

UCLA

UCLA Electronic Theses and Dissertations

Title

Super Responders: Predicting Expressive Language Gains Among Minimally Verbal Children with Autism Spectrum Disorder

Permalink

<https://escholarship.org/uc/item/4hf453hc>

Author

Panganiban, Jonathan Luke

Publication Date

2018

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Super Responders: Predicting Expressive Language Gains Among Minimally Verbal Children
with Autism Spectrum Disorder

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of
Philosophy in Education

by

Jonathan Luke Panganiban

2018

© Copyright by

Jonathan Luke Panganiban

2018

ABSTRACT OF THE DISSERTATION

Super Responders: Predicting Expressive Language Gains Among Minimally Verbal Children
with Autism Spectrum Disorder

by

Jonathan Panganiban

Doctor of Philosophy in Education

University of California, Los Angeles, 2018

Professor Connie L. Kasari, Chair

Much research in autism spectrum disorders (ASD) has focused on the development of efficacious interventions to address the core deficits of ASD. However, the heterogeneous nature of ASD complicates the development of such interventions. With great heterogeneity in the expression of ASD's core deficits, it is unlikely that there is a one size fits all intervention. It is important for researchers to understand for whom an intervention works. Advancements in data analytics, in particular machine learning, provide new methods to identify subgroups among a given population, and can potentially help to identify for whom intervention works best. Of

particular interest are minimally verbal individuals. A targeted social communication intervention known as JASPER (Joint Attention, Symbolic Play, Engagement, and Regulation) has shown promise for improving language outcomes among minimally verbal children with ASD and may provide the context to examine the question of for whom an intervention benefits. This study aims to develop a model predicting expressive language gains among minimally verbal, preschool aged children with ASD that received a targeted social communication intervention.

Classification and regression tree (CART) analysis was used to explore the relationship between child characteristics and gains in expressive language. Secondary data analysis was conducted on a sample of 99 minimally verbal, preschool age children with ASD, collected from participants across five previous intervention studies. Expressive language gains (outcome) were calculated using expressive language age equivalents from the Mullen Scales for Early Learning. Predictors for the analyses were taken from child demographics and behavioral assessments completed prior to intervention. The initial list of predictors included race, gender, ASD severity, visual reception age equivalent, fine motor age equivalent, joint attention gestures, requesting gestures, and play skills. Using expressive language age equivalent change scores, 47% (n = 47) of the sample were identified as “super responders,” children that exceeded expressive language gains typically expected through maturation. To predict responder status, all initial predictors were used to generate conditional inference forest, from which the most important variables would be chosen for the final model. Conditional inference results identified three variables to be fitted into the final model; play diversity, requesting gestures, and fine motor age equivalent. A final conditional inference tree was created, with play diversity being the only significant predictor of responder status. Participants with an entry play diversity score above 23 predicted

super response while scores of 23 or below predicted slow response. The overall model accuracy was 67%, with a specificity of 55% and sensitivity of 78%. As a comparison, stepwise logistic regression was run, and play diversity was again the only significant predictor of responder status ($\chi^2(1) = 10.686, p = .001$). Receiver operating characteristic curves were generated to compare model performance, and comparison of area under the curves for the two models showed no statistical difference ($p = .82$).

Overall accuracy of the conditional inference tree was moderate, and performed similarly to the more traditional logistic regression analysis. However, the conditional inference tree provides a cutoff point that may provide clinical utility over the regression results. Both models identify play diversity as an important predictor of expressive language gains from JASPER, which is a play based social communication intervention. Additionally, our model appears to be more sensitive to identifying slow responders. The role of play diversity and expressive language gains in JASPER is discussed.

The dissertation of Jonathan Luke Panganiban is approved.

Sandra H. Graham

Jennie Katherine Grammer

Sheryl Harumi Kataoka Endo

Connie L. Kasari, Committee Chair

University of California, Los Angeles

2018

Contents

| | |
|---|----|
| Introduction | 1 |
| Comprehensive Interventions..... | 2 |
| Targeted Interventions | 3 |
| Determining the Evidence Base | 3 |
| For Whom Interventions Work..... | 4 |
| Moderators of Response..... | 5 |
| Profiling Responders to Interventions..... | 7 |
| Methodological Advances for Determining Responders | 8 |
| Classification and Regression Trees for Responder Status..... | 9 |
| Aims..... | 10 |
| Methods..... | 11 |
| Participants | 11 |
| Design..... | 12 |
| The Intervention – JASPER | 12 |
| Measures | 13 |
| Analysis | 15 |
| Results..... | 17 |
| Participant Demographics | 17 |
| Identification of Responders..... | 18 |
| Conditional Inference Forest for Variable Selection..... | 18 |
| Building the Conditional Inference Tree (CIT) | 19 |
| Building the Logistic Regression Model..... | 19 |
| Comparison of CIT and Logistic Regression Models..... | 19 |
| Building a Second Conditional Inference Tree | 20 |
| Discussion | 20 |
| Personalized Intervention..... | 20 |
| Identifying Super Responders | 22 |
| Predicting Response Status in JASPER..... | 23 |
| Play Diversity | 26 |
| Machine Learning in ASD Intervention Research | 28 |
| Limitations..... | 29 |
| Conclusions | 29 |
| Appendix..... | 31 |
| References | 41 |

| | |
|---|----|
| Table 1 Descriptives | 31 |
| Table 2 Sample Characteristics | 32 |
| Table 1 Stepwise Logistic Regression Results | 38 |
| Figure 1 Conditional Inference Forest Variable Importance Plot #1 | 33 |
| Figure 2 Conditional Inference Forest Variable Importance Plot #2 | 34 |
| Figure 3 Conditional Inference Forest Variable Importance Plot #3 | 35 |
| Figure 4 Correlation matrix for predictors in final model..... | 36 |
| Figure 5 Conditional Inference Tree of Play Diversity Predicting Super vs. Slow Responders, confidence level of .95..... | 37 |
| Figure 6 ROC Curves Comparing Predictive Performance Between the Conditional Inference Tree Model and the Logistic Regression Model. | 39 |
| Figure 7 Conditional Inference Tree of Play Diversity and Fine Motor Predicting Super vs. Slow Responders, confidence level of .90. | 40 |

Curriculum Vitae

EDUCATION

- 2017 M.A., Education, University of California, Los Angeles
Advisor: Connie L. Kasari, Ph.D.
- 2003 B.A., Psychology, University of California, Los Angeles

RESEARCH EXPERIENCE

- 2017–Present University of California, Los Angeles- The Goldman Foundation
- *Position:* Study Coordinator, Trainer, and Graduate Research Assistant
 - *Grant:* JASPER in the classroom: Implementation of a train the trainer model
 - *PI:* Connie L. Kasari, Ph.D.
 - *Roles:*
 - Recruit LAUSD Preschool for All Learners (PAL) teachers
 - Manage data collection, entry, and cleaning
 - Supervise administration of the intervention
- 2013–Present University of California, Los Angeles- NIMH Autism Center of Excellence
- *Position:* Lead Interventionist and Graduate Research Assistant
 - *Grant:* Adaptive Interventions for Children with Autism in the Community (AIM-ASD)
 - *PIs:* Connie L. Kasari, Ph.D., Ann P. Kaiser, Ph.D., Tristram H. Smith, Ph.D., & Catherine Lord, Ph.D.
 - *Roles:*
 - Administer social communication intervention
 - Rate interventionist fidelity
 - Provide clinical feedback for intervention team
- 2012–Present University of California, Los Angeles- NIMH: Collaborative R01
- *Position:* Interventionist and Graduate Research Assistant
 - *Grant:* Interventions for Communication in Autism Network (ICAN)
 - *PIs:* Connie L. Kasari, Ph.D., Rebecca Landa, Ph.D., & Tristram H. Smith, Ph.D.
 - *Roles:* Administer social communication and discrete trial training interventions

TEACHING EXPERIENCE

- 2018 Fall **California State University, Los Angeles**
Adjunct Professor, graduate course
- *EDSP 5800: Research Methods in Special Education*
- 2018 Spring **California State University, Los Angeles**
Adjunct Professor, graduate course
- *EDSP 5886: Educating Students with ASD*
- 2015 Summer **University of California, Los Angeles**
Teaching Assistant, undergraduate course
- *Education 132: Autism: Mind, Brain, and Education*
 - *Professor:* Connie L. Kasari, Ph.D.
- 2015 Winter **University of California, Los Angeles**
Teaching Assistant, undergraduate course
- *Education 132: Autism: Mind, Brain, and Education*
 - *Professor:* Connie L. Kasari, Ph.D.

AWARDS AND HONORS

| | |
|------|---|
| 2018 | University Fellowship (\$1,244.50), University of California, Los Angeles |
| 2017 | University Fellowship (\$1,660), University of California, Los Angeles |
| 2016 | University Fellowship (\$1000), University of California, Los Angeles |
| 2013 | University Fellowship (\$5,000), University of California, Los Angeles |
| 2013 | Graduate Summer Research Mentorship Program Fellowship (\$6,000), University of California, Los Angeles |
| 2012 | Graduate Division Fellowship (\$20,000), University of California, Los Angeles |

CONFERENCE ACTIVITY

| | |
|------|--|
| 2017 | International Meeting for Autism Research, Poster presentation, <i>Teacher Perceptions and the Implementation of Jasper for Students with ASD</i> |
| 2017 | International Meeting for Autism Research, Poster presentation, <i>A Community-Partnered Intervention in South Los Angeles for Young Children at-Risk for ASD</i> |
| 2016 | International Meeting for Autism Research, Poster presentation, <i>Understanding the Communication Skills in Verbal and Minimally Verbal Children with Autism Spectrum Disorder</i> |
| 2015 | International Meeting for Autism Research, Presentation, <i>Autism in the African American Community of South Los Angeles: A Community Partnered Participatory Research Approach</i> |
| 2015 | International Meeting for Autism Research, Poster presentation, <i>Characterizing Play in Children with ASD: Differences in Joint Attention and Requesting Across Play Levels</i> |
| 2015 | International Meeting for Autism Research, The Effects of Teacher Perceptions on Fostering Engagement during Dyadic Play Interactions with Students with ASD |
| 2015 | UC Center for Special Education and Development and Developmental Disabilities Research Conference, Presentation, <i>Addressing Autism in the African-American Community: Establishing a Partnership to Promote Progress</i> |
| 2014 | International Meeting for Autism Research, Poster presentation, <i>Measuring Joint Attention in Children with Autism Spectrum Disorder through Structured and Unstructured Play</i> |
| 2013 | UC Center for Special Education and Developmental Disabilities Research Conference, Poster presentation, <i>Measuring Joint Attention in Children with Autism Spectrum Disorder Across Multiple Assessment Structures</i> |

PRESENTATIONS

| | |
|------|--|
| 2018 | Third International Conference “Autism: World of Opportunities,” Almaty, Kazakhstan Presentation, <i>The Foundation of JASPER: Joint Engagement and Play</i> |
| 2017 | Second International Conference “Autism: World of Opportunities,” Astana, Kazakhstan Presentation, <i>Early Intervention: The JASPER Model</i> |
| 2015 | Brown Bag Symposium “Autism in Our Community – It’s OK to Talk About It”, Presentation, <i>Autism 101 – What We Know, What We Are Still Learning</i> |

PROFESSIONAL MEMBERSHIPS

| | |
|-----------|---|
| 2013-2015 | Council Chair, UC Center for Special Education and Development and Developmental Disabilities Research Doctoral Student Advisory Committee |
| 2012 | Vice Council Chair, UC Center for Special Education and Development and Developmental Disabilities Research Doctoral Student Advisory Committee |

Introduction

Autism Spectrum Disorders (ASD), as we know them today, were first described by Leo Kanner in 1943. In this seminal paper, Dr. Kanner described 11 children (eight boys and three girls). His early descriptions of autism not only laid the groundwork for the entire field of study, but have remained consistent with ASD as it is known today. While ASD was once considered a rare condition, we have seen steady increases in population rates. We are now realizing that ASD is more common in the population than once thought. Recent estimates put the prevalence of ASD among children in the U.S. as high as 1 in 45 (Zablotsky, Black, Maenner, Schieve, & Blumberg, 2015). While the direct cause of ASD, and its resulting increase in prevalence, still confound researchers, much has been learned about ASD and those it affects. We know that there is a strong genetic component to ASD, shown through high concordance rates among identical twins (Ronald et al., 2006; Taniai, Nishiyama, Miyachi, Imaeda, & Sumi, 2008; Rosenberg et al., 2009; Hallmayer et al., 2011) and increased risk among biological siblings (Sumi, Taniai, Miyachi, & Tanemura, 2006; Ozonoff et al., 2011). We also know that ASD more commonly affects males than it does females (Baio, 2012). While no direct cause for ASD has been found, recent work does point to strong associations with several known genetic disorders (Abrahams & Geschwind, 2008). This has led to some researchers referring to ASD as the “autisms” (Coleman & Gillberg, 2012).

As our ability to identify individuals with ASD has broadened, so too has our understanding of the great variety in the phenotypic expression of ASD. All individuals diagnosed with ASD share the same core characteristics, difficulties with social communication and the presence of restricted and repetitive behaviors. However, their expression of these characteristics can vary greatly. For instance, individuals with ASD have shown a wide range of

expressive language development with as many as 30% of individuals with ASD only achieving minimal verbal abilities (Tager-Flusberg, Paul, & Lord, 2005). We also see a wide spectrum of cognitive ability. While intellectual disability is often associated with ASD, estimates show that as much as 44% of individuals with ASD have average to above average IQ (Baio, 2012). In addition to this, it is also estimated that as many as 10% of individuals with ASD express savant level skills (Rimland, 1978). It is this great diversity among individuals with ASD that has inspired the saying “if you know one person with autism, then you know one person with autism.”

Because of this great diversity of skills and challenges among the ASD population, researchers and interventionists face many questions on how best to provide support and services to people with ASD. Early intervention is key, as starting earlier in development can have greater cascading benefits. The method and targets of early intervention form the basis of a large section of the ASD research literature. Intervention styles can be discussed in two broad categories, comprehensive interventions and targeted interventions.

Comprehensive Interventions

The challenges associated with ASD can have far reaching effects on development. These effects are often felt throughout all aspects of an individual’s life. Therefore, intervention programs for individuals with ASD were developed to address the many situations and developmental pathways one may experience. These comprehensive approaches are based on a set a core principles that are used to intervene across developmental, cognitive, and social domains. The most prominent of these comprehensive programs are based on the work of Dr. Lovaas (1987), who used behavioral principles to develop a comprehensive teaching program to address the deficits associated with ASD. This work serves as the foundation for much of the

intervention we see now for individuals with ASD. More recently, others have attempted to develop more updated comprehensive approaches targeting earlier intervention (Dawson et al., 2010) as well as combining developmental and behavioral principles (Koegel & Koegel, 2006). In general, these comprehensive programs seek to develop a curriculum that can be used to address any or all of the difficulties individuals with ASD face.

Targeted Interventions

Another intervention approach focuses more on specific areas of difficulty or developmental domains, known as targeted interventions. The goal of targeted interventions is to address specific core deficits of ASD with the hope that improvements in core deficits can have beneficial cascading effects in other areas of development. One example is an intervention designed to target a specific core deficit of children with ASD, joint attention. Joint attention is widely acknowledged to be a preverbal skill leading to the development of language among typically developing individuals (Tomasello & Farrar, 1986). However, children with ASD typically have impaired and atypical development of joint attention skills (Paparella, Goods, Freeman, & Kasari, 2011). It is presumed then that targeting joint attention skills through intervention could lead to improved language development. This is exactly what researchers found. Targeting joint attention skills in children with ASD not only led to immediate improvements in joint attention (Kasari, Freeman, & Paparella, 2006), but also had cascading effects on language development (Kasari, Paparella, Freeman, & Jahromi, 2008; Kasari, Gulsrud, Freeman, Paparella, & Helleman, 2012).

Determining the Evidence Base

While both comprehensive and targeted intervention approaches have generally shown positive results, the great variation in abilities and challenges among individuals with ASD poses

a significant problem for intervention researchers. Traditional intervention research is based on the ideas of establishing efficacy and effectiveness of any given intervention. It is the responsibility then of ASD researchers to rigorously test interventions for efficacy and effectiveness. To do this, Chambless and Hollon (1998) suggest three requirements. First, the treatment must be shown to be beneficial in controlled study. Second, the treatment must be useful for a clearly defined group and circumstances in applied settings. Finally, the treatment must be more efficient than other alternative interventions. Achieving these goals has been historically difficult for ASD intervention researchers (Kasari, 2002), with critics noting the lack of rigorous testing methods used in many intervention studies (Warren et al., 2011). Efforts have been made to develop a model for the development, validation, and dissemination of ASD interventions (Smith et al., 2007), leading to considerable progress in the development of evidence based practices (Wong et al., 2015). However, the heterogeneous nature of ASD makes it difficult for intervention researchers to progress beyond the efficacy stage. This can be especially difficult for comprehensive intervention approaches testing effectiveness in applied settings where practitioners are likely to encounter individuals with a wide variety of skills and challenges. Furthermore, without the highly controlled conditions of efficacy research, it is difficult to assume that similar results can be expected when interventions are introduced in settings where conditions can vary widely. Considering these challenges, it is likely that there is no one-size-fits-all intervention (Lord & McGee, 2001).

For Whom Interventions Work

To truly develop effective and impactful interventions for ASD, researchers must do a better job of understanding intervention, and for whom it works. This includes a better understanding of who may benefit the most from a given intervention, and ways to predict

response. Indeed, researchers have found several common predictors of response to ASD intervention. Generally, age at entry, IQ, symptom severity, language level, adaptive skills, imitation skills, comprehension, and non-verbal IQ predict outcomes in intervention trials (Harris & Handelman, 2000; Howlin, Magiati, & Charman, 2009; Itzhak & Zachor, 2011; Schreibman, Dufek, & Cunningham, 2011). However, these data are complicated by inconsistency in outcome measurement across studies. When examining these findings based on the measured outcome of each design, it is clear that these findings become far more nuanced. Examining expressive language, age at entry (Kasari, Gulsrud, Freeman, Paparella, & Helleman, 2012), joint attention skills (Kasari et al., 2012; Paul, Campbell, Gilbert, & Tsiouri, 2013), play skills (Kasari et al., 2012), imitation skills (Sallows & Graupner, 2005), expressive language at entry (Gordon et al., 2011), social avoidance (Ingersoll, Schreibman, & Stahmer, 2001), and ASD symptomatology (Vivanti, Dissanayake, Zierhut, Rogers, & Victorian ASELCC Team, 2013) all predict expressive language gains associated with intervention. For adaptive skills, age at entry (Baker-Ericzen, Stahmer, & Burns, 2007) and interest in toys (Klintwall, Macari, Eikeseth, & Chawarska, 2014) predicts better outcomes associated with intervention. While IQ is a common measure among intervention studies, predictors of change in IQ were not as easily identified. Not surprisingly, IQ at the start of intervention and ASD symptomatology were predictive of later IQ gains (Sallows & Graupner, 2005).

Moderators of Response

While the previous studies looked at factors that may predict global developmental outcomes, another approach is to look at predictors in terms of intervention specific outcomes. Looking more specifically at intervention outcomes maybe be beneficial in developing and adapting a given intervention. In one example, better expressive language at entry was associated

with improved use of a picture exchange communication system (PECS), which was the main target of the intervention (Gordon et al., 2011). While this approach is certainly informative, focusing on intervention specific outcomes can quickly become complex. It is not as simple as finding out what types of individuals do well. Instead it becomes a question of which individuals do well for which outcome, under what circumstances. Identifying predictors of response to intervention can be just as variable as the individuals being studied.

Studies that have compared two different interventions highlight the complexities of predictors on intervention outcomes. In one study comparing PECS with a play based language intervention, levels of initiations of joint attention had different effects on the two intervention groups. Children with higher levels of joint attention at entry showed more joint attention improvements in the play based intervention while children with lower levels of joint attention at entry show more requesting improvements with PECS (Yoder & Stone, 2006). The authors also looked at the novel predictor of object exploration, defined as the number of unique toys touched during a structured play assessment at entry. They found that the play based intervention was better for children entering with lower object exploration and PECS was better for children entering with higher object exploration (Yoder & Stone, 2006). The authors concluded that the play based intervention benefitted children that were less motivated by toys at the start by helping to develop interest in toys, while children that were already motivated by toys were ready to utilize PECS for communication. Another study examined two different language interventions, a naturalistic intervention and a more traditional discrete trial training intervention. Both groups benefitted from intervention with joint attention skills moderating change in both groups. However, receptive language abilities affected each group differently. Children with higher receptive language abilities did better in the naturalistic intervention while children with

lower receptive language abilities did better with the discrete trials intervention (Paul et al., 2013).

While moderator studies provide valuable information for whom the given intervention may benefit the most, their findings are difficult to interpret. That difficulty becomes even more apparent when attempting to utilize this information in applied settings. In general, studies clearly define their outcome variables, but these outcomes can vary greatly between studies. This type of variability can make it difficult for clinicians or educators to make use of this information. Measurement and analysis using standardized measures can be more interpretable.

A similar problem occurs when considering the variables used. Information such as age, symptom severity, and other measures gathered from standardized assessments can be more readily interpreted in applied settings. However, more novel variables, such as social avoidance and object exploration, may be more difficult for clinicians and educators to reliably quantify. Identifying and quantifying these characteristics in applied settings will likely require further training and expertise.

Profiling Responders to Interventions

Aside from identifying predictors of response to intervention, researchers have also attempted to behaviorally profile responders. That is, retrospectively identifying individuals that have completed an intervention program, shown significant progress, and then using those individuals to develop a profile to predict response to the intervention among future participants. In a series of studies examining pivotal response training, researchers looked back at participants from a previous study to identify individuals that exhibited the most favorable and least favorable outcomes. Videotaped structured intake assessments were reviewed to develop a behavioral profile for each of the two groups, referred to as the responders and non-responders.

Responders were found to exhibit more interest in toys, more tolerance of another person in close proximity, lower rates of non-verbal self-stimulatory behavior, and higher rates of verbal self-stimulatory behavior. Conversely, non-responders exhibited the opposite profile. These profiles were then used to identify new participants to test their predictive ability. Six participants that fit the responder profile were matched (by language, age, and IQ) with five participants that fit the non-responder profile. As predicted, children matching the responder profile showed positive outcomes while the non-responders did not (Sherer & Schreibman, 2005). While these findings are very encouraging, they are somewhat limited. The behavior profiles were based off of coding from videotaped assessments of 11 children (6 responders and 5 non-responders). Although this is a good first step, this number is far too small to generalize these findings. However, the positive results and method can serve as an example of the type of research needed to better understand intervention and the individuals for whom they work.

Methodological Advances for Determining Responders

Developing a behavioral profile to predict responder status for a given intervention can be time consuming, particularly when working with larger participant pools and datasets. More recent developments of statistical prediction models have the potential to address some of these difficulties. In particular, the use of machine learning paradigms can be a promising way of developing predictive models of responder status for ASD interventions. Machine learning is a term that refers to the use of computer algorithms to recognize patterns in data, and in turn make predictions on data. This strong predictive power and classification ability have been useful in several areas of ASD research. Machine learning techniques have been used to improve diagnosis of ASD by looking at predictors of ASD designation across different sites in the United States (Lord et al., 2012). Furthermore, the predictive power of machine learning

paradigms were used to analyze specific item level scores to propose a more streamlined model for the two gold standard assessments for ASD, the Autism Diagnostic Interview (ADI; Wall, Dally, Luyster, Jung, & DeLuca, 2012) and the Autism Diagnostic Observation Scale (ADOS; Wall, Kosmicki, DeLuca, Harstad, & Fusaro, 2012). Leveraging the predictive power of machine learning, and its ability to handle large and complex data sets, researchers have used it to examine early predictors of ASD diagnosis. In a reanalysis of event-related potential data gathered from electroencephalograms of infants, researchers were able to correctly identify the at risk infants with 64% accuracy, representing an improvement over the original analysis (Stahl, Pickles, Elsabbagh, Johnson, & BASIS Team, 2012). Genetic data is similarly complex, and the use of machine learning can be particularly useful. A recent study has had some success in developing a model predicting ASD diagnosis based on gene profiles, achieving a 71.7% accuracy (Skafidas et al., 2012).

Classification and Regression Trees for Responder Status

While the term machine learning encompasses a range of statistical methods, one method particularly well suited for ASD intervention research is the use of classification and regression trees (CART, Breiman, Friedman, Stone, & Olshen, 1984). CART methods have been used for a variety of purposes. Most notably, CART analysis has been used to make key decisions about individuals, including the categorization of emergency room intakes for heart attack risk to determining the decision to accept or reject an application for credit. Furthermore, the CART method and its extensions are equipped to efficiently analyze the complex data associated with individuals with ASD and the intervention studies they participate in. One of the strengths of the CART method is its ability to quickly and easily describe associations in data (Lemon, Roy, Clark, Friedmann, & Rakowski, 2003). It is also particularly suited to identify subsets within a

given population (Breiman et al., 1984). Researchers have leveraged the strengths of CART analysis to examine temporal lobe differences among children with ASD (Neeley et al., 2006) and predictors of outcome among 18 month old high risk sibling of children with ASD (Chawarska et al., 2014). The findings by Chawarska et al. (2014) highlight the utility of CART analysis in identifying subgroups. Their findings identified three specific behavior grouping among high risk infant siblings related to later ASD diagnosis. These groupings included having poor eye contact with low communicative gestures and giving; poor eye contact with low imaginative play; and intact eye contact paired with lack of giving and repetitive behaviors. As more ASD research continues to take advantage of CART techniques, intervention analysis has typically not made use of these techniques. The only current example utilizes CART methods to analyze participants mid-study to identify subgroups of children that are responding differently to the intervention (Shih, Patterson, & Kasari, 2014). This study highlights the great potential in using CART methods for intervention research. The findings from this study suggest measurements collected mid-treatment can be used to identify children that may need altered treatment, thus providing decision points to personalize their intervention plan.

Aims

The overall goal of this study is to contribute to the research in personalizing ASD interventions by identifying individuals that are more or less likely to benefit from a particular intervention. This study takes a retrospective approach by examining a sample of children that have received a targeted social communication. The first aim is to define and identify subgroups of responders. To do so, the study examines expressive language gains over the course of the intervention, as measured by a standardized measure of cognitive development. The intervention examined here has shown consistent expressive language gains, and it is expected that subgroups

of responders can be identified based on expressive language gains. It is hypothesized that a group of super responders will be identified. These individuals will have experienced gains in expressive language above and beyond what would be expected through maturation. The second aim of the study is to develop a model to predict responder status. Advanced machine learning techniques are used to develop a decision tree based prediction model. Predictors will include demographic and assessment information collected before the start of intervention. All predictors will be quantifiable variables collected from standardized assessments and instruments. It is hypothesized that a decision tree based predictive model will accurately predict responder status. The primary predictors will be frequency of joint attention gestures and mastered play level. This study's third aim is to compare the machine learning technique with more traditional model building through logistic regression. It is hypothesized that the machine learning model will outperform the logistic regression in predicting responder status in addition to providing more clinically relevant information.

Methods

Participants

This study collects data from participants across five different JASPER projects. The primary outcome of interest is expressive language gains through JASPER, and participants are included based on age and expressive language level at entry. Participants selected are between 24 months and 60 months of age at entry. While it is estimated that 30% of individuals with ASD only develop minimal expressive language skills (Tager-Flusberg, Paul, & Lord, 2005), it is still unclear if children under 60 months of age are indeed minimally verbal or pre-verbal. It may be the case that children at this age can still develop expressive language skills with appropriate intervention. This project only included children that were assessed using the ADOS 2, module

1, which is designed for children whose spontaneous language is primarily single words or through behavioral means (Lord, DiLavore, & Gotham, 2012). Overall the sample consists of 99 minimally verbal, preschool age children with ASD that received the JASPER intervention.

Design

The current study is a secondary analysis using data gathered from participants across five separate JASPER intervention studies. Each of the studies collected common measures of ASD assessment, demographics, cognitive development, social communication skills, and play development.

The Intervention – JASPER

This study examines an intervention for children with ASD developed to target joint attention, symbolic play, engagement, and regulation called JASPER. This intervention is backed by an evidence base spanning over a decade. In the original study, targeting joint attention and symbolic play separately both led to gains in expressive language for children with ASD. Targeting joint attention was particularly effective for children entering with the lowest language levels (Kasari, Paparella, Freeman, & Jahromi, 2008). A five year follow up of these children found that the effects of the two intervention groups maintained (Kasari et al., 2012). Following these findings, the two interventions were combined to form the method now known as JASPER. In the years since, JASPER has shown to be effective for toddlers (Kasari et al., 2014), as a parent mediated intervention (Kasari, Gulsrud, Wong, Kwon, & Locke, 2010), in preschools (Goods, Ishijima, Chang, & Kasari, 2013; Chang, Shire, Shih, Gelfand, & Kasari, 2016), and for minimally verbal children with ASD (Kasari et al., 2014). As the evidence for JASPER continues to grow, and efforts to bring it to the community increase, this study hopes to provide practitioners a method to better understand which children may benefit the most from JASPER.

Measures

Demographics: Prior to entering intervention, caregivers of the participants completed a demographic form. For the current study child age at entry, reported gender, and reported race/ethnicity are used as predictors in the analysis.

ADOS (Lord, DiLavore, & Gotham, 2012): The ADOS is a standardized semi-structured assessment used to evaluate individuals suspected of having an ASD. The participant is engaged in a variety of activities designed to elicit social and communication behaviors associated with the diagnosis of ASD's. Cutoff scores allow for diagnosis of autism or autism spectrum disorder. The ADOS was administered to all children across all studies to confirm diagnosis. As noted above, scores from the ADOS are used to identify minimally verbal children for inclusion in the current sample. ADOS comparison scores are also used as a proxy measure for ASD symptom severity.

Mullen Scales of Early Learning (MSEL; Mullen, 1995): The MSEL is an assessment of early intellectual development and school readiness. It is appropriate for the ages from birth to five years eight months. Child development is assessed across five domains; gross motor, visual reception, fine motor, expressive language, and receptive language. Age equivalents can be calculated for each of the five domains. Non-verbal and verbal developmental quotients can also be calculated using the visual reception and fine motor subscales. An overall standardized composite score can also be calculated for a measure of overall cognitive development. For the current study, expressive language age equivalents will be used to calculate the outcome measure and age equivalents for fine motor and visual reception development will be used as predictors.

Early Social Communication Scales (ESCS; Mundy et al., 2003): A structured table top assessment, the ESCS is designed to elicit joint attention and requesting behavior. Sitting across

from a child at a table, an experimenter engages the child with a variety of play activities including wind-up toys, a balloon, a car, a ball, and tickle songs for 15 to 20 minutes. The experimenter will also periodically point to pictures in a book and around the room to elicit a child response to joint attention. ESCS sessions are videotaped and later coded for specific joint attention skills. The specific skills coded are coordinated looks, alternating gazes, points, gives, shows, and language. Skills can also be coded in combination with language and/or eye contact (eg. a point with language, or a give with eye contact and language). The ESCS provides a total score of joint attention and requesting skills. For the current study, joint attention and requesting gestures are used as predictors.

Structured Play Assessment (SPA; Ungerer & Sigman, 1981): The SPA is a play-based assessment designed to measure a child's level of play. An experimenter sits across from the child and presents one of five different toy sets individually until the child has had an opportunity to play with each set. The five sets include the following toys: set one contains a puzzle, shape sorter, and nesting cups; set two contains a tea set, dolls, and blocks; set three contains a phone, a brush, a mirror, and dolls; set four contains a furniture set and dolls; set five contains animal figures, blocks, a car, and a barn. Altogether, the SPA lasts approximately 10 to 15 minutes. SPA sessions are videotaped and later coded for play levels displayed by the child. Individual play acts are tracked and used to assess the child's play level. There are 16 levels of play that fall under five broad categories; indiscriminate acts, discriminate acts, combination play, pre-symbolic play, and symbolic play. The SPA is coded by tracking the individual play types the child exhibits, and that type's frequency. For example, if child plays with the puzzle and completes all six pieces, that is coded as one play type with a frequency of six. Total number of play types is scored for play diversity. Mastered play level is determined by the highest level

of play in which a child exhibits two different types, with a frequency of at least five. Play diversity and mastered play level are included as predictors for analysis.

Analysis

The current study examines the utility of a novel statistical method to look at predictors of language gains achieved while receiving a targeted intervention (JASPER). Specifically, the analysis is based on classification and regression trees (CART) to develop predictive models of intervention gains, and will compare these models with those developed using more traditional regression techniques.

CART methods, and its derivatives, provide several potential benefits for ASD intervention research. CART analysis is particularly suited to identify subsets within a given population (Breiman et al., 1984; Lemon et al., 2003). Furthermore, where traditional regression methods are adept at exploring the effects of a single variable or set of variables on an outcome measure, CART methods can more effectively describe associations in the data (Lemon et al., 2003). Using these methods, we will be able to identify subgroups of children and how the predictor variables are associated within those subgroups. From a more practical perspective, CART results are presented as decision trees with clearly defined decision points. These findings are more easily interpretable, particularly in a practical setting. Decision trees can allow interventionists to more quickly assess the potential benefit of a specific intervention, like JASPER, for a specific child. CART analyses also provide some benefits from a statistical standpoint. CART methods are non-parametric, and not subject to the assumptions necessary in traditional parametric analyses. This is particularly useful for the current study as the data involves minimally verbal children with ASD. For these individuals, data may be positively

skewed due to the low frequency of certain measured behaviors. Such data is handled easily by CART methods.

The ease of use and interpretation of CART analyses can also lead to potential drawbacks. Lemon et al. (2003) note the temptation of “data dredging” by simply adding as many variables into the analysis as possible. Without setting a priori parameters, CART analysis will fit its prediction model as thoroughly and completely as possible, resulting in over fitting. While this can result in a model with more predictive power, it sacrifices interpretability. Resulting subgroups may not make much clinical or practical sense. It is recommended then that a priori hypotheses be made of variable importance and direction when using CART analyses (Marshall, 2001). Furthermore, CART analyses have been shown to display statistical biases to certain types of data (Hothorn, Hornik, & Zeileis, 2006). In particular, when analyzing datasets with a mix of categorical and continuous variables, continuous variables are often favored when determining decision points in the resulting trees. It is recommended that an alternative form of CART, known as conditional inference trees (CIT), is used when analyzing such data sets (Hothorn, Hornik, & Zeileis, 2006). CIT analysis is well suited to handle the mix of continuous and categorical data in the current study’s data set.

CIT analysis forms the foundation for the current study’s analytic plan. Preliminary analysis is conducted using an extension of CIT analysis called conditional inference forests. These forests are used to generate variable importance values for large sets of predictors which will guide variable selection for the final model. Once the final variables are selected, they are then run through CIT analysis to develop the final tree model. In addition to the CIT analyses, traditional regression analyses are performed to compare results from the two different statistical

approaches and to explore the utility of using the machine learning techniques. All statistical analyses were conducted using R version 3.5.1 (R Core Team, 2013).

Results

Participant Demographics

The sample for the current study consists of minimally verbal, preschool age children. Descriptives (see Table 1) show this is reflected in the sample with a mean age just under 40 months. The average expressive language age equivalent for this sample is 17 months, showing significant language delays. Table 2 displays demographic data, and this sample represents a more diverse population with 65% of the subjects identifying as belonging to racially diverse groups. The high percentage of male participants (86%) is as expected, and in line with the gender differences in prevalence rates among individuals with ASD.

As this sample consists of participants from five different studies, preliminary analysis was conducted to examine any significant differences between participants in the different studies. Kruskal-Wallis and Fisher's Exact tests were run to look for group differences (see Table 2). The different studies differed in several variables: expressive language age equivalents at entry, $\chi^2(4, N = 99) = 11.27, p = .023$, age, $\chi^2(4, N = 99) = 64.85, p = .023$, severity, $\chi^2(4, N = 99) = 10.91, p = .027$, joint attention gestures, $\chi^2(4, N = 99) = 13.919, p = .007$, requesting gestures, $\chi^2(4, N = 99) = 21.88, p < .001$, play diversity, $\chi^2(4, N = 99) = 15.25, p = .004$, fine motor age equivalency, $\chi^2(4, N = 99) = 15.03, p = .004$, and race/ethnicity, $p = .007$. To account for these differences in the analysis, a variable was created to represent the studies to control for these differences.

Identification of Responders

Response status was determined by calculating change scores in expressive language age equivalents (ELAE) from the MSEL. The MSEL is administered prior to the start of intervention and again after completion. To determine slow and super responders, change scores in ELAE were calculated and then divided by the duration (in months) between administration of pre and post MSEL tests. Scores greater than one were categorized as super responders. In our sample 47% (n = 47) of individuals were identified as super responders and 53% (n = 52) were identified as slow responders. Responder status was then used as a binary outcome in all subsequent analyses.

Conditional Inference Forest for Variable Selection

Table 1 displays descriptives for the outcome and predictor variables used in the following analyses. To identify variable importance and selection for the final model, conditional inference forests were generated with responder status as the outcome. Due to the random nature of forest building, three separate forests were generated, each containing 10,000 trees. The result of each forest provides variable importance scores of the predictors based on conditional permutation of the predictor variables as described by Strobl et al. (2008). Positive scores indicate predictors that increase model efficiency, negative scores indicate predictors that decrease model efficiency, and scores of zero indicate a neutral influence. Figure 1 shows the results for forest one. Overall, seven predictors are identified as increasing model efficiency, with three (requesting gestures, play diversity, and fine motor skills) showing particularly strong importance for the model. Figure 2 shows the results for forest two. Again, we find requesting gestures, play diversity, and fine motor skills to have the strongest influence on model efficiency. As expected, the third forest (shown in figure 3) identifies similar rankings for variable

importance. While forest models are adept at identifying variable importance, the exact significance and how they influence the model is less clear. To explore this further, a single tree model based on the identified predictors will help us better interpret the predictive model. Final variable selection is greatly dependent on the specific data, and without an established significance test for variable importance scores, predictors were chosen based strength of influence, parsimony, and clinic relevance. As such, the final predictors chosen for CIT analysis were requesting gestures, play diversity, and fine motor skills.

Building the Conditional Inference Tree (CIT)

Results from the CIT analysis are depicted in Figure 4. The CIT model was built using a .95 confidence level, meaning only predictors with a p value less than .05 would be selected for the model. Using requesting gestures, play diversity, and fine motor skills as predictors, the resulting tree included a single significant split with two terminal nodes. The single significant predictor of responder status was play diversity, with a splitting point of a play diversity score of 23 ($p = .004$). The overall performance of the CIT model to accurately predict responder status is 67%, with a specificity of 55% and sensitivity of 78%.

Building the Logistic Regression Model

Stepwise logistic regression analysis was also used to explore the relationship between requesting gestures, play diversity, fine motor skills and responder status. Similar to the CIT, results of the stepwise logistic regression show play diversity as the only significant predictor of responder status, $\chi^2 (1) = 10.686, p = .001$.

Comparison of CIT and Logistic Regression Models

To compare the predictive power of the CIT model with the logistic regression model, receiver operating characteristic (ROC) curves were generated, and the area under the curve

(AUC) values were calculated for each model. ROC curves are presented in Figure 5. AUC is often used as quality indicator and method for comparison classification models (Bradley, 1997). AUC values for the CIT model (.68) and for the logistic regression model (.69) were compared using a DeLong test and found not to be significant ($p = .82$). There was no difference in predictive performance between the CIT and logistic regression models.

Building a Second Conditional Inference Tree

To address the moderate predictive accuracy of the first CIT analysis, a second analysis was run. This time the confidence level was reduced to .90, so that predictors with p values lower than .10 would be included in the model. The CIT analysis was repeated with the same three predictors, resulting in a new tree. Results from the second CIT analysis are depicted in Figure 6. The primary split remained play diversity, with a score of 23 identifying the two groups. However, now fine motor age equivalents of 26 months further distinguished the low play diversity group. Overall accuracy of the new model saw a slight increase to 69%, with a sensitivity of .72 and a specificity of .67.

Discussion

Personalized Intervention

Individuals with ASD often present with a wide range of abilities and challenges. While many intervention models and techniques have been developed, providing interventionists with more tools, there is likely no single technique that can appropriately address every individual's needs. Therefore it is imperative that intervention researchers focus on ways to individualize treatment. As intervention techniques continue to evolve, parallel efforts must be made to identify the individuals for whom interventions will benefit the most.

One approach to identify for whom an intervention may work is to retrospectively examine individuals that have received intervention. Beyond looking at intervention effects, we can examine predictors of success in a given intervention. This is not a novel idea, and much work has already been dedicated to predicting optimal outcomes. This research has provided us with general ideas of predictors such as IQ, language ability, and age. However, seeing such variety in skills and difficulties among individuals with ASD, there is still much more we can do to help personalize intervention on a more individual level.

The current study attempts to address this by looking at an established, evidence based intervention for children with ASD. Known as JASPER, this intervention has an extensive evidence base, showing consistent improvement in joint attention, play, and engagement for children with ASD. More interestingly, these findings have also led to gains in expressive language. Furthermore, the researchers behind JASPER have collected consistent and thorough pre-intervention data on all of their participants. This wealth of data, along with consistent outcomes, provides an excellent opportunity to explore potential predictors of responders to JASPER.

As noted earlier, predicting intervention outcomes is not a new idea. However, with advancements in statistical techniques such as machine learning, there is an opportunity to provide more nuanced predictive models. In the current study, machine learning techniques are applied to examine potential predictors of response to JASPER. Specifically, decision tree based models are used. These techniques have the potential to provide clear subgroups of responders, with specific information that can aid in important intervention decisions.

Identifying Super Responders

In order to predict response, it is important to first identify responders. This study's first aim was to identify "super" responders among children that received JASPER. To meet this aim the study attempted to satisfy three conditions; identify an outcome that is clinically relevant, quantifiable, and utilizes a broadly accepted standardized measure. One important response measure for individuals with ASD is expressive language. Development of functional expressive language is an important predictor of future outcomes (Magiati, Tay, & Howlin, 2014). Clinically, functional communication is a primary intervention goal, with use of functional expressive language being the optimal outcome. JASPER has specifically shown to lead to expressive language gains in children with ASD (Kasari et al., 2012). Therefore, identifying super responders based on expressive language would be clinically significant, and an important factor in identifying for whom JASPER would work the best.

Measuring expressive language ability can be complicated, and the use of a reliable and standardized measure is necessary. Additionally, among young children with ASD, expressive language development is extremely variable. Making gains at a young age, particularly before the age of five, can have important consequences for all phases of development. The MSEL is a standardized cognitive measure, widely used in ASD research, shown to be sensitive to developmental delays (Burns, King, & Spencer, 2013), and particularly useful in the assessment of children with ASD (Bishop, Guthrie, Coffing, & Lord, 2011). With regards to expressive language development, the MSEL provides T scores, percentile ranks, and age equivalents (in months). Due to the expressive language and behavioral challenges seen in young children with ASD, MSEL age equivalents may be a better indicator of language ability for these children and a potential source for measuring intervention effects (Akshoomoff, 2006).

By calculating change scores in expressive language age equivalents and comparing them to the duration between assessments, each participant was given a responder score. A responder score less than one indicated that their expressive language development between administrations of the assessments was lower than what we would expect due to maturation. Any responder scores greater than one would indicate expressive language gains above and beyond what would be expected through maturation. As expected, this measure of response status did identify two subgroups in our sample with about 47% being identified as super responders. Looking at the average age of our sample at just over 3 years at entry ($M = 39.7$ months, $SD = 9.9$) with an average expressive language age equivalent at entry of just over 17 months ($SD = 8.9$), our sample was significantly language delayed. With this delay, they appear to have been on a depressed developmental trajectory. Identifying close to half of the current sample as meeting the criteria for super response is encouraging. Almost half of the children that received JASPER had experienced accelerated growth in expressive language as measured by the MSEL.

However, with the current analysis, it is unclear exactly what the nature of expressive language growth is in this sample. Even for the children categorized as slow responders, nearly all of them made some expressive language gains, as measured by the MSEL. Considering their likely depressed developmental trajectory, any improvement may be clinically significant. While the current study emphasized identifying super responders, future work should focus on overall response, and identify what meaningful changes may occur as a result of intervention.

Predicting Response Status in JASPER

Utilizing the subgroups of responders identified, the current study applied CIT modeling in two ways to develop a predictive model. First, conditional random forests were used to identify the most important predictors associated with responder status. While seven predictors

overall were identified as positively affecting prediction power, three were chosen for the final model. Currently, there is no established standard to measure significance of variable importance, so variable selection involves assessment of the data, goals of analysis, and parsimony. For the current project, three predictors stood out as influential, and clinically relevant to predicting response to JASPER. The top three variables chosen were play diversity, requesting gestures, and fine motor skills. In the context of JASPER, a play based naturalistic intervention, these three predictors have particular clinical relevance. Play diversity and fine motor skills may indicate the motivation and ability to play with a varied set of toys, which would be beneficial in a play based intervention such as JASPER. The use of requesting gestures as a predictor was somewhat unexpected. However, considering these children were identified as being minimally verbal, gesture use would be important. Increased gesture use for any communicative intent could indicate a child is comfortable communicating with others, albeit without expressive language. Because of the random nature of forest building and the fact that each predictive model is built from a whole series of trees, interpretation of the significance and effects of predictors is less clear. Running the predictors through a single tree can then shed light on the relationships of the predictors and responder status.

In the first CIT analysis, the split criterion was set at a confidence level of .95, so that only predictors with a p value below .05 would be included in the tree. The result was a tree with a single splitting point and two terminal nodes (figure 4). The significant predictor was play diversity, with a splitting point at a score of 23. This model indicates that children that started the intervention with a play diversity score higher than 23 were predicted to be super responders at exit. However a closer look at the predictive power shows a more nuanced picture. The overall accuracy of the model was 67%. So while play diversity is a significant predictor of responder

status, it is only marginally accurate. However, this level of accuracy is similar to other ASD studies using machine learning paradigms. One group looking at EEG data to predict early risk reports predictive accuracy of 64% (Stahl et al., 2012), and another group reported a 71.7% accuracy for a model predicting ASD diagnosis from genetic markers (Skafidas et al., 2012). A further examination of the current study's model characteristics shows us what the model is particularly effective for. Looking at the model's sensitivity (.55) and specificity (.78) we see that the model is actually a very poor predictor of super responders. Conversely, it does a much better job of correctly identifying slow responders. So, it may be more accurate to say that children with a play diversity score below 23 are less likely to be super responders in JASPER.

To compare the CIT analysis with a more traditional method, the predictors were re-run in a step-wise regression analysis. The resulting model again found play diversity as the only significant predictor, $\chi^2(1) = 10.68, p = .001, 95\% \text{ CI } [1.02, 1.11]$. These findings confirm the results from the CIT analysis, and that both techniques confirm the significance of play diversity. To compare the respective predictive power of the two models, ROC curves were generated and AUC scores were calculated. AUC scores are a measure of quality for predictive models by providing a probability score that given a set of two options, the model will correctly identify the two; in our case super vs. slow responders. Analysis of AUC scores showed no significance between the two models. In contrast with the current study's hypothesis, both models perform equally well when predicting responder status. It is still interesting that play diversity continues to be a significant predictor of responder status across all of the analyses. However, the respective scores of .68 for the CIT model and .69 of the logistic regression both indicate only marginal predictive accuracy.

One way to improve model performance is to sacrifice inferential power for predictive power. To that end, a second CIT analysis was run, relaxing the confidence level to .90. Any predictors with a p value under .10 were now to be included in the model. In this scenario, reducing the confidence level sacrifices some inferential power for predictors, but we are able to potentially gain clinically relevant information. With the new split criteria, the CIT analysis provides a tree with two splits and three terminal nodes (see Figure 6). The primary split is still play diversity with a score of 23, but now the low play diversity group is further divided by a fine motor age equivalent of 26 months. However, the increase in predictive accuracy is only marginally improved, approaching 70%. There is also an increase in sensitivity, now up to 72%. Still, the clinical information this provides paints a slightly clearer picture of slow responders. Children that come in with low play diversity and low fine motor skills may not do so well with JASPER. This further reinforces the strength of this model to rule out super response.

Play Diversity

It was hypothesized that joint attention gestures and play mastery would be predictive of expressive language gains, and play diversity was unexpected. However, in the context of a play based intervention such as JASPER, play diversity may very well be important. For the current project, play diversity is defined as the total amount of play types a child engages in during the structured play assessment. A higher score in play diversity indicates a child that engaged in more play types across a variety of the toys during the assessment. This may be pointing at a child that can play flexibly with different toys, and is motivated to play with toys. Both traits would be beneficial in JASPER. It can be noted that although play mastery was not a significant predictor, it is closely related to play diversity. Children that can play at a higher level are able to play with more of the toys in more diverse ways. Children at lower play levels are somewhat

limited in what they can do with the toys. What may be happening here is that a higher play level is not predictive, but reaching a certain play level will allow for more diverse play. At the preschool ages, we can expect children to be playing comfortably at a combination level and increasingly at a pre-symbolic play level. At these levels, children can essentially play with any toy in any number of ways. Even higher level toys such as dolls or figures can be treated as combination objects rather than pretend objects. Looking at our sample, over 80% of our participants had mastered combination (27%) or pre-symbolic play (54%). Considering the model's most accurate predictions are to rule out super responders, children with low play diversity (and low fine motor skills in the second model) may indicate children that don't yet have the skills, motivation, or flexibility to excel in a play based intervention such as JASPER.

The play diversity score may also be representing other skills or challenges affecting a child's ability to do well in JASPER. As noted before, a high score in play diversity requires the child to be flexible when playing with toys. Considering many children with ASD exhibit extreme repetitive and restricted behaviors, this may be the reason why low play diversity children in this sample were less likely to be categorized as super responders. It may be that extremely repetitive play negatively affects SPA performance and can be difficult to work through in JASPER sessions. Low play diversity may also be an indicator of child's ability to self-regulate. The structured play assessment (SPA), from which the play diversity score is derived, requires a child to sit at a table and play with different sets of toys, usually with an unknown assessor. Children that have difficulty regulating themselves to the point where they are unable to sit through the assessment may also have difficulty getting through a JASPER session. Another potential impact on play diversity may be socio-economic status. Socio-economic status has been shown to influence availability of a physical play space and play

materials (Freitas, Gabbard, Cacola, Montebelo, & Santos, 2013). Without access to these in the home environment, some children may have little opportunity to develop the type of play diversity needed to score well in the SPA.

Machine Learning in ASD Intervention Research

The current study also explores the utility of machine learning paradigms in ASD intervention research. Methods like CART analysis are becoming increasingly popular for their flexibility, relative ease of use, and ease of interpretation. However, they can be deceptively easy. Inherent biases in the variable selection and model fitting process plague CART analysis (Strobl, Boulesteix, Zeileis, & Hothorn, 2007), and researchers must look closely at their data to determine the appropriateness of any statistical analysis. In the current study, traditional CART analysis would be ill advised due to the complexity of the data. While traditional CART analysis is non-parametric, it tends to favor variables with multiple splitting points. Therefore there is an inherent bias towards continuous variable or categorical variables with multiple levels. The current study analyzes various predictors that include continuous variables, binary categorical variables, and multi-level categorical variables. Using traditional CART analysis with the current data set would raise many questions about the validity of the results. Fortunately, advancements in CART analyses detailed by Hothorn et al. (2006) and Strobl et al. (2008) have established the CIT framework to address these issues of bias. Utilizing this new framework provides confidence that the findings reported here are valid representations of the data. As statistical analyses continue to improve and grown in strength, it will benefit ASD intervention researchers to take advantage of these advancements. However, with this great power comes an even greater responsibility for researchers to ensure that their implementation is appropriate and thorough.

Limitations

There are several limitations to this current study. The CIT analysis used for this study is traditionally tested on a second test sample. To achieve this, researchers will often split their sample into a training set and a test set. However, this study wanted to take advantage of the entire sample to maximize the power to identify significant predictors. Furthermore, the sample was collected from participants across five different studies. There were several systematic differences across the studies, and although they were accounted for by creating a “study” variable, there may be underlying differences influencing the analysis. In all three of the forest models, the study predictor was identified as a predictor that contributed to model efficiency. However, it was not ranked as highly as the three main predictors, and subsequent CIT analysis shows that it would likely not be significant. Another limitation was the inability to examine effects of dose and intensity. Because dose and intensity varied systematically by study, it would be difficult to tease apart any effect due to other potential systematic differences. Future studies should consider ways to effectively examine the effects of dose and intensity.

Conclusions

This study highlights the importance and challenges of understanding for whom an intervention works best. For JASPER, it is encouraging to see that there are identifiable super responders, even across several years of research studies. Although the prediction model developed here was more effective in ruling out super responders, the better we understand our interventions and for whom they work or don’t work, the better of the children receiving intervention will be.

Intervention researchers should continue to look ahead to better predict for whom an intervention will benefit. As statistical methods and data collection techniques continue to

improve, there will be new opportunities and new methods to take advantage of. The current study was reliant on quantifiable, standardized measures that could be utilized and compared across studies. However, this a major challenge in the field of ASD research (Zwaigenbaum et al., 2015).

Future studies can continue to take advantage of novel approaches, and design studies to identify responder status much sooner, and systematically. Studies that employ prospective approaches like sequential multiple assignment randomized trials (SMART) can explore the idea of augmenting treatments in real time. Combined with retrospective work, like the current project, novel prospective designs can be very effective ways for researchers to study personalization and achieve truly individualized treatment.

Appendix

Table 2 Descriptives

| <i>Outcome Variable</i> | | |
|---|---------------|--------------|
| Expressive Language M (SD) | 17.77 (8.92) | months |
| <i>Predictor Variables</i> | | |
| Age <i>M (SD)</i> | 39.79 (9.9) | months |
| Fine Motor Age Equivalent <i>M (SD)</i> | 25.83 (7.56) | |
| Visual Reception Equivalent <i>M (SD)</i> | 25.6 (8.46) | |
| Joint Attention gestures <i>M (SD)</i> | 4.79 (6.3) | frequency |
| Requesting gestures <i>M (SD)</i> | 12.65 (9.61) | |
| Play Diversity <i>M(SD)</i> | 20.81 (11.34) | |
| Male | 86 (86%) | <i>n (%)</i> |
| Play Mastery | | |
| Simple | 7 (7%) | <i>n (%)</i> |
| Combination | 27 (27%) | |
| Pre-symbolic | 54 (54%) | |
| Symbolic | 11 (11%) | |
| Race | | |
| African American | 4 (4%) | <i>n (%)</i> |
| White | 35 (35%) | |
| Latino | 14 (14%) | |
| Asian | 22 (22%) | |
| Other | 24 (24%) | |

Table 3 Sample Characteristics

| <i>Variable M (SD)</i> | <i>Total^a</i> | <i>Study 1^b</i> | <i>Study 2^c</i> | <i>Study 3^d</i> | <i>Study 4^e</i> | <i>Study 5^f</i> | <i>p</i> |
|---------------------------|--------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------|
| Male <i>n</i> (%) | 86 (86%) | 11 (84%) | 27 (87%) | 13 (87%) | 4 (80%) | 31 (89%) | .94 |
| Age | 39.79 (9.9) | 45.1 (5.39) | 47.73 (7.37) | 30.29 (3.37) | 52.6 (12.01) | 38.62 (10) | <.001*** |
| ELAE | 17.77 (8.92) | 17.45 (6.67) | 23.47 (9.3) | 17.17 (9.49) | 8.8 (6.3) | 17 (9.66) | .02* |
| Race <i>n</i> (%) | | | | | | | .006** |
| African American | 4 (4%) | 2 (6%) | 0 (0%) | 0 (0%) | 0 (0%) | 2 (15%) | |
| White | 35 (35%) | 8 (25%) | 3 (20%) | 21 (60%) | 1 (20%) | 2 (15%) | |
| Latino | 14 (14%) | 2 (6%) | 2 (13%) | 3 (8%) | 2 (40%) | 5 (38%) | |
| Asian | 22 (22%) | 9 (29%) | 6 (40%) | 4 (11%) | 1 (20%) | 2 (15%) | |
| Other | 24 (24%) | 10 (32%) | 4 (26%) | 7 (20%) | 1 (20%) | 2 (15%) | |
| Severity | 6.82 (1.93) | 6.86 (1.96) | 7.2 (1.21) | 7 (2.07) | 8.6 (1.67) | 5.54 (1.66) | .02* |
| Joint Attention Gestures | 4.79 (6.3) | 4.16 (5.27) | 8.93 (8.98) | 3.54 (5.12) | 0.4 (0.89) | 6.54 (6.89) | <.01** |
| Requesting Gestures | 12.65 (9.61) | 16.87 (10.04) | 16.47 (9.52) | 8.31 (7.76) | 5.3 (1.92) | 12.69 (9.07) | <.001*** |
| Play Mastery <i>n</i> (%) | | | | | | | .42 |
| Simple | 7 (7%) | 1 (3%) | 1 (7%) | 4 (11%) | 1 (20%) | 0 (0%) | |
| Combo | 27 (27%) | 7 (23%) | 4 (27%) | 11 (31%) | 3 (60%) | 2 (15%) | |
| Pre-Symbolic | 54 (54%) | 19 (61%) | 8 (53%) | 17 (49%) | 1 (20%) | 9 (69%) | |
| Symbolic | 11 (11%) | 4 (13%) | 2 (13%) | 3 (9%) | 0 (0%) | 2 (15%) | |
| Play Diversity | 20.81 (11.34) | 24 (11.52) | 21.2 (11.34) | 16.74 (9.93) | 11.8 (10.47) | 27.15 (10) | .004** |
| VRAE | 25.6 (8.46) | 27.45 (8.68) | 27.4 (9.4) | 23.51 (7.77) | 23 (3.54) | 25.69 (9.47) | .48 |
| FMAE | 25.83 (7.56) | 28.77 (8.08) | 29.07 (9.05) | 22.31 (4.96) | 22.8 (2.59) | 25.69 (7.79) | .004** |

Note. ELAE = Expressive Language Age Equivalents; VRAE = Visual Reception Age Equivalents; FMAE = Fine Motor Age Equivalents

^a*N* = 99. ^b*n* = 31. ^c*n* = 15. ^d*n* = 35. ^e*n* = 5. ^f*n* = 13.

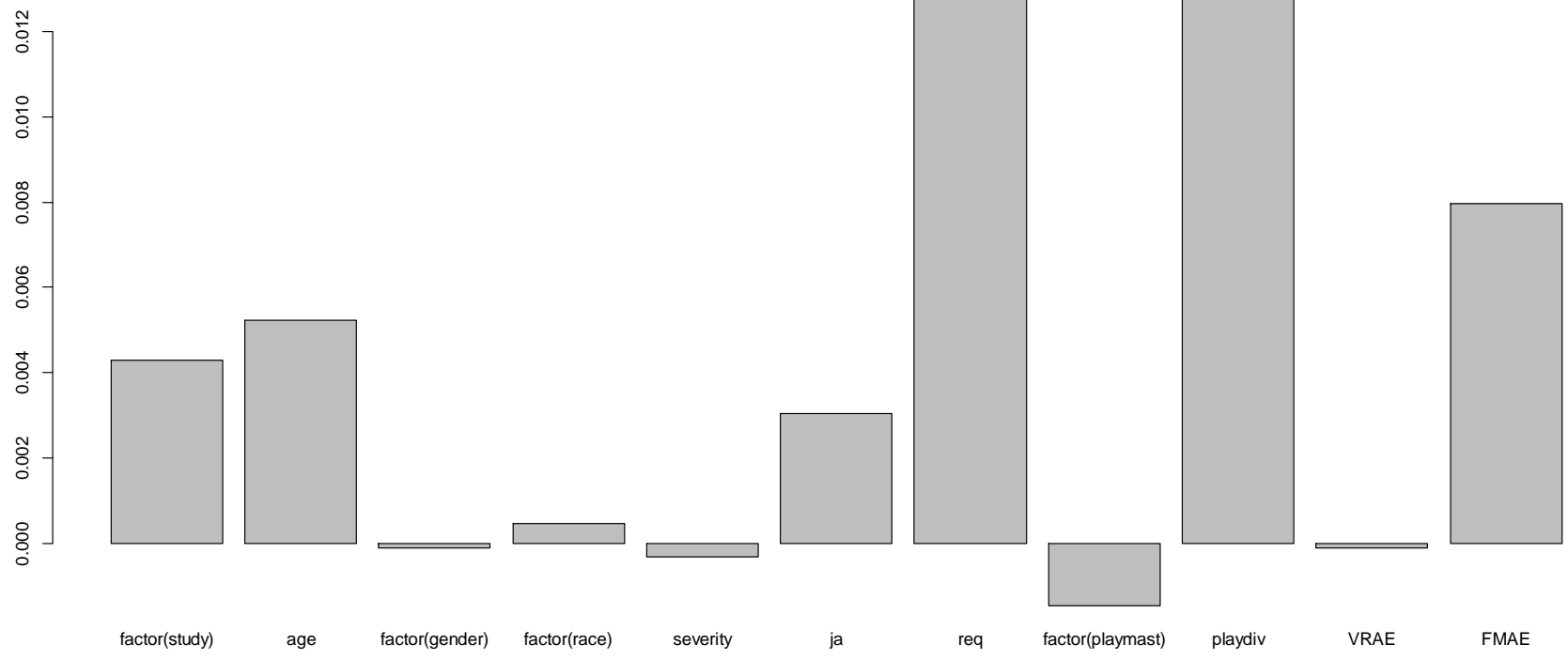


Figure 1 Conditional Inference Forest Variable Importance Plot #1

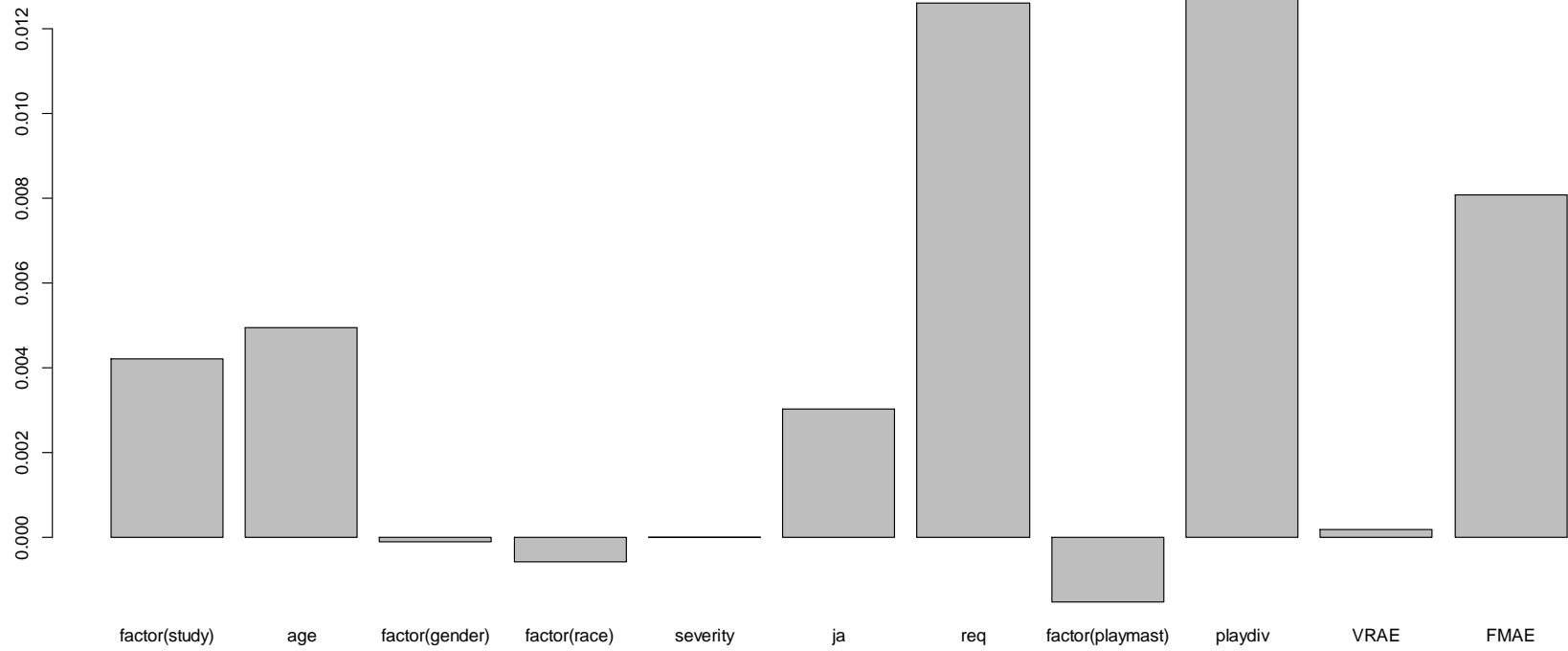


Figure 2 Conditional Inference Forest Variable Importance Plot #2

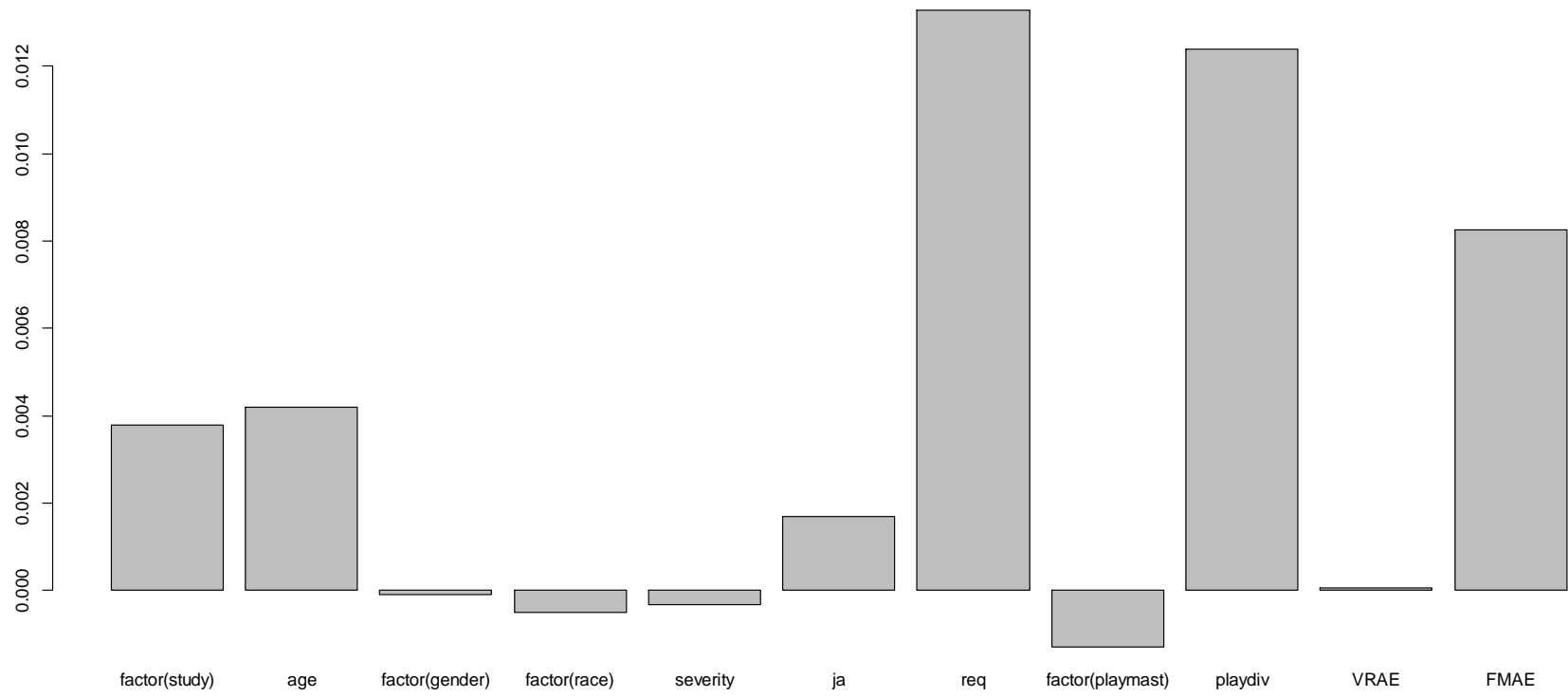


Figure 3 Conditional Inference Forest Variable Importance Plot #3

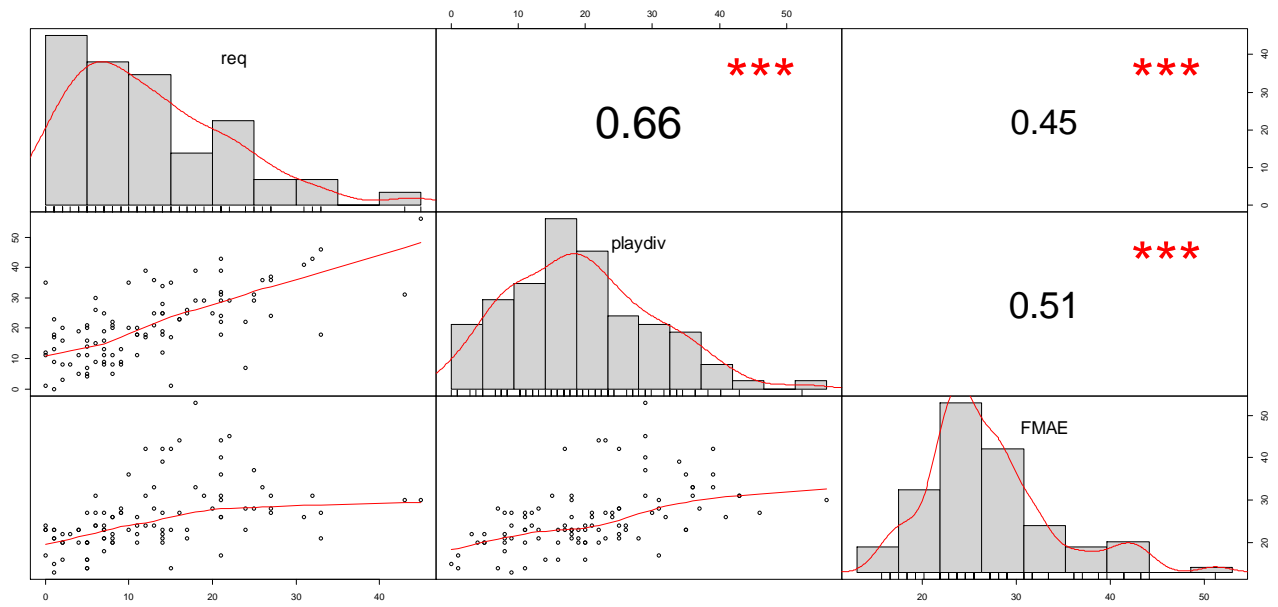


Figure 4 Correlation matrix for predictors in final model

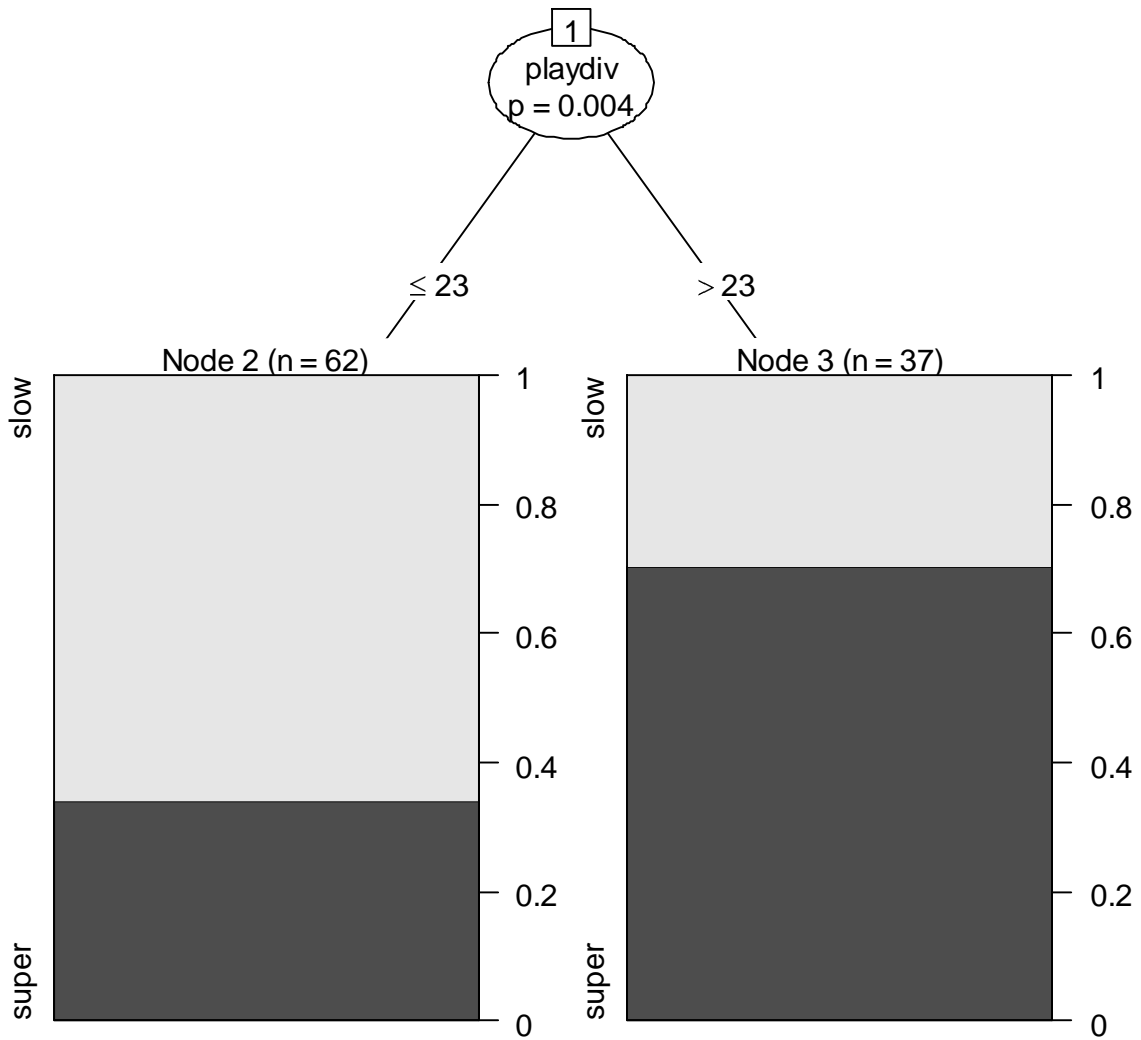


Figure 5 Conditional Inference Tree of Play Diversity Predicting Super vs. Slow Responders, confidence level of .95.

Table 4 Stepwise Logistic Regression Results

| | | <i>95% CI for Odds Ratio</i> | | |
|----------------|---------------|------------------------------|-------------------|--------------|
| | B (SE) | <i>Lower</i> | <i>Odds Ratio</i> | <i>Upper</i> |
| Included | | | | |
| Intercept | -1.41 (0.48) | | | |
| Play Diversity | 0.06 (0.02) | 1.02 | 1.06 | 1.11 |

$R^2 = .08$ (Hosmer & Lemeshow), $.10$ (Cox & Snell), $.13$ (Nagelkerke), Model $\chi^2(1) = 10.686$, $p = .001^{**}$

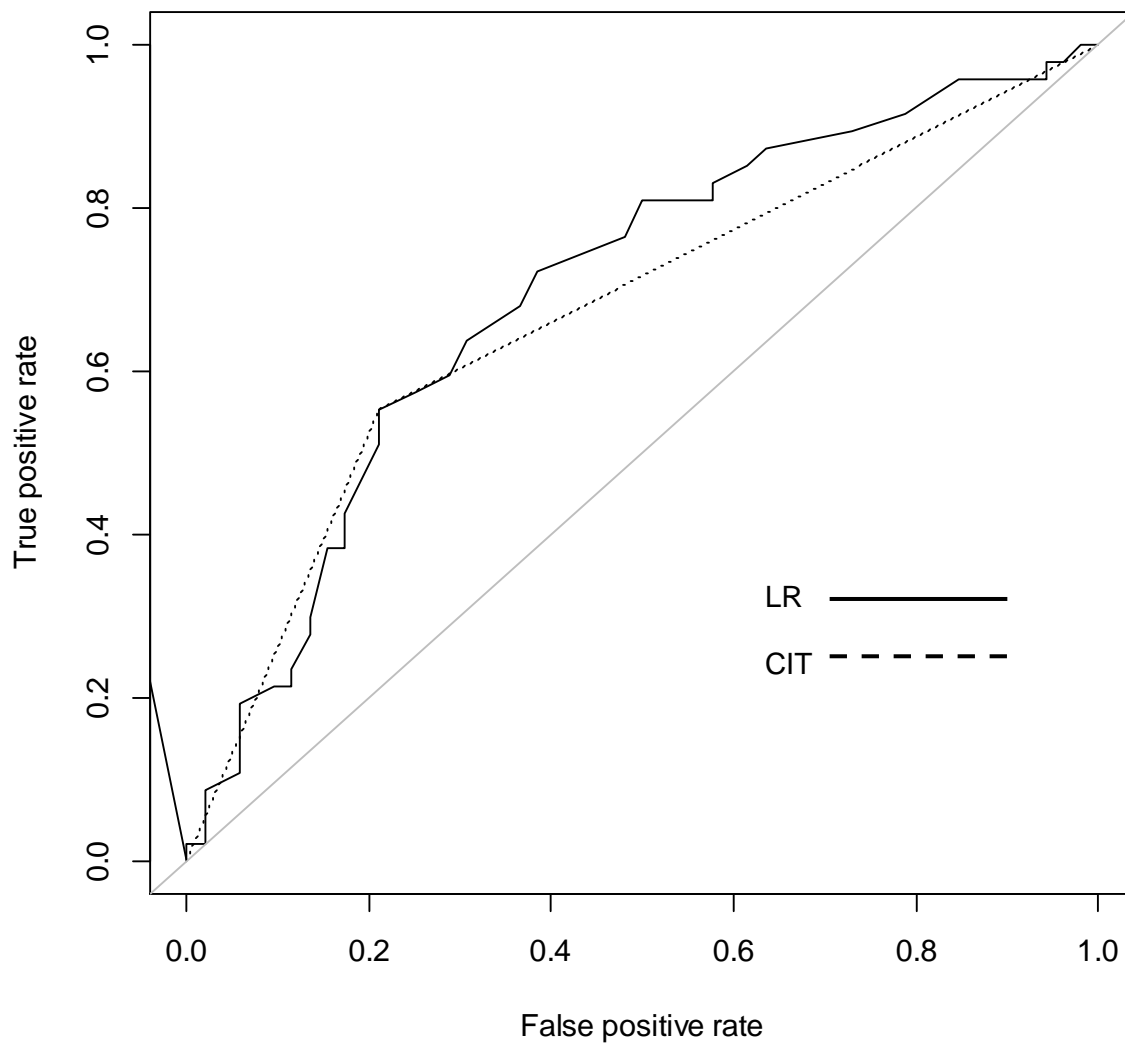


Figure 6 ROC Curves Comparing Predictive Performance Between the Conditional Inference Tree Model and the Logistic Regression Model.

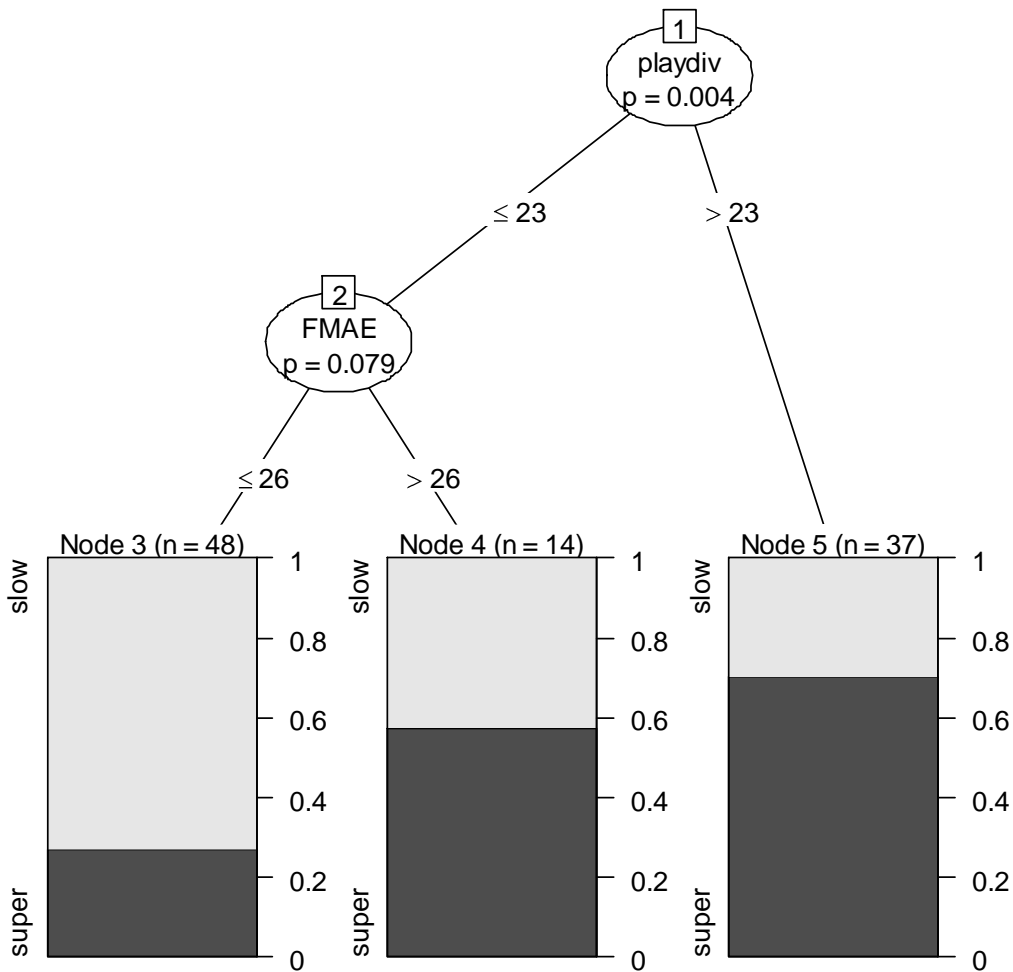


Figure 7 Conditional Inference Tree of Play Diversity and Fine Motor Predicting Super vs. Slow Responders, confidence level of .90.

References

- Abrahams, B. S., & Geschwind, D. H. (2008). Advances in autism genetics: on the threshold of a new neurobiology. *Nature Reviews Genetics*, 9(5), 341.
- Akshoomoff, N. (2006). Use of the Mullen Scales of Early Learning for the assessment of young children with autism spectrum disorders. *Child Neuropsychology*, 12(4-5), 269-277.
- Baio, J. (2012). Prevalence of Autism Spectrum Disorders: Autism and Developmental Disabilities Monitoring Network, 14 Sites, United States, 2008. Morbidity and Mortality Weekly Report. *Surveillance Summaries*. Volume 61, Number 3. Centers for Disease Control and Prevention.
- Baker-Ericzén, M. J., Stahmer, A. C., & Burns, A. (2007). Child demographics associated with outcomes in a community-based pivotal response training program. *Journal of Positive Behavior Interventions*, 9(1), 52-60.
- Bishop, S. L., Guthrie, W., Coffing, M., & Lord, C. (2011). Convergent validity of the Mullen Scales of Early Learning and the differential ability scales in children with autism spectrum disorders. *American journal on intellectual and developmental disabilities*, 116(5), 331-343.
- Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and regression trees*. CRC press.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32.
- Burns, T. G., King, T. Z., & Spencer, K. S. (2013). Mullen Scales of Early Learning: The utility in assessing children diagnosed with autism spectrum disorders, cerebral palsy, and epilepsy. *Applied Neuropsychology: Child*, 2(1), 33-42.

- Chambless, D. L., & Hollon, S. D. (1998). Defining empirically supported therapies. *Journal of consulting and clinical psychology*, 66(1), 7.
- Chang, Y. C., Shire, S. Y., Shih, W., Gelfand, C., & Kasari, C. (2016). Preschool Deployment of Evidence-Based Social Communication Intervention: JASPER in the Classroom. *Journal of autism and developmental disorders*, 46(6), 2211-2223.
- Chawarska, K., Shic, F., Macari, S., Campbell, D. J., Brian, J., Landa, R., ... & Young, G. S. (2014). 18-month predictors of later outcomes in younger siblings of children with autism spectrum disorder: a baby siblings research consortium study. *Journal of the American Academy of Child & Adolescent Psychiatry*, 53(12), 1317-1327.
- Coleman, M., & Gillberg, C. (2012). *The autisms*. OUP USA.
- Dawson, G., Rogers, S., Munson, J., Smith, M., Winter, J., Greenson, J., ... & Varley, J. (2010). Randomized, controlled trial of an intervention for toddlers with autism: the Early Start Denver Model. *Pediatrics*, 125(1), e17-e23.
- Dawson, G., Rogers, S., Munson, J., Smith, M., Winter, J., Greenson, J., ... & Varley, J. (2010). Randomized, controlled trial of an intervention for toddlers with autism: the Early Start Denver Model. *Pediatrics*, 125(1), e17-e23.
- Goods, K. S., Ishijima, E., Chang, Y. C., & Kasari, C. (2013). Preschool based JASPER intervention in minimally verbal children with autism: Pilot RCT. *Journal of autism and developmental disorders*, 43(5), 1050-1056.
- Freitas, T. C., Gabbard, C., Caçola, P., Montebelo, M. I., & Santos, D. C. (2013). Family socioeconomic status and the provision of motor affordances in the home. *Brazilian journal of physical therapy*, 17(4), 319-327.

- Goods, K. S., Ishijima, E., Chang, Y. C., & Kasari, C. (2013). Preschool based JASPER intervention in minimally verbal children with autism: Pilot RCT. *Journal of autism and developmental disorders*, 43(5), 1050-1056.
- Gordon, K., Pasco, G., McElduff, F., Wade, A., Howlin, P., & Charman, T. (2011). A communication-based intervention for nonverbal children with autism: What changes? Who benefits?. *Journal of consulting and clinical psychology*, 79(4), 447.
- Gresham, F. M., & MacMillan, D. L. (1998). Early intervention project: Can its claims be substantiated and its effects replicated?. *Journal of autism and developmental disorders*, 28(1), 5-13.
- Hallmayer, J., Cleveland, S., Torres, A., Phillips, J., Cohen, B., Torigoe, T., ... & Lotspeich, L. (2011). Genetic heritability and shared environmental factors among twin pairs with autism. *Archives of general psychiatry*, 68(11), 1095-1102.
- Harris, S. L., & Handleman, J. S. (2000). Age and IQ at intake as predictors of placement for young children with autism: A four-to six-year follow-up. *Journal of autism and developmental disorders*, 30(2), 137-142.
- Hothorn, T., Hornik, K., & Zeileis, A. (2006). Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical statistics*, 15(3), 651-674.
- Howlin, P., Magiati, I., & Charman, T. (2009). Systematic review of early intensive behavioral interventions for children with autism. *American journal on intellectual and developmental disabilities*, 114(1), 23-41.
- Hresko, W. P., Reid, D. K., & Hammill, D. D. (1981). *The test of early language development*. Austin, TX: Pro-ed.

- Ingersoll, B., Schreibman, L., & Stahmer, A. (2001). Brief report: Differential treatment outcomes for children with autistic spectrum disorder based on level of peer social avoidance. *Journal of Autism and Developmental Disorders*, 31(3), 343-349.
- Itzhak, E. B., & Zachor, D. A. (2011). Who benefits from early intervention in autism spectrum disorders?. *Research in Autism Spectrum Disorders*, 5(1), 345-350.
- Kanner, L. (1943). Autistic disturbances of affective contact.
- Kasari, C. (2002). Assessing change in early intervention programs for children with autism. *Journal of autism and developmental disorders*, 32(5), 447-461.
- Kasari, C., Freeman, S., & Paparella, T. (2006). Joint attention and symbolic play in young children with autism: A randomized controlled intervention study. *Journal of Child Psychology and Psychiatry*, 47(6), 611-620.
- Kasari, C., Paparella, T., Freeman, S., & Jahromi, L. B. (2008). Language outcome in autism: randomized comparison of joint attention and play interventions. *Journal of consulting and clinical psychology*, 76(1), 125.
- Kasari, C., Gulsrud, A. C., Wong, C., Kwon, S., & Locke, J. (2010). Randomized controlled caregiver mediated joint engagement intervention for toddlers with autism. *Journal of autism and developmental disorders*, 40(9), 1045-1056.
- Kasari, C., Gulsrud, A., Freeman, S., Paparella, T., & Hellemann, G. (2012). Longitudinal follow-up of children with autism receiving targeted interventions on joint attention and play. *Journal of the American Academy of Child & Adolescent Psychiatry*, 51(5), 487-495
- Kasari, C., Kaiser, A., Goods, K., Nietfeld, J., Mathy, P., Landa, R., ... & Almirall, D. (2014). Communication interventions for minimally verbal children with autism: A sequential

- multiple assignment randomized trial. *Journal of the American Academy of Child & Adolescent Psychiatry*, 53(6), 635-646.
- Kasari, C., Siller, M., Huynh, L. N., Shih, W., Swanson, M., Hellemann, G. S., & Sugar, C. A. (2014). Randomized controlled trial of parental responsiveness intervention for toddlers at high risk for autism. *Infant Behavior and Development*, 37(4), 711-721.
- Klintwall, L., Macari, S., Eikeseth, S., & Chawarska, K. (2014). Interest level in 2-year-olds with autism spectrum disorder predicts rate of verbal, nonverbal, and adaptive skill acquisition. *Autism*, 1362361314555376.
- Koegel, R. L., & Kern Koegel, L. (2006). *Pivotal Response Treatments for Autism: Communication, Social, and Academic Development*. Brookes Publishing Company. PO Box 10624, Baltimore, MD 21285.
- Lemon, S. C., Roy, J., Clark, M. A., Friedmann, P. D., & Rakowski, W. (2003). Classification and regression tree analysis in public health: methodological review and comparison with logistic regression. *Annals of behavioral medicine*, 26(3), 172-181.
- Lord, C., & McGee, J. P. (2001). *Educating children with autism*. Washington DC: National Academic Press.
- Lord, C., DiLavore, P. C., & Gotham, K. (2012). *Autism diagnostic observation schedule*. Western Psychological Services.
- Lord, C., Petkova, E., Hus, V., Gan, W., Lu, F., Martin, D. M., ... & Algermissen, M. (2012). A multisite study of the clinical diagnosis of different autism spectrum disorders. *Archives of general psychiatry*, 69(3), 306-313.
- Lovaas, O. I. (1987). Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal of consulting and clinical psychology*, 55(1), 3.

- Luiselli, J. K., Cannon, B. O. M., Ellis, J. T., & Sisson, R. W. (2000). Home-Based Behavioral Intervention for Young Children with Autism/Pervasive Developmental Disorder A Preliminary Evaluation of Outcome in Relation to Child Age and Intensity of Service Delivery. *Autism*, 4(4), 426-438.
- Magiati, I., Tay, X. W., & Howlin, P. (2014). Cognitive, language, social and behavioural outcomes in adults with autism spectrum disorders: a systematic review of longitudinal follow-up studies in adulthood. *Clinical Psychology Review*, 34(1), 73-86.
- Marshall, R. J. (2001). The use of classification and regression trees in clinical epidemiology. *Journal of clinical epidemiology*, 54(6), 603-609.
- Mullen, E. M. (1995). Mullen scales of early learning (pp. 58-64). Circle Pines, MN: AGS. Chicago
- Mundy, P., Delgado, C., Block, J., Venezia, M., Hogan, A., & Seibert, J. (2003). Early social communication scales (ESCS). Coral Gables, FL: University of Miami.
- Neeley, E. S., Bigler, E. D., Krasny, L., Ozonoff, S., McMahon, W., & Lainhart, J. E. (2007). Quantitative temporal lobe differences: autism distinguished from controls using classification and regression tree analysis. *Brain and Development*, 29(7), 389-399.
- Ozonoff, S., Young, G. S., Carter, A., Messinger, D., Yirmiya, N., Zwaigenbaum, L., ... & Hutman, T. (2011). Recurrence risk for autism spectrum disorders: a Baby Siblings Research Consortium study. *Pediatrics*, 128(3), e488-e495.
- Paparella, T., Goods, K. S., Freeman, S., & Kasari, C. (2011). The emergence of nonverbal joint attention and requesting skills in young children with autism. *Journal of communication disorders*, 44(6), 569-583.

- Paul, R., Campbell, D., Gilbert, K., & Tsiouri, I. (2013). Comparing spoken language treatments for minimally verbal preschoolers with autism spectrum disorders. *Journal of autism and developmental disorders*, 43(2), 418-431
- R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
- Rimland, B. (1978). Savant capabilities of autistic children and their cognitive implications.
- Ronald, A., Happe, F., Bolton, P., Butcher, L. M., Price, T. S., Wheelwright, S., ... & Plomin, R. (2006). Genetic heterogeneity between the three components of the autism spectrum: a twin study. *Journal of the American Academy of Child & Adolescent Psychiatry*, 45(6), 691-699.
- Rosenberg, R. E., Law, J. K., Yenokyan, G., McGready, J., Kaufmann, W. E., & Law, P. A. (2009). Characteristics and concordance of autism spectrum disorders among 277 twin pairs. *Archives of pediatrics & adolescent medicine*, 163(10), 907-914.
- Sallows, G. O., & Graupner, T. D. (2005). Intensive behavioral treatment for children with autism: Four-year outcome and predictors. *American Journal on Mental Retardation*, 110(6), 417-438.
- Schreibman, L., Dufek, S., & Cunningham, A. B. (2011). Identifying moderators of treatment outcome for children with autism. In *International handbook of autism and pervasive developmental disorders* (pp. 295-305). Springer New York.
- Sheinkopf, S. J., & Siegel, B. (1998). Home-based behavioral treatment of young children with autism. *Journal of autism and developmental disorders*, 28(1), 15-23.

- Sherer, M. R., & Schreibman, L. (2005). Individual behavioral profiles and predictors of treatment effectiveness for children with autism. *Journal of consulting and clinical psychology, 73*(3), 525.
- Shih, W., Patterson, S. Y., & Kasari, C. (2014). Developing an adaptive treatment strategy for peer-related social skills for children with autism spectrum disorders. *Journal of Clinical Child & Adolescent Psychology, 1-11*.
- Skafidas, E., Testa, R., Zantomio, D., Chana, G., Everall, I. P., & Pantelis, C. (2012). Predicting the diagnosis of autism spectrum disorder using gene pathway analysis. *Mol Psychiatry, 19*, 504-510.
- Smith, T., Scahill, L., Dawson, G., Guthrie, D., Lord, C., Odom, S., ... & Wagner, A. (2007). Designing research studies on psychosocial interventions in autism. *Journal of autism and developmental disorders, 37*(2), 354-366.
- Sparrow, S. S., Cicchetti, D. V., & Balla, D. A. (2005). Vineland adaptive behavior scales: (Vineland II), survey interview form/caregiver rating form. Livonia, MN: Pearson Assessments.
- Stahl, D., Pickles, A., Elsabbagh, M., Johnson, M. H., & BASIS Team. (2012). Novel machine learning methods for ERP analysis: a validation from research on infants at risk for autism. *Developmental neuropsychology, 37*(3), 274-298.
- Strobl, C., Boulesteix, A. L., Zeileis, A., & Hothorn, T. (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC bioinformatics, 8*(1), 25.
- Strobl, C., Boulesteix, A. L., Kneib, T., Augustin, T., & Zeileis, A. (2008). Conditional variable importance for random forests. *BMC bioinformatics, 9*(1), 307.

- Sumi, S., Taniai, H., Miyachi, T., & Tanemura, M. (2006). Sibling risk of pervasive developmental disorder estimated by means of an epidemiologic survey in Nagoya, Japan. *Journal of human genetics*, 51(6), 518-522.
- Tager-Flusberg, H., Paul, R., & Lord, C. (2005). Language and communication in autism. *Handbook of autism and pervasive developmental disorders*, 1, 335-364.
- Tager-Flusberg, H., & Kasari, C. (2013). Minimally verbal school-aged children with autism spectrum disorder: the neglected end of the spectrum. *Autism Research*, 6(6), 468-478.
- Taniai, H., Nishiyama, T., Miyachi, T., Imaeda, M., & Sumi, S. (2008). Genetic influences on the broad spectrum of autism: Study of proband-ascertained twins. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 147(6), 844-849.
- Tomasello, M., & Farrar, M. J. (1986). Joint attention and early language. *Child development*, 1454-1463.
- Ungerer, J. A., & Sigman, M. (1981). Symbolic play and language comprehension in autistic children. *Journal of the American Academy of Child Psychiatry*, 20(2), 318-337.
- Virués-Ortega, J. (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.
- Vivanti, G., Dissanayake, C., Zierhut, C., Rogers, S. J., & Victorian ASELCC Team. (2013). Brief report: Predictors of outcomes in the Early Start Denver Model delivered in a group setting. *Journal of Autism and Developmental Disorders*, 43(7), 1717-1724.
- Warren, Z., McPheeters, M. L., Sathe, N., Foss-Feig, J. H., Glasser, A., & Veenstra-VanderWeele, J. (2011). A systematic review of early intensive intervention for autism spectrum disorders. *Pediatrics*, 127(5), e1303-e1311.

- Wall, D. P., Dally, R., Luyster, R., Jung, J. Y., & DeLuca, T. F. (2012). Use of artificial intelligence to shorten the behavioral diagnosis of autism. *PloS one*, 7(8), e43855.
- Wall, D. P., Kosmicki, J., Deluca, T. F., Harstad, E., & Fusaro, V. A. (2012). Use of machine learning to shorten observation-based screening and diagnosis of autism. *Translational psychiatry*, 2(4), e100.
- Wong, C., Odom, S. L., Hume, K., Cox, A. W., Fettig, A., Kucharczyk, S., & Schultz, T. R. (2013). Evidence-based practices for children, youth, and young adults with autism spectrum disorder. Chapel Hill: The University of North Carolina.
- Yoder, P., & Stone, W. L. (2006). A randomized comparison of the effect of two prelinguistic communication interventions on the acquisition of spoken communication in preschoolers with ASD. *Journal of Speech, Language, and Hearing Research*, 49(4), 698-711.
- Yoder, P., & Stone, W. L. (2006). Randomized comparison of two communication interventions for preschoolers with autism spectrum disorders. *Journal of consulting and clinical psychology*, 74(3), 426.
- Zablotsky, B., Black, L. I., Maenner, M. J., Schieve, L. A., & Blumberg, S. J. (2015). Estimated Prevalence of Autism and Other Developmental Disabilities Following Questionnaire Changes in the 2014 National Health Interview Survey. *National health statistics reports*, (87), 1-21.
- Zwaigenbaum, L., Bauman, M. L., Choueiri, R., Kasari, C., Carter, A., Granpeesheh, D., ... & Pierce, K. (2015). Early intervention for children with autism spectrum disorder under 3 years of age: recommendations for practice and research. *Pediatrics*, 136(Supplement 1), S60-S81.