2017

Effectiveness of Community-Based EIBI Treatment: A Longitudinal Analysis of Adaptive Behavior and Language Outcomes

John Kuntz
University of South Carolina

Effectiveness of Community-Based EIBI Treatment: A Longitudinal Analysis of Adaptive Behavior and Language Outcomes

by

John Kuntz

Bachelor of Arts
LeMoyne College, 1983

Master of Science
State University of New York at Stony Brook, 1986

Submitted in Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy in

Epidemiology

The Norman J. Arnold School of Public Health

University of South Carolina

2016

Accepted by:

Melinda Forthofer, Major Professor

Robert McKeown, Committee Member

Suzanne McDermott, Committee Member

Jia-Jia Zhang, Committee Member

Robert Hock, Committee Member

Lacy Ford, Senior Vice Provost and Dean of Graduate Studies

Acknowledgements

I would like to express my gratitude and appreciation to the staff of the South Carolina Department of Disabilities and Special Needs Pervasive Development Disorders Program. They were responsible for initiating and supporting this research. I would like to also express my thanks to Dr. Melinda Forthofer (dissertation advisor), who provided extensive advice, encouragement, and patience and Dr. Robert McKeown, who guided me throughout. Dr. Robert Hock was instrumental in identifying the research opportunity and working closely with me. My other committee members, Dr. Jia-Jia Zhang and Dr. Susan McDermott provided invaluable direction and feedback throughout. I would like to thank my employer, my managers and my coworkers, who supported me and encouraged me in this pursuit. Finally, I would like to thank my family, particularly my sister Bernadette, who showed the way. Most importantly, I would like to thank my amazing children, Stephanie, Corinne, Andrew, Sarah, and Evan who motivate and inspire me every day.

Abstract

Autism spectrum disorders (ASDs) are a group of lifelong, neurodevelopmental disorders characterized by deficits in social interaction, communication, relationship development and by the presence of repetitive or stereotypical behaviors including restricted interests. Continued advances in understanding treatment outcomes and broadening access to effective treatment is critical to improving the quality of life of children with autism and their families and minimizing the cost associated with care. The overall aim of this study is to assess the effectiveness of a community-based implementation of an Early Intensive Behavioral Intervention (EIBI) treatment program through a large-sample, longitudinal secondary analysis of administrative data. Additionally, it identifies baseline characteristics that predict improvements in adaptive behaviors and language. It also assesses the impact of data collection and data management on the internal and external validity of those findings. Using historical data from the South Carolina Department of Disabilities and Special Needs (SC DDSN) Pervasive Development Disorder (PDD) Program, this retrospective cohort study analyzed 615 children, aged 3-8, who had completed two years of EIBI treatment, with treatment beginning in years 2007 through 2011. This study demonstrated statistically significant average gains in adaptive behavior, expressive and receptive language after two years of EIBI treatment. It showed that gains were achieved in each

of the first two years for adaptive behaviors and expressive language while receptive language only showed gains in the first year. It showed that 40% of children experienced gains equivalent to a medium effect size in adaptive behaviors. It demonstrated that age at entry and baseline measures of adaptive behavior and language moderated gains. Lastly, it showed that missing data and incomplete records did not impact the validity of results. As more children are diagnosed with autism and treated through large, community-based programs, the administrative data collected provides a potentially rich source of research data. Given the findings, reinforced here, that only a subset of children are benefitting from EIBI treatment, larger research samples are needed to better explore the moderators of outcomes. By improving data management, data quality and data retention, large, multi-year studies can provide sufficient statistical power to better understand relationships that have a direct impact on program costs.

Table of Contents

List of Tables

List of Figures

Introduction

Autism spectrum disorders (ASDs) are a group of lifelong, neurodevelopmental disorders characterized by deficits in social interaction, communication, relationship development and by the presence of repetitive or stereotypical behaviors including restricted interests (American Psychiatric Association 2013, Johnson, Myers, and the Council on Children with Disabilities 2007). Estimates of autism prevalence have risen from 1 in 2000 in the early 1980s, to 1 in 68 currently (Baio 2014, Newschaffer et al. 2007). While specific ASD symptoms may vary throughout the lifetime, and there is some evidence of non-treatment related improvements in IQ and specific behaviors (Dietz et al. 2007), ASDs are not curable and require chronic management over the lifetime (Myers, Johnson, and the Council on Children with Disabilities 2007). ASDs pose a significant burden for families and society in general. Healthcare expenditures are estimated at up to nine times the lifetime costs of care received by other Medicaid-eligible children and three times those of children with Intellectual Disability (Newschaffer et al. 2007, Bouder, Spielman, and Mandell 2009, Lavelle et al. 2014). Families caring for a child diagnosed with ASD are more likely to face large, out-of-pocket expenditures, stop or reduce work hours, and spend more than 10 hours per week coordinating treatment and care (Kogan et al. 2008, Zablotsky et al. 2014).

Continued advances in understanding treatment outcomes and broadening access to effective treatment is critical to improving the quality of life of children with autism and their families and minimizing the cost associated with care. Conventional treatment focuses on the acquisition of skills commonly lacking in those with ASDs. The most broadly implemented and studied of these methods are based on Applied Behavioral Analysis (ABA) (Foxx 2008, Granpeesheh, Tarbox, and Dixon 2009). Early Intensive Behavioral Intervention (EIBI) is the application of ABA principles for the treatment of young children with ASDs (Peters-Scheffer et al. 2011, Granpeesheh, Tarbox, and Dixon 2009). Since 2007, EIBI has been the preferred treatment for children with autism according to the American Academy of Pediatrics and the National Institute of Child Health and Human Development (Myers, Johnson, and the Council on Children with Disabilities 2007, Harvey et al. 2010). There is general consensus that efficacy of EIBI for children with autism is well established (Eikeseth et al. 2012, Healy and Lydon 2013, Perry et al. 2008). Despite consistent findings regarding overall EIBI outcomes, heterogeneity of results at the individual level is commonly reported with only 25-50% of children receiving EIBI achieving desired outcomes, as defined within the individual studies (Ben-Itzchak and Zachor 2007, Lovaas 1987, Perry et al. 2011). As prevalence of autism has risen, demand for EIBI services and funding for those services has risen as well (Harvey et al. 2010, Reichow 2012, Baio 2014, Newschaffer, Falb, and Gurney 2005). Insurance reform and expansion of coverage has further contributed to increased demand and utilization of EIBI services. This convergence of factors has heightened the need to understand the effectiveness of EIBI in 'real-world' settings

where most children now receive treatment (Perry et al. 2008) and to address research gaps related to EIBI program implementation. The overall aim of this study is to assess the effectiveness of a publically-funded implementation of an EIBI treatment program through a large-sample, longitudinal secondary analysis of administrative data. Using historical data from the South Carolina Department of Disabilities and Special Needs (SC DDSN) Pervasive Development Disorder (PDD) Program, this analysis assesses treatment-related gains throughout a two-year treatment period. Additionally, it identifies baseline characteristics that predict improvements in adaptive behaviors and language. It also assesses the impact of data collection and data management on the internal and external validity of those findings. Finally, it discusses the implications for EIBI program implementation and future research.

Background

The Diagnostic and Statistical Manual of Mental Disorders (DSM) is the manual used by clinicians and researchers to classify mental disorders. In 2013, *The Diagnostic and Statistical Manual of Mental Disorders (DSM), Fifth Addition (DSM-5[1])* was published, replacing DSM-IV TR, which had been in use since 2000 (American Psychiatric Association. 2000, American Psychiatric Association 2013). Due to changes in diagnostic criteria published in DSM-5, individuals who would have previously received a distinct diagnoses for autistic disorder, Asperger's syndrome, or pervasive development disorder not otherwise specified (PDD-NOS) under DSM-IV TR, are likely to receive a

[1] In the latest edition, The Diagnostic and Statistical Manual of Mental Disorders has shifted from using Roman numerals to Arabic numbers

diagnosis of ASD under DSM-5 with specific descriptors that distinguish the individual characteristics associated with Asperger's syndrome or PDD-NOS. For instance, many individuals who would have received a diagnosis of Asperger's syndrome under DSM-IV would receive a diagnosis of autism spectrum disorder "without language or intellectual impairment" (APA, 2013; p. 32) under DSM-5. By collapsing the previously distinct diagnoses, DSM-5 more clearly recognizes that they are part of a spectrum. Despite the collapsed diagnosis, it is important to recognize that intellectual disability remains "perhaps the most common co-occurring disorder" (Matson and Shoemaker, 2009, p.1111). Children with a diagnosis of autism and comorbid intellectual disability show different core-symptoms and respond more poorly to treatment than children with autism without comorbid intellectual disability (Matson and Shoemaker 2009).

Epidemiology

Estimates of the prevalence of ASDs have risen dramatically since the 1980s when they were considered rare disorders with a prevalence of fewer than 5 per 10,000 (Newschaffer et al. 2007). Continued changes in diagnostic criteria, diagnostic substitution, availability of special education services and overall increases in awareness have contributed to the increased prevalence estimates over time (Fombonne 2009, Shattuck 2006, Wing and Potter 2002). Despite explanations for increased prevalence, it is not possible to rule out an overall increase in the incidence of ASDs. ASDs are biologically based with a genetic and heritable component (Newschaffer et al. 2007). Risk of occurrence is 50 to 100 times greater in siblings of those already diagnosed with ASDs (Prater and Zylstra 2002). There are no biological markers of the disorder

(Newschaffer et al. 2007, Johnson, Myers, and and the Council on Children with Disabilities 2007). Substantial heterogeneity exists in presentation and development among children with ASDs (Charman et al. 2011). Coexisting Intellectual Disability (IQ less than or equal to 70) is seen in approximately 31% of the children with an ASD (Baio 2014). This decline from estimates of 90% before the 1990s may be attributable to the increased diagnoses of milder forms of ASDs (Johnson, Myers, and and the Council on Children with Disabilities 2007). Other coexisting conditions may include anxiety, depression, obsessive-compulsive behaviors, phobias and medical conditions including gastrointestinal symptoms and "immune system dysregulation" (Newschaffer et al. 2007 p.238) (Matson and Nebel-Schwalm 2007). Boys are four to five times as likely as girls to receive a diagnosis of ASD. Beyond the genetic component, little consistent evidence has been found regarding the causes of ASDs, although genetic interactions with environmental, immunological and hormonal factors are being investigated (Newschaffer et al. 2007).

Treatment

Due to the lifelong nature of the disorder, the variability of individual deficits, and the sometimes conflicting information regarding potential causes, treatment options for individuals with ASDs have been as varied as the underlying symptoms. Despite no conclusive evidence of physiologic mechanisms, some treatments, including chelation, diet modification, and pharmacological interventions purport to repair or alter the underlying biological mechanisms that result in the symptoms of ASDs (Levy and Hyman 2005, Davis et al. 2013, Goin-Kochel, Mackintosh, and Myers 2009). More

conventional treatments focus on the acquisition of skills commonly lacking in those with ASDs. The most broadly implemented and studied of these methods are based on Applied Behavioral Analysis (ABA) (Granpeesheh, Tarbox, and Dixon 2009, Foxx 2008). ABA refers broadly to a set of treatments based upon behavior analytic methods which trace back to the work of B.F. Skinner (Morris, Smith, and Altus 2005). Behavioral techniques of reinforcement, extinction, generalization and stimulus control are used to obtain desired behaviors (Morris, Smith, and Altus 2005, Granpeesheh, Tarbox, and Dixon 2009). In the treatment of autism, ABA methods are used across age ranges, in a variety of settings (home, community, school), and with few restrictions on intensity, duration, or the requirement of 1:1 interventions. ABA methods could be used to facilitate the acquisition and development of specific skills or improvements on a narrow set of outcomes such as attention, cognitive development, behaviors, or social skills (Williams White, Keonig, and Scahill 2007, Llaneza et al. 2010, Patten and Watson 2011, Patterson, Smith, and Jelen 2010, Ben Itzchak et al. 2008).

Early Intensive Behavioral Intervention (EIBI) is the application of ABA principles for the treatment of young children with ASDs (Peters-Scheffer et al. 2011, Granpeesheh, Tarbox, and Dixon 2009). While there are different models of EIBI interventions, common core features of the treatment include (Virués-Ortega 2010, Foxx 2008, Green, Brennan, and Fein 2002): (1) intensive interventions of between 25 and 40 hours per week (with guidelines recently reduced to lessen the burden on families (Spreckley and Boyd 2009)); (2) treatment that is provided for at least two years; (3) treatment that begins at the earliest possible age, ideally in the pre-school

years; (4) treatment that is individualized, comprehensive and therefore, administered in one-on-one settings; (5) treatment that begins in the home and gradually transitions to other environments (school, community) after the acquisition of appropriate skills; (6) parents who are trained and active in the treatment process; (7) qualified and highly trained staff to manage the treatment. A key principal of these methods is the demonstration that interventions are responsible for observed improvements in behavior. As such, rigorous baseline and ongoing measurement is essential to EIBI programs.

In foundational research by Lovaas, numerous individual studies, and more recent comprehensive evaluations, positive improvements have been attributed to EIBI across the domain of deficits present in autism (Peters-Scheffer et al. 2011, Virués-Ortega 2010, Reichow 2012, Eldevik et al. 2009, Howlin, Magiati, and Charman 2009, Lovaas 1987, Sallows, Graupner, and MacLean Jr 2005, Remington et al. 2007, Howard et al. 2005, Ben Itzchak et al. 2008, Eldevik et al. 2012, Healy and Lydon 2013). Across studies, a variety of outcome metrics are used. In a review of 32 studies, Matson and Goldin (Matson and Goldin 2014) found that adaptive behavior outcomes using Vineland Adaptive Behavior Scales (VABS) (Sparrow, Balla, and Cicchetti 1984, 2005) are measured in 66% of the selected studies. Standardized IQ is measured in 41% of those studies (although instruments differ). In the 22 studies included in Virues-Ortega's meta-analysis (Virués-Ortega 2010), outcome measures include IQ (82%), receptive language (45%), expressive language (41%) and adaptive behaviors (64%). While methodological differences exist within and across studies, there is general consensus

that evidence for the efficacy of EIBI for children with autism is well established (Eikeseth et al. 2012, Healy and Lydon 2013). However, inconsistencies exist regarding the specific outcomes achieved, the percentage of children who experience improvement, and the child-specific factors which predict outcomes with direct implications for managing large community-based programs.

A convergence of factors has led to increased opportunity and increased need to study EIBI effectiveness in large-scale, community-based programs. The dramatic rise in autism prevalence estimates has increased awareness and demand for treatment (Baio 2014, Newschaffer et al. 2007). Concurrently, the cost to society associated with autism is skyrocketing. Accounting for medical care, non-medical care and lost productivity of family member, Leigh and Du estimated the 2015 economic burden of autism in the United States to be $268.3 billion, increasing to $460.8 billion in 2025 (Leigh and Du 2015). The 2015 costs are comparable to the 2012 costs of diabetes ($245 billion). While expensive, EIBI treatment services have shown to be effective in lowering the overall costs to society. In a 1998 study, Jacobson et al. demonstrated a cost avoidance through age 22 of approximately $200,000 and a savings approaching $1.1 million per child through age 55, associated with 3 years of EIBI treatment. This is in spite of annual cost of treatment of more than $33,000 (Jacobson, Mulick, and Green 1998). Chasson et al. saw similar results in 2007, estimating annual savings of $208,000 through 18 years of special education, compared to receiving 3 years of EIBI services (Chasson, Harris, and Neely 2007). EIBI is now considered a well-established intervention, producing outcomes consistent with the 'highest levels of evidence-based treatments'

(Reichow 2012 p. 518, Matson and Jang 2013). Accordingly, demand for EIBI services and funding for those services has risen as well (Harvey et al. 2010, Reichow 2012, Baio 2014, Newschaffer, Falb, and Gurney 2005). Insurance reform and expansion of coverage have further contributed to increased demand and utilization of EIBI services. As of March, 2015, 43 states plus the District of Columbia have legislation mandating insurance coverage for the diagnosis and treatment of autism compared to 15 states with mandated coverage in December, 2009 (National Conference of State Legislatures 2015, Autism Speaks). Finally, beginning in 2014, The Affordable Care Act (ACA) has closed some loopholes and inconsistencies in private insurance and state laws by requiring coverage of behavioral health treatment including EIBI services for autism (National Autism Network). In turn, EIBI research needs have moved beyond establishing efficacy, to demonstrating effectiveness and explaining the heterogeneity of outcomes in real-world settings. This research needs to better inform specific recommendations regarding the duration and intensity of treatment and it needs to better identify the characteristics of the children who experience positive outcomes.

Compared to the small, closely managed settings where efficacy has been widely established, community and state-run programs introduce greater heterogeneity of children enrolled, training and supervision of staff, treatment provided, and consistency and controls of data collection needed to evaluate outcomes (Perry et al. 2008). Turnover of EIBI therapists, lower levels of experience among therapists and supervisors, adherence to EIBI protocols and lower supervision of staff and parents have been identified as obstacles to the success of EIBI treatment in non-clinical settings

(Eikeseth et al. 2012). At the same time, the substantially larger number of children who are treated through these community-managed programs provides a unique opportunity to address existing gaps in the understanding of EIBI treatment effectiveness. A limited number of large-scale studies that assessed effectiveness of EIBI in community settings (Perry et al. 2008, Granpeesheh et al. 2009, Fernell et al. 2011) were found.

Objectives

This study expands the knowledge regarding community-based EIBI interventions through analysis of another large-scale community-based program. Using historical data from the South Carolina Department of Disabilities and Special Needs (SC DDSN) Pervasive Development Disorder (PDD) Program, this retrospective cohort study assesses the impact of data collection and data management on program evaluation for 6 years of enrollment data on over 600 children. It identifies baseline characteristics that predict improvements in adaptive behaviors and language. It assesses treatment-related gains throughout the treatment period. Lastly, it analyzes and assesses the impact of data availability and data management on overall study findings.

This analysis addresses three specific questions:

1. Does completion of two full years of treatment improve child outcomes?
    a. Is there an overall improvement in adaptive behaviors?
    b. Is there overall improvement in expressive and receptive language?

c. Does treatment intensity, measured by weekly treatment hours received predict gains in adaptive behaviors and language? If so, how?

d. Do improvements in adaptive behavior and language vary by length of treatment in months? If so, how?

e. What proportion of children achieve outcomes greater than or equal to a medium effect size in adaptive behavior and language?

2. What are the factors that moderate effects of treatment on change in adaptive behaviors and language?

   a. Do baseline child characteristics (age at enrollment, gender) moderate treatment effects on language and adaptive behavior outcomes? If so, how?

   b. Does baseline assessment of language and adaptive behaviors moderate treatment effects on language and adaptive behavior outcomes? If so, how?

   c. Do program factors (treatment intensity, treatment duration, and cohort) moderate effects on language and adaptive behaviors? If so, how?

3. What baseline child and program factors (cohort, EIBI provider) are correlated with availability and completeness of data for program evaluation among those children who completed at least 2 years of treatment? What are the implications for overall study validity?

Recently, EIBI curricula and associated studies have placed a greater emphasis on measuring changes in adaptive functioning, shifting away from measuring changes in

intellectual and cognitive skills (Healy and Lydon 2013, Eldevik et al. 2009, Matson and Smith 2008, Howlin, Magiati, and Charman 2009). Eldevik et al. suggested that outcomes of adaptive behaviors "tell us more about the children's skills in daily life" (Eldevik et al. 2009 p. 448). Matson and Smith concluded that IQ "may not be an appropriate dependent measure" of EIBI research due to difficulty in obtaining reliable and valid IQ data, and the relative stability of the IQ scores (Matson and Smith 2008 p. 69). The study analyzes outcomes in adaptive behaviors, receptive and expressive language using the following measures:

- Receptive language using the Peabody Picture Vocabulary Test IV (PPVT) (Dunn and Dunn 1981)
- Expressive language using the Expressive Vocabulary Test (EVT) (Williams 1997)
- Adaptive behaviors using the Vineland Adaptive Behavior Scales-II (VABS) (Sparrow, Balla, and Cicchetti 2005).

Chapter 1

Effectiveness of Community-Based EIBI Treatment: A Longitudinal Analysis of Overall Adaptive Behavior and Language Outcomes[2]

[2] Kuntz, J.M. To be submitted to *Research in Autism Spectrum Disorders.*

1.1 Introduction

Autism spectrum disorders (ASDs) are a group of lifelong, neurodevelopmental disorders characterized by deficits in social interaction, communication, relationship development, and by the presence of repetitive or stereotypical behaviors including restricted interests (American Psychiatric Association 2013, Johnson, Myers, and and the Council on Children with Disabilities 2007). Recent prevalence estimates of the disorder are 14.7 per 1000 (1 in 68) (Baio 2014). Coexisting Intellectual Disability (IQ less than or equal to 70) is seen in approximately 31% of the children with an ASD (Baio 2014). The co-occurrence of ASD and ID has declined since the 1990s when it was estimated to be 90%. It is likely that diagnosis of milder forms of ASD including Asperger's account for this change (Johnson, Myers, and and the Council on Children with Disabilities 2007). While specific ASD symptoms may vary throughout the lifetime, and there is some evidence of non-treatment related improvements in IQ and specific behaviors (Dietz et al. 2007), ASDs are not curable and require chronic management (Myers, Johnson, and the Council on Children with Disabilities 2007). Much of the research on EIBI has established the efficacy of the treatment if a research environment (Reichow 2012, Healy and Lydon 2013, Eikeseth et al. 2012, Perry et al. 2008). Continued advances are needed in understanding treatment outcomes in community settings to improve the quality of life of children with autism and their families and to minimize the cost associated with care.

1.1.1 Treatment

Due to the lifelong nature of the disorder, the variability of individual deficits, and the sometimes conflicting information regarding potential causes, treatments for individuals with ASDs have been as varied as the underlying symptoms. Despite no conclusive evidence of physiologic mechanisms, some treatments, including chelation, diet modification, and pharmacological interventions purport to repair or alter the underlying biological mechanisms that result in the symptoms of ASDs (Levy and Hyman 2005, Davis et al. 2013, Goin-Kochel, Mackintosh, and Myers 2009). More conventional treatments focus on the acquisition of skills commonly lacking in those with ASDs. The most broadly implemented and studied of these methods are based on Applied Behavioral Analysis (ABA) (Granpeesheh, Tarbox, and Dixon 2009, Foxx 2008). ABA methods are generally aimed at the acquisition and development of specific skills or improvements on a narrow set of outcomes such as attention, cognitive development, behaviors, or social skills (Williams White, Keonig, and Scahill 2007, Llaneza et al. 2010, Ben-Itzchak et al. 2008, Patten and Watson 2011, Patterson, Smith, and Jelen 2010). Early Intensive Behavioral Intervention (EIBI) is the application of ABA principles for the treatment of young children with ASDs (Peters-Scheffer et al. 2011, Granpeesheh, Tarbox, and Dixon 2009). EIBI programs are designed to address the variety of deficits commonly observed among children with ASDs, through early, intensive, structured intervention. While there are different models of EIBI interventions (Foxx 2008, Green, Brennan, and Fein 2002, Virués-Ortega 2010), common core features of the treatment include: (1) intensive interventions of between 25 and 40 hours per week (with

guidelines recently reduced to lessen the burden on families (Spreckley and Boyd 2009)); (2) treatment that is provided for at least two years; (3) treatment that begins at the earliest possible age, ideally in the pre-school years; (4) treatment that is individualized, comprehensive and therefore, administered in one-on-one settings; (5) treatment that begins in the home and gradually transitions to other environments (school, community) after the acquisition of appropriate skills; (6) parents who are trained and active in the treatment process; and (7) qualified and highly trained staff to manage the treatment.

EIBI is one of the most studied forms of treatment of ASDs (Reichow 2012). From initial work by Lovaas (Lovaas 1987) to more recent comprehensive evaluations, positive improvements have been attributed to EIBI across the domain of deficits present in autism (Peters-Scheffer et al. 2011, Virués-Ortega 2010, Reichow 2012, Eldevik et al. 2009, Howlin, Magiati, and Charman 2009). Despite methodological differences within and across studies, there is general consensus that evidence for the efficacy of EIBI for children with autism is well established (Eikeseth et al. 2012, Healy and Lydon 2013, Perry et al. 2008). In Lovaas' foundational research on EIBI treatment (Lovaas 1987) among children with a pre-treatment age of less than 40 months, 9 out of 19 children (47%) who received high intensity treatment for 2 years achieved normal intellectual and educational functioning, compared to just 2% of children across 2 control groups (n=40). Since then, numerous individual studies evaluating the efficacy of EIBI in comparison with an alternative-treatment control group have generated similarly positive findings (Sallows, Graupner, and MacLean Jr 2005, Remington et al.

2007, Healy and Lydon 2013, Ben-Itzchak et al. 2008, Howard et al. 2005, Eldevik et al. 2012).

Across individual studies, a variety of outcome metrics are used. In their review of 11 studies, Howlin et al. (2009) discuss variation in outcomes measured and inconsistent tests from 'child to child and from baseline to follow-up within the same study' (Howlin, Magiati, and Charman 2009 p. 29). In a review of 32 studies, Matson and Goldin (Matson and Goldin 2014) found that adaptive behavior outcomes using Vineland Adaptive Behavior Scales (VABS) (Sparrow, Balla, and Cicchetti 1984, 2005) are measured in 66% of the selected studies. Standardized IQ is measured in 41% of those studies (although instruments differ). 35 other measures are used infrequently, including measures to assess the core symptoms of ASDs. In the 22 studies included in Virues-Ortega's meta-analysis (Virués-Ortega 2010), outcome measures include IQ (82%), receptive language (45%), expressive language (41%) and adaptive behaviors (64%). The format used to deliver the EIBI program also can vary widely and is often not clearly described (Howlin, Magiati, and Charman 2009, Reichow 2012, Matson and Jang 2013). Initial EIBI programs based on Lovaas' original description use discrete trials training (DTT), where hundreds of individual learning opportunities were leveraged throughout the day (Granpeesheh, Tarbox, and Dixon 2009). Differences in treatment formats, and lack of clarity regarding the specific methods used also inhibit the ability to compare across studies.

These differences limit the ability to generalize and compare individual study findings (Healy and Lydon 2013), but four descriptive review papers (Howlin, Magiati, and Charman 2009, Granpeesheh, Tarbox, and Dixon 2009, Matson and Smith 2008, Rogers and Vismara 2008) and six meta-analyses (Spreckley and Boyd 2009, Eldevik et al. 2009, Reichow and Wolery 2009, Virués-Ortega 2010, Makrygianni and Reed 2010, Peters-Scheffer et al. 2011) have attempted to bridge these differences to draw conclusions regarding EIBI outcomes. Virues-Ortega's meta-analysis across 22 qualifying studies (2010), including 323 subjects found positive treatment effects across outcomes including IQ, receptive and expressive language skills, and multiple adaptive behaviors, such as daily living skills, motor skills, and socialization. The analysis of Eldevik et al. (2009) included 9 studies with 153 treatment subjects and concludes that the growing body of evidence supports the effectiveness of EIBI in improving intellectual, social, language, and adaptive functioning of young children. Reichow and Wolery (2009) analyzed 11 studies with 251 treatment subjects concluding that on average EIBI is an effective treatment, with significant gains seen in average IQ. Peters-Scheffer et al. (2011) reviewed 11 studies with 344 participants. They found improvements associated with EIBI treatment in IQ, adaptive behaviors, and language. Makrygianni and Reed (2010) included 14 studies with 303 subjects. They found EIBI to be "very effective in improving intellectual, language, communication and social abilities of children with ASDs" (Makrygianni and Reed 2010 p. 588). Reichow (2012) provided a comprehensive review of five meta-analyses including important discussions of the inclusion criteria and characteristics of the individual studies comprising each of the meta-analyses (Reichow

and Wolery 2009, Eldevik et al. 2009, Makrygianni and Reed 2010, Virués-Ortega 2010, Spreckley and Boyd 2009, Reichow 2012). Despite overlap in individual studies and methodological differences between meta-analyses, Reichow (2012) concluded that EIBI can produce, on average "large gains in IQ and/or adaptive behaviors for many children with ASDs" (p517) consistent with "the highest levels of evidence-based treatments" (p518). The weighted mean effect sizes for IQ and adaptive behavior ranged from .38–1.19 and .30–1.09, respectively across the individual meta-analyses contained in the study.

Efficacy studies are designed to evaluate interventions under optimum conditions while effectiveness studies assess impacts under "real-world" conditions (Flay 1986). Much of the research on EIBI has established the efficacy of the treatment. Compared to the small, closely managed settings where efficacy has been widely established, community and state-run programs introduce greater heterogeneity of children enrolled, training and supervision of staff, treatment provided, and consistency and controls of data collection needed to evaluate outcomes (Perry et al. 2008). Turnover of EIBI therapists, lower levels of experience among therapists and supervisors, adherence to EIBI protocols and lower supervision of staff and parents have been identified as obstacles to the success of EIBI treatment in non-clinical settings (Eikeseth et al. 2012). There are few studies which test the effectiveness of EIBI treatment in a natural community setting, where the 'vast majority of children' now receive treatment (Perry et al. 2008 p. 623). Many of these effectiveness studies have had small samples with treatment occurring in a limited time frame (Sallows, Graupner,

and MacLean Jr 2005, Hayward et al. 2009, Sheinkopf and Siegel 1998, Eldevik et al. 2012, Magiati, Charman, and Howlin 2007, Cohen, Amerine-Dickens, and Smith 2006, Howard et al. 2005). Therefore, even these effectiveness studies do not address the unique challenges of a large-scale implementation, including quality control, data management, program adherence and subject variability.

Several larger studies assessing the effectiveness of EIBI in community settings (Perry et al. 2008, Fernell et al. 2011, Smith, Klorman, and Mruzek 2015, Bibby et al. 2002, Flanagan, Perry, and Freeman 2012, Magiati, Charman, and Howlin 2007) were found, including 2 using data from the same Ontario, Canada EIBI implementation (Flanagan, Perry, and Freeman 2012, Perry et al. 2008). Adaptive behaviors using Vineland were the most consistently assessed outcomes across these studies. Bibby et al. (2002) found a significant mean increase of 8.9 points in the Vineland composite standard score after approximately 2 years of EIBI treatment for a limited group (n=21) of children for whom pre-treatment measures were available. Fernell et al. (2011) compared 93 children who received a high-intensity EIBI intervention to 105 children who received less-intensive intervention based upon other behavioral methods. They found a small but significant increase in Vineland composite standard scores across groups. Flanagan, Perry, & Freeman (2012) show a significant difference in pre-post average Vineland composite standard scores compared to an individually matched waitlist control. Importantly, they saw limited gains in the EIBI treatment group, while the control group average scores decreased over the study period. Magiati et al., (2007) also saw gains in Vineland age-equivalent scores and in raw scores of receptive and

expressive language, but they observed reductions in Vineland composite standard scores. In a study of 71 children who received 2 years of treatment in a publically-funded, community-based EIBI treatment program, Smith, Klorman, & Mruzek (2015) found improvements in the Vineland composite score standardized for the sample's mean and SD from intake through Year 2 but not on the overall composite standard score. Perry et al. (2008) analyzed the effectiveness of EIBI treatment in a community-based sample of 332 children who received 20-40 hours of weekly EIBI treatment for 18 months. The environmental factors discussed in the Perry et al. article are much more representative of the current study, making these results particularly relevant. They showed statistically significant gains in Vineland age equivalents across all domains and small but significant gains in the composite standard score.

Despite compelling evidence of improvements in adaptive behaviors, inconsistencies exist regarding the specific outcomes achieved and the percentage of children who achieve positive outcomes. More than two decades after Lovaas' initial conclusions regarding the benefits of EIBI, Goin-Kochel, Mackintosh & Myers (2009) conclude that "we are currently unable to predict which children will respond to particular treatments, what intensity of treatment might make a difference, and what behaviors the treatments might affect" (Goin-Kochel, Mackintosh, and Myers 2009 p. 529). Effects of treatment intensity and duration have direct implications for managing costs in large community-based programs.

1.1.1.1 Treatment intensity and duration

Virues-Ortega (2010) provided evidence of a dose-response effect associated with treatment intensity and duration "that is not obvious from the simple examination of individual studies" (Virués-Ortega 2010 p. 397). Specifically, he found that improvement in adaptive behaviors varied by treatment intensity, but not by treatment duration. He also found that treatment gains in receptive and expressive language varied by treatment duration, but not by treatment intensity. For both language and adaptive behaviors, he found that a dose-response existed for total treatment duration (intensity times duration). Makrygianni & Reed (2010) found that treatment intensity of greater than 25 hours per week resulted in greater improvements in adaptive behaviors compared to a lower-intensity group. Their conclusions regarding the duration of treatment were mixed, with the suggestion that several years of treatment did not necessarily maintain progress and that "program effectiveness varies independently from program duration" (Makrygianni and Reed 2010 p. 586). They did not find differences in impacts on language associated with varying intensity or duration. Using participant data from multiple studies, Eldevik et al. (2010) concluded that intensity of greater than 36 hours weekly delivers superior outcomes on adaptive behaviors. Fernell et al. (2011) found that a low-intensity EIBI treatment group actually outperformed a high-intensity group on the VABS composite standard score of adaptive behaviors. In their large community-based study, Perry et al., (2008) did not look at intensity of treatment. They did find that the 11% of the children in their study who achieved the

best outcome did have longer average treatment duration (25.9 months versus 17.7 months) than those children with lesser progress.

There is a dearth of research evaluating duration effects across multiple measurement periods using Vineland standard scores. Most studies discussed thus far evaluated change from baseline to a single follow-up point or reported outcomes on raw scores or age-equivalent scores. Eikeseth et al. (2012) observed significant gains in VABS composite score during the first year of treatment, but no significant gains from the end of year 1 to the end of year 2. Bibby et al. (2002) saw no change in Vineland composite scores from end of year 1 to the end of year2, despite continued improvements in age-equivalent composite score. Similarly, Cohen, Dickens & Smith (2006) showed gains in Vineland composite scores in the first year, but no improvement in the second year. Understanding how improvement varies through time across treatment domains can have a substantial impact on managing program costs. This study will advance the very limited knowledge of EIBI treatment duration impact across multiple measurement periods.

1.1.1.2 Range of improvement

Just as variability exists in outcomes used to assess overall program improvement, there is little consistency in determining whether individual children experience meaningful change. Remington et al. (2007) used the Reliable Change Index (Jacobson and Truax 1991, Remington et al. 2007), which classifies a change as reliable only if that change is greater than that which would be expected due to sampling and/or

measurement error (Eldevik et al. 2010). Perry et al. classified children into 7 categories of change based on multiple criteria including VABS Adaptive Behavior Composite Scores and found that more than half of the children experienced 'very successful' outcomes (25%) or clinically significant improvement (30%), which was calculated based on change in development rate (Perry et al. 2008, Perry et al. 2009, Freeman and Perry 2010). Bibby et al. reported that 7 out of 21 children improved their VABS Composite Standard Score by more than 15 pts from baseline to the end of year 2 (Bibby et al. 2002). No additional studies were found which describe and classify the range of outcomes observed on adaptive behavior or language outcomes.

Collectively, these studies provide some support for EIBI effectiveness in the most common settings for implementation. As shown, few studies can be considered a large-scale implementation of EIBI in a non-clinical setting. This study seeks to expand the knowledge regarding community-based EIBI interventions through analysis of another large-scale community-based program. This study will advance the knowledge of the overall effectiveness and the range of outcomes across multiple measurement periods.

1.1.2 SC EIBI program and the current study

During the 2006 legislative session of the South Carolina General Assembly, $3 million was appropriated to the South Carolina Department of Disabilities and Special Needs (DDSN) to develop the Pervasive Developmental Disorder (PDD) Program by January 2007. During the 2007 session, the General Assembly appropriated an

additional $4.5 million demonstrating their commitment to the treatment of autism and other PDDs. As of March, 2013, 1526 children have received PDD services under this program. The following description is excerpted from the DDSN manual (South Carolina Department of Disabilities and Special Needs 2013).

> The purpose of the PDD Program is to provide intensive in-home intervention to children ages 3 through 10 years diagnosed with a Pervasive Developmental Disorder, which includes Autism, Asperger's and PDD – NOS (Not Otherwise Specified) (prior to changes in diagnostic criteria introduced in DSM-5). Children must be eligible to receive Medicaid or have documentation of financial ineligibility. Children must meet the Level of Care (LOC) assessment requiring the degree of care that would be required in an Intermediate Care Facility for the Intellectually Disabled (ICF/ID). Children who meet these criteria, and receive a PDD diagnosis before age eight may receive Early Intensive Behavioral Intervention (EIBI) services for three years or until their 11th birthday, whichever comes first. The EIBI services are designed to develop skills in cognition, behavior, communications and social interaction. They are provided face-to-face in the child's natural environment, which may include the home or community locations. This environment specifically excludes any educational setting where educational services are simultaneously provided during school hours.

The overall aim of this retrospective cohort study is to assess the effectiveness of a community-based EIBI treatment program through a large-sample, longitudinal analysis of administrative data. Using historical data from the South Carolina Department of Disabilities and Special Needs (SC DDSN) Pervasive Development Disorder (PDD) Program, this analysis answers three questions:

1. Were there statistically significant improvements in adaptive behavior, expressive language and receptive language after two years of EIBI treatment?
2. How does treatment duration and intensity impact adaptive behavior and language outcomes?
3. What was the numerical range of outcomes experienced by individual children in adaptive behavior and language?

1.2 Method

1.2.1 Participants

This study examines program participants who began EIBI treatment in program cohort years 2006-2011. During the study time period, 948 children were enrolled with 615 having completed at least two years of EIBI treatment as of March, 2013. Data were obtained from stored paper files only for those children who had completed at least two years of treatment and entered into an Excel spreadsheet by two graduate assistants under the direction of the study author and the direct supervision of SC DDSN personnel. Additional paper files containing the original diagnostic measures for these children were obtained from the 4 Regional DDSN Care Centers. Similar to the study

conducted by Perry et al. (2008), individual files may have had some incomplete data, but the final study population consisted of those children who had data from at least two measurement periods for any of the 3 primary outcomes: adaptive behaviors, expressive language or receptive language. One observation was removed from the study based on a visual inspection of the data which showed an unreasonably high baseline VABS Composite Standard Score (140) and high baseline Vineland Communications domain score (137) compared to extremely low baseline language scores (EVT and PPVT equal to 23 and 29 respectively). This resulted in a final study population of 419 children. This is almost two-thirds of the population of 615 children who had completed two years of treatment and 42% of the 948 children who had entered the program during the study period. Of the 948 children, 333 children did not complete two years of treatment due to aging, moving or otherwise leaving the program voluntarily. Inclusion in the study was based solely on the availability of the data. Because the final sample represents a substantial proportion of those who completed treatment and inclusion in the sample was not based on specific child characteristics, these results provide insight into treatment effectiveness for the full population of children completing treatment.

1.2.2 Measures

Recently, EIBI curricula and associated studies have placed a greater emphasis on measuring changes in adaptive functioning, shifting away from measuring changes in intellectual and cognitive skills (Healy and Lydon 2013, Eldevik et al. 2009, Matson and

Smith 2008, Howlin, Magiati, and Charman 2009). Eldevik et al. (2009) suggested that outcomes of adaptive behaviors "tell us more about the children's skills in daily life" (p448). Matson and Smith (2008) concluded that IQ "may not be an appropriate dependent measure" of EIBI research due to difficulty in obtaining reliable and valid IQ data, and the relative stability of the IQ scores (p69). This study focuses on improvements in adaptive behaviors, receptive and expressive language. Upon acceptance into the program, the assessment conducted by the EIBI consultant includes the following measures which are the focus of the analyses in this dissertation:

- Receptive language using the Peabody Picture Vocabulary Test IV (PPVT) (Dunn and Dunn 2012)
- Expressive language using the Expressive Vocabulary Test (EVT) (Williams 1997, Williams White, Keonig, and Scahill 2007)
- Adaptive behaviors using the Vineland Adaptive Behavior Scales-II (VABS) (Sparrow, Balla, and Cicchetti 2005)

PPVT assesses vocabulary knowledge and understanding and is determined by having a respondent identify a picture which best represents a word provided by the examiner (Hoffman, Templin, and Rice 2012). EVT measures vocabulary and word retrieval by having the child provide a synonym for a given word, or label a picture for the word provided by an examiner (Roberts et al. 2007). VABS is used to assess daily personal and social functioning and includes standardized scores (M=100.0; S.D. = 15) on Communication, Socialization and Daily Living Skills as well as an Adaptive Behavior Composite (ABC) across those dimensions. It also includes standardized scores for

Motor Skills and Maladaptive Behaviors. Additionally, the available data includes sub-domain-level age equivalent and v-scale scores.

EIBI consultants are required to submit annual updated assessments to the case manager of each of the above measures. Beyond the assessment data, weekly treatment hours allocated and actual treatment hours delivered by the EIBI consultant, the lead therapist and the line therapist were captured. In addition to assessment and treatment data, additional child-specific data captured from the SC DDSN files included age-at-intake, gender and race and year of enrollment (cohort).

Each child enrolled in the program received a pre-treatment diagnosis of autism with evidence from at least two of the following:

- Autism Diagnostic Observation Schedule (ADOS) (Lord et al. 1999)
- Autism Diagnostic Interview (ADI) (Lord and Rutter 1994)
- Social Communication Questionnaire (SCQ) (Rutter, Bailey, and Lord 2003)
- Childhood Autism Rating Scale (CARS) (Schopler, Reichler, and Renner 1988)

These measurements were conducted only at initial diagnosis and are not captured on a pre-post treatment basis. IQ was also infrequently and inconsistently captured only in the pre-treatment timeframe. Therefore, changes in these measures cannot be evaluated as part of this evaluation. Tables 1.1 and 1.2 provide baseline measures and descriptions of the final sample. Table 1.3 provides average scores across each measurement period.

Table 1.1 Baseline measures - means

Measure	N	Mean	Std Dev	Range Min	Range Max
Age at Enrollment (years)	418	5.9	1.2	3.3	8.8
Assessment					
Expressive Vocabulary Test (EVT)	370	66.0	30.5	20.0	126.0
Receptive Vocabulary Test (PPVT)	364	67.1	32.1	20.0	133.0
Vineland Adaptive Behavior (VABS) Standard Scores					
Communications	392	69.4	17.8	22.0	117.0
Daily Living Skills	391	69.0	15.2	25.0	114.0
Socialization	392	65.9	12.7	23.0	122.0
Motor Skills	373	73.6	14.1	37.0	121.0
Adaptive Behavior Composite (ABC)	375	67.0	12.8	42.0	103.0
ADOS Mod1 Communication Total	167	5.5	2.1	0.0	18.0
IQ	55	73.3	18.1	26.0	109.0
Social Communication Questionnaire (SCQ) Total Score	109	20.2	6.1	5.0	34.0
Childhood Autism Rating Scale (CARS)Total Score	169	35.9	5.2	21.0	49.0

Table 1.2 Baseline measures - frequencies

Measure		Frequency	% of Total
Age at Enrollment			
	3	9	2.1%
	4	114	27.2%
	5	109	26.0%
	6	94	22.4%
	7	66	15.8%
	8	26	6.2%
	Missing	1	0.2%
Total		419	100%
Gender			
	Female	56	13.4%
	Male	318	75.9%
	Missing	45	10.7%
Total		419	100%
Enrollment Year			
	2007	62	14.8%
	2008	70	16.7%
	2009	81	19.3%
	2010	141	33.7%
	2011	65	15.5%
Total		419	100%
Diagnosis			
	Asperger's	13	3.1%
	Autism	304	72.6%
	PDD/Other	46	11.0%
	Rett's	1	0.2%
	Missing	55	13.1%
Total		419	100%

Table 1.3 Average score by measurement period

	Baseline			End of Year 1			End of Year 2		
Measure	N	Mean	S.D	N	Mean	S.D	N	Mean	S.D
EVT	364	67.05	32.11	379	72.35	32.34	377	74.35	32.09
PPVT	370	65.97	30.50	384	71.78	31.16	380	72.32	31.10
ABC	375	66.99	12.84	387	71.50	15.08	392	73.99	15.96
Communication	392	69.35	17.85	390	73.97	18.89	402	75.48	18.89
Daily Living Skills	391	69.00	15.20	390	74.08	16.99	402	76.79	17.17
Socialization	392	65.88	12.73	389	70.54	15.08	405	72.41	15.98
Motor Skills	373	73.55	14.10	360	76.95	15.78	355	80.57	17.10

1.2.3 Intervention

EIBI service providers are selected by the child's parents. Authorized service providers, who are individually vetted by SC DDSN, are responsible for recruiting, hiring, retaining, and terminating employees. EIBI services have the oversight of an EIBI Consultant. Five separate components are included:

1. Assessment of the child's current needs, conducted by the EIBI consultant.

2. Program Development and Training, provided by the EIBI consultant, involving the development of an individualized treatment plan and provides training to family members and therapists who implement the individual interventions.

3. Plan implementation, which is also provided by the EIBI Consultant and involves implementation of the plan, monthly monitoring of the effectiveness of the plan and supervision of the Lead and Line therapists who deliver the individual plan.

4. Lead therapy, which is provided by a Lead Therapist and involves oversight and weekly monitoring of the implementation and effectiveness of the plan and review of all recorded data.

5. Line therapy, which is provided by a Line Therapist who is responsible for carrying out the treatment plan as designed and recording data associated with monitoring and measuring outcomes.

Providers are required to ensure adherence to the Essential Practice Elements of ABA (per the Behavior Analyst Certification Board, Inc. Guidelines) through all phases of assessment and treatment as described in the SCDDSN PDD Manual.

1.3 Results

Changes in baseline scores through the end of two years of treatment were assessed using the t-tests and the Signed Rank test from the SAS Proc Univariate procedure. EVT and PPVT baseline scores were not normally distributed with more than 10% of the children receiving a baseline standardized score of 20. Therefore, non-parametric tests are used for these measures. VABS Composite Scores and domain-specific scores had slight divergence from the normal distribution based on visual inspection of Q-Q plots. For these measures both parametric and non-parametric results will be presented. Results are shown in Table 1.4.

1.3.1 Question 1: Were there overall improvements in language and adaptive behaviors?

Statistically significant improvements were seen in overall adaptive behaviors levels and individual domains ($p < .001$) as measured by VABS standard scores, based upon non-parametric tests of the median differences between the assessment at the end of year 2, and the assessment at intake. Statistically significant improvements were seen in expressive language (EVT: $p < .001$) and receptive language (PPVT: $p < .001$) based upon non-parametric tests of the median differences between the assessment at

Table 1.4 Change in baseline measures after 2 years of treatment

	Baseline			End of Year 2			Difference								Effect Size
Measure	N	Mean	S.D	N	Mean	S.D	N	Mean	S.D	t	Pr >= |t|	Median	S	Pr >= |S|	
EVT	335	66.8	32.3	335	74.2	32.1						3.0	10357	<.0001	0.23
PPVT	339	65.6	30.4	339	72.1	31.2						3.0	9801	<.0001	0.21
ABC	353	67.0	13.0	353	73.7	16.1	353	6.7	10.0	12.6	<.0001	6.0	20016	<.0001	0.46
Communication	374	69.0	18.0	374	75.2	18.9	374	6.2	11.0	10.9	<.0001	6.0	19588	<.0001	0.34
Daily Living Skill	372	68.8	15.2	372	76.3	16.8	372	7.5	13.0	11.2	<.0001	6.5	19448	<.0001	0.47
Socialization	376	65.8	12.8	376	72.1	16.1	376	6.3	11.8	10.4	<.0001	5.0	19540	<.0001	0.44
Motor Skills	324	73.5	14.3	324	80.2	16.9	324	6.6	13.3	8.9	<.0001	5.0	11825	<.0001	0.42

the end of year 2, and the assessment at intake (baseline) (Table 1.4). Effect sizes were small for Expressive and Receptive language and the Communications subdomain of Vineland and larger for overall Adaptive Behavior Composite Score and the Daily Living Skills and Socialization Domains (Fritz, Morris, and Richler 2012).

1.3.2. Question 2: How does treatment duration and intensity impact adaptive behavior and language outcomes?

1.3.2.1 Treatment Duration

All children in the study received at least two years of treatment so analysis of the effects of treatment duration is based upon differences in changes observed between the first and second years of treatment. Table 1.5 provides estimates of individual changes during the first and second year of treatment for any child who has either pair of measures. Significant changes are seen during Year 1 across all measures. No significant change is observed in Year 2 for Receptive language and Communication Domain. For those children with all 3 measures (Table 1.6) a significant difference is seen on paired comparisons of Year 1 and Year 2 changes on all measures except Motor Skills. Unlike the changes within individual years, the differences in changes between years are approximately normally distributed, so non-parametric results are not included.

1.3.2.2 Treatment intensity

Treatment intensity is defined as the total number of hours per week of therapy hours billed by the EIBI consultant, the lead therapist, and the line therapist. At the direction of SC DDSN personnel, hours billed during the first complete calendar year of

Table 1.5 Changes in outcomes: year 1 versus year 2

			Year 1 Difference					
Measure	N	Mean	S.D	t	Pr >= |t|	Median	S	Pr >= |s|
EVT	335	5.44	13.18	7.56	<.0001	2.0	9685	<.0001
PPVT	340	5.66	13.21	7.90	<.0001	3.0	10836.5	<.0001
ABC	350	4.57	9.07	9.43	<.0001	4.0	16161.5	<.0001
Communication	364	5.00	10.24	9.32	<.0001	4.0	15904	<.0001
Daily Living Skills	363	5.25	11.97	8.35	<.0001	4.0	14668.5	<.0001
Socialization	363	4.87	10.20	9.09	<.0001	4.0	15208	<.0001
Motor Skills	329	3.59	11.75	5.54	<.0001	3.0	7887.5	<.0001

			Year 2 Difference					
Measure	N	Mean	S.D	t	Pr >= |t|	Median	S	Pr >= |s|
EVT	351	1.87	9.90	3.53	0.000	0.0	4506	0.001
PPVT	353	0.39	12.16	0.61	0.543	0.0	-623	0.676
ABC	366	2.23	7.78	5.49	<.0001	2.0	10735.5	<.0001
Communication	375	1.16	9.44	2.37	0.018	1.0	6363.5	0.000
Daily Living Skills	374	2.46	10.74	4.43	<.0001	2.0	9179	<.0001
Socialization	376	1.67	8.99	3.60	0.000	2.0	8859.5	<.0001
Motor Skills	317	3.76	9.33	7.18	<.0001	3.0	8145.5	<.0001

Table 1.6 Comparison of changes between year 1 and year 2

	Year 1 Difference		Year 2 Difference		Difference			
Measure	Mean	S.D	Mean	S.D	Mean	S.D	t	Pr >= |t|
EVT (n=310)	5.70	13.41	1.61	9.44	-4.1	16.7	-4.3	<.0001
PPVT (n=313)	6.02	13.59	0.59	11.89	-5.4	19.7	-4.9	<.0001
ABC (n=330)	4.68	9.15	2.15	7.98	-2.5	13.8	-3.3	0.002
Communication (n=348)	5.21	10.39	1.21	9.65	-4.0	16.7	-4.5	<.0001
Daily Living Skills (n=346)	5.36	12.16	2.29	10.69	-3.1	18.8	-3.0	0.005
Socialization (n=349)	4.97	10.35	1.56	9.15	-3.4	15.6	-4.1	<.0001
Motor Skills (n=291)	3.20	11.74	3.45	9.12	0.3	16.0	0.3	0.672

treatment were deemed to be representative of the weekly treatment hours received during the first two years of treatment. Table 1.7 provides differences in two-year outcomes from baseline, presented by average weekly treatment hours received.

Differences are observed in the amount of improvement seen in baseline measures corresponding to increases in treatment intensity, particularly with the highest levels of treatment intensity. However, using SAS' Proc Mixed to conduct a repeated measures analysis with EVT, PPVT and ABC Composite as the outcome variables, there are no statistically significant differences associated with changes in average weekly therapy hours on any of the outcome measures where hours was analyzed as a continuous variable. Figures 1.1 - 1.3 show the range of outcomes in changes to EVT, PPVT and ABC by average weekly hours. As shown, despite observed increases in scores, substantial variability exists within categories of weekly hours. Additional Post-hoc analysis also showed no significant differences in outcomes comparing those who averaged more than 27 hours of total weekly treatment to those who received less than 27 hours of total weekly treatment or to those who received less than 16 hours of total weekly treatment. Using SAS' Proc GLMPOWER, it was determined that a sample size of approximately 5600 would have been required in order to find that the observed ABC outcomes were statistically different by treatment hours received. Alternately, at the existing sample sizes, average two-year gains in ABC among those with the highest treatment intensity (weekly hours > 27) needed to be approximately 11 points higher than those with the lowest treatment intensity (weekly hours < 16). With an observed difference in change in ABC score of just over 2 points

Table 1.7 Change in baseline measures after 2 years of treatment by average total treatment hours

	Weekly Hours < 16			Weekly Hours 16-27			Weekly Hours >27		
Measure	N	Mean	S.D	N	Mean	S.D	N	Mean	S.D
EVT[a]	100	5.24	15.76	147	8.05	15.96	88	8.50	19.26
PPVT[b]	101	6.46	16.74	147	5.97	13.66	91	7.29	18.74
ABC[c]	102	5.75	10.37	163	6.75	10.18	88	7.80	9.18
Communication	112	4.83	10.05	168	6.62	10.25	98	6.76	12.87
Daily Living Skills	110	7.24	12.83	169	7.36	13.03	97	7.74	13.00
Socialization	112	5.61	12.40	170	6.14	11.18	98	7.23	12.01
Motor Skills	97	6.62	14.32	144	6.79	12.32	84	6.14	13.78

[a] Repeated Measures Analysis using Proc Mixed - No time x treatment hours interaction (F = .49; p = .61)

[b] Repeated Measures Analysis using Proc Mixed - No time x treatment hours interaction (F = .98; p = .37)

[c] Repeated Measures Analysis using Proc Mixed - No time x treatment hours interaction (F = .45; p = .64)

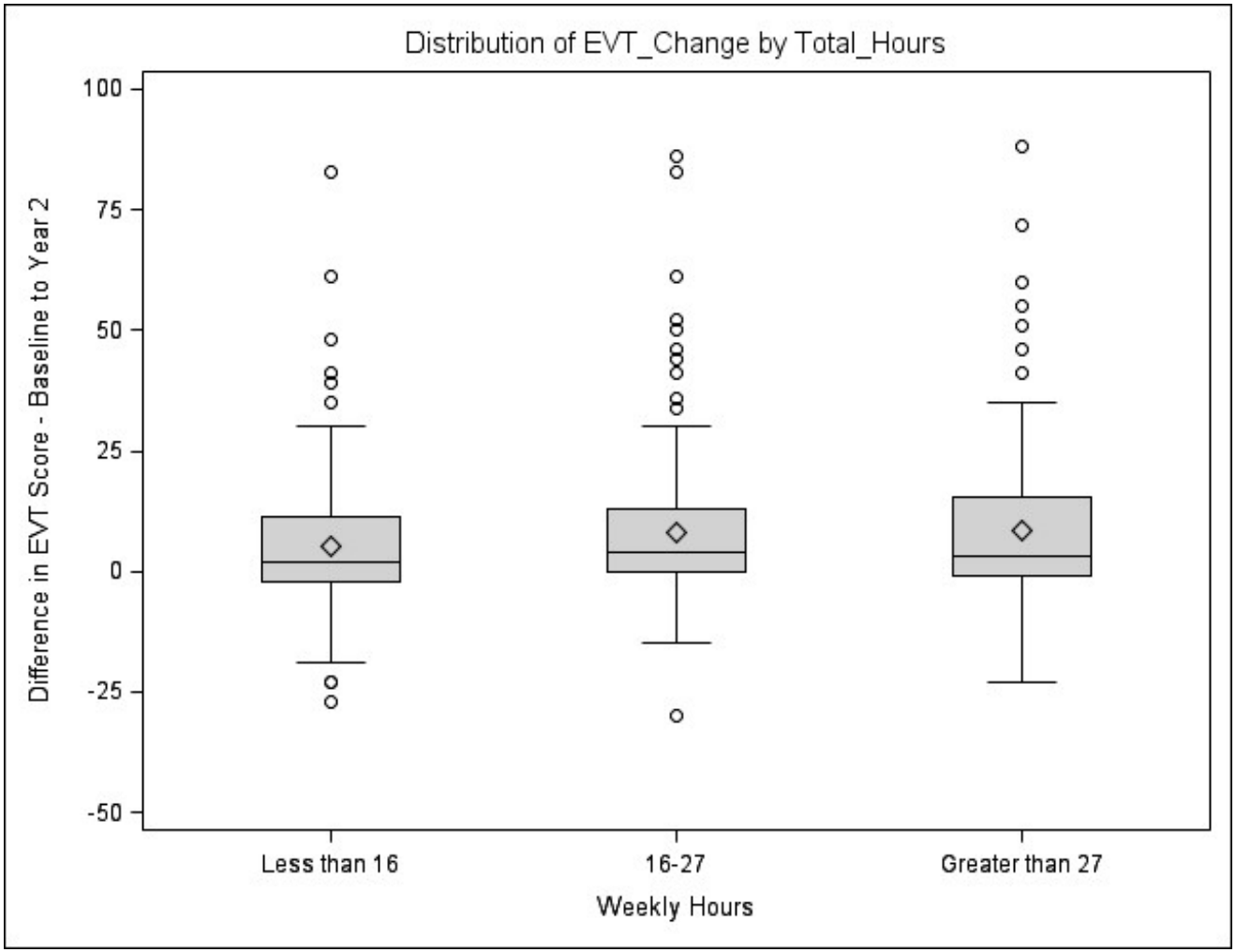

Figure 1.1 Individual changes in EVT scores from baseline through end of year 2, grouped by weekly treatment hours received.

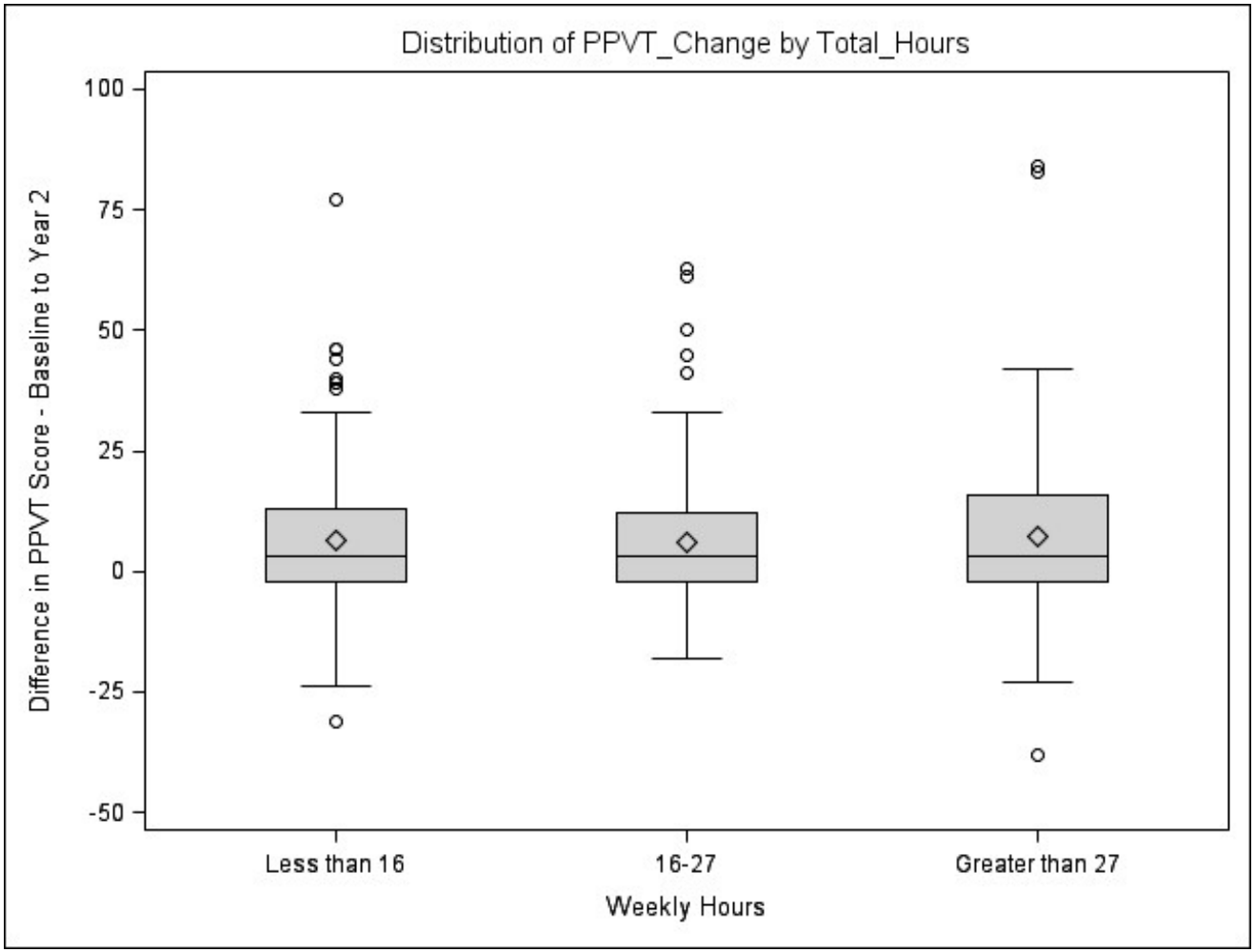

Figure 1.2 Individual changes in PPVT scores from baseline through end of year 2, grouped by weekly treatment hours received.

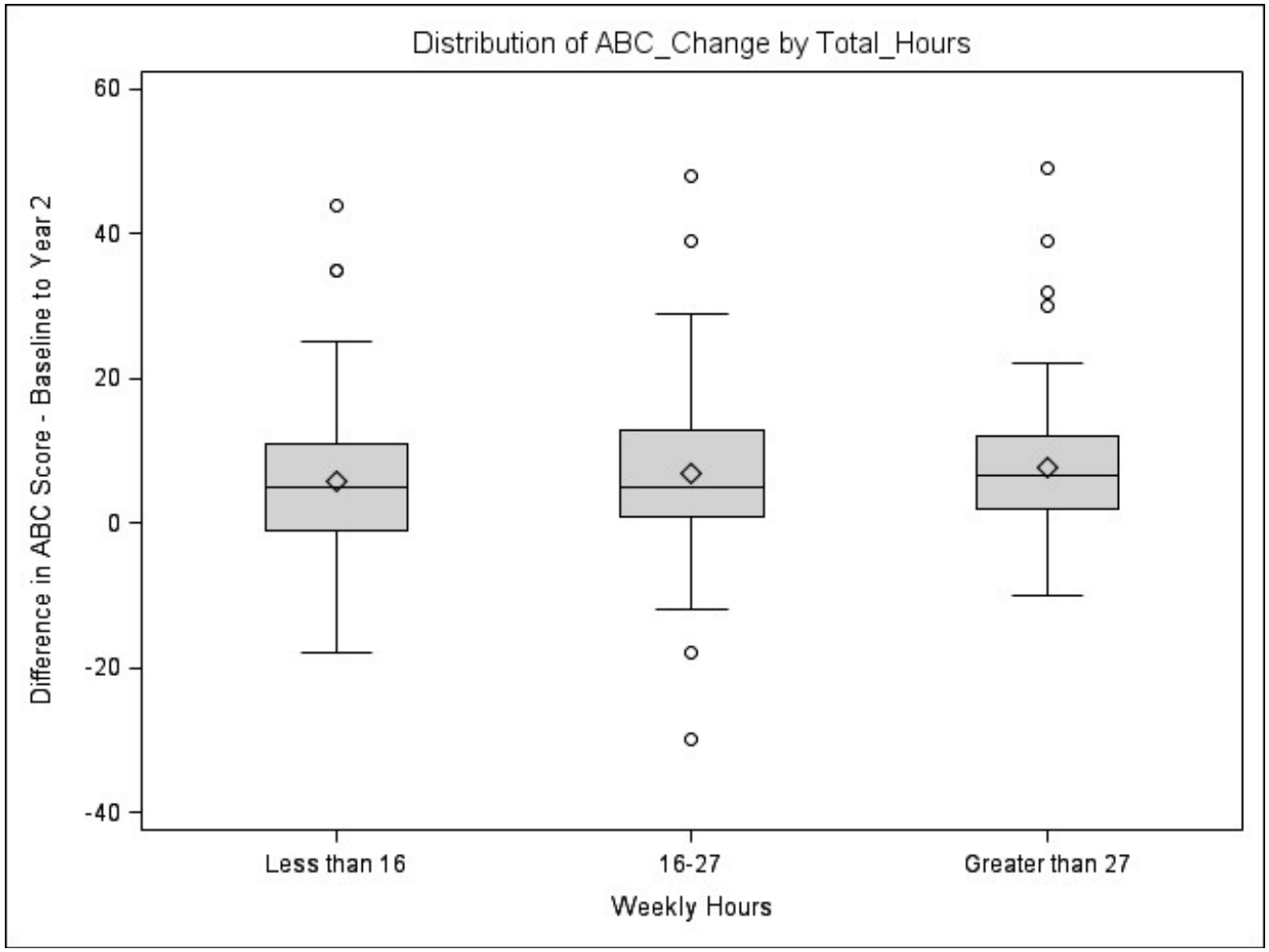

Figure 1.3. Individual changes in ABC scores from baseline through end of year 2, grouped by weekly treatment hours received.

between these groups, there was insufficient statistical power to detect smaller effects between groups (Castelloe 2014). Due to larger variance in language outcomes (EVT and PPVT), statistical power to detect differences by treatment intensity is smaller.

1.3.3 Question 3: What is the range of outcomes observed for adaptive behaviors and language?

Previous analysis showed that statistically significant average gains were achieved across all outcome measures. Table 1.8 presents the range of outcomes observed from baseline to the end of Year 2. As shown, approximately 25% of the children experienced no gains, while more than 50% of the children showed gains of more than 5 points on each of the adaptive behavior outcomes. Using medium (.5) and large (.8) effect sizes as a reference point, Table 1.9 shows that 41% of children achieved an improvement in Adaptive Behavior Composite score that would be classified as a medium effect size (a 7.3 point gain), while 27% experienced a large effect size (an 11.7 point gain). Similar results are seen for all adaptive behavior outcomes, with smaller gains observed on the Communication domain.

The distribution of gains in Receptive and Expressive language scores were comparable to gains observed in adaptive behaviors, consistent with the changes seen in overall average gains reported earlier. Table 1.8 shows that the bottom quartile of children experienced no change (EVT) or a slight reduction (PPVT) in their baseline language scores, while 25% achieved gains of 13 points or more on EVT and 12 points or more on PPVT. Due to higher variability in the baseline scores, effect size thresholds are

Table 1.8 Range of outcomes: Distribution of gains in specified measure

		Change in Score Observed at Given Percentile				
Measure	N	10%	25%	50%	75%	90%
EVT	335	-7	0	3	13	24
PPVT	339	-10	-2	3	12	26
Vineland Adaptive Behavior (VABS)						
ABC (Composite)[a]	353	-4	1	6	12	18
Communication	378	-6	0	6	11	20
Daily Living Skills	376	-7	-1	6	15	24
Socialization	380	-5	0	5	12	21
Motor Skills	325	-7	0	5	14	23

[a]50% of children had a change in ABC Composite Score of 6 or less from baseline to end of year 2

Table 1.9 Range of outcomes: Percent of children achieving specified gains in measure

	EVT	PPVT	ABC	Comm.	DLS	Social	Motor
N	315	318	335	358	356	360	307
Medium Effect Size threshold	16.1	15.4	7.3	9.2	8.0	7.3	7.8
% of children achieving	18%	22%	41%	33%	42%	42%	42%
Large Effect Size threshold	25.7	24.6	11.7	14.8	12.8	11.6	12.5
% of children achieving	10%	13%	27%	16%	32%	26%	29%

higher for the language outcomes, reducing the percentage of children who've experienced medium or large gains, as shown in Table 1.9.

1.4 Discussion

This study is one of the largest studies to date that reports on the effectiveness of EIBI in a community setting. Looking at 419 children who received EIBI through South Carolina's Department of Disabilities and Special Needs Pervasive Development Disorders Program, this study demonstrated overall improvements in adaptive behavior and language outcomes and resulted in 'medium-sized' gains in adaptive behaviors for 41% of the children who completed two years of treatment. These outcomes exceed those observed in the only other comparably sized study to date (Perry, 2008), which observed only limited gains in adaptive behavior composite standard scores. Average gains in ABC standard scores were comparable to those observed by Bibby et al. (2002) and Eikeseth et al. (2012) on smaller samples but effect sizes were smaller than those observed by Eikeseth. As discussed by Perry et al. (2008), evaluation of standard scores is a "more stringent test" (p632) of improvement which cannot be used to assess absolute growth, but instead, reflects growth beyond that which would be expected among equivalently-aged peers (Hoffman, Templin, and Rice 2012).

This study is one of few studies that explicitly compares outcomes achieved during the first year of treatment to the outcomes achieved during the second year of treatment. Importantly, overall statistically significant changes in average scores seen in the first year of treatment continue through the second year of treatment on all

outcome measures except for receptive language (PPVT). All outcomes show a statistically significant difference in changes between the first and second year, with the exception of motor skills, which continues to show the same magnitude of improvement through the second year.

No association was seen between the average weekly number of treatment hours received and changes in language and adaptive behavior outcomes. As discussed, average treatment hours from the first full calendar year of treatment was used for each child, which may not have been reflective of the treatment hours received in the initial months of treatment, where gains were greatest. Additionally, treatment hours allocated and received were somewhat based on baseline severity, making it particularly challenging to find an association between hours received and treatment outcomes. Finally, even with a large sample, there was limited statistical power to detect whether differences in treatment outcomes varied based upon the number of treatment hours (Castelloe 2014). Therefore, these findings should not be interpreted as an indication that treatment intensity does not matter.

Literature regarding the range of outcomes achieved is difficult to compare due to the use of different outcome measures (cognitive, autism symptomology, language and adaptive behaviors) and differing classifications regarding a positive or desired outcome. This study simply presents the range of outcomes achieved and includes an assessment of the size of that outcome. Rather than focus on attainment of a specific score indicative of a return to a 'normal' range of functioning, this analysis looks at the

overall distribution of outcomes achieved. Since the VABS standard scores are used a 5 point gain by a child with a baseline Adaptive Behavior Composite Score of 45 is the same magnitude of change as a 5 point gain by a child with a baseline score of 77. In that case, classifying only one of those children as having experienced a return to normal functioning (based on an outcome above 80) masks the underlying impact of the treatment, despite the potential clinical differences between the two scenarios. This study found that approximately 40% of children achieved moderate gains on adaptive behavior outcomes while 25% achieved large gains. Results for receptive and expressive language outcomes were smaller. As noted by Warren et al. (2011), even small improvements in standardized scores 'may translate into large, meaningful improvements in quality of life for children and their families' (Warren et al. 2011 p. 1309).

This study adds to a growing body of literature demonstrating that EIBI can be effective in large-scale community settings. Like prior studies, a wide range of individual outcomes are observed. This study clearly shows that second year gains are smaller than those achieved in the first year, with potential implications for management of program costs. Given the variability of gains observed, it is important to continue to understand the child-specific factors that predict positive outcomes.

1.4.1 Strengths and Limitations

A major strength of this study is the large sample, obtained from multiple cohorts (2007-2011) of children which encompassed the launch of the program in 2006

and associated start-up challenges of recruiting etc. (Perry, 2008). In addition to more precise estimates of outcomes, the sample size allowed a robust comparison of the differences in first and second year gains and the range of outcomes achieved. It also allows further large-sample analysis of differences in outcomes by age at enrollment, gender, cohort and baseline levels scores on adaptive behavior and language.

While lack of a control group is an obvious limitation of the study, both efficacy and effectiveness of EIBI has been established relative to non-treatment controls and alternative treatment controls. A second limitation is the inability to examine changes in autism symptomology since CARS scores were only available at baseline. A third limitation is the generalization of treatment hours that was used to study the impact of treatment intensity. Finally, as with any large program, issues of dropout and missing data could impact the internal validity of the findings presented. However, with two-thirds of the eligible population included in the study and no evidence or likelihood that missing data would be related to program outcomes, it is likely that these results have internal validity with respect to the population of children receiving treatment.

Given the implications of treatment intensity on costs and burden to the families, it would have been interesting and relevant to more extensively analyze the impact, particularly given the inability to observe differences in this study. However, this would be difficult to do, due to data availability and the complexity of the analysis. While total weekly hours billed does reflect the intensity of the treatment, it may not adequately capture the changing focus of treatment as new skills are acquired. This

analysis used a static measure of treatment intensity for the two-year treatment period, when there is variability week-to-week in the intensity and in the treatment focus. Furthermore, treatment intensity can vary through time based on whether or not improvements are observed. So, analysis of treatment intensity impacts may require a more granular set of outcomes that are not reflected in standardized scores of high level outcome measures.

Chapter 2

Effectiveness of Community-Based EIBI Treatment: Analysis of Predictors of Change in Adaptive Behavior and Language[3]

[3] Kuntz, J.M. To be submitted to *Research in Autism Spectrum Disorders.*

2.1 Introduction

Autism spectrum disorders (ASDs) are a group of lifelong, neurodevelopmental disorders characterized by deficits in social interaction, communication, relationship development and by the presence of repetitive or stereotypical behaviors including restricted interests (American Psychiatric Association 2013, Johnson, Myers, and the Council on Children with Disabilities 2007). While specific ASD symptoms may vary throughout the lifetime, and there is some evidence of non-treatment related improvements in IQ and specific behaviors (Dietz et al. 2007), ASDs are not curable and require chronic management over the lifetime (Myers, Johnson, and the Council on Children with Disabilities 2007). ASDs pose a significant burden for families and society in general. Recent prevalence estimates of the disorder are 14.7 per 1000 (1 in 68) (Baio 2014). Families caring for a child diagnosed with ASD are more likely to face large, out-of-pocket expenditures, stop or reduce work hours, and spend more than 10 hours per week coordinating treatment and care (Kogan et al. 2008). Continued advances in understanding treatment outcomes and broadening access to effective treatment is critical to improving the quality of life of children with autism and their families and minimizing the cost associated with care.

2.1.1 Early Intensive Behavioral Intervention

Applied Behavioral Analysis (ABA) refers broadly to a set of treatments based upon behavior analytic methods which trace back to the work of B.F. Skinner (Morris, Smith, and Altus 2005). In the treatment of autism, ABA methods are used across age

ranges, in a variety of settings (home, community, school), and with few restrictions on intensity, duration or the requirement of 1:1 interventions. EIBI is the application of ABA principles for the treatment of young children with autism (Granpeesheh, Tarbox, and Dixon 2009, Foxx 2008, Peters-Scheffer et al. 2011). Treatment is comprehensive, and intensive, usually entailing 20-40 hours per week of highly-individualized treatment for two or more years (Lovaas 1987, Peters-Scheffer et al. 2012, Green, Brennan, and Fein 2002, Foxx 2008, Virués-Ortega 2010, Jensen and Spannagel 2011). The efficacy and effectiveness of EIBI treatment is well-established through individual studies (Lovaas 1987, Sallows, Graupner, and MacLean Jr 2005, Remington et al. 2007, Eikeseth et al. 2007, Perry et al. 2008, Granpeesheh, Tarbox, and Dixon 2009) and meta-analyses (Virués-Ortega 2010, Reichow 2012, Eldevik et al. 2009, Makrygianni and Reed 2010).

Across EIBI studies, inconsistencies exist regarding the specific outcomes achieved, the percentage of children who experience improvement, and the factors which predict outcomes. Despite consistent findings regarding the overall effectiveness of EIBI, heterogeneity of results at the individual level is commonly reported. EIBI studies have looked at a variety of outcome measures including cognitive skills, adaptive behaviors and autism symptomology (Howlin, Magiati, and Charman 2009, Matson and Goldin 2014). Additionally, different classifications of positive outcomes are seen, including placement in mainstream settings (Lovaas 1987, Bibby et al. 2002, Harris and Handleman 2000), attainment of average functioning or clinically significant improvement (Jacobson and Truax 1991, Remington et al. 2007, Perry et al. 2008, Eikeseth et al. 2007, Bibby et al. 2002). Regardless of the outcome measure or the

criteria to classify positive outcomes, only 25-50% of children receiving EIBI are reported to achieve good outcomes (Ben-Itzchak and Zachor 2007, Perry et al. 2011, Lovaas 1987).

Given the intensity and duration of the treatment, provision of EIBI services is costly, with estimated annual cost of treatment as high as $10,000 per child per month (Smith, Klorman, and Mruzek 2015) with typical estimates of between $20,000 to $60,000 per child annually (Chasson, Harris, and Neely 2007, Payakachat et al. 2012) . Several studies have demonstrated the overall cost-benefit of EIBI treatment (Jacobson, Mulick, and Green 1998, Chasson, Harris, and Neely 2007, Peters-Scheffer et al. 2012), but given the expense and the heterogeneity of individual child outcomes, it is important to identify which children are likely to benefit from EIBI services and the factors which predict positive outcomes (Perry et al. 2011, Reichow and Wolery 2009, Goin-Kochel, Mackintosh, and Myers 2009, Smith, Klorman, and Mruzek 2015). More than two decades after Lovaas' initial conclusions regarding the benefits of EIBI, Goin-Kochel, Mackintosh and Myers (2009) conclude that "we are currently unable to predict which children will respond to particular treatments, what intensity of treatment might make a difference, and what behaviors the treatments might affect" (2009, 529). Smith et al. have recently concluded that reliable predictors of EIBI outcomes still do not exist (Smith, Klorman, and Mruzek 2015, Eldevik et al. 2010). Important, but sometimes inconsistent findings have emerged regarding the importance of baseline language and adaptive behavior scores and other child-specific characteristics such as age on changes in language and adaptive behaviors

2.1.2 Outcomes and predictors

2.1.2.1 Baseline functioning

Eldevik (2010) et al., Virues-Ortega (2010), and Makrygianni and Reed (Makrygianni and Reed 2010) investigated the relationship between effect sizes and baseline measures of adaptive behavior composite scores (ABC), with different results. Eldevik et al. found that baseline measures of ABC predicted subsequent gains in ABC, with lower baseline behavioral scores predicting larger gains. Virues-Ortega, found no associations of gains in ABC with baseline measures. Makrygianni and Reed (2010) found that higher adaptive behavior scores are associated with better adaptive behavior outcomes. Using the previously discussed measure of 'reliable change', Remington et al. found that higher baseline adaptive behavior scores were associated with greater likelihood of achieving 'reliable change' in adaptive behaviors. In the large community study, Perry et al. (2011) found that baseline adaptive functioning explained a small but significant portion of the variability in changes in adaptive behaviors with a positive correlation between baseline adaptive behavior composite and the Year 2 outcome. So, findings regarding the impact of baseline adaptive behavior scores on changes in adaptive behaviors are inconsistent.

2.1.2.2 Age-at-entry

While early intervention is one of the foundational components of EIBI treatment (Foxx 2008, Green, Brennan, and Fein 2002), findings regarding age-at-entry as a predictor of treatment outcomes are surprisingly 'equivocal' (Perry et al. 2011)

(p593). In their initial research among children who were less than 40 months of age at treatment inception, Lovaas et al. found no association between age and outcomes (Lovaas 1987). These findings have been corroborated among similarly younger populations (Hayward et al. 2009, Ben Itzchak and Zachor 2011). In an older population of 4 to 7 year-olds, Eikeseth et al. found no correlation between age and change in adaptive behavior scores. In each of these studies, small samples and narrow age ranges within study may prevent associations from emerging (Perry et al. 2011). Additionally, each of these studies reflect a clinical, rather than community setting. In Virues-Ortegas meta-analysis, pre-treatment age did not affect treatment outcomes on adaptive behaviors or language across studies where this assessment was possible. In a meta-analysis focused on studies of younger children (less than 54 months at entry), Makrygianni and Reed (2010) found no significant association between age-at-entry and adaptive behavior outcomes. In their large, community-based study, Perry et al. (2011) found that age at entry was negatively correlated with gains in VABS adaptive behaviors standard scores across all domains. Furthermore, they found that those children under 48 months at entry achieved greater gains than those children older than 48 months on the same adaptive behavior domains.

2.1.3 Current Study

The overall aim of this study is to identify the predictors of overall outcomes and child-specific changes in a publically-funded implementation of an EIBI treatment program through a large-sample, secondary analysis of longitudinal outcome data.

Using historical data from the South Carolina Department of Disabilities and Special Needs (SC DDSN) Pervasive Development Disorder (PDD) Program, this analysis identifies baseline characteristics that predict improvements in adaptive behaviors and language. An earlier chapter demonstrated the overall effectiveness of this program and examined whether outcomes varied by treatment duration and intensity. Briefly, the previous analysis showed that overall average gains were achieved on receptive language, expressive language and adaptive behavior outcomes through 2 years of EIBI treatment. Gains were greater through the first 12 months of treatment compared to the second 12 months of treatment, but statistically significant average improvement continued in the second year on adaptive behaviors and expressive language. There were no differences in observed outcomes based on treatment intensity, defined as the number of weekly treatment hours received. Finally, as in other studies, there was substantial heterogeneity of results, with approximately 25% of the children achieving no gains on each of the language and adaptive behavior outcomes and approximately 25% of children achieving gains in adaptive behaviors that would be classified as a large effect size (Fritz, Morris, and Richler 2012). Smaller gains were seen in expressive and receptive language.

2.2 Method

2.2.1 Participants

This study examines program participants who began EIBI treatment in program cohort years 2006-2011. Data were obtained from stored paper files only for those

children who had completed at least two years of treatment and entered into an Excel spreadsheet by two graduate assistants under the direction of the study author and the direct supervision of SC DDSN personnel. Additional paper files containing the original diagnostic measures for these children were obtained from the 4 Regional DDSN Care Centers. Similar to the study conducted by Perry et al. (2008), individual files may have had some incomplete data, but the final study population consisted of those children who had data from at least two measurement periods for any of the 3 primary outcomes: adaptive behaviors, expressive language or receptive language. This resulted in a final study population of 419 children. This is more than two-thirds of the population of 615 children who had completed two years of treatment. Table 2.1 provides baseline characteristics of those children included in this study. Inclusion was based solely on the availability of the data. Because the final sample represents a substantial proportion of those who completed treatment and inclusion in the sample was not based on specific child characteristics, these results provide insight into treatment effectiveness for the full population of children completing treatment.

2.2.2 Measures

Each child enrolled in the program received a pre-treatment diagnosis of autism with evidence from at least two of the following:

- Autism Diagnostic Observation Schedule (ADOS) (Lord et al. 1999)
- Autism Diagnostic Interview (ADI) (Lord and Rutter 1994)

Table 2.1 Baseline measures

Measure	N	Mean	Std	Range	
				Min	Max
Age at Enrollment (years)	418	5.9	1.2	3.3	8.8
Assessment					
Expressive Vocabulary Test (EVT)	364	67.1	32.1	20.0	133.0
Receptive Vocabulary Test (PPVT)	370	66.0	30.5	20.0	126.0
Vineland Adaptive Behavior (VABS) Standard Scores					
Communications	392	69.4	17.8	22.0	117.0
Daily Living Skills	391	69.0	15.2	25.0	114.0
Socialization	392	65.9	12.7	23.0	122.0
Motor Skills	373	73.6	14.1	37.0	121.0
Adaptive Behavior Composite (ABC)	375	67.0	12.8	42.0	103.0
Childhood Autism Rating Scale (CARS)Total Score	169	35.9	5.2	21.0	49.0
Social Communication Questionnaire (SCQ) Total Score	109	20.2	6.1	5.0	34.0
IQ	55	73.3	18.1	26.0	109.0

- Social Communication Questionnaire (SCQ) (Rutter, Bailey, and Lord 2003)
- Childhood Autism Rating Scale (CARS) (Schopler, Reichler, and Renner 1988)

These measurements were conducted only at initial diagnosis and are not captured on a pre-post treatment basis. IQ was also infrequently and inconsistently captured only in the pre-treatment timeframe. Therefore, changes in these measures will not be evaluated as part of this evaluation.

The proposed study will focus on factors that predict improvements in adaptive behaviors, receptive and expressive language. Specific factors considered are baseline performance of adaptive behaviors and language. Child's age at entry, diagnosis and gender are also examined. Upon acceptance into the program, the assessment conducted by the EIBI consultant includes the following measures which are the focus of this analysis:

- Receptive language using the Peabody Picture Vocabulary Test IV (PPVT) (Dunn and Dunn 1981)
- Expressive language using the Expressive Vocabulary Test (EVT) (Williams 1997)
- Adaptive behaviors using the Vineland Adaptive Behavior Scales-II (VABS) (Sparrow, Balla, and Cicchetti 2005)

2.3 Results

2.3.1 Question 1: Do baseline measures of outcome variables vary by individual child-specific attributes?

Prior to conducting analyses of moderators of language and adaptive behavior outcomes, variability in baseline measures of those language and adaptive behaviors was examined across different levels of child-specific attributes. For continuous variables IQ and CARS Total Score, the sample was split at the mean of each of those variables. Baseline Expressive and Receptive language are highly correlated (r=.92), as are baseline language and adaptive behaviors (r=.74(EVT), r=.75(PPVT)) so no further analysis is shown. As seen in Table 2.2, significant differences are seen in baseline measures of adaptive behavior composite (p=.015) expressive language (p=.018) and receptive language (p=.0098) by enrollment cohort. Changes in overall program funding varied at different points in the study period, affecting the number of 'slots' available, and therefore, the characteristics of the children who started treatment in each year. No significant differences are observed in baseline language and adaptive behavior composite scores by gender or diagnosis. While not the focus of the overall analysis due to incomplete data, baseline adaptive behavior and language scores vary by baseline measures of IQ and CARS score.

Table 2.2 Baseline score by subgroup

		Adaptive Behavior Composite			Expressive Vocabulary Test			Peabody Picture Vocabulary Test		
Variable	n	Mean	Std	p[a]	Mean	Std	p[a]	Mean	Std	p[a]
Overall	419	66.99	12.84		67.05	32.11		65.97	30.5	
Enrollment Cohort				0.015			0.0183			0.0098
2007	62	66.03	12.95		58.88	32.59		56.68	33.24	
2008	70	63.49	11.31		61.65	31.2		61.06	28.04	
2009	81	70.78	15.43		70.77	35.2		68.21	32.5	
2010	141	67.43	11.35		72.89	29.49		72.17	28.21	
2011	65	65.8	12.55		62.08	31.27		63.95	29.74	
Gender				0.076			0.446			0.403
Female	56	63.96	11.52		63.31	32.86		62.02	32.11	
Male	318	67.37	13.29		67.12	32.28		66.02	30.72	
Diagnosis				0.23			0.335			0.227
Asperger's	13	73.5	16.44		78.91	27.35		79.64	22.94	
Autism	304	66.57	12.53		65.94	31.75		64.96	30.42	
PDD/Other	47	68.78	13.12		73.58	30.94		72.24	28.81	
NA	55	66.23	13.28		65.66	36.57		63.54	33.54	
Age at Enrollment				0.288			0.121			0.164
<= 5.9	224	67.32	13.18		70.21	33.1		68.58	31.23	
> 5.9	194	66.55	12.46		63.74	30.7		62.87	29.47	
CARS Total Score				<.0001			<.0001			<.0001
<= 35.9	89	69.94	11.89		77.92	27.22		75.21	26.9	
>35.9	80	62.81	12.55		52.41	30.7		50.66	29.65	
IQ				<.0001			<.0001			<.0001
<= 73.3	27	65.76	10.07		63.91	29.39		57.12	27.58	
>73.3	28	75.17	11.66		91.58	18.17		92.15	17.73	

[a] SAS Proc GLM used for ANOVA test of differences in average baseline score between categories for each measure (ABC, EVT, PPVT)

2.3.2 Question 2: Do language and adaptive behavior outcomes variables vary by age and baseline language and adaptive behavior performance?

In order to assess whether outcomes varied by baseline characteristics, average two-year difference in language and adaptive behavior were analyzed. Two sample t-tests were used to assess whether language and adaptive behaviors two-year outcomes differed for those children above and below the mean of age at entry, baseline language score and baseline adaptive behavior score. Despite a much smaller sample, similar analysis was also conducted for baseline measures of IQ and CARS Total Score. Table 2.3 shows statistically significant gains in average scores after two years of treatment on outcomes of adaptive behaviors (mean = 6.72, $p<.0001$), expressive language (mean = 7.33, $p<.0001$) and receptive language (mean = 6.47, $p<.0001$).

T-tests comparing outcomes of older children to those of younger children show significantly different outcomes on adaptive behaviors ($p=.01$). Age was not a moderator of differences in expressive language ($p=.84$) or expressive language ($p=.077$. Lower expressive vocabulary score at baseline is associated with smaller gains in adaptive behavior ($p=.0003$), but larger gains in expressive vocabulary ($p<.0001$). Similar results are seen when looking at receptive language at baseline, where lower baseline scores predict greater improvements in receptive language ($p=.003$), but lower gains in adaptive behaviors ($p<.0001$).

Table 2.3 Change in outcome measures by baseline differences

	Two-Year Difference in ABC				Two-Year Difference in EVT				Two-Year Difference in PPVT			
Baseline Measure	n	mean	std	p[a]	n	mean	std	p[a]	n	mean	std	p[a]
Overall Difference	353	6.72	10.00		335	7.33	16.84		339	6.47	16.04	
Age at Enrollment				0.01				0.84				0.0772
<= 5.9	192	7.90	11.14		175	7.27	17.63		179	7.92	18.19	
> 5.9	160	5.31	8.28		159	6.91	14.85		159	4.88	13.13	
EVT				0.00				<.0001				0.1670
<= 67	136	4.60	8.08		141	12.30	21.84		141	7.83	19.01	
> 67	184	8.46	10.83		194	3.71	10.64		192	5.25	13.15	
PPVT				<.0001				0.13				0.0030
<= 66	161	4.32	7.79		157	8.90	19.88		164	9.17	18.97	
>66	166	8.86	11.15		176	6.03	13.56		175	3.94	12.22	
ABC				0.84				0.07				0.3387
<= 69	213	6.81	9.16		184	8.54	18.72		188	7.13	17.23	
> 69	140	6.58	11.18		129	5.29	13.41		130	5.49	13.17	
CARS				0.03				0.72				0.5788
<= 35.9	76	7.64	11.34		76	6.13	14.23		77	5.14	14.29	
>35.9	66	4.06	7.91		64	7.00	13.94		66	6.64	17.79	
IQ				0.01				0.39				0.2508
<= 73.3	23	1.30	7.93		22	1.50	16.78		22	1.45	11.99	
>73.3	23	9.83	12.60		23	5.22	11.61		23	6.22	15.19	

[a]SAS Proc Ttest used to test differences in average two-year change between categories for each measure (ABC, EVT, PPVT)

2.3.3 Question 3: Do language and adaptive behavior outcomes variables vary by program cohort, gender or diagnosis?

Table 2.4 provides differences in outcomes for children who entered the program in different years (cohort), by gender, and by initial diagnosis. No significant differences in outcomes are observed by gender and diagnosis across language and adaptive behavior outcomes. No differences are observed in adaptive behavior outcomes by enrollment cohort, but two-year differences are seen in expressive vocabulary (p=.0499) and receptive vocabulary (p=.01) by enrollment cohort. All of these potential predictors are explored in multi-variate analysis in the next section.

2.3.3.1 Simultaneous Analysis of moderators across 3 Measurement Periods

Repeated Measures analysis was conducted using SAS Proc Mixed to simultaneously assess moderators of change in expressive language, receptive language and adaptive behavior composite score. For each outcome variable, pre-treatment scores (baseline) and post-treatment scores were obtained after approximately 12 and 24 months of treatment. As discussed earlier, a previous chapter assessed overall outcomes associated with EIBI treatment. Statistically significant gains were observed for each of the outcome variables after two years of treatment, after controlling for all the baseline measures discussed.

Table 2.4 Two-year change in baseline scores across subgroups

		Change in Adaptive Behavior Composite			Change in Expressive Vocabulary Test			Change in Peabody Picture Vocabulary Test		
Variable	n	Mean	Std	p[a]	Mean	Std	p[a]	Mean	Std	p[a]
Total Sample	419	6.72	10.00		7.33	16.84		6.47	16.04	
Enrollment Cohort				0.449			0.050			0.011
2007	62	6.96	10.65		12.41	18.54		12.32	22.72	
2008	70	7.85	10.17		9.00	18.01		6.07	18.49	
2009	81	7.06	10.58		7.00	13.29		7.82	11.71	
2010	141	6.81	8.89		4.45	15.66		3.66	11.36	
2011	65	4.36	10.70		6.32	21.54		3.63	17.07	
Gender				0.07			0.12			0.06
Female	56	4.42	9.17		4.02	11.91		2.61	11.88	
Male	318	7.16	10.23		8.12	17.28		7.39	16.68	
Diagnosis				0.98			0.95			0.87
Asperger's	13	4.00	9.93		7.57	9.52		1.25	11.61	
Autism	304	6.91	10.15		7.70	16.77		6.99	15.47	
PDD/Other	47	6.50	11.93		4.44	15.03		2.65	15.59	
NA	55	6.45	6.81		7.38	19.99		7.43	20.20	

[a]SAS Proc GLM used for ANOVA test of differences in average two-year change between categories for each measure (ABC, EVT, PPVT)

2.3.3.2 Adaptive Behavior Outcomes

This analysis looks at the impact of individual baseline measures on changes in adaptive behaviors. Table 2.5 shows that continuous variables, age at entry (p=.01), enrollment cohort (p=.004), baseline adaptive behaviors (p<.0001) and baseline receptive language (PPVT) (p<.0001) moderate average change in adaptive behavior across two years of treatment. As age at entry increases, lower average gains are seen in all outcome measures. Importantly, as shown in Figure 2.1, there is no age effect through the first 12 months of treatment (p=.96), but older children are less likely to continue to make improvements through the second measurement period (p=.016). Lower adaptive behavior scores at baseline correspond to higher average change in adaptive behaviors. Higher baseline receptive language scores predict greater gains in adaptive behaviors. Given the high correlation between baseline receptive and expressive language, these variables were not simultaneously evaluated as moderators of change in any of the outcome measures. While individual analysis of enrollment cohort showed no impact on adaptive behavior outcomes, simultaneously controlling for age, baseline language and adaptive behaviors does show a different outcome effect by cohort (p=.004). Gender and diagnosis were not significant explanatory variables after controlling for other variables.

2.3.3.3 Language Outcomes

Age at entry (p=.0001), baseline adaptive behaviors (p<.0001) and baseline receptive language (p<.0001) also moderate average change in receptive language. In

Table 2.5 Moderators of change over time in outcomes

	ABC		EVT		PPVT	
	F	p	F	p	F	p
Variable						
Age at Enrollment	5.04	0.01	0.38	0.69	9.04	0.0001
Enrollment Cohort	2.88	0.004	1.66	0.11	0.55	0.82
Gender	0.62	0.54	1.75	0.17	1.82	0.16
Baseline ABC	27.94	<.0001	15.73	<.0001	10.99	<.0001
Baseline EVT			34.42	<.0001		
Baseline PPVT	37.07	<.0001			24.43	<.0001

SAS Proc Mixed used to test time*variable interaction

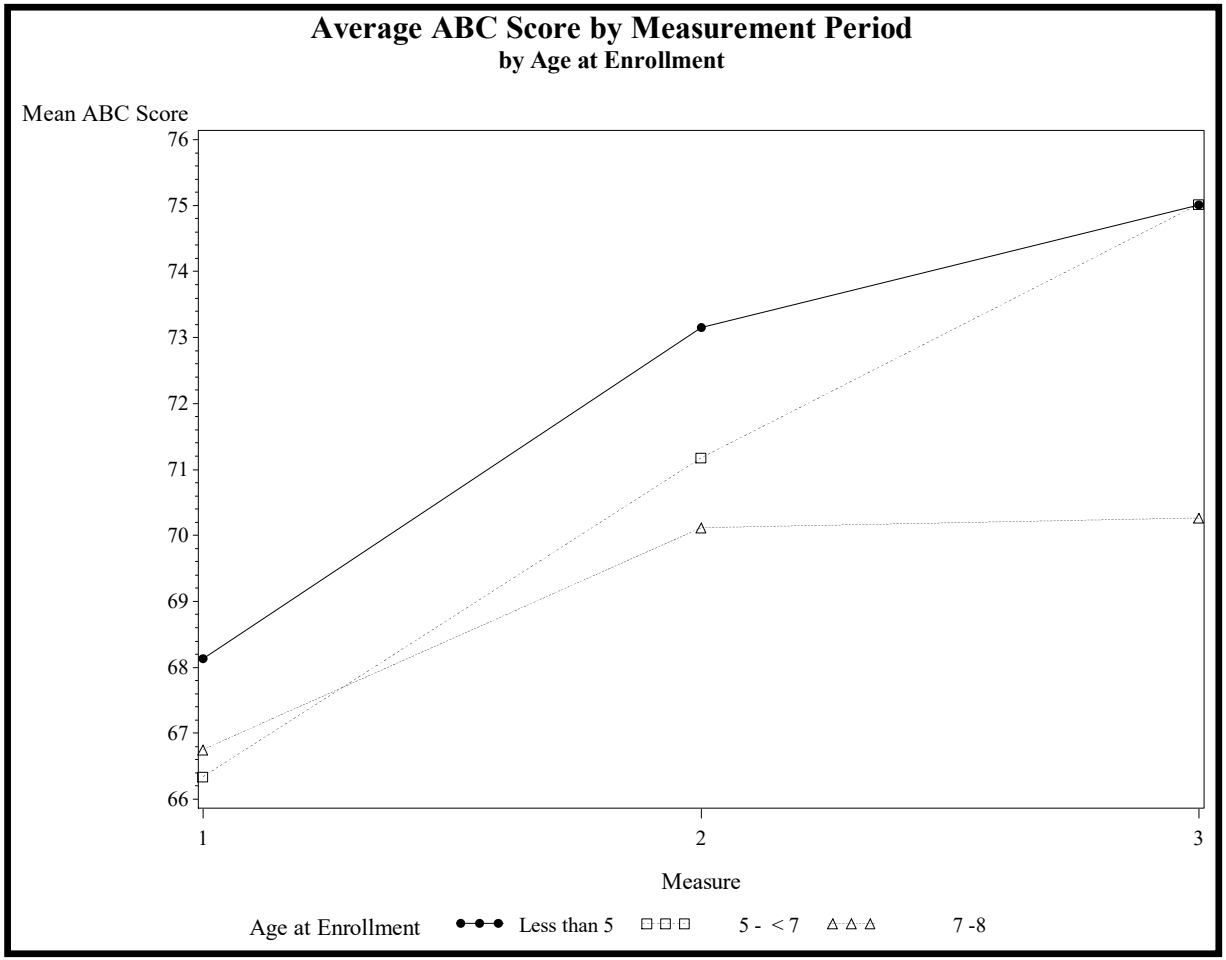

Figure 2.1 Average ABC score by measurement period by age at enrollment

this case, younger children show greater gains in each of the subsequent measurement periods compared to adaptive behavior outcomes, where different age impacts were found only in the 2nd year of treatment. Only baseline expressive language (EVT) and baseline adaptive behavior scores moderate changes in expressive language. After controlling for baseline differences in language and adaptive behaviors, enrollment cohort was no longer a moderator of differences in language outcomes. Gender and diagnosis were not significant explanatory variables after controlling for other variables.

2.3.4 Predictors of individual child outcomes

The range of outcomes observed across all of the outcome variables is shown in Table 2.6. Top quartile of two-year change is used to classify a positive outcome. As shown, 25% of the children have shown changes from baseline scores of 12 or 13 points for expressive language, receptive language, and the adaptive behavior composite. This generally corresponds with a 'large' treatment effect size for all outcomes. Table 2.7 provides differences in baseline measures between those children who achieved top-quartile gains and those children who did not. This initial analysis shows that younger age ($p=.001$), higher language scores (EVT: $p<.0001$, PPVT: $p<.0001$) and higher adaptive behavior composite score ($p=.02$) at baseline shows greater likelihood of being in the top quartile of adaptive behavior gains. Younger average age is also seen in the top quartile of gains in receptive language (PPVT, $p=.0001$), but not in that of expressive language. Consistent with earlier findings, lower baseline expressive language scores

Table 2.6 Range of outcomes: Distribution of gains in specified measure

	Change in Score Observed at Given Percentile					
	N	10%	25%	50%	75%	90%
Variable						
EVT	335	-7	0	3	13	24
PPVT	339	-10	-2	3	12	26
Vineland Adaptive Behavior (VABS) Standard Scores						
ABC[a]	353	-4	1	6	12	18
Communication	378	-6	0	6	11	20
Daily Living Skills	376	-7	-1	6	15	24
Socialization	380	-5	0	5	12	21
Motor Skills	325	-7	0	5	14	23

[a]50% of children had a change in ABC Composite Standard Score of 6 or less from baseline to end of year 2

Table 2.7 Baseline Measures: Differences between top quartile and remaining sample

Measure	Change in ABC					Change in EVT					Change in PPVT				
	Bottom 75% n=270		Top 25% n=83			Bottom 75% n=256		Top 25% n=79			Bottom 75% n=254		Top 25% n=84		
	mean	std	mean	std	p	mean	std	mean	std	p	mean	std	mean	std	p
Age at Enrollment (years)	6.02	1.26	5.55	1.05	0.001	5.95	1.23	5.99	1.29	0.822	6.1	1.21	5.51	1.23	0.000
Assessment															
Expressive Vocabulary Test (EVT)	61.1	32.19	82.61	26.89	<.0001	69.56	31.88	58.05	32.14	0.005	66.99	32.16	64.69	32.51	0.577
Receptive Vocabulary Test (PPVT)	60.21	30.68	82.04	24	<.0001	66.67	30.05	65.03	32.1	0.676	67.95	30.37	58.55	29.65	0.014
Vineland Adaptive Behavior (VABS)															
Communications	67.22	18.04	75.58	16.82	0.000	69.65	17.78	70.92	17.73	0.582	69.54	18.14	70.19	16.2	0.771
Daily Living Skills	68.24	15.1	71.27	16.26	0.121	69.2	15.09	71.27	13.9	0.282	69.55	14.97	69.81	14.59	0.892
Socialization	64.95	12.74	69	13.21	0.013	66.21	12.83	67.53	11.62	0.419	66.52	12.83	66.02	11.03	0.753
Motor Skills	73.4	14.34	73.85	13.95	0.807	73.2	14.33	75.22	12.13	0.275	73.29	14.11	74.01	13.2	0.683
ABC	66.12	12.86	69.9	13.14	0.020	67.18	13.05	68.77	12.23	0.353	67.33	13.11	67.53	11.88	0.903
CARS Total Score	36.11	5.09	34.52	5.36	0.157	35.89	5.03	35.76	5	0.902	35.92	5.29	36.17	4.64	0.812
SCQ Total Score	20.4	6.71	20.23	3.39	0.877	19.62	6.13	22.19	4.18	0.116	20.07	5.56	19.69	7.69	0.833
IQ	70.54	18.13	83.44	9.19	0.006	73.12	17.75	78.25	11.96	0.361	74.03	16.66	76.33	16.32	0.711

Lower n for CARS, SCQ, IQ

are seen among those in the top quartile of gains on expressive language, while lower receptive language scores are seen in the top quartile of receptive language gains.

2.3.4.1 Multiple logistic regression of predictors of child outcomes

SAS Proc Logistic was used to perform logistic regression to predict top quartile outcomes for each of the outcome measures, controlling for age, enrollment cohort, gender and baseline adaptive behavior and language score. Due to high correlations between baseline expressive and receptive language, only one of these potential moderators was considered in each of the models. As shown in Table 2.8 and 2.9, best outcomes on adaptive behaviors are associated with younger age at entry ($p=.01$, $OR=.72$), lower baseline adaptive behaviors ($p=.02$, $OR=.961$) and higher baseline receptive language ($p<.0001$, $OR=1.04$). Higher baseline adaptive behavior scores ($p<.0001$, $OR=1.075$) and lower expressive language baseline ($p<.0001$, $OR=.97$) predicted top quartile outcomes on expressive language. Age at entry was not a predictor of top quartile outcomes. Similar to findings regarding expressive language, higher baseline adaptive behavior score ($p=.005$, $OR=1.048$), and lower baseline expressive language ($p=.0004$, $OR=.98$). Lower age-at-entry also predicted top quartile outcomes on expressive language. Cohort, gender and diagnosis did not predict top quartile outcomes for any of the outcome measures.

2.3.5 Additional Analysis

While baseline measures of IQ and autism severity (CARS Total Score) were inconsistently captured for the study sample, and no post-treatment measures were

Table 2.8 Predictors of top quartile gains

	ABC		EVT		PPVT	
	Wald Chi-Sq	p	Wald Chi-Sq	p	Wald Chi-Sq	p
Age at Enrollment	6.68	0.01	0.39	0.53	13.72	0.0002
Baseline ABC	5.41	0.02	17.36	<.0001	7.95	0.0048
Baseline EVT			21.54	<.0001		
Baseline PPVT	24.11	<.0001			12.48	0.0004

SAS Proc Logistic used to predict likelihood of being in the top quartile of gains for each outcome

Table 2.9 Odds ratio (OR) estimates for predictors of top quartile gains

	ABC			EVT			PPVT		
	OR Estimate	L95	U95	OR Estimate	L95	U95	OR Estimate	L95	U95
Age at Enrollment	0.72	0.56	0.92				0.65	0.52	0.82
Baseline ABC	0.96	0.93	0.99	1.08	1.04	1.11	1.05	1.01	1.08
Baseline EVT				0.97	0.95	0.98			
Baseline PPVT	1.04	1.02	1.06				0.98	0.96	0.99

SAS Proc Logistic used to predict likelihood of being in the top quartile of gains for each outcome

captured, these have been among the most common outcome measures and moderators studied. So, additional analysis is provided to understand how baseline IQ and CARS score moderate overall average gains observed and predict individual child outcomes. As seen in Table 2.7, those children who experience top quartile gains in adaptive behaviors had significantly higher baseline IQ (p=.006) than those who experienced smaller gains or no gains. No significant differences in baseline IQ are seen between top and lower quartile performance on expressive or receptive language outcomes. There is no significant difference in baseline autism severity (CARS) on any of the outcome measures between top quartile and lower quartile performance.

2.3.5.1 Initial Cognitive Level

These observations are corroborated through logistic regression where IQ is the only independent variable used to explain top quartile performance on each of the outcomes. Using a lower alpha level in consideration of substantially smaller samples sizes, baseline measure of IQ is significant predictor of top quartile outcomes on adaptive behaviors (n=46, p=.057, OR=1.057), with higher baseline IQ associated with better outcomes. When controlling for baseline adaptive behavior score and age-at-entry, IQ is still a significant predictor of top quartile adaptive behavior outcomes, with higher baseline IQ predicting top quartile gains (p=.077, OR=1.061). No significant differences were found for expressive language (n=45, p=.354) or receptive language (n=45, p=.7).

Similar results are seen when predicting average gains across each of the outcome variables using SAS Proc Mixed for repeated measures regression analysis. Initially including only IQ in the model, average gains in adaptive behaviors are moderated by baseline IQ, with higher IQ predicting greater gains (F=3.58, p=.03). However, among those children who have a baseline IQ score, IQ is not significant when included in a model with baseline adaptive behaviors and baseline receptive language and it results in a poorer-fitting model when either of the other variables are eliminated.

2.3.5.2 Initial Autism Severity

For those children who had a baseline CARS score, baseline autism severity does not predict top quartile gains on any of the outcome measures studied when CARS is the only independent variable in a logistic regression model. Similarly, it was not a predictor of average gains in any of the outcome measures across two measurement periods.

2.3.5.3 Non-normality of data

Baseline data for expressive language (EVT) and receptive language (PPVT) were not normally distributed, with more than 10% of the children receiving a baseline score of 20 on either of these measures. This also resulted in slightly non-normal data when differences were calculated between baseline and final measures after two years of treatment. To address the non-normality of language outcomes, several approaches were taken. As stated earlier, repeated measures analysis using Proc Mixed was used since it is robust to violations of assumptions of non-normal error distributions (Jacqmin-Gadda et al. 2007, Zhang and Davidian 2001) and has the advantage of being

able to use more observations. To validate the use of Proc Mixed, results were replicated using Proc GLM, which is not subject to the same normality assumptions, to conduct a repeated measures MANOVA, treating the differences between baseline and year one, and between year one and year two as the repeated outcome. Using Proc GLM decreases available sample since observations with any missing data are excluded, compared to using Proc Mixed, which is able to leverage observations with any 2 measures in calculating overall mean effects. However, when comparing Proc Mixed and Proc GLM results among a common sample, results were quite consistent. Finally, a third analysis was conducted using Proc Mixed where the sample was limited to only those observations with a baseline PPVT or EVT score greater than 20. This analysis yielded the same conclusions regarding the effects of age, baseline adaptive behavior composite score and baseline language score on Expressive and Receptive language outcomes.

2.4 Discussion

This study reports on the predictors of average gains and individual child outcomes on adaptive behavior and language following two years of EIBI treatment. With a sample of 419 children who received EIBI through South Carolina's Department of Disabilities and Special Needs Pervasive Development Disorders Program, this study examined child-specific factors of gender, age-at-entry, baseline adaptive behavior composite score and baseline language score (either expressive or receptive). We also investigated treatment cohort (year of enrollment) to assess whether challenges of program start-up (Perry et al. 2008), and ongoing changes in funding and eligibility

affect treatment outcomes. Additional analysis was conducted investigating baseline cognitive level (IQ) and autism severity (CARS). Data limitations prevented analysis of the EIBI providers. Four providers were identified comprising 209 of the 419 children. However, one of these represented 150 children while the other three served less than 25 children each, limiting the ability to make comparisons, particularly while controlling for other key variables. Furthermore, providers captured for the remaining 210 children either served few children (less than 10) or were primarily responsible for case management, rather than EIBI services. Therefore, analysis of the provider network was not conducted.

Consistent with findings from Perry et al., earlier age at entry is associated with better outcomes on adaptive behaviors (Perry et al. 2011). The current study shows that higher average gains and top quartile gains are associated with younger age at entry. There is evidence that the effect of age on adaptive behavior outcomes varies by treatment duration, where gains are similar in the first 12 months of treatment, but greater diminishment in gains is seen in older age groups in the second year of treatment. Younger age at entry is also associated with positive average gains and top quartile gains in receptive language (PPVT), but not in receptive language (EVT).

Baseline language and adaptive behavior scores are each predictive of better outcomes. Higher adaptive behavior scores at baseline predict greater average gains in language outcomes and greater likelihood for top quartile gains in language outcomes. Similarly, higher receptive language scores at baseline predict greater average gains in

adaptive behaviors and higher likelihood for top quartile gains. Because receptive and expressive language were highly correlated (r>.9), they were not simultaneously considered in predicting adaptive behavior outcomes nor language outcomes. On all outcome measures, lower baseline performance on that measure was predictive of both greater average gains, and higher likelihood for top quartile membership. This could be reflective of regression to the mean, or it could represent the outcome of targeted intervention against specific deficits.

IQ has been frequently studied both as an outcome variable and as a predictor of outcomes. While the current study shows differences in outcomes associated with baseline IQ for the limited subset of children with an IQ measure, adaptive behavior and language performance at baseline yield a better predictive model. Autism severity as measured by CARS Total score did not moderate adaptive behavior or language outcomes.

2.4.1 Strengths

Strengths of this study include a substantially larger sample size than most comparable studies and an ability to explore results through 2 post-treatment measurement periods. Not only does this enable simultaneous analysis of predictors of positive outcomes, but it also allows for exploration of differing impacts across measurement periods. As seen with the changing impact of age at entry through post-treatment measurement periods, the finding that younger age is associated with positive outcomes is more nuanced. This study suggests that that impact varies by

measurement period. The large sample size also highlighted the common finding that children improved most on areas of greatest deficit, whether that was in language or adaptive behavior domains. It also allowed simultaneous exploration of program-related factors, in this case, year of enrollment, while controlling for baseline language and adaptive behavior levels. Finally, it allowed for assessment of the factors moderating average gains and the factors predicting top quartile performance. As seen, findings were consistent whether assessing moderators of average gains or the predictors of top quartile gains.

2.4.2 Limitations

Since this study is focused on predictors of change, limitations with the overall study discussed in a previous chapter are not as relevant to this study (Perry et al. 2008, Perry et al. 2011). Overall issues include a lack of a control group and limitations on the availability of outcome measures of IQ and autism symptoms (CARS Total Score), but that is not relevant for assessing predictors of gains in language and adaptive behaviors.

As discussed earlier, violation of normality assumptions for some of the analyses was a potential issue, particularly when exploring language outcomes. However, by conducting multiple analyses, including non-parametric tests, with identical conclusions, this issue was addressed.

It would have been interesting to explore the effect that different provider networks have on the outcomes. In addition to the sample size issues mentioned previously, further understanding of the providers would have been required in order to

undertake this aspect of the analysis. Referrals to individual providers may be related to geographic and socio-economic differences in the families and children served. Additionally, it would be useful to better understand differences observed by enrollment cohort and whether that was related to changes in providers over time, or, whether overall funding, referral sources and capacity impacted treatment impacts across enrollment cohorts.

Chapter 3

Missing Data in a Community-Based EIBI Program: Findings and Implications[4]

[4] Kuntz, J.M. To be submitted to a journal to be determined.

3.1 Background

Autism spectrum disorders (ASDs) are a group of lifelong, neurodevelopmental disorders characterized by deficits in social interaction, communication, relationship development and by the presence of repetitive or stereotypical behaviors including restricted interests (American Psychiatric Association 2013, Johnson, Myers, and the Council on Children with Disabilities 2007). Early Intensive Behavioral Intervention (EIBI) programs are designed to address the variety of deficits commonly observed among children with ASDs, through early, intensive, structured intervention. From initial work by Lovaas (Lovaas 1987) to more recent comprehensive evaluations, positive improvements have been attributed to EIBI across the domain of deficits present in autism (Peters-Scheffer et al. 2011, Virués-Ortega 2010, Reichow 2012, Eldevik et al. 2009, Howlin, Magiati, and Charman 2009). A key principal of these methods is the demonstration that interventions are responsible for observed improvements in behavior. As such, rigorous baseline and ongoing measurement is essential to EIBI programs.

Efficacy studies are designed to evaluate interventions under optimum conditions while effectiveness studies assess impacts under "real-world" conditions (Flay 1986). Much of the research on EIBI has established the efficacy of the treatment in small, closely-managed settings. There are few studies which test the effectiveness of EIBI treatment in a natural community setting where the 'vast majority of children' now receive treatment (Perry et al. 2008) (p623). Community and state-run programs introduce greater heterogeneity of children enrolled, training and supervision of staff,

treatment provided, and consistency and controls of data collection needed to evaluate outcomes (Wells 2014, Perry et al. 2008, Love et al. 2009). Turnover of EIBI therapists, lower levels of experience among therapists and supervisors, adherence to EIBI protocols and lower supervision of staff and parents have been identified as obstacles to the success of EIBI treatment in non-clinical settings (Eikeseth et al. 2012). Additionally, while ongoing measurement is foundational to EIBI programs, research is not the primary focus of data collection and data storage efforts. Most of these effectiveness studies have had small samples (less than 30) with treatment occurring in a single location or for a duration of less than 24 months (Sallows, Graupner, and MacLean Jr 2005, Hayward et al. 2009, Sheinkopf and Siegel 1998, Eldevik et al. 2012, Magiati, Charman, and Howlin 2007, Cohen, Amerine-Dickens, and Smith 2006, Howard et al. 2005). Therefore, even these effectiveness studies do not fully inform large-scale implementation, as they do not adequately consider issues such as quality control, data management, program adherence and subject variability.

A convergence of factors has led to increased opportunity to study EIBI effectiveness in large-scale, community-based programs. A rise in autism prevalence estimates, from 1 in 2000 in the early 1980s, to 1 in 68 currently, has increased awareness and demand for treatment (Baio 2014, Newschaffer et al. 2007). EIBI is considered a well-established intervention, producing outcomes consistent with the 'highest levels of evidence-based treatments' (Reichow 2012 p. 518, Matson and Jang 2013). Since 2007, EIBI has been the preferred treatment for children with autism according to the American Academy of Pediatrics and the National Institute of Child

Health and Human Development (Myers, Johnson, and the Council on Children with Disabilities 2007, Harvey et al. 2010). Accordingly, demand for EIBI services and funding for those services has risen as well (Harvey et al. 2010, Reichow 2012, Baio 2014, Newschaffer, Falb, and Gurney 2005). Insurance reform and expansion of coverage has further contributed to increased demand and utilization of EIBI services. As of March, 2015, 43 states plus the District of Columbia have legislation mandating insurance coverage for the diagnosis and treatment of autism compared to 15 states with mandated coverage in December, 2009 (National Conference of State Legislatures 2015, Autism Speaks). Finally, beginning in 2014, The Affordable Care Act (ACA) has closed some loopholes and inconsistencies in private insurance and state laws by requiring coverage of behavioral health treatment including EIBI services for autism (National Autism Network). In turn, EIBI research has moved beyond establishing efficacy, to demonstrating effectiveness and explaining heterogeneity of outcomes in real-world settings through observational studies often involving previously collected administrative data.

Increasingly, existing electronic medical records and other administrative data are being used for public health surveillance and epidemiologic research (Virnig and McBean 2001). These data sources are widely available, cover large groups of individuals in health care and education settings, and may be longitudinal in nature (Iezzoni 2002). However, administrative records and databases were not established for epidemiologic research and therefore, their use for that purpose has limitations which must be understood and managed (Grimes 2010, Iezzoni 2002). Research involving

analysis of administrative data tends to preclude investigator control over the timing of data collection as well as collection procedures themselves with possible implications for the completeness, level of detail and accuracy of the data collected (Schneeweiss and Avorn 2005). Those entering the data may do so without familiarity with how the data will be or could be used for research purposes. Additionally, there remains extensive use of pen and paper data collection in the provision of EIBI services (Tarbox et al. 2010). For EIBI services, where individual children have assessment and treatment data extending across several years, there is significant opportunity for the loss of individual paper files as well as inconsistent data availability through lengthy time periods. Further data loss may occur when disparate data collection processes are used for eligibility and enrollment into treatment programs, billing, and treatment outcomes. Despite this, there is little discussion in existing effectiveness research regarding the completeness of data and its impact on internal validity of the findings.

In their assessment of Ontario, Canada's EIBI implementation, Perry et al. explain that their analysis included data for 332 children, representing 'one-third of the entire population of children served in the program during that time period' (Perry et al. 2008 pp 625-626). They indicated that inclusion in the data set was not based on progress, child or family factors, and that it was therefore likely to be representative of the treatment population. Additionally, as evidence of the obstacles discussed earlier, inclusion in the final sample was based on having one outcome measure at two time periods, but they explained that not all data within the final data set was complete.

Actual sample sizes presented in their analyses generally varied between 273 and 292 observations, with more limited availability of IQ data.

In a study of children who received EIBI treatment in community-based programs, Smith et al. discuss dropout among the 206 children who were referred for inclusion in the study to a final sample of 71. Although 36 children were excluded due to specific study eligibility criteria, 91 others were excluded due to limited data availability, and 8 more refused to participate (Smith, Klorman, and Mruzek 2015). There is no further discussion of whether the final study sample of 71 were similar to the 99 who were dropped for reasons beyond study criteria. Smaller effectiveness studies conducted in one location across narrower timeframes are less impacted by the data management issues discussed. Eikeseth et al. conducted an EIBI effectiveness study among a group of children who were enrolled in an EIBI treatment program between March of 2008 and May of 2010 and received one year of EIBI treatment through one location (Eikeseth et al. 2012). Of the 38 children receiving services, 35 were included in the study, with dropout related to study eligibility criteria, not data availability.

There is still an urgent need to understand child-specific and treatment-related factors that explain the heterogeneity of outcomes associated with EIBI services. The increased utilization of EIBI services in large, community-based settings has dramatically increased the availability of data to conduct this research. In earlier chapters, I demonstrated overall improvements in language and adaptive behavior outcomes

among 419 children who received two years of EIBI treatment through the South Carolina Department of Disability and Special Needs (SC DDSN) Pervasive Development Disorders Program. Additionally, I demonstrated that age-at-entry and baseline language and adaptive behavior performance moderate outcomes. The purpose of this paper is to examine whether overall data availability impacts the internal validity of those findings. Missing and incomplete files from 32% (196) of the 615 children who received EIBI services in the study time frame (2007-2011) prevented inclusion in the final analysis sample. This paper examines the characteristics of children included in the study, compared to those excluded based on data completeness and discusses the impact on study validity and the implications for conducting EIBI research with administrative data in a loosely-managed environment.

3.1.1 Setting

During the 2006 legislative session of the South Carolina General Assembly, $3 million was appropriated to the South Carolina Department of Disabilities and Special Needs (DDSN) to develop the Pervasive Developmental Disorder (PDD) Program by January 2007. During the 2007 session, the General Assembly appropriated an additional $4.5 million demonstrating their commitment to the treatment of autism and other PDDs. The following description is excerpted from the DDSN manual (South Carolina Department of Disabilities and Special Needs 2013):

> The purpose of the PDD Program is to provide intensive in-home intervention to children ages 3 through 10 years diagnosed with a Pervasive Developmental

Disorder, which includes Autism, Asperger's and PDD – NOS (Not Otherwise Specified) (prior to changes in diagnostic criteria introduced in DSM-5). Children must be eligible to receive Medicaid or have documentation of financial ineligibility. Children must meet the Level of Care (LOC) assessment requiring the degree of care that would be required in an Intermediate Care Facility for the Intellectually Disabled (ICF/ID). Children who meet these criteria, and receive a PDD diagnosis before age eight may receive Early Intensive Behavioral Intervention (EIBI) services for three years or until their 11th birthday, whichever comes first. They also receive Case Management services. The EIBI services are designed to develop skills in cognition, behavior, communications and social interaction. They are provided face-to-face in the child's natural environment, which may include the home or community locations. This environment specifically excludes any educational setting where educational services are simultaneously provided during school hours. Case management services assist the children and their families with all aspects of the PDD program.

3.1.2 EIBI Provider Services

EIBI service providers are selected by the child's parents. Authorized service providers, who are individually vetted by SC DDSN, are responsible for recruiting, hiring, retaining, and terminating employees. EIBI services have the oversight of an EIBI Consultant. Five separate components are included:

1. Assessment of the child's current needs, conducted by the EIBI consultant.

2. Program Development and Training, provided by the EIBI consultant, involving the development of an individualized treatment plan and provides training to family members and therapists who implement the individual interventions.

3. Plan implementation, which is also provided by the EIBI Consultant and involves implementation of the plan, monthly monitoring of the effectiveness of the plan and supervision of the Lead and Line therapists who deliver the individual plan.

4. Lead therapy, which is provided by a Lead Therapist and involves oversight and weekly monitoring of the implementation and effectiveness of the plan and review of all recorded data.

5. Line therapy, which is provided by a Line Therapist who is responsible for carrying out the treatment plan as designed and recording data associated with monitoring and measuring outcomes.

Providers are required to ensure adherence to the Essential Practice Elements of ABA (Per the Behavior Analyst Certification Board, Inc. Guidelines) through all phases of assessment and treatment.

3.2 Methods

3.2.1 Measures

Each child enrolled in the program received a pre-treatment diagnosis of autism with evidence from at least two of the following:

- Autism Diagnostic Observation Schedule (ADOS) (Lord et al. 1999)
- Autism Diagnostic Interview (ADI) (Lord and Rutter 1994)
- Social Communication Questionnaire (SCQ) (Rutter, Bailey, and Lord 2003)
- Childhood Autism Rating Scale (CARS) (Schopler, Reichler, and Renner 1988)

Upon acceptance into the program, the assessment conducted by the EIBI consultant includes the following measures:

- Receptive language using the Peabody Picture Vocabulary Test IV (PPVT) (Dunn and Dunn 1981)
- Expressive language using the Expressive Vocabulary Test (EVT) (Williams 1997)
- Adaptive behaviors using the Vineland Adaptive Behavior Scales-II (VABS) (Sparrow, Balla, and Cicchetti 2005)

EIBI consultants are required to submit annual updated assessments to the case manager of each of the above measures. The adaptive behavior assessment data using Vineland includes standardized scores on Communication, Socialization and Daily Living Skills as well as an Adaptive Behavior Composite (ABC) across those dimensions. It also includes standardize scores for Motor Skills and Maladaptive Behaviors. Additionally, the available data includes sub-domain-level age equivalent and v-scale scores.

Beyond the assessment data, weekly treatment hours allocated and actual treatment hours delivered by the EIBI consultant, the lead therapist and the line therapist were captured. In addition to assessment and treatment data, additional child-specific data captured from the SC DDSN files included age-at-intake, gender and race and year of enrollment (cohort).

3.2.2 Participants

As of March, 2013, 1526 children had received PDD services under this program. This study examines program participants who completed two years of EIBI treatment in program cohort years 2007-2011. During the study time period 615 children met the inclusion criteria through March, 2013.

3.2.3 Data Collection

Data were obtained from stored paper files and entered into an Excel spreadsheet by two graduate assistants under the direction of the study author and the direct supervision of SC DDSN personnel. Additional paper files containing the original diagnostic measures for these children were obtained from the 4 Regional DDSN Care Centers. Data collection sought to have complete data from all 3 input sources:

- Enrollment and billing data stored electronically in DDSN internal systems
- Diagnostic measurement preceding enrollment in the DDSN program obtained in paper files from Regional DDSN Care Centers for this effort

- Assessment data collected throughout the EIBI treatment program and stored in DDSN paper files

The purpose of the data collection was to evaluate the overall effectiveness of the SC DDSN EIBI treatment program. From the overall population of 615 children where a complete record was sought, an analysis population was created that included those children who had at least 2 out of 3 measures (baseline, Year 1, Year 2) for any of the primary 3 outcome metrics: Expressive Language (EVT), Receptive Language (PPVT) and the Vineland Adaptive Behavior Composite Score (ABC). This criteria provides the largest possible population to study changes in average score for the children receiving treatment and resulted in a population of 419 children.

3.2.4 Analysis

The focus of this analysis is to understand the data availability, validity and implications for ongoing program evaluation. Focusing on the 615 children where a complete record was sought, this analysis addresses the following:

- What is the overall completeness of the data?
- Is the final analysis sample representative of the starting population?
- What factors are associated with missing data from the final analysis sample of 419 children?

3.3 Results

3.3.1 Question 1: How complete is the data for the 615 children who received two-years of treatment?

Table 3.1 shows the overall completeness of the final DDSN dataset for static measures. As seen, paper files stored at DDSN identified 615 children with enrollment date (100%) and birthdate (99.8%) nearly complete. Other enrollment data, including name of the EIBI treatment provider (93.7%), source of funding (89.9%), diagnosis (86.7%), and allocated treatment hours per week (66.8%) were less complete. Finally, diagnostic data obtained from the Regional Centers was even more limited. While 63.7% had at least one of the four measures available, fewer than 43% of the children had a score from the same standard diagnostic instrument. CARS Total Score was obtained for 42.3% of the children and only 12.2% of the children had an IQ score.

Table 3.2 provides completeness of time-dependent measures, including baseline and outcome data, as well as treatment hours provided by year (line, lead, and plan). Incomplete data is seen in the assessment data collected annually with generally fewer than 400 of the 615 children having baseline and subsequent measures from VABS, PPVT and EVT.

3.3.2 Question 2: Are children with complete records representative of program participants?

Given the objective of assessing overall program effectiveness, a final analytical sample was constructed by including those children who had at least 2 measures for at least

Table 3.1 Data completeness – Time independent variables

Source	Variable	N	Missing N	% Complete
Enrollment	Consumer #	615	0	100.0%
	Date of Birth	614	1	99.8%
	Enrollment Date	615	0	100.0%
	Diagnosis	533	82	86.7%
	Gender	426	189	69.3%
	Funding Source (State, Waiver)	553	62	89.9%
	EIBI Treatment Provider	576	39	93.7%
	Allocated Treatment Hours per Week	411	204	66.8%
Pre-Enrollment Diagnostic Data	Availability of any initial diagnostic measure	392	223	63.7%
	ADOS Mod1 Communication Total	228	387	37.1%
	IQ or IQ Equivalence Score	75	540	12.2%
	Social Comm Questionnaire (SCQ)	141	474	22.9%
	CARS Total Score	260	355	42.3%

Table 3.2 Data completeness – Time dependent variables

		Baseline			Year 1			Year 2		
Source	Variable	N	N Missing	% Complete	N	N Missing	% Complete	N	N Missing	% Complete
Enrollment	Line Therapy Hours	NA	NA	NA	388	227	63.1%	388	227	63.1%
	Lead Therapy Hours	NA	NA	NA	390	225	63.4%	389	226	63.3%
	Plan Therapy Hours	NA	NA	NA	388	227	63.1%	389	226	63.3%
Assessment and Outcomes	Expressive Vocabulary Test	372	243	60.5%	386	229	62.8%	382	233	62.1%
	Peabody Picture Vocabulary Test	377	238	61.3%	392	223	63.7%	384	231	62.4%
	VABS Adaptive Behavior Composite	381	234	62.0%	392	223	63.7%	398	217	64.7%
	VABS Communication	399	216	64.9%	398	217	64.7%	408	207	66.3%
	VABS Daily Living Skills	398	217	64.7%	398	217	64.7%	408	207	66.3%
	VABS Socialization	399	216	64.9%	396	219	64.4%	411	204	66.8%
	VABS Motor Skills	380	235	61.8%	367	248	59.7%	360	255	58.5%
	Maladaptive Behaviors	312	303	50.7%	315	300	51.2%	324	291	52.7%

one of the primary outcomes. This resulted in a final analysis sample of 419 children. One of the obvious challenges in determining whether the final sample is representative of the entire treatment population is the limited availability of data to make the comparison. Tables 3.3 and 3.4 provides differences in data completeness based upon inclusion or exclusion in the final sample. This will identify which variables can be used to assess differences in the two groups (included and excluded). As seen earlier, enrollment year and child's age-at-enrollment are generally available for both samples. Diagnosis is missing equally in both samples with 86.2% of the excluded sample and 86.9% of the analysis sample having a diagnosis in the final data set. EIBI Treatment provider (81.6%), funding source (69.9%), gender (26.5%) and allocated treatment hours (20.9%) were more likely to be missing in the excluded sample, compared to the analysis sample which was approximately 90% complete for each of those variables. Despite limited availability overall, pre-enrollment diagnostic data (CARS, ADOS, SCQ, and IQ) has similar availability in the final analysis sample and the excluded sample. Because the samples were defined based upon the availability of assessment data, this data is generally unavailable in the excluded sample with fewer than 5% of this sample having assessment data available. Based on this view of data availability, comparisons between samples can be made where the data is sufficiently available in each sample.

To assess whether this analysis sample is representative of the population of 615 children who completed the program, two analyses were conducted. The first analysis looked at frequency distributions of categorical variables that were available in both samples and conducted chi-square tests of differences. Table 3.5 compares the

Table 3.3 Comparison of data completeness between final sample and excluded sample – Time independent variables

		Excluded Sample (n=196)		Analysis Sample (n=419)	
Source	Variable	N	% Complete	N	% Complete
Enrollment	Date of Birth	195	99.5%	419	100.0%
	Enrollment Date	196	100.0%	419	100.0%
	Diagnosis	169	86.2%	364	86.9%
	Gender	52	26.5%	374	89.3%
	Funding Source (State, Waiver)	137	69.9%	416	99.3%
	EIBI Treatment Provider	160	81.6%	416	99.3%
	Allocated Treatment Hours per Week	41	20.9%	370	88.3%
Pre-Enrollment Diagnostic Data	Availability of any initial diagnostic measure	77	39.3%	146	34.8%
	ADOS Mod1 Communication Total	61	31.1%	167	39.9%
	IQ or IQ Equivalence Score	20	10.2%	55	13.1%
	Social Communication Questionnaire (SCQ)	32	16.3%	109	26.0%
	CARS Total Score	91	46.4%	169	40.3%

Table 3.4 Comparison of data completeness between final sample and excluded sample – Time dependent variables

Source	Variable	Baseline N	Baseline % Complete	Year 1 N	Year 1 % Complete	Year 2 N	Year 2 % Complete
Excluded Sample (n=196)							
Enrollment	Line Therapy Hours	NA	NA	13	6.6%	13	6.6%
	Lead Therapy Hours	NA	NA	13	6.6%	13	6.6%
	Plan Therapy Hours	NA	NA	13	6.6%	13	6.6%
Assessment and Outcomes	Expressive Vocabulary Test	8	4.1%	7	3.6%	5	2.6%
	Peabody Picture Vocabulary Test	7	3.6%	8	4.1%	4	2.0%
	VABS Adaptive Behavior Composite	6	3.1%	5	2.6%	6	3.1%
	VABS Communication	7	3.6%	8	4.1%	6	3.1%
	VABS Daily Living Skills	7	3.6%	8	4.1%	6	3.1%
	VABS Socialization	7	3.6%	7	3.6%	6	3.1%
	VABS Motor Skills	7	3.6%	7	3.6%	5	2.6%
	Maladaptive Behaviors	3	1.5%	4	2.0%	4	2.0%
Analysis Sample (n=419)							
Enrollment	Line Therapy Hours	NA	NA	375	89.5%	376	89.7%
	Lead Therapy Hours	NA	NA	377	90.0%	376	89.7%
	Plan Therapy Hours	NA	NA	375	89.5%	376	89.7%
Assessment and Outcomes	Expressive Vocabulary Test	364	86.9%	379	90.5%	377	90.0%
	Peabody Picture Vocabulary Test	370	88.3%	384	91.6%	380	90.7%
	VABS Adaptive Behavior Composite	375	89.5%	387	92.4%	392	93.6%
	VABS Communication	392	93.6%	390	93.1%	402	95.9%
	VABS Daily Living Skills	391	93.3%	390	93.1%	402	95.9%
	VABS Socialization	392	93.6%	389	92.8%	405	96.7%
	VABS Motor Skills	373	89.0%	360	85.9%	355	84.7%
	Maladaptive Behaviors	309	73.7%	311	74.2%	320	76.4%

Table 3.5 Comparison of final analysis sample to excluded sample: Frequency distribution

		Excluded Sample % of row	Analysis Sample % of row	p
Cohort				
	2007 (n=128)	51.6%	48.4%	<.0001
	2008 (n=131)	46.6%	53.4%	
	2009 (n=124)	34.7%	65.3%	
	2010 (n=161)	12.4%	87.6%	
	2011 (n=71)	8.5%	91.6%	
	Total(n=615)	31.9%	68.1%	
Diagnosis				
	Asperger's (n=14)	7.1%	92.9%	0.018
	Autism (n=463)	34.3%	65.7%	
	PDD/Other (n=46)	17.4%	82.6%	
	Rett's Syndrome (n=1)	0.0%	100.0%	
	Total (n=524)	32.1%	67.9%	
Funding Source				
	State (n=60)	26.7%	73.3%	0.719
	Medicaid Waiver (n=493)	24.5%	75.5%	
	Total (n=553)	24.8%	75.2%	
Gender				
	Female (n=65)	13.9%	86.2%	0.661
	Male (n=361)	11.9%	88.1%	
	Total (n=426)	12.2%	87.8%	

distributions by enrollment cohort, diagnosis, funding source, and gender. Because of the number of distinct EIBI providers, this variable was not considered. As seen, more recent program years were more likely to have sufficient data to be included in the analysis sample ($p<.0001$). While autism as a diagnosis comprised 88% of the combined samples, those in the final analysis sample were slightly more likely to receive some other diagnosis such as Asperger's ($p=.018$). No differences were seen with respect to source of funding or gender between samples.

A second analysis was conducted to assess mean difference between samples on available continuous data. T-tests were performed to assess differences in mean age-at-enrollment, ADOS, CARS, SCQ and IQ. As seen in Table 3.6, the analysis population is younger on average by approximately 5 months ($p=.0002$). Additionally, average performance on the Social Communications Questionnaire (SCQ) is lower in the analysis population than in the excluded population ($p=.047$). This pattern is suggestive of lower autism symptomology in the study population compared to the excluded population. No significant differences are seen in measures of CARS Total Score, ADOS, or IQ. Importantly, younger age is a cornerstone of EIBI treatment and a critical variable in research of moderators of EIBI outcomes. In the previous chapter, younger age was a significant moderator of average gains in adaptive behaviors and expressive language. Similarly, enrollment cohort moderated average gains in adaptive behaviors, with later cohorts (2010, 2011) showing smaller gains after controlling for other factors. Last, baseline measures of adaptive behaviors and language were significant moderators of language and adaptive behavior outcomes. While comparisons between samples could

Table 3.6 Comparison of final analysis sample to excluded sample: Mean values

	n	mean	std	p
Age at Enrollment				
Not Contained in Analysis Population	194	6.359	1.530	0.0002
Contained in Analysis Population	418	5.929	1.228	
ADOS				0.154
Not Contained in Analysis Population	61	5.902	1.795	
Contained in Analysis Population	167	5.467	2.108	
Cars Total Score				0.695
Not Contained in Analysis Population	91	35.687	4.763	
Contained in Analysis Population	169	35.944	5.167	
IQ				0.101
Not Contained in Analysis Population	20	65.450	18.360	
Contained in Analysis Population	55	73.327	18.119	
SCQ				0.047
Not Contained in Analysis Population	32	22.594	5.506	
Contained in Analysis Population	109	20.193	6.082	

not be performed for these measures, baseline differences in diagnosis and SCQ score may be suggestive of underlying differences in language and adaptive behavior as SCQ is negatively correlated with baseline measures of Adaptive Behavior Composite (r=-.214, p=.03), Expressive Language (r=-.283, p=.004) and Receptive Language (r=-.25, p=.013).

Despite some differences in age-at-enrollment, SCQ score and diagnosis, many other variables show similarities between the analysis sample and the excluded sample. Additionally, while significant, differences in age are small and differences in diagnosis will have limited impact on internal validity, given the large proportion of those with a diagnosis of 'autism' in both samples.

3.3.3 Question 3: How complete is the data for the final analysis population?

Additional analysis was done to assess factors associated with missing data within the final analysis sample. Previously, Table 3.4 showed the availability of key outcome data for the 419 children in the final sample at baseline, Year 1, and Year 2. For this analysis, an indicator was created to measure whether any of the key outcome metrics (ABC, PPVT, and EVT) had any missing data across the 3 periods. Of the 419 records in the final analysis population, 144 (34%) had at least one missing measurement as just described. Table 3.7 compares average scores between the sub-population that has complete data, and those that are missing at least one outcome measure for at least one measurement period. As shown in Table 3.7, there are no statistically significant differences in any of the variables analyzed between the two sub-populations.

Table 3.7 Association between missing any outcome data (ABC, PPVT, EVT) and baseline measures in analysis sample

Baseline Measure		n	mean	std	p
Receptive Language Score (PPVT)					0.295
	Complete data for all outcomes	275	64.99	30.85	
	Missing data for at least one measure	95	68.80	29.46	
Expressive Language Score (EVT)					0.196
	Complete data for all outcomes	275	65.81	32.39	
	Missing data for at least one measure	89	70.89	31.08	
Vineland Adaptive Behavior Composite (ABC)					0.474
	Complete data for all outcomes	275	67.28	13.02	
	Missing data for at least one measure	100	66.20	12.34	
Vineland Communications					0.916
	Complete data for all outcomes	273	69.29	17.88	
	Missing data for at least one measure	119	69.50	17.84	
Vineland Daily Living Skills					0.367
	Complete data for all outcomes	272	69.46	15.31	
	Missing data for at least one measure	119	67.95	14.96	
Vineland Motor Skills					0.861
	Complete data for all outcomes	263	73.64	14.07	
	Missing data for at least one measure	110	73.35	14.25	
Vineland Socialization Skills					0.237
	Complete data for all outcomes	273	66.38	12.72	
	Missing data for at least one measure	119	64.72	12.75	
Treatment Hours Approved					0.598
	Complete data for all outcomes	266	28.78	3.29	
	Missing data for at least one measure	104	28.60	2.73	
Age-at-Enrollment					0.423
	Complete data for all outcomes	275	5.96	1.26	
	Missing data for at least one measure	143	5.86	1.17	
ADOS					0.799
	Complete data for all outcomes	114	5.44	2.24	
	Missing data for at least one measure	53	5.53	1.81	
CARS Total Score					0.415
	Complete data for all outcomes	116	35.72	5.00	
	Missing data for at least one measure	53	36.42	5.53	
IQ					0.334
	Complete data for all outcomes	38	74.92	16.33	
	Missing data for at least one measure	17	69.76	21.72	
Social Communication Questionnaire (SCQ)					0.915
	Complete data for all outcomes	69	20.14	5.94	
	Missing data for at least one measure	40	20.28	6.40	

In a previous chapter, SAS Proc Mixed was used to estimate changes in adaptive behavior and language scores (ABC, EVT, and PPVT) across the 3 time periods. One benefit of using Proc Mixed is that observations with some missing data can contribute to the overall mean estimation. The analysis presented here shows that those observations in the final analysis data set with some missing data are similar to those observations that had complete data, increasing the likelihood that estimates obtained from SAS Proc Mixed are reflective of the full sample and not impacted by the 'missingness' of the data.

Thus far, this analysis has shown that the final analysis population is generally representative of the original treatment population. In addition, it has shown that, despite missing outcome data for some time periods, the analysis sub-population with 100% complete data is similar to the sub-population with some missing outcome data. This is important when drawing conclusions regarding the average change. For evaluating change at the child level, baseline measures of performance are critical. In an earlier chapter, logistic regression was used to identify the factors associated with top quartile gains from baseline to the end of year 2, for each of the outcome variables: EVT, PPVT and ABC. This last analysis looks at the availability of baseline (pre-treatment) measures in the final analysis population and assesses whether those missing baseline data for the key outcome measures differed from those who had baseline data available. As seen in Table 3.8, across a variety of variables, there is limited evidence that the presence or absence of missing baseline data is related to independent measures which may be associated with an individual child's likelihood to

Table 3.8 Mean measurement differences by presence of specified baseline variables (ABC, PPVT, EVT)

		Have Baseline Data			Missing Baseline Data				
Baseline Variable	Measure	n	mean	std	n	mean	std	t	p
Adaptive Behavior Composite Baseline (ABC)	Receptive Language Score (PPVT)	346	65.5	30.4	24	72.8	31.9	-1.13	0.258
	Expressive Language Score (EVT)	339	66.5	32.1	25	74	32.3	-1.12	0.263
	Vineland Adaptive Behavior Composite (ABC)								
	Vineland Communications	372	69.2	17.9	20	72.3	17.6	-0.76	0.449
	Vineland Daily Living Skills	371	68.9	15.2	20	71.1	15.7	-0.62	0.536
	Vineland Motor Skills	359	73.5	14	14	76	16.8	-0.66	0.509
	Vineland Socialization Skills	372	65.9	12.8	20	66.2	11.5	-0.1	0.921
	Treatment Hours Approved	346	28.8	3.1	24	27.3	3.3	2.32	0.021
	Age-at-Enrollment	374	5.9	1.2	44	6	1.1	-0.26	0.796
	ADOS	148	5.3	1.9	18	5.8	1.4	-1.06	0.291
	CARS Total Score	154	35.9	5.2	15	36.1	5.3	-0.15	0.882
	IQ	49	72.2	17.4	6	82.3	22.9	-1.3	0.200
	Social Communication Questionnaire (SCQ)	100	20.2	6.1	9	19.7	6.5	0.27	0.788
Receptive Language Baseline (PPVT)	Receptive Language Score (PPVT)								
	Expressive Language Score (EVT)	362	67.1	32.1	2	59	55.2	0.36	0.723
	Vineland Adaptive Behavior Composite (ABC)	346	67.3	12.8	29	62.8	13.4	1.82	0.069
	Vineland Communications	363	69.9	17.5	29	62.8	20.4	2.08	0.039
	Vineland Daily Living Skills	362	69.6	14.9	29	61	17.3	2.99	0.003
	Vineland Motor Skills	346	73.7	14	27	72.1	15.4	0.55	0.582
	Vineland Socialization Skills	363	66.3	12.6	29	59.9	13.2	2.63	0.009
	Treatment Hours Approved	341	28.7	3.2	29	29	2.6	-0.55	0.582
	Age-at-Enrollment	369	6	1.2	49	5.6	1.1	1.83	0.068
	ADOS	144	5.3	1.9	22	5.8	1.3	-1.54	0.133
	CARS Total Score	155	35.8	5.2	14	38	4.7	-1.56	0.120
	IQ	51	73	17.6	4	78	26.5	-0.53	0.597
	Social Communication Questionnaire (SCQ)	99	20.1	6.1	10	21.1	5.6	-0.49	0.623
Expressive Language Baseline (EVT)	Receptive Language Score (PPVT)	362	66.5	30.5	8	40.3	16.2	2.43	0.016
	Expressive Language Score (EVT)								
	Vineland Adaptive Behavior Composite (ABC)	339	67.4	12.9	36	63.5	12.3	1.73	0.084
	Vineland Communications	356	69.9	17.7	36	63.6	18.7	2.05	0.041
	Vineland Daily Living Skills	355	69.5	14.9	36	63.6	17.4	2.23	0.026
	Vineland Motor Skills	339	73.8	14	34	71	14.8	1.11	0.269
	Vineland Socialization Skills	356	66.4	12.7	36	60.7	12.1	2.59	0.010
	Treatment Hours Approved	336	28.7	3.2	34	29	2.6	-0.48	0.634
	Age-at-Enrollment	363	6	1.2	55	5.7	1.1	1.78	0.075
	ADOS	143	5.3	1.9	23	6	1.5	-1.69	0.094
	CARS Total Score	153	35.7	5.2	16	38	5	-1.68	0.094
	IQ	49	74.4	16.2	6	64.5	30.2	0.79	0.463
	Social Communication Questionnaire (SCQ)	99	20.1	6.1	10	21.1	5.6	-0.49	0.623

benefit from the treatment. Among those missing baseline expressive language scores (EVT), there is lower performance on baseline receptive language (p=.016), and Vineland Communication (p=.041), Vineland Daily Living Skills (p=.026) and Vineland Socialization Skills (p=.01). Similar results were found for observations missing baseline receptive language scores (PPVT), where lower Vineland Communications, Daily Living Skills, and Socialization Skills were seen. No differences were seen between samples based upon the presence or absence of baseline Adaptive Behavior Composite Score.

3.4 Discussion

This study provides a comprehensive analysis of the completeness of data in an observational study of the effectiveness of EIBI and the moderators of outcomes in a community-based EIBI treatment program. As more children are diagnosed with autism and treated through large, community-based programs, the administrative data collected provides a potentially rich source of research data. Larger treatment populations in these settings provide greater statistical power to investigate child-specific moderators and predictors of the heterogeneity of outcomes seen in EIBI research such as age, treatment intensity, and treatment duration. Furthermore, such settings permit the investigation of program implementation and program management variables that can translate more directly into local policy decisions. These include the selection of providers of EIBI services, the impact of qualifications and training of EIBI consultants and further study into the effects of treatment duration and intensity, which have implications for costs and family burden.

Nonetheless, to better align research and treatment program objectives, improvements in data collection and data management are warranted. As part of these improvements, use of data collection tools to reduce reliance on paper and manual data entry are needed. In turn, training consultants and service providers in the use of these tools and the potential benefits of improved data collection is also needed. Rigorous measurement and demonstration of improvement is a key tenet of EIBI treatment. This study shows the overall occurrence of missing data in a non-research setting. It highlighted some potential threats to the internal validity of results presented in earlier chapters regarding program effectiveness based upon data completeness. To a large extent, this study demonstrated that the final analytical sample was generally representative of the overall treatment population. At the same time, it highlights the potential of a more rigorous data management focus. One-third of the analysis sample was lost to missing data. At an individual variable level, proportions of missing data were even higher. Based on the limited number of large-scale effectiveness studies, these data issues are not unique. Now that EIBI services are covered by insurance and widely sought, improvements to data collection and data management are warranted to enhance the research application of existing program data.

Discussion

This study is one of the most rigorous studies to date on the effectiveness of EIBI in a community setting. Taking advantage of existing data collected through South Carolina's Department of Disabilities and Special Needs Pervasive Development Disorders Program, this study assessed overall effectiveness, determined predictors of positive outcomes and examined the impact of data management and data availability on the validity of findings with a sample of more than 600 children with ASDs who received EIBI services. This study corroborates and advances findings regarding age at which EIBI treatment should begin and the impact of treatment duration on treatment outcomes. It provides important new results on baseline child characteristics that predict positive outcomes. It replicates other published findings that fewer than 50% of children experience desired outcomes. Finally, it provides results of the only known analysis of data availability and its potential impact on study validity in a large-scale, multi-year EIBI treatment program. Together, these findings have important implications for the implementation and management of community-based EIBI programs.

Overall outcomes

Looking at complete assessment data for 419 children, this study demonstrated overall average improvements in adaptive behavior (ABC) and language outcomes (EVT, PPVT) associated with two years of EIBI treatment, and 'medium-sized' gains in adaptive behaviors for 40% of the children. These outcomes exceed those observed in the only other comparably sized study to date (Perry, 2008), which observed only limited gains in adaptive behavior composite standard scores. Average gains in ABC standard scores were comparable to those observed by Bibby et al. (2002) and Eikeseth et al. (2012) on smaller samples but effect sizes were smaller than those observed by Eikeseth (2012). This study is one of few studies that explicitly compares outcomes achieved during the first year of treatment to the outcomes achieved during the second year of treatment. Statistically significant improvements were observed in each of the 12 month periods for adaptive behavior composite outcomes and expressive language (EVT) outcomes. Receptive language (PPVT) did not show statistically significant gains in the second 12-month period. This follows larger gains seen in the first year on all outcomes. No association was seen between the average weekly number of treatment hours received and changes in language and adaptive behavior outcomes. This study found that approximately 40% of children achieved moderate gains (calculated as an increase of 7.4 or greater on standard score on adaptive behavior outcomes while 25% achieved large gains. Results for receptive and expressive language outcomes were smaller.

Predictors of outcomes

This study reports on the predictors of average gains and individual child outcomes on adaptive behavior and language following two years of EIBI treatment. Consistent with findings from Perry et al., earlier age at entry is associated with better outcomes on adaptive behaviors (Perry et al. 2011). The current study shows that higher average gains and top quartile gains are associated with younger age at entry. There is evidence that the effect of age on adaptive behavior outcomes varies by treatment duration, where gains are similar in the first 12 months of treatment, but greater diminishment in gains is seen in older age groups in the second year of treatment. Younger age at entry is also associated with positive average gains and top quartile gains in receptive language (PPVT), but not in receptive language (EVT).

Baseline language and adaptive behavior scores are each predictive of better outcomes. Higher adaptive behavior scores at baseline predict greater average gains in language outcomes and greater likelihood for top quartile gains in language outcomes. Similarly, higher receptive language scores at baseline predict greater average gains in adaptive behaviors and higher likelihood for top quartile gains. Because receptive and expressive language were highly correlated ($r>.9$), they were not simultaneously considered in predicting adaptive behavior outcomes nor language outcomes. On all outcome measures, lower baseline performance on that measure was predictive of both greater average gains, and higher likelihood for top quartile membership. This could be reflective of regression to the mean, or it could represent the outcome of targeted intervention against specific deficits.

Data availability and impact

The final set of analyses investigated the overall completeness of data and its potential impact on overall study findings. Missing and incomplete files from 32% (196) of the 615 children who received EIBI services in the study time frame (2007-2011) prevented inclusion in the final analysis sample of 419 children. Data availability for this study exceeded that of the Perry study, which reported availability for approximately one third of the children served by the program during the study period (Perry et al. 2008). In this study, earlier cohort years were more likely to have missing data. The average age of children included in the study was approximately 5 months lower than those excluded due to missing data (5.9 years in the analysis sample versus 6.3 years in the excluded sample). The analysis sample was slightly more likely to have children with a diagnosis of Asperger's (3.6%) compared to those excluded (.5%). Missing data for baseline language and adaptive behavior scores prevented comparisons between samples for these measures. However, the analysis sample showed higher performance on The Social Communications Questionnaire (SCQ). This may be suggestive of underlying differences in language and adaptive behavior as SCQ is negatively correlated with baseline measures of Adaptive Behavior Composite ($r=-.214$, $p=.03$), Expressive Language ($r=-.283$, $p=.004$) and Receptive Language ($r=-.25$, $p=.013$). Despite some differences in age-at-enrollment, SCQ score and diagnosis, most other variables showed similarities between the analysis sample and the excluded sample. Additionally, while significant, differences in age are small and differences in diagnosis will have limited

impact on internal validity, given the large proportion (88%) of those with a diagnosis of 'autism' in the overall population.

Strengths

This is one of an extremely limited number of studies to evaluate EIBI outcomes in a large-scale, non-research setting. Using previously collected data, this study provides a real-world assessment of EIBI effectiveness through the limitations of non-research data. Research involving analysis of administrative data tends to preclude investigator control over the timing of data collection as well as collection procedures themselves with possible implications for the completeness, level of detail and accuracy of the data collected (Schneeweiss and Avorn 2005). However, an associated strength of using administrative data is a substantially larger sample size than most comparable studies and an ability to explore results through 2 post-treatment measurement periods. Not only does this enable simultaneous analysis of predictors of positive outcomes, but it also allows for exploration of differing impacts across measurement periods. As seen with the changing impact of age at entry through post-treatment measurement periods, the finding that younger age is associated with positive outcomes is more nuanced. This study suggests that that impact varies by measurement period. The large sample size also highlighted the common finding that children improved most on areas of greatest deficit, whether that was in language or adaptive behavior domains. It also allowed simultaneous exploration of program-related factors, in this case, year of enrollment, while controlling for baseline language and adaptive behavior levels. It also allowed for assessment of the factors moderating

average gains and the factors predicting top quartile performance. As seen, findings were consistent whether assessing moderators of average gains or the predictors of top quartile gains. Last, this is the only study of its kind to examine the potential drawbacks of using historical administrative data through analysis of study inclusion and the associated factors.

Limitations

Despite the benefits of large samples, the study has limitations. Despite specific analysis to quantify and assess missing data, one-third of those treated were excluded from the final analysis sample due to missing data. Additionally, the finding that EIBI outcomes did not vary by treatment intensity was also subject to data availability. Based on discussion with program administrators, treatment hours for the second year of treatment were considered most representative due to inconsistent data capture in the first and third year of treatment. More extensive use of billing data may have provided a more accurate year-over-year representation of actual hours received in each year of treatment, but this was unavailable in the data collection effort. In several analyses, violations of normality assumptions were a potential issue, particularly when exploring language outcomes. This was addressed by conducting multiple analyses, including the use of non-parametric tests, with identical conclusions.

Recommendations

This study provided a comprehensive analysis of the program effectiveness, the moderators of outcomes and completeness of data in an observational study of a

community-based EIBI treatment program. As more children are diagnosed with autism and treated through large, community-based programs, the administrative data collected provides a potentially rich source of research data. Given the findings, reinforced here, that only a subset of children are benefitting from EIBI treatment, larger research samples are needed to better explore the moderators of outcomes. Larger treatment populations in these settings provide greater statistical power to investigate child-specific moderators and predictors of the heterogeneity of outcomes seen in EIBI research such as age, treatment intensity, and treatment duration. These settings are uniquely positioned to explore rarely-analyzed factors such as the family and social environment that may affect child-specific outcomes. Furthermore, such settings permit the investigation of program implementation and program management variables that can translate more directly into local policy decisions. These include the selection of providers of EIBI services, the impact of qualifications and training of EIBI consultants and further study into the effects of treatment duration and intensity, which have implications for costs and family burden.

However, to explore these factors and to better align research and treatment program objectives, improvements in data collection and data management are warranted. This study shows the overall occurrence of missing data in a non-research setting. At the same time, it highlights the potential of a more rigorous data management focus. One-third of the analysis sample was lost to missing data. At an individual variable level, proportions of missing data were even higher. Based on the limited number of large-scale effectiveness studies, these data issues are not unique. By

improving data management, data quality and data retention, large, multi-year studies can provide sufficient statistical power to better understand relationships that have a direct impact on program costs. Leveraging up-to-date electronic data collection methods can save time and cost in data collection while minimizing manual re-entry of data. Improvements in data collection will enable matching across multiple systems and databases, enabling potential access to socio-economic data and geo-demographic data including availability and distance to EIBI services. Linkages to other administrative data may allow long-term follow up studies several years after EIBI treatment has ended.

The landscape of autism and autism treatment continues to rapidly evolve. Prevalence has continued its dramatic increase. But advances in the acceptance of EIBI as an effective treatment, combined with the expansion of insurance coverage for diagnosis and treatment now enables the autism community to study thousands of children who are receiving EIBI services in a natural setting, compared to only ten years ago, when efficacy research was being conducted on tens of children. Given the projected growth in autism-related costs over the next ten years, it is imperative that research and administration join together to dramatically improve our understanding of EIBI treatment effectiveness.

References

American Psychiatric Association, American Psychiatric Association D. S. M. Task Force. 2013. "Diagnostic and statistical manual of mental disorders : DSM-5." http://dsm.psychiatryonline.org/book.aspx?bookid=556.

American Psychiatric Association. 2000. *Diagnostic and statistical manual of mental disorders : DSM-IV-TR*. 4th ed. Washington, DC: American Psychiatric Association.

Autism Speaks. "Autism Insurance Reform." Accessed 02/17/2016. https://www.autismspeaks.org/state-initiatives.

Baio, Jon. 2014. "Prevalence of autism spectrum disorder among children aged 8 years-autism and developmental disabilities monitoring network, 11 sites, United States, 2010." *Morbidity and mortality weekly report. Surveillance summaries (Washington, DC: 2002)* 63:1.

Ben-Itzchak, Esther, Eliezer Lahat, Ruth Burgin, and A. Ditza Zachor. 2008. "Cognitive, behavior and intervention outcome in young children with autism." *Research in Developmental Disabilities* 29 (5):447-458.

Ben-Itzchak, Esther, and Ditza A. Zachor. 2007. "The effects of intellectual functioning and autism severity on outcome of early behavioral intervention for children with autism." *Research in Developmental Disabilities* 28 (3):287-303.

Ben Itzchak, Esther, and Ditza A Zachor. 2011. "Who benefits from early intervention in autism spectrum disorders?" *Research in Autism Spectrum Disorders* 5 (1):345-350.

Bibby, Peter, Svein Eikeseth, Neil T. Martin, Oliver C. Mudford, and David Reeves. 2002. "Progress and outcomes for children with autism receiving parent-managed intensive interventions." *Research in Developmental Disabilities* 23 (1):81-104.

Bouder, James N, Stuart Spielman, and David S Mandell. 2009. "Brief report: Quantifying the impact of autism coverage on private insurance premiums." *Journal of Autism and Developmental Disorders* 39 (6):953-957.

Castelloe, John. 2014. "Power and Sample Size for MANOVA and Repeated Measures with the GLMPOWER Procedure." Proceedings of the SAS Global Forum 2014 Conference, Cary, NC: SAS Institute Inc. URL http://support. sas. com/resources/papers/proceedings14/SAS030-2014. pdf.

Charman, Tony, CRG Jones, Andrew Pickles, Emily Simonoff, Gillian Baird, and Francesca Happé. 2011. "Defining the cognitive phenotype of autism." *Brain research* 1380:10-21.

Chasson, Gregory, Gerald Harris, and Wendy Neely. 2007. "Cost comparison of Early Intensive Behavioral Intervention and special education for children with autism." *Journal of Child and Family Studies* 16 (3):401-413. doi: 10.1007/s10826-006-9094-1.

Cohen, Howard, Mila Amerine-Dickens, and Tristram Smith. 2006. "Early intensive behavioral treatment: Replication of the UCLA model in a community setting." *Journal of Developmental & Behavioral Pediatrics* 27 (2):S145-S155.

Davis, Tonya N, Mark O'Reilly, Soyeon Kang, Russell Lang, Mandy Rispoli, Jeff Sigafoos, Giulio Lancioni, Daelynn Copeland, Shanna Attai, and Austin Mulloy. 2013. "Chelation treatment for autism spectrum disorders: A systematic review." *Research in Autism Spectrum Disorders* 7 (1):49-55.

Dietz, Claudine, Sophie Swinkels, Jan Buitelaar, Emma van Daalen, and Herman van Engeland. 2007. "Stability and change of IQ scores in preschool children diagnosed with autistic spectrum disorder." *European Child & Adolescent Psychiatry* 16 (6):405-410. doi: 10.1007/s00787-007-0614-3.

Dunn, Lloyd M, and DM Dunn. 2012. "Peabody Picture Vocabulary Test, (PPVT™-4)." *Johannesburg: Pearson Education Inc*.

Dunn, Lloyd M, and Leota M Dunn. 1981. *PPVT: Peabody Picture Vocabulary Test, Revised*: American Guidance Service.

Eikeseth, Svein, Lars Klintwall, Erik Jahr, and Peter Karlsson. 2012. "Outcome for children with autism receiving early and intensive behavioral intervention in mainstream preschool and kindergarten settings." *Research in Autism Spectrum Disorders* 6 (2):829-835. doi: http://dx.doi.org/10.1016/j.rasd.2011.09.002.

Eikeseth, Svein, Tristram Smith, Erik Jahr, and Sigmund Eldevik. 2007. "Outcome for Children with Autism who Began Intensive Behavioral Treatment Between Ages 4 and 7." *Behavior Modification* 31 (3):264-278. doi: 10.1177/0145445506291396.

Eldevik, Sigmund, Richard P Hastings, Erik Jahr, and J Carl Hughes. 2012. "Outcomes of behavioral intervention for children with autism in mainstream pre-school settings." *Journal of Autism and Developmental Disorders* 42 (2):210-220.

Eldevik, Sigmund, Richard P. Hastings, J. Carl Hughes, Erik Jahr, Svein Eikeseth, and Scott Cross. 2009. "Meta-analysis of Early Intensive Behavioral Intervention for children with autism." *Journal Of Clinical Child And Adolescent Psychology: The Official Journal For The Society Of Clinical Child And Adolescent Psychology, American Psychological Association, Division 53* 38 (3):439-450.

Eldevik, Sigmund, Richard P. Hastings, J. Carl Hughes, Erik Jahr, Svein Eikeseth, and Scott Cross. 2010. "Using participant data to extend the evidence base for Intensive Behavioral Intervention for children with autism." *American Journal on Intellectual and Developmental Disabilities* 115 (5):381-405. doi: 10.1352/1944-7558-115.5.381.

Fernell, Elisabeth, Åsa Hedvall, Joakim Westerlund, Lotta Höglund Carlsson, Mats Eriksson, Martina Barnevik Olsson, Anette Holm, Fritjof Norrelgen, Liselotte Kjellmer, and Christopher Gillberg. 2011. "Early intervention in 208 Swedish preschoolers with autism spectrum disorder. A prospective naturalistic study." *Research in Developmental Disabilities* 32 (6):2092-2101.

Flanagan, Helen E, Adrienne Perry, and Nancy L Freeman. 2012. "Effectiveness of large-scale community-based intensive behavioral Intervention: a waitlist comparison study exploring outcomes and predictors." *Research in Autism Spectrum Disorders* 6 (2):673-682.

Flay, Brian R. 1986. "Efficacy and effectiveness trials (and other phases of research) in the development of health promotion programs." *Preventive medicine* 15 (5):451-474.

Fombonne, Eric. 2009. "Epidemiology of Pervasive Developmental Disorders." *Pediatr Res* 65 (6):591-598.

Foxx, Richard M. 2008. "Applied Behavior Analysis Treatment of Autism: The State of the Art." *Child and Adolescent Psychiatric Clinics of North America* 17 (4):821-834.

Freeman, N, and A Perry. 2010. "Outcomes of intensive behavioural intervention in the Toronto Preschool Autism Service." *Journal on Developmental Disabilities* 16 (2):17-32.

Fritz, Catherine O, Peter E Morris, and Jennifer J Richler. 2012. "Effect size estimates: current use, calculations, and interpretation." *Journal of Experimental Psychology: General* 141 (1):2.

Goin-Kochel, Robin P., Virginia H. Mackintosh, and Barbara J. Myers. 2009. "Parental reports on the efficacy of treatments and therapies for their children with autism spectrum disorders." *Research in Autism Spectrum Disorders* 3 (2):528-537.

Granpeesheh, Doreen, Dennis R. Dixon, Jonathan Tarbox, Andrew M. Kaplan, and Arthur E. Wilke. 2009. "The effects of age and treatment intensity on behavioral intervention outcomes for children with autism spectrum disorders." *Research in Autism Spectrum Disorders* 3 (4):1014-1022. doi: http://dx.doi.org/10.1016/j.rasd.2009.06.007.

Granpeesheh, Doreen, Jonathan Tarbox, and Dennis R Dixon. 2009. "Applied behavior analytic interventions for children with autism: a description and review of treatment research." *Annals of clinical psychiatry* 21 (3):162-173.

Green, Gina, Lynn C Brennan, and Deborah Fein. 2002. "Intensive behavioral treatment for a toddler at high risk for autism." *Behavior Modification* 26 (1):69-102.

Grimes, David A. 2010. "Epidemiologic research using administrative databases: garbage in, garbage out." *Obstetrics & Gynecology* 116 (5):1018-1019.

Harris, S. L., and J. S. Handleman. 2000. "Age and IQ at intake as predictors of placement for young children with autism: a four- to six-year follow-up." *Journal of Autism & Developmental Disorders* 30 (2):137-142.

Harvey, A Celeste, Mark T Harvey, Mary Beth Kenkel, and Dennis C Russo. 2010. "Funding of applied behavior analysis services: Current status and growing opportunities." *Psychological Services* 7 (3):202.

Hayward, Diane, Svein Eikeseth, Catherine Gale, and Sally Morgan. 2009. "Assessing progress during treatment for young children with autism receiving intensive behavioural interventions." *Autism* 13 (6):613-633. doi: 10.1177/1362361309340029.

Healy, Olive, and Sinéad Lydon. 2013. "Early Intensive Behavioural Intervention in Autism Spectrum Disorders."

Hoffman, Lesa, Jonathan Templin, and Mabel L Rice. 2012. "Linking outcomes from Peabody Picture Vocabulary Test forms using item response models." *Journal of Speech, Language, and Hearing Research* 55 (3):754-763.

Howard, Jane S, Coleen R Sparkman, Howard G Cohen, Gina Green, and Harold Stanislaw. 2005. "A comparison of intensive behavior analytic and eclectic treatments for young children with autism." *Research in Developmental Disabilities* 26 (4):359-383.

Howlin, Patricia, Iliana Magiati, and Tony Charman. 2009. "Systematic review of early intensive behavioral interventions for children with autism." *Journal Information* 114 (1).

Iezzoni, Lisa I. 2002. "Using Administrative Data to Study Persons with Disabilities." *The Milbank Quarterly* 80 (2):347-379.

Jacobson, John W., James A. Mulick, and Gina Green. 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. doi: 10.1002/(sici)1099-078x(199811)13:4<201::aid-bin17>3.0.co;2-r.

Jacobson, Neil S., and Paula Truax. 1991. "Clinical significance: A statistical approach to defining meaningful change in psychotherapy research." *Journal of Consulting and Clinical Psychology* 59 (1):12-19. doi: 10.1037/0022-006x.59.1.12.

Jacqmin-Gadda, Hélène, Solenne Sibillot, Cécile Proust, Jean-Michel Molina, and Rodolphe Thiébaut. 2007. "Robustness of the linear mixed model to misspecified error distribution." *Computational Statistics & Data Analysis* 51 (10):5142-5154.

Jensen, Vanessa, and Sarah Spannagel. 2011. "The Spectrum of Autism Spectrum Disorder: A Spectrum of Needs, Services, and Challenges." *Journal of Contemporary Psychotherapy* 41 (1):1-9. doi: 10.1007/s10879-010-9161-1.

Johnson, Chris Plauche, Scott M. Myers, and the Council on Children With Disabilities. 2007. "Identification and evaluation of children with autism spectrum disorders." *Pediatrics* 120 (5):1183-1215. doi: 10.1542/peds.2007-2361.

Kogan, Michael D., Bonnie B. Strickland, Stephen J. Blumberg, Gopal K. Singh, James M. Perrin, and Peter C. van Dyck. 2008. "A national profile of the health care experiences and family impact of autism spectrum disorder among children in the United States, 2005–2006." *Pediatrics* 122 (6):e1149-e1158. doi: 10.1542/peds.2008-1057.

Lavelle, Tara A, Milton C Weinstein, Joseph P Newhouse, Kerim Munir, Karen A Kuhlthau, and Lisa A Prosser. 2014. "Economic burden of childhood autism spectrum disorders." *Pediatrics* 133 (3):e520-e529.

Leigh, J Paul, and Juan Du. 2015. "Brief Report: Forecasting the Economic Burden of Autism in 2015 and 2025 in the United States." *Journal of autism and developmental disorders* 45 (12):4135-4139.

Levy, Susan E., and Susan L. Hyman. 2005. "Novel treatments for autistic spectrum disorders." *Mental Retardation & Developmental Disabilities Research Reviews* 11 (2):131-142.

Llaneza, Danielle C., Susan V. DeLuke, Myra Batista, Jacqueline N. Crawley, Kristin V. Christodulu, and Cheryl A. Frye. 2010. "Communication, interventions, and scientific advances in autism: A commentary." *Physiology & Behavior* 100 (3):268-276.

Lord, C Rutter, and M Rutter. 1994. "M. & Le Couteur, A.(1994) Autism Diagnostic Interview-Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders." *Journal of Autism and Developmental Disorders* 24 (5):659-685.

Lord, Catherine, Michael Rutter, Pamela C DiLavore, and Susan Risi. 1999. "Autism diagnostic observation schedule-WPS (ADOS-WPS)." *Los Angeles, CA: Western Psychological Services*.

Lovaas, O. Ivar. 1987. "Behavioral treatment and normal educational and intellectual functioning in young autistic children." *Journal of Consulting and Clinical Psychology* 55 (1):3-9. doi: 10.1037/0022-006x.55.1.3.

Love, Jessa R., James E. Carr, Season M. Almason, and Anna Ingeborg Petursdottir. 2009. "Early and intensive behavioral intervention for autism: A survey of clinical practices." *Research in Autism Spectrum Disorders* 3 (2):421-428.

Magiati, Iliana, Tony Charman, and Patricia Howlin. 2007. "A two-year prospective follow-up study of community-based early intensive behavioural intervention and specialist nursery provision for children with autism spectrum disorders." *Journal Of Child Psychology And Psychiatry, And Allied Disciplines* 48 (8):803-812.

Makrygianni, Maria K., and Phil Reed. 2010. "A meta-analytic review of the effectiveness of behavioural early intervention programs for children with Autistic Spectrum Disorders." *Research in Autism Spectrum Disorders* 4 (4):577-593. doi: http://dx.doi.org/10.1016/j.rasd.2010.01.014.

Matson, Johnny L, and Rachel L Goldin. 2014. "Early Intensive Behavioral Interventions: Selecting behaviors for treatment and assessing treatment effectiveness." *Research in Autism Spectrum Disorders* 8 (2):138-142.

Matson, Johnny L, and Jina Jang. 2013. "Autism spectrum disorders: Methodological considerations for early intensive behavioral interventions." *Research in Autism Spectrum Disorders* 7 (7):809-814.

Matson, Johnny L, and Marie S Nebel-Schwalm. 2007. "Comorbid psychopathology with autism spectrum disorder in children: An overview." *Research in Developmental Disabilities* 28 (4):341-352.

Matson, Johnny L., and Mary Shoemaker. 2009. "Intellectual disability and its relationship to autism spectrum disorders." *Research in Developmental Disabilities* 30 (6):1107-1114. doi: http://dx.doi.org/10.1016/j.ridd.2009.06.003.

Matson, Johnny L., and Kimberly R. M. Smith. 2008. "Current status of intensive behavioral interventions for young children with autism and PDD-NOS." *Research in Autism Spectrum Disorders* 2 (1):60-74. doi: http://dx.doi.org/10.1016/j.rasd.2007.03.003.

Morris, E. K., N. G. Smith, and D. E. Altus. 2005. "B. F. Skinner's contributions to applied behavior analysis." *Behavior Analyst* 28 (2):99-131.

Myers, Scott M., Chris Plauché Johnson, and the Council on Children With Disabilities. 2007. "Management of children with autism spectrum disorders." *Pediatrics* 120 (5):1162-1182. doi: 10.1542/peds.2007-2362.

National Autism Network. "Affordable Care Act (ACA) and Autism." Accessed 02/17/2016. http://nationalautismnetwork.com/legislation/federal-initiatives/affordable-care-act-and-autism.

National Conference of State Legislatures. 2015. "Insurance Coverage for Autism." Accessed 02/17/2016. http://www.ncsl.org/research/health/autism-and-insurance-coverage-state-laws.aspx.

Newschaffer, Craig J., Lisa A. Croen, Julie Daniels, Ellen Giarelli, Judith K. Grether, Susan E. Levy, David S. Mandell, Lisa A. Miller, Jennifer Pinto-Martin, Judy Reaven, Ann M. Reynolds, Catherine E. Rice, Diana Schendel, and Gayle C. Windham. 2007. "The Epidemiology of autism spectrum disorders*." *Annual Review of Public Health* 28 (1):235-258. doi: 10.1146/annurev.publhealth.28.021406.144007.

Newschaffer, Craig J., Matthew D. Falb, and James G. Gurney. 2005. "National Autism Prevalence Trends From United States Special Education Data." *Pediatrics* 115 (3):e277-e282. doi: 10.1542/peds.2004-1958.

Patten, Elena, and Linda R. Watson. 2011. "Interventions targeting attention in young children with autism." *Am J Speech Lang Pathol* 20 (1):60-69. doi: 10.1044/1058-0360(2010/09-0081).

Patterson, Stephanie Y., Veronica Smith, and Michaela Jelen. 2010. "Behavioural intervention practices for stereotypic and repetitive behaviour in individuals with autism spectrum disorder: a systematic review." *Developmental Medicine & Child Neurology* 52 (4):318-327. doi: 10.1111/j.1469-8749.2009.03597.x.

Payakachat, Nalin, J. Mick Tilford, Erica Kovacs, and Karen Kuhlthau. 2012. "Autism spectrum disorders: a review of measures for clinical, health services and cost–effectiveness applications." *Expert review of pharmacoeconomics & outcomes research* 12 (4):485-503. doi: 10.1586/erp.12.29.

Perry, A., H. E. Flanagan, J. D. Geier, and N. L. Freeman. 2009. "Brief report: the Vineland Adaptive Behavior Scales in young children with autism spectrum disorders at different cognitive levels." *Journal of Autism & Developmental Disorders* 39 (7):1066-1078. doi: 10.1007/s10803-009-0704-9.

Perry, Adrienne, Anne Cummings, Jennifer Dunn Geier, Nancy L. Freeman, Susan Hughes, Louise LaRose, Tom Managhan, Jo-Ann Reitzel, and Janis Williams. 2008. "Effectiveness of Intensive Behavioral Intervention in a large, community-based program." *Research in Autism Spectrum Disorders* 2 (4):621-642. doi: http://dx.doi.org/10.1016/j.rasd.2008.01.002.

Perry, Adrienne, Anne Cummings, Jennifer Dunn Geier, Nancy L. Freeman, Susan Hughes, Tom Managhan, Jo-Ann Reitzel, and Janis Williams. 2011. "Predictors of outcome for children receiving intensive behavioral intervention in a large, community-based program." *Research in Autism Spectrum Disorders* 5 (1):592-603.

Peters-Scheffer, Nienke, Robert Didden, Hubert Korzilius, and Johnny Matson. 2012. "Cost comparison of early intensive behavioral intervention and treatment as usual for children with autism spectrum disorder in the Netherlands." *Research in Developmental Disabilities* 33 (6):1763-1772. doi: http://dx.doi.org/10.1016/j.ridd.2012.04.006.

Peters-Scheffer, Nienke, Robert Didden, Hubert Korzilius, and Peter Sturmey. 2011. "A meta-analytic study on the effectiveness of comprehensive ABA-based early intervention programs for children with Autism Spectrum Disorders." *Research in Autism Spectrum Disorders* 5 (1):60-69. doi: http://dx.doi.org/10.1016/j.rasd.2010.03.011.

Prater, Christopher D., and Robert G. Zylstra. 2002. "Autism: a medical primer." *American Family Physician* 66 (9):1667-1674.

Reichow, Brian. 2012. "Overview of Meta-Analyses on Early Intensive Behavioral Intervention for Young Children with Autism Spectrum Disorders." *Journal of*

Autism and Developmental Disorders 42 (4):512-520. doi: 10.1007/s10803-011-1218-9.

Reichow, Brian, and Mark Wolery. 2009. "Comprehensive synthesis of early intensive behavioral interventions for young children with autism based on the UCLA young autism project model." *Journal of Autism and Developmental Disorders* 39 (1):23-41.

Remington, Bob, Richard P Hastings, Hanna Kovshoff, Francesca degli Espinosa, Erik Jahr, Tony Brown, Paula Alsford, Monika Lemaic, Nicholas Ward, and Jr MacLean, William E. 2007. "Early intensive behavioral intervention: Outcomes for children with autism and their parents after two years." *American Journal on Mental Retardation* 112 (6):418-438.

Roberts, Joanne, Johanna Price, Elizabeth Barnes, Lauren Nelson, Margaret Burchinal, Elizabeth A Hennon, Lauren Moskowitz, Anne Edwards, Cheryl Malkin, and Kathleen Anderson. 2007. "Receptive vocabulary, expressive vocabulary, and speech production of boys with fragile X syndrome in comparison to boys with Down syndrome." *Journal Information* 112 (3).

Rogers, Sally J, and Laurie A Vismara. 2008. "Evidence-based comprehensive treatments for early autism." *Journal of Clinical Child & Adolescent Psychology* 37 (1):8-38.

Rutter, Michael, Anthony Bailey, and Cathrine Lord. 2003. *The Social Communication Questionnaire: Manual*: Western Psychological Services.

Sallows, Glen O., Tamlynn D. Graupner, and William E. MacLean Jr. 2005. "Intensive Behavioral Treatment for Children With Autism: Four-Year Outcome and Predictors." *American Journal on Mental Retardation* 110 (6):417-438. doi: 10.1352/0895-8017(2005)110[417:ibtfcw]2.0.co;2.

Schneeweiss, Sebastian, and Jerry Avorn. 2005. "A review of uses of health care utilization databases for epidemiologic research on therapeutics." *Journal of clinical epidemiology* 58 (4):323-337.

Schopler, E, RJ Reichler, and BR Renner. 1988. "Child Autism Rating Scale." *Western Psychological Services Corporation*.

Shattuck, Paul T. 2006. "The Contribution of Diagnostic Substitution to the Growing Administrative Prevalence of Autism in US Special Education." *Pediatrics* 117 (4):1028-1037. doi: 10.1542/peds.2005-1516.

Sheinkopf, Stephen J., and Bryna Siegel. 1998. "Home-based behavioral treatment of young children with autism." *Journal of Autism and Developmental Disorders* 28 (1):15-23. doi: 10.1023/a:1026054701472.

Smith, Tristram, Rafael Klorman, and DanielW Mruzek. 2015. "Predicting Outcome of Community-Based Early Intensive Behavioral Intervention for Children with Autism." *Journal of Abnormal Child Psychology*:1-12. doi: 10.1007/s10802-015-0002-2.

South Carolina Department of Disabilities and Special Needs. 2013. "Pervasive Developmental Disorder Waiver/State Funded Program Manual For Case Managers and Early Intensive Behavioral Intervention Providers." Accessed 07/11/2014. http://www.ddsn.sc.gov/providers/medicaidwaiverservices/pdd/Pages/PDDManual.aspx.

Sparrow, Sara S, David A Balla, and Domenic V Cicchetti. 1984. *Vineland adaptive behavior scales: Interview edition, survey form manual*: American Guidance Service Circle Pines, MN.

Sparrow, Sara S, David A Balla, and Domenic V Cicchetti. 2005. *Vineland II: Vineland Adaptive Behavior Scales*: American Guidance Service.

Spreckley, Michèle, and Roslyn Boyd. 2009. "Efficacy of applied behavioral intervention in preschool children with autism for improving cognitive, language, and adaptive behavior: a systematic review and meta-analysis." *The Journal of pediatrics* 154 (3):338-344. doi: citeulike-article-id:9524707.

Tarbox, Jonathan, Arthur E Wilke, Rachel S Findel-Pyles, Ryan M Bergstrom, and Doreen Granpeesheh. 2010. "A comparison of electronic to traditional pen-and-paper data collection in discrete trial training for children with autism." *Research in Autism Spectrum Disorders* 4 (1):65-75.

Virnig, Beth A, and Marshall McBean. 2001. "Administrative data for public health surveillance and planning." *Annual Review of Public Health* 22 (1):213-230.

Virués-Ortega, Javier. 2010. "Applied behavior analytic intervention for autism in early childhood: Meta-analysis, meta-regression and dose–response meta-analysis of multiple outcomes." *Clinical Psychology Review* 30 (4):387-399.

Warren, Zachary, Melissa L. McPheeters, Nila Sathe, Jennifer H. Foss-Feig, Allison Glasser, and Jeremy Veenstra-VanderWeele. 2011. "A Systematic Review of Early Intensive Intervention for Autism Spectrum Disorders." *Pediatrics* 127 (5):e1303-e1311. doi: 10.1542/peds.2011-0426.

Wells, Kenneth B. 2014. "Treatment research at the crossroads: the scientific interface of clinical trials and effectiveness research." *American Journal of Psychiatry*.

Williams, Kathleen T. 1997. *Expressive vocabulary test*: American Guidance Service.

Williams White, Susan, Kathleen Keonig, and Lawrence Scahill. 2007. "Social skills development in children with autism spectrum disorders: A review of the intervention research." *Journal of Autism and Developmental Disorders* 37 (10):1858-1868. doi: 10.1007/s10803-006-0320-x.

Wing, Lorna, and David Potter. 2002. "The epidemiology of autistic spectrum disorders: is the prevalence rising?" *Mental Retardation and Developmental Disabilities Research Reviews* 8 (3):151-161.

Zablotsky, Benjamin, Luther G Kalb, Brian Freedman, Roma Vasa, and Elizabeth A Stuart. 2014. "Health care experiences and perceived financial impact among families of children with an autism spectrum disorder." *Psychiatric Services*.

Zhang, Daowen, and Marie Davidian. 2001. "Linear mixed models with flexible distributions of random effects for longitudinal data." *Biometrics* 57 (3):795-802.