disabilities, 26*(2), 105-118.

Rogers, S. J. (2000). Interventions that facilitate socialization in children with autism. *Journal of Autism and Developmental Disorders, 30*(5), 399-409.

Rogers, S. J., & Vismara, L. A. (2008). Evidence-Based Comprehensive Treatments for Early Autism. *Journal Of Clinical Child & Adolescent Psychology, 37*(1), 8-38.

Rumberger, R. W. & Palardy, G. J. (2005). Does segregation still matter? The impact of student composition on academic achievement in high school. *Teachers College Record, 107*(9), 1999-2045.

Sallows, G. O. & Graupner, T. D. (2005). Intensive behavioral treatment for children with autism: four-year outcome and predictors. *American Journal of Mental Retardation, 110*(6), 417-438.

Sanetti, L. M., Gritter, K. L., & Dobey, L. M. (2011). Treatment integrity of interventions with children in the school psychology literature from 1995 to 2008. *School Psychology Review, 40*(1), 72-84.

Schulte, A. C., Easton, J. E., & Parker, J. (2009). Advances in treatment integrity research: multidisciplinary perspectives on the conceptualization, measurement, and enhancement of treatment integrity. *School Psychology Review, 38*(4), 460-475.

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.

Smith, T. (1999). Outcome of early intervention for children with autism. Clinical Psychology: Research and Practice, 6, 33-49.

Smith, T. (2001). Discrete trial teaching in the treatment of autism. Clinical Psychology: Research and Practice, 6, 33-49.

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.

Southworth, S. (2010). Examining the effects of school composition on North Carolina student achievement over time. *Education Policy Analysis Archives, 18*(29), 1-42.

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.

Stahmer, A. C. (1999). Using pivotal response training to facilitate appropriate play in children with autistic spectrum disorders. *Child Language Teaching and Therapy*, 29-40.

Stahmer, A. C., Suhrheinrich, J., Reed, S., Bolduc, C., & Schreibman, L. (2011). Pivotal response training in the classroom setting. *Preventing School Failure, 54*(4), 265-274.

Steege, M. W., Mace, C., Perry, L., & Longenecker, H. (2007). Applied behavior analysis: beyond discrete trial teaching. *Psychology in Schools, 44*(1), 91-99.

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.

Suhrheinrich, J. (2011). Training teachers to use pivotal response training with children with autism: coaching as a critical component. *Teacher Education and Special Education, 34*(4), 339-349.

Suhrheinrich, J., Stahmer, A. C., Reed, S., Schreibman, L., Reisinger, E., & Mandell, D. (2013). Implementation challenges in translating pivotal response training into community settings. *Journal of Autism and Developmental Disorders, 43,* 2970-2976.

Sundberg, M. L. (2008) *Verbal behavior milestones assessment and placement program: The VB-MAPP*. Concord, CA: AVB Press.

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

Swanson, E., Wanzek, J., Haring, C., Ciullo, S., & McCully, L. (2011). Intervention fidelity in special and general education research journals. *Journal of Special Education, 47*(1), 3-13.

Thrupp, M. & Lupton, R. (2006). Taking school contexts more seriously: the social justice challenge. *British Journal of Educational Studies, 54*(3), 308-328.

Tincani, M. (2007). Beyond consumer advocacy: Autism spectrum disorders, effective instruction, and public schools. *Intervention in School and Clinic, 43*(1), 47-51.

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.

Volkmar, F. R. & Pauls, D. (2003). Autism. *The Lancet, 362*(9390), 1133-1141.

Wallace, J. M., Goodkind, S., Wallace, C. M., & Bachman, J. G. (2008). Racial, ethnic, and gender differences in school discipline among U.S. high school students: 1991-2005. *The Negro Educational Review, 59*(1), 47-61.

Welch, K. & Payne, A. A. (2010). Racial threat and punitive school discipline. *Social Problems, 57*(1), 25-48.

Wenz-Gross, M. & Upshur, C. (2012). Implementing a primary prevention social skills intervention in urban preschools: factors associated with quality and fidelity. *Early Education and Development, 23*(4), 427-450.

Willms, J. D. (2010). School composition and contextual effects on student outcomes. *Teachers College Record, 112*(4), 1008-1037.

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

Wong, C., Odom, S. L., Hume, K. Cox, A. W., Fettig, A., Kucharczyk, S., … Schultz, T. R. (2013). *Evidence-based practices for children, youth, and young adults with Autism Spectrum Disorder.* Chapel Hill: The University of North Carolina, Frank Porter Graham Child Development Institute, Autism Evidence-Based Practice Review Group.

Wrigley, T. (2006). Schools and poverty: questioning the Effectiveness and Improvement paradigms. *Improving Schools, 9*(3), 273-290.

Young, J. M., Krantz, P. J., McClannahan, L. E., & Poulson, C. L. (1994). Generalized imitation and response-class formation in children with autism. *Journal of Applied Behavior Analysis, 27,* 685–698.

APPENDIX

CORRELATION MATRIX

Measure	1	2	3	4	5
1. School Composition					
2. Academic Performance	-.603***				
3. Disciplinary Practices	.308	-.589***			
4. Fidelity to STAR	-.167	.034	-.087		
5. Number of Years Teaching	-.230	.233	-.149	.069	

*** p <.001