Examination of the Factor Structure of the Autism Diagnostic Interview-Revised (ADI-R) within a Simplex Population

---

A Dissertation Presented to

The Faculty of the Curry School of Education

University of Virginia

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

---

by

Jordan L. Wade, M.Ed.

August 2016

Department of Human Services
Curry School of Education
University of Virginia
Charlottesville, Virginia

APPROVAL OF THE DISSERTATION

This dissertation, “Examination of the Factor Structure of the Autism Diagnostic Interview-Revised (ADI-R) within a Simplex Population”, has been approved by the Graduate Faculty of the Curry School of Education in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

____
Dr. Ronald E. Reeve (Chair)

____
Dr. Timothy Konold

____
Dr. Jane Hilton

____
Dr. Peter Patrick

______________________Date

DEDICATION

This work is dedicated to my husband, Kevin. Without his support, encouragement, and patience, I would not have been able to complete this research. Throughout my time at Curry, he has kept me grounded and provided both mental and emotional support, giving me the strength to finish.

ACKNOWLEDGEMENTS

This work would not be possible without the support and guidance of my advisor, Dr. Ronald Reeve. Thank you for your mentorship and wisdom these past four years. Your patience and encouragement fostered my development as a clinician and a researcher, for which I will always be grateful.

Thank you to my dissertation committee, Timothy Konold, Jane Hilton, and Peter Patrick. Your guidance has been invaluable throughout this process.

I am forever indebted to the Curry School of Education for funding my training and research while at the University of Virginia. The financial support of the Curry School has made this dissertation possible.

To my fellow lab members, Tiffany Torigoe, Stephany Cox, and Neill Broderick: I am incredibly grateful to have had the opportunity to work alongside each of you. I continue to be astounded by your breadth of knowledge and clinical expertise related to children with developmental disabilities. More than fantastic colleagues, you are all amazing friends. Thank you for all of your support throughout my time at Curry.

Lastly, I would like to thank my family: my husband, Kevin, my son, Atticus, and my parents, Len and Rene Horton. To Kevin and Atticus: thank you for your seemingly endless patience and for bringing me incredible happiness even on the most frustrating of days. To my parents, thank you for your unwavering support and encouragement. You always believed in me, even, and especially, when my confidence wavered.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Study 2

Study 3

Overview of Three Manuscript Dissertation

This line of research utilized data from the Simons Simplex Collection (SSC), a large repository including both genetic and phenotypic data on simplex families (i.e., one individual carries an ASD diagnosis – all other family members are unaffected). The three papers described within this document utilized phenotypic data, with the goal of gaining greater clarity regarding the use of parental maladaptive coping mechanisms, specifically substance use, as related to both parental and child characteristics, in addition to examining the factor structure of several widely used diagnostic measures. This dissertation adheres to the parameters set forth by the Curry School of Education Guidelines for Manuscript Style Dissertations. As required by the guidelines, I am the lead author on the first and third papers, and contributed substantially to the second paper as the second author. The first paper, *Impact of Child Problem Behaviors and Parental Broad Autism Phenotype Traits on Substance Use among Parents of Children with ASD*, was published in the Journal of Autism and Developmental Disorders (Wade, Cox, Reeve, & Hull, 2014). The second study, *Model Invariance across Genders of the Broad Autism Phenotype Questionnaire*, was also published in the Journal of Autism and Developmental Disorders (Broderick, Wade, Meyer, Hull, & Reeve, 2015). Springer, the publisher, provided written permission for both articles to be submitted to Libra. The third study, *Examination of the Factor Structure of the Autism Diagnostic Interview-Revised (ADI-R) within a Simplex*

Population, is being prepared for submission to the Journal of Autism and Developmental Disorders.

Linking Document:

Examination of the Factor Structure of the Autism Diagnostic Interview-Revised (ADI-R) within a Simplex Population

This line of research affords greater clarity regarding the phenotypic traits of simplex families (i.e., child carries an ASD diagnosis but no first- through third-degree relatives are affected), by utilizing data from the Simons Simplex Collection (SSC). Parental wellbeing and their role in the diagnostic process were of particular interest. The first study examined parental maladaptive coping mechanisms, specifically substance use, whereas the last two papers explored the factor structure of widely used diagnostic measures, including the Broad Autism Phenotype Questionnaire (BAPQ; Piven et al., 1997) and the Autism Diagnostic Interview-Revised (ADI-R; Le Couteur, Lord, & Rutter, 2003; Lord, Rutter, & Le Couteur, 1994), both of which were completed by parents in the SSC sample.

The demands associated with having a child with ASD are multifaceted and can affect parental wellbeing, as well as that of other family members. Research suggests that this group of parents report lower marital happiness, poor family cohesion, and poor adaptability (Higgins, Bailey, & Pearce, 2005). Given these findings, it is critical that researchers and clinicians seek to understand the coping mechanisms utilized by parents in order to inform treatments that support and strengthen the entire family unit.

The first paper, *Impact of Child Problem Behaviors and Parental Broad Autism Phenotype Traits on Substance Use Among Parents of Children with ASD* (Wade, Cox, Reeve, & Hull, 2014), explored factors related to parental endorsement of maladaptive coping mechanisms, specifically substance use, including the impact of child externalizing behavior and parental broad autism phenotype (BAP) traits. Although there is substantial research on the BAP as well as the effects of child externalizing behaviors on parents, there is little on parental substance use as related to parents of individuals with ASD. Results indicated that child externalizing behaviors and BAP traits were predictors of substance use, although the specific factors and substances varied among mothers and fathers. For both parents, child externalizing behaviors predicted tobacco use, whereas parental rigidity increased risk of tobacco use for fathers but not mothers. Additionally, among mothers, child externalizing behaviors increased risk of illegal substance use, whereas maternal rigidity decreased risk of alcohol use. These findings suggest that clinicians should be watchful of parental substance use, particularly if a child is exhibiting externalizing behaviors or the parents endorse BAP traits.

In order to better understand the BAP, the author worked closely with a colleague in studying a commonly utilized assessment administered to capture the BAP – the Broad Autism Phenotype Questionnaire (BAPQ; Piven et al., 1997). The BAP refers to traits that are similar, but milder, to those associated with ASD that are present in undiagnosed relatives of individuals on the spectrum (Sucksmith, Roth, & Hoekstra, 2011). The BAP grants researchers the opportunity to better understand the genetic and phenotypic basis of ASD. This study, *Model Invariance across Genders of the Broad Autism Phenotype Questionnaire* (Broderick, Wade, Meyer, Hull, & Reeve, 2015), which was published in the *Journal of Autism and Developmental Disorders*, conducted a series of confirmatory factor analyses to determine whether or not the factors were upheld across genders. Model invariance was upheld at each level of parameter constraint; however, model fit indices suggested limited goodness-of-fit between the proposed model and the sample. Exploratory analyses were conducted to investigate alternate factor structure model; the proposed three-factor structure model was ultimately supported, as it was the most parsimonious. However, limited goodness-of-fit suggests that the measure is capturing other latent variables that are not accounted for by the three-factor model. These findings suggest that while model invariance was upheld across genders, additional research examining the factor structure of the BAPQ (Piven et al., 1997) is needed.

The BAPQ is of particular interest in its potential to increase understanding of the broader spectrum. Other measures have been developed with the intent of collecting a detailed developmental history, which is a crucial component of a comprehensive evaluation (Ozonoff, Goodlin-Jones, & Solomon, 2005). This information is often obtained through parental report. Although parents provide critical information, there is literature that suggests that parental bias may yield an incomplete picture (Chawarska, Klin, Paul, & Volkmar, 2006; Mildenberger, Sitter, Noterdaeme, & Amorosa, 2001; Noterdaeme, Mildenberger, Sitter, & Amorosa, 2002). Given the importance of parents in the diagnostic process and the need to better understand the measures commonly utilized in the assessment of ASD, the culmination of this line of research is the third paper, *Examination of the Factor Structure of the Autism Diagnostic Interview-Revised (ADI-R) within a Simplex Population*.

The ADI-R (Le Couteur et al., 2003; Lord et al., 1994) is a comprehensive interview administered to parents/caregivers of individuals with ASD and is considered, in tandem with the ADOS (Lord et al., 2000; Lord, Rutter, DiLavore, & Risi, 2001), the "gold standard" of ASD assessment tools (Ozonoff et al., 2005). Within the literature, the factor structure of commonly used diagnostic tools is frequently employed in an effort to more precisely define the underlying symptomatology of ASD (Norris, Lecavalier, & Edwards, 2012). Findings suggest that the factor structure based on DSM-IV-TR diagnostic criteria does not provide the best fit, instead proposing two-factor models as an alternative (Frazier et al., 2008; Snow et al., 2009). However, much of the existing research has utilized rather small sample sizes, thus constituting a need for further analysis of the factor structure of the ADI-R using larger samples (Frazier et al., 2008). The third paper within this document utilized the Simons Simplex Collection, which includes data on over 2,500 children and adolescents, to provide further information regarding the factor structure of the ADI-R within a large, simplex population.

In addition to providing further insight into the factor structure of the ADI-R, the third study utilized multiple-group analyses to better understand the interrelationship between the ADI-R and the ADOS. The two measures were originally developed for research purposes, with the intention of being administered together (Le Couteur et al., 2003; Lord et al., 2000; Lord, et al., 2001; Lord et al., 1994). However, despite this intent, there are few studies directly comparing the measures (Falkmer, Anderson, Falkmer, & Horlin, 2013). To address this gap, the third study utilized multiple-group analyses to study whether the ADI-R subdomains demonstrated invariance across ADOS modules 2 and 3. These findings provided insight into the interrelationship of the ADI-R and the ADOS, as well as offering additional support for model fit differing from that proposed by the DSM-IV-TR.

References

Broderick, N., Wade, J.L., Meyer, J.P., Hull, M., & Reeve, R.E. (2015). Model invariance across genders of the Broad Autism Phenotype Questionnaire. *Journal of Autism and Developmental Disorders*, *45*(10), 3133-3147.

Chawarska, K., Klin, A., Paul, R., & Volkmar, F. (2007). Autism spectrum disorder in the second year: Stability and change in syndrome expression. *Journal of Child Psychology and Psychiatry*, *48*(2), 128-138.

Falkmer, T., Anderson, K., Falkmer, M., & Horlin, C. (2013). Diagnostic procedures in autism spectrum disorders: A systematic literature review. *European Child and Adolescent Psychiatry*, *22*, 329-340.

Frazier, T. W., Youngstrom, E. A., Kubu, C. S., Sinclair, L., & Rezai, A. (2008). Exploratory and confirmatory factor analysis of the Autism Diagnostic Interview-Revised. *Journal of Autism and Developmental Disorders*, *38*(3), 474-480.

Higgins, D.J., Bailey, S.R., & Pearce, J.C. (2005). Factors associated with functioning style and coping strategies of families with a child with an autism spectrum disorder. *Autism*, *9*(2), 125-137.

Le Couteur, A., Lord, C., & Rutter, M. (2003). *The Autism Diagnostic Interview – Revised* (ADI-R). Los Angeles, CA: Western Psychological Services.

Lord, C., Risi, S., Lambrecht, L., Cook, H.E., Leventhal, B.L., DiLavore, P.C. … Rutter, M. (2000). The Autism Diagnostic Observation Schedule – Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*, *30*, 205-223.

Lord, C., Rutter, M., DiLavore, P.D., & Risi, S. (2001). *Autism Diagnostic Observation Schedule* (ADOS). Los Angeles, CA: Western Psychological Services.

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

Mildenberger, K., Sitter, S., Noterdaeme, M., & Amorosa, H. (2001). The use of the ADI-R as a diagnostic tool in the differential diagnosis of children with infantile autism and children with a receptive language disorder. *European Child & Adolescent Psychiatry*, *10*, 248-255.

Norris, M., Lecavalier, L., Edwards, M.C. (2012). The structure of autism symptoms as measured by the Autism Diagnostic Observation Schedule. *Journal of Autism and Developmental Disorders, 42* (6), 1075-1086.

Noterdaeme, M., Mildenberger, K., Sitter, S., & Amorosa, H. (2002). Parent information and direct observation in the diagnosis of pervasive and specific developmental disorders. *Autism*, *6*(2), 159-168.

Ozonoff, S., Goodlin-Jones, B.L., & Solomon, M. (2005). Evidence-based assessment of autism spectrum disorders in children and adolescents. *Journal of Clinical Child and Adolescent Psychology*, *34*(3), 523-540.

Piven, J., Palmer, P., Jacobi, D., Childress, D., & Arndt, S. (1997). Broader autism phenotype: Evidence from a family history study of multiple-incidence families. *American Journal of Psychiatry, 154*(2), 185-190.

Snow, A.V., Lecavalier, L., & Houts, C. (2009). The structure of the Autism Diagnostic Interview – Revised: Diagnostic and phenotypic implications. *Journal of Child Psychology and Psychiatry*, *50*(6), 734-742.

Sucksmith, E., Roth, I., Hoekstra, R.A. (2011). Autistic traits below the clinical threshold: Re-examining the Broader Autism Phenotype in the 21st century. *Neuropsychology Review*, *21*, 360-389.

Wade, J., Cox, N.B., Reeve, R.E., & Hull, M. (2014). Brief report: Impact of child problem behaviors and parental Broad Autism Phenotype traits on substance use among parents of children with ASD. *Journal of Autism and Developmental Disorders*, *44*(10), 2621-2627.

Impact of Child Problem Behaviors and Parental Broad Autism Phenotype Traits on Substance Use among Parents of Children with ASD

Jordan L. Wade, Neill B. Cox, Ronald E. Reeve, & Michael Hull

University of Virginia

Published in *Journal of Autism and Developmental Disorders* (2014)

Wade, J., Cox, N.B., Reeve, R.E., & Hull, M. (2014). Brief report: Impact of child problem behaviors and parental Broad Autism Phenotype traits on substance use among parents of children with ASD. *Journal of Autism and Developmental Disorders, 44*(10), 2621-2627.

Abstract

Using data from the Simons Simplex Collection, the present study examined the impact of child externalizing behavior and parental broad autism phenotype traits on substance use among parents of children with Autism Spectrum Disorder (n = 2,388). For both fathers and mothers, child externalizing behaviors predicted tobacco use (OR = 1.01 and OR = 1.02, respectively), whereas rigidity increased risk of tobacco use for fathers (OR = 1.29) but not mothers. Additionally, among mothers, child externalizing behaviors increased risk of illegal substance use (OR = 1.04), whereas maternal rigidity decreased risk of alcohol use (OR = .83). Collectively, results suggest that child externalizing behaviors and parental rigidity may have differing impacts on the types of substances used by parents.

Impact of Child Problem Behaviors and Parental Broad Autism Phenotype Traits on Substance Use among Parents of Children with ASD

Children with Autism Spectrum Disorder (ASD) often exhibit problem behaviors secondary to, or exacerbating, the core symptoms of the syndrome (Centers for Disease Control and Prevention [CDC], 2012a). Research indicates that the severity of child problem behavior is correlated to parental stress (Beck, 2003; Davis & Carter, 2008; Higgins, 2005; Ingersoll & Hambrick, 2011; Ingersoll, Meyer, & Becker, 2011) and depressive symptoms (Benson, 2006; Benson & Karlof, 2009; Davis & Carter, 2008; Ingersoll & Hambrick, 2011; Ingersoll, Meyer, & Becker, 2011). The literature also indicates that alcohol use is one way of coping with children who exhibit challenging behaviors (Pelham et al., 1997). Additionally, research suggests that first- and second-degree relatives of children with ASD exhibit greater rates of alcohol use (DeLong & Dwyer, 1998; Lobascher et al., 1970; Miles, Takahashi, Haber, & Hadden, 2003; Piven et al., 1991; Smalley, McCracken, & Tanguay 1995) and that parents who abuse substances report higher stress and more negative views of their children's emotional and behavioral functioning (Killeen & Brady, 2000). Finally, the presence of traits such as aloofness, rigidity, and pragmatic language difficulties (BAP traits) may increase vulnerability to stress and internalizing mood symptoms in parents of children with ASD (Ingersoll et al., 2011).

Child Externalizing Behaviors

Davis and Carter (2008) reported that a substantial portion of parents of toddlers with ASD endorsed clinically significant levels of stress and depressive symptoms. Externalizing behaviors (e.g., hitting, kicking, biting) were correlated with higher stress levels among fathers, whereas a lack of self-regulation skills (e.g., eating, sleeping, emotion regulation), were associated with higher stress among mothers. Other studies looking specifically at mothers have found that severity of problem behavior is related to maternal stress level and depressive symptoms (Beck, 2003; Benson & Karlof, 2009; Ingersoll & Hambrick, 2011). Overall, the literature suggests that child externalizing behaviors are correlated to parental stress levels and depressive symptomatology.

Parental Mental Health and Substance Use

The limited research that exists suggests that first- and second-degree relatives of individuals with ASD demonstrate higher incidence rates of alcoholism (DeLong & Dwyer, 1998; Lobascher et al., 1970; Miles, Takahashi, Haber, & Hadden, 2003; Piven et al., 1991; Smalley, McCracken, & Tanguay 1995). Among parents of children with ASD, Lobascher et al. (1970) found higher rates of alcoholism and psychiatric illness when compared to parents of typically developing children. Piven et al. (1991) reported that the lifetime prevalence rate of anxiety disorders was statistically greater among parents of children with ASD when compared to parents of children with Down syndrome (23.5% versus 2.9%). Additionally, results identified a trend towards greater rates of alcoholism (12.3% versus 0%) and major depressive disorder (27.2% versus 14.8%) in parents of children with ASD (Piven et al., 1991). More recent studies have found similar results, documenting increased rates of alcoholism among first- and second-degree relatives of individuals with ASD when compared to those of Down syndrome families (Miles et al., 2003). Taken together, these results suggest that parents of children with ASD may be at increased risk of psychiatric illness and alcohol use.

Parental substance use is especially concerning because research suggests that it may alter the way that parents perceive their children's emotional and behavioral problems (Killeen & Brady, 2000). Killeen and Brady (2000) found that mothers entering a residential treatment facility for alcohol and drug abuse reported higher levels of emotional and behavioral problems in their children at entry to the program than at six and 12 months post-discharge. These findings indicate that mothers' perceptions of their children's behavioral and emotional functioning may have improved as a function of substance use treatment, although the extent to which this finding reflects changes in parenting practices, perceptions, and/or child behaviors is less clear.

Broad Autism Phenotype

Studying the relatives of individuals diagnosed with ASD allows for better understanding of the genetic underpinnings of the disorder. Twin and family studies of ASD have demonstrated significant recurrence and heritability rates, supporting the notion that the disorder has a strong genetic basis (Bailey et al., 1995; Folstein & Rutter, 1977; Ozonoff et al., 2011). Folstein and Rutter (1977) were the first to report that some family members of individuals with ASD demonstrated similar traits. The term broad autism phenotype (BAP) emerged to define the set of personality and language characteristics that are similar to, but milder than, traits found in individuals diagnosed with ASD (Bailey et al., 1995; Bolton et al., 1994; Losh, et al., 2008; Murphy et al., 2000; Piven et al., 1997a; Piven et al., 1997b).

The literature clearly supports the concept of the BAP, leading to the emergence of studies examining the functioning of parents endorsing BAP traits. Ingersoll and

Hambrick (2011) found that parents of children with ASD with higher scores on the BAP were more likely to report maladaptive coping strategies, less social support, and greater depressive symptomatology. Moreover, Ingersoll et al. (2011) suggest that the confluence of exhibiting BAP traits and having a child with ASD may increase susceptibility to depression. However, other studies have not found a relationship between affective disorders and BAP traits in relatives of children with ASD (Bolton, Pickles, Murphy, & Rutter, 1998), demonstrating the need for further research investigating the impact of these traits on vulnerability to psychiatric disorders and stress.

Research Questions

Based on the current literature, parents of children with ASD who exhibit BAP traits and have children with elevated externalizing behavior problems may be particularly susceptible to detrimental levels of stress, and thus maladaptive coping mechanisms, including substance use. However, there is a paucity of research regarding substance use within this population. As such, this study examined two interrelated research questions. First, what is the prevalence of substance use among parents of children with ASD? Second, do child behavior problems and parental BAP characteristics confer increased risk for parental substance use? Given prior research, we hypothesized that parents of children with ASD would endorse higher rates of alcohol use when compared to national estimates. No hypothesis was offered with regards to tobacco and illegal drug use due to the lack of research in this area. In regard to the second question, we hypothesized that severity of child externalizing behavior and presence of broad autism phenotype traits in parents would be related to parental substance use.

Methods

Participants

The current study utilized data from the Simons Simplex Collection (SSC) Version 14, a project of the Simons Foundation Autism Research Initiative (SFARI). Permission to use the data was obtained from SFARI, and the Institutional Review Board at the University of Virginia approved the study. The SSC is a nationwide sample of 2,643 simplex families who have one child between the ages of four and 18 with an ASD diagnosis. To confirm the diagnosis, both the Autism Diagnostic Observation Schedule (ADOS; Lord, Rutter, DiLavore, & Risi, 1999) and the Autism Diagnostic Interview-Revised (ADI-R; Rutter, Le Couteur, & Lord, 2003), two of the leading autism diagnostic tools, were administered to each proband. Children with first- through third-degree relatives with an ASD diagnosis were excluded, and the participation of both biological parents was required for study inclusion. Data for the SCC were collected through 12 university-affiliated research clinics located throughout the United States. More specific information regarding SSC data collection can be found in a previous article (Fischbach & Lord, 2010).

Due to missing data, the sample for the current study contained 2,388 of the 2,643 probands and their families. Of the 2,388 families, 1,961 (82.1%) were quads (comprised of mother, father, proband, and an unaffected sibling) and 427 (17.9%) were triads (comprised of mother, father, and proband). Most probands were male (86.9%), which is highly similar to the 5:1 ratio reported by the CDC (2012b). Based on SSC diagnostic criteria (a compilation of ADOS, ADI-R, and the clinician's Best Estimate diagnosis), probands in the sample represented the following diagnoses: 90.3% Autism, 7.7% Autism Spectrum Disorder, and 2.0% Asperger's Disorder. On average, the fathers in the sample were 42.5 years old (SD = 6.4), mothers were 40.4 years old (SD = 5.7), and probands were 9.0 years old (SD = 3.6). Racial and ethnic composition of the sample closely matched United States Census data (Table 1), with the exception of an underrepresentation of African Americans (3.9% of probands compared to 13.1% of the U.S. population) and Hispanics (11.3% of probands compared to 16.7% of U.S. population) (United States Department of Commerce, 2013). Ninety percent of parents in the sample reported being married (Table 2), and 72.6% reported household incomes of $66,000 or higher (Table 3), which is somewhat higher than the $52,762 median household income reported by the United States Department of Commerce (2013).

Measures

Child Behavior Checklist. Severity of child problem behavior was assessed using the Externalizing subscale of the Child Behavior Checklist (CBCL). The CBCL (Achenbach, 1991) is a parent-report questionnaire that measures a child's emotional and behavioral functioning. The CBCL has been used with multiple populations, including children with ASD (Hartley, Sikora, & McCoy, 2008). As reported in previous literature, the internal consistency of the Externalizing score falls in the .92 to .96 range (Achenbach, 1991). The mean CBCL Externalizing Problem T-score for this sample was 56.49 (SD = 10.59; range: 32-97), which falls within the non-clinical range.

Broad Autism Phenotype Questionnaire (BAPQ). The Broad Autism Phenotype Questionnaire (BAPQ) was used to assess parental BAP traits. The BAPQ is a

36-item self-report questionnaire that is intended to measure the presence of broad personality and language characteristics associated with the autism spectrum, also referred to as the BAP (Piven et al., 1997a). Although both self- and informant-report versions of the BAPQ exist, only self-reports were gathered by the SSC (e.g., mothers were asked to respond about themselves and fathers were asked to respond about themselves).

Participants answer questions based on the frequency with which a statement applies to them, and responses are measured on a six-point Likert scale ranging from one (“very rarely”) to six (“very often”). Three subscales are attained (Aloof, Rigid, and Pragmatic Language), in addition to a Total score. Hurley, Losh, Parlier, Reznick, and Piven (2007) found that the BAPQ has good inter-item reliability; Cronbach’s alpha coefficient is .94 for the Aloof subscale, .91 for the Rigid subscale, and .85 for the Pragmatic Language subscale, and inter-item reliability for all 36 items is .95.

For comparison purposes, Hurley and colleagues (2007) calculated the following cut-off scores as indicators of significant BAP traits: for males, Aloof = 3.25, Rigid = 3.65, Pragmatic Language = 2.95, and Total = 3.35, and for females, Aloof = 3.00, Rigid = 3.25, Pragmatic Language = 2.70, and Total = 3.25 (Table 4). Within the sample used for this study, the following means were found for the maternal BAPQ scores: Aloof = 2.38 (SD = .78), Rigid = 2.67 (SD = .72), Pragmatic Language = 2.10 (SD = .60), and Total = 2.38 (SD = .57). The means for the paternal BAPQ scores were as follows: Aloof = 2.82 (SD = .86), Rigid = 2.88 (SD = .74), Pragmatic Language = 2.34 (SD = .64), and Total = 2.68 (SD = .60).

SSC Parent Substance Use History Form. The Parent Substance Use History Form, used to measure parental substance use, is a self-report questionnaire that assesses current and past use of tobacco, alcohol, amphetamines, cocaine, ecstasy, heroin, hallucinogens, inhalants, marijuana, and methadone. All of the questions are based on past and current substance use and are answered with a "yes" or "no." For this study, current substance use data were utilized; data pertaining to past substance use were not analyzed. Additionally, few mothers and fathers endorsed using illegal substances; thus, these categories were combined to form one variable – Illegal Substances. Tobacco and alcohol were kept as distinct categories. Among mothers, 62.5% (n = 1,493) endorsed alcohol use, 8.9% (n = 212) indicated tobacco use, and 2.6% (n = 63) endorsed current illegal substance use. Among fathers, 70.6% (n = 1,687) indicated that they drink alcohol, 17.8% (n = 426) endorsed using tobacco products, and 4.9% (n = 117) indicated current illegal substance use.

Results

Analyses

Descriptive statistics and binary logistic regression were used to analyze the data. Six logistic regression models were tested, three for mothers and three for fathers. In each logistic regression, substance use (yes, no) was the dependent variable. The predictors included the BAPQ subscales (Aloof, Rigid, and Pragmatic Language) and the CBCL Externalizing composite score. To determine differences among the groups regarding annual household income, race, and ethnicity, chi-square tests were analyzed. Differences in income, race, and ethnicity were significant among both mothers and fathers using alcohol (Table 5). Differences among income were significant among fathers and mothers using tobacco, as well as fathers using illegal substances. Among mothers using illegal substances, no differences among the variables were found. The variables that were found to be significant were entered into each model in the first step as predictors; thus, the variables in the first block varied by model. The variables of interest (BAPQ subscale scores and CBCL Externalizing score) were then entered in step two for all models (see Impact of Child Problem Behaviors and the BAP on Parental Substance Use section).

Prevalence of Substance Use

According to Schiller, Lucas, and Peregoy (2012), in 2011, 19% of adults, categorized as 18 or older, in the United States reported current cigarette use. When analyzed by gender, 21% of men and 17% of women classified themselves as current smokers (Table 6). Within the SSC sample, both mothers and fathers reported lower tobacco prevalence rates than the national average (8.9% and 17.8%, respectively). In regard to alcohol consumption, according to Schiller and colleagues (2012), 52% of adults reported regular drinking. When analyzed by gender, 60% of males and 44% of females classified themselves as regular drinkers (Schiller et al., 2012). The SSC sample reported higher rates of alcohol use among both mothers and fathers (62.5% and 70.6%, respectively); however, no data regarding frequency of alcohol consumption were available, making it impossible to determine whether the SSC sample were regular or infrequent drinkers. Lastly, per National Center for Health Statistics (2012) data, 8.9% of individuals aged twelve and over endorsed using illegal substances within the past month.

In the SSC sample, both mothers and fathers reported substantially lower illegal substance use (2.6% and 4.9%, respectively).

Impact of Child Problem Behaviors and the BAP on Parental Substance Use

In reporting results, for all models, change in chi-square from the first block model to the full model containing the variables of interest are reported. Among fathers, the change in model chi-square fit was not significant for alcohol (χ^2 (4) = 8.71, $p = .07$) or illegal substance use (χ^2 (4) = 8.33, $p = .08$). However, the model change in fit was significant for paternal tobacco use (χ^2 (4) = 14.82, $p = .005$). Among the predictors tested, the CBCL Externalizing composite score was significant (Wald = 4.12, $p = .043$, OR = 1.01), as was the Rigid subscale of the BAPQ (Wald = 8.56, $p = .003$, OR = 1.29). Both the CBCL Externalizing composite score and BAPQ Rigid score predicted tobacco use even when controlling for other variables. Fathers of children with externalizing behaviors were 17.2% more likely to report tobacco use, and fathers with higher BAPQ Rigid scores were 20.9% more likely to use tobacco.

Among mothers, the change in model chi-square fit was significant for alcohol (χ^2 (4) = 14.53, $p = .006$), tobacco (χ^2 (4) = 12.83, $p = .012$), and illegal substances (χ^2 (4) = 10.67, $p = .031$). For maternal alcohol use, among the covariates tested, the Rigid subscale of the BAPQ was significant when other variables were controlled (Wald = 6.52, $p = .011$, OR = .828), indicating that mothers with elevated Rigid subscale scores were 45% less likely to report using alcohol. For maternal tobacco use, among the covariates tested, the CBCL Externalizing composite score was significant (Wald = 8.198, $p = .004$, OR = 1.021), suggesting that mothers with children with more severe externalizing behaviors were 10% more likely to report using tobacco. Lastly, for maternal illegal substance use, the CBCL Externalizing composite score was significant when other variables were controlled (Wald = 7.762, p = .005, OR = 1.035), indicating that mothers with children with more severe externalizing behaviors were only 0.31% more likely to report using illegal substances.

Discussion

Results of this study suggest that parents of children with ASD report greater rates of alcohol use and lower rates of tobacco and illegal substances when compared to national estimates. Results of binary logistic regressions indicate that children's externalizing behaviors and parental BAP traits impact substance use differently among mothers and fathers. Notably, children's externalizing behaviors predicted tobacco use among both fathers and mothers of children with ASD. This relationship also held true for mothers and illegal substance use, with mothers of children with more externalizing behavior problems more likely to report using illegal substances. However, this group was only 0.31% more likely, giving it a relatively small predictive value.

With regard to parental broad autism phenotype traits, rigidity, but not aloofness or pragmatic language difficulties, predicted both increased and decreased parental substance use. Specifically, fathers endorsing higher levels of rigidity were more likely to report current tobacco use, whereas higher levels of rigidity predicted decreased alcohol use among mothers. Fathers endorsing rigidity were 20.9% more likely to report using tobacco, and mothers who reported rigidity were 45% *less* likely to report alcohol use.

Although the model for paternal alcohol use was not significant, fathers with higher rigidity scores were also less likely to report alcohol use.

Nicotine and alcohol are legal substances that are often used to reduce tension and frequently result in a calming effect. Given that rigidity is conceptualized as struggling to adapt to change (Hurley et al. 2007), tobacco use may help to temporarily increase one's flexibility in response to unexpected events or circumstances. However, as stated earlier, mothers endorsing rigidity traits were less likely to use alcohol. Because alcohol use can result in feeling uninhibited, individuals endorsing rigid traits may find the effects of alcohol unsettling given their desire for control and consistency. Alternatively, the differences found in parental substance use related to rigidity may indicate a lack of an association between BAP traits (specifically rigidity) and parental substance use. Additional research is needed to further explore potential associations between BAP traits and parental substance use before any definitive conclusions are drawn.

The current results suggest that both child externalizing behaviors and parental rigidity influence substance use among parents of children with ASD. Although the mechanisms by which this relationship operates is beyond the scope of this study, it is proposed that particular child and parental characteristics interact, creating elevated stress levels that are decreased, at least temporarily, by the use of substances. Additional research is needed to verify this potential relationship.

Limitations and Future Directions

There are several limitations of this study that should be acknowledged. First, parental substance use was measured through the SSC Substance Use History Form, which does not have established reliability or validity. Parents could only respond “yes” or “no” to current substance use; therefore, no conclusions can be drawn about the frequency or quantity of use. This makes it impossible to discriminate between parents who are *abusing* or *consuming* substances within appropriate limits, as well as precludes our ability to detect relations between child/parent characteristics and amount of substance use. Thus, the individuals endorsing substance use in the sample may vary considerably. Additionally, inherent limitations exist in self-report measures, most notably the potential for social desirability bias, which may have impacted parental responses.

Another limitation of this study is that although the literature suggests that it is best to collect both self- and informant-reports (Hurley et al., 2007), only BAPQ self-report data were available. Lastly, related to measures, the mean CBCL score (t = 56.49) fell well within the average range; thus our conclusions regarding the effect of child externalizing behavior on parental substance use should be interpreted with caution. Furthermore, data regarding who completed the CBCL questionnaire (mother/father) was not available. Given that perceptions of problem behaviors differ among mothers and fathers (Davis & Carter, 2008), this information should be accounted for in future analyses.

Additionally, the sample for this study was not representative in regards to annual household income and marital status. Compared to the U.S. population, this sample was more affluent and more likely to be married. Additionally, African Americans and Hispanics were underrepresented. Future research should examine paternal and maternal substance use among parents of children with ASD with a more diverse sample, particularly in regards to socioeconomic status, marital status, race, and ethnicity.

Despite these limitations, this study is among the first to examine prevalence and predictors of maternal and paternal substance use among parents of children with ASD. Thus, although it is largely an exploratory study, results suggest that parental rigidity and child externalizing behavior problems may affect substance use in this population. A notable exception, however, is for maternal alcohol use, for which rigidity appears to decrease consumption. These findings suggest that clinicians working with parents of children with ASD should be attuned to the potential use of substances as a coping mechanism. In general, clinicians should be more cognizant of the functioning of the family system as a whole.

Future research should examine the frequency of substance use, as well as assess additional suspected predictors (e.g., parental stress) and other maladaptive and adaptive coping strategies. Data measuring parental stress were not available and thus not included in the tested models; however, it may have served as a mediator linking parent and child factors with parental substance use, warranting further research.

References

Achenbach, T. M. (1991). *Manual for the Child Behavior Checklist/4-18 and the 1991 profile*. Burlington, VT: University of Vermont, Department of Psychiatry.

Bailey, A., Le Couteur, A., Gottesman, I., Bolton, P., Simonoff, E., Yuzda, E., & Rutter, M. (1995). Autism as a strongly genetic disorder: Evidence from a British twin study. *Psychological Medicine*, *25*(1), 5-17. doi: 10.1017/S0033291700028099

Beck, A., Hastings, R.P., & Daley, D. (2004). Pro-social behaviour and behaviour problems independently predict maternal stress. *Journal of Intellectual and Developmental Disability*, *29*(4), 339-349, doi: 10.1080/13668250400014509

Benson, P.R. (2006). The impact of child symptom severity on depressed mood among parents of children with ASD: The mediating role of stress proliferation. *Journal of Autism and Developmental Disorders*, *36*, 685-695, doi: 10.1007/s10803-006-0112-3.

Benson, P.R. & Karlof, K.L. (2009). Anger, stress proliferation, and depressed mood among parents of children with ASD: A longitudinal replication. *Journal of Autism and Developmental Disabilities*, *39*, 350-362, doi: 10.1007/s10803-008-0632-0.

Bolton, P., Macdonald, H., Pickles, A., Rios, P., Goode, S., Crowson, M., Bailey, A., & Rutter, M. (1994). A case-control family history study of autism. *Journal of Child Psychology and Psychiatry, 35*(5)*,* 877-900.

Bolton, P.F., Pickles, A., Murphy, M., & Rutter, M. (1998). Autism, affective and other psychiatric disorders: patterns of familial aggregation. *Psychological Medicine*, *28*, 385-395, doi: 10.1017/S0033291797006004

Centers for Disease Control and Prevention. (2012a). Autism Spectrum Disorders(s). Retrieved from http://www.cdc.gov/ncbddd/autism/signs.html

Centers for Disease Control and Prevention. (2012b). Prevalence of Autism Spectrum Disorders - autism and developmental disabilities monitoring network, 14 sites, United States, 2008. *Surveillance Summaries, 61*(SS03), 1-19. Retrieved from http://www.cdc.gov/mmwr/preview/mmwrhtml/ss6103a1.htm?s_cid=ss6103a1_w

Davis, N.O. & Carter, A.S. (2008). Parenting stress in mothers and fathers of toddlers with autism spectrum disorders: Associations with child characteristics. *Journal of Autism and Developmental Disorders*, *38*, 1278-129, doi: 10.1007/s10803-007-0512-z.

DeLong, G.R. & Dwyer, J.T. (1998). Correlation of family history with specific autistic subgroups: Asperger's syndrome and bipolar affective disease. *Journal of Autism and Developmental Disorders*, *18*, 593-600.

Fischbach, G. & Lord, C. (2010). The Simons Simplex Collection: A resource for identification of autism genetic risk factors. *Neuron, 68*(2), pp. 192-195, doi: 10.1016/j.neuron.2010.10.006

Folstein, S. & Rutter, M. (1977). Infantile autism: a genetic study of 21 twin pairs. *Journal of Child Psychology and Psychiatry, 18*(4), 297-321.

Gerkensmeyer, J.E., Perkins, S.M., Scott, E.L., & Wu, J. (2008). Depressive symptoms among primary caregivers of children with mental health needs: Mediating and moderating variables. *Archives of Psychiatric Nursing*, *22*(3), 135-146. doi:10.1016/j.apnu.2007.06.016

Hartley, S.L., Sikora, D.M., & McCoy, R. (2008). Prevalence and risk factors of maladaptive behaviour in young children with Autistic Disorder. *Journal of Intellectual Disability Research*, *49*, 819-829, doi: 10.1111/j.1365-2788.2008.01065.x

Higgins, D.J., Bailey, S.R., & Pearce, J.C. (2005). Factors associated with functioning style and coping strategies of families with a child with an autism spectrum disorder. *Autism*, *9*(2), 125-137, doi: 10.1177/1362361305051403

Hurley, R., Losh, M., Parlier, M., Reznick, J., & Piven, J. (2007). The Broad Autism Phenotype Questionnaire. *Journal of Autism and Developmental Disorders*, *37*(9), 1679-1690, doi: 10.1007/s10803-006-0299-3

Ingersoll, B. & Hambrick, D.Z. (2011). The relationship between the broader autism phenotype, child severity, and stress and depression in parents of children with autism spectrum disorders. *Research in Autism Spectrum Disorders*, *5*, 337-344. doi:10.1016/j.rasd.2010.04.017

Ingersoll, B., Meyer, K., & Becker, M.W. (2011). Increased rates of depressed mood in mothers of children with ASD associated with the presence of the broader autism phenotype. *Autism Research*, *4*(2), 143-148, doi: 10.1002/aur.170

Killeen, T., & Brady, K.T. (2000). Parental stress and child behavioral outcomes following substance abuse residential treatment follow-up at 6 and 12 months. *Journal of Substance Abuse Treatment*, *19*(1), 23-29, doi:10.1016/S0740-5472(99)00078-1.

Lobascher, M.E., Kingerlee, P.E., & Gubbay, S.S. (1970). Childhood Autism: An investigation of aetiological factors in twenty-five cases. *The British Journal of Psychiatry*, *117*, 525-529.

Losh, M., Childress, D., Lam, K., & Piven, J. (2008). Defining key features of the broad autism phenotype: a comparison of multiple- and single-incidence families. *American Journal of Medical Genetics, Part B (Neuropsychiatric Genetics)*, *147B*, 424-433.

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

Miles, J.H., Takahashi, T.N., Haber, A., & Hadden, L. (2003). Autism families with high incidence of alcoholism. *Journal of Autism and Developmental Disorders*, *33*(4), 403-415. doi: 10.1023/A:1025010828304

Murphy, M., Bolton, P.F., Pickles, A., Fombonne, E., Piven, J., and Rutter, M. (2000). Personality traits of the relatives of autistic probands. *Psychological Medicine, 30*, 1411-1424. doi: 10.1017/S0033291799002949

National Center for Health Statistics. (2012). Health, United States, 2012: With special feature on emergency care. Hyattsville, MD. Retrieved from http://www.cdc.gov/nchs/data/hus/hus12.pdf#058

Ozonoff, S., Young, G.A., Carter, A., Messinger, D., Yirmiya, N., Zwaigenbaum, L., et al. (2011). Recurrence risk for autism spectrum disorders: a baby siblings research consortium study. *Pediatrics, 128*, e488-e495.

Pelham, W.E., Lang, A.R., Atkeson, B., Murphy, D.A., Gnagy, E.M., Greiner, A.R., Vodde-Hamilton, M., & Greenslade, K.E. (1997). Effects of deviant child behavior on parental distress and alcohol consumption in laboratory interactions. *Journal of Abnormal Child Psychology*, *25*(5), 413-424. doi: 10.1023/A:1025789108958

Piven, J., Chase, G.A., Landa, R., Wzorek, M., Gayle, J., Cloud, D., & Folstein, S. (1991). Psychiatric disorders in the parents of autistic individuals. *Journal of the American Academy of Child and Adolescent Psychiatry*, *30*(3), 471-478. doi: 10.1097/00004583-199105000-00019

Piven, J., Palmer, P., Jacobi, D., Childress, D., & Arndt, S. (1997a). Broader autism phenotype: Evidence from a family history study of multiple-incidence families. *American Journal of Psychiatry, 154* (2), pp. 185-190.

Piven, J., Palmer, P., Landa, R., Santangelo, S., Jacobi, D., & Childress, D. (1997b). Personality and language characteristics in parents from multiple-incidence autism families. *American Journal of Medical Genetics (Neuropsychiatric Genetics), 74,* 398-411.

Rutter, M., Le Couteur, A., & Lord, C. (2003). *Autism Diagnostic Interview, Revised (ADI-R)*. Los Angeles: Western Psychological Services.

Schiller, J.S., Lucas, J.W., Peregoy, J.A. (2012). Summary health statistics for U.S. adults: National Health Interview Survey, 2011. National Center for Health Statistics. *Vital and Health Statistics*, *10*(256). Retrieved from http://www.cdc.gov/nchs/data/series/sr_10/sr10_256.pdf

Smalley, S.L., McCracken, J., & Tanguay, P. (1995). Autism, affective disorders, and social phobia. *American Journal of Human Genetics*, *60*, 19-26.

United States Department of Commerce. (2013). United States Census Bureau. Retrieved from http://quickfacts.census.gov/qfd/states/00000.html

Appendix

Table 1

Race and Ethnicity of Probands, Mothers, and Fathers

	Probands	Mothers	Fathers
Race			
White	78.5%	80.7%	81.7%
Asian	4.4%	5.2%	4.8%
More-than-one-race	7.7%	4.0%	2.9%
Other	4.6%	3.9%	3.9%
African American	3.9%	4.0%	4.8%
Not specified	0.7%	1.8%	1.4%
Native American	0.3%	0.2%	0.2%
Native Hawaiian	0.1%	0.2%	0.3%
Ethnicity			
Non-Hispanic	88.6%	90.7%	91.7%
Hispanic	11.4%	9.3%	8.3%

Table 2

Parent Relationship Status

Married	90.3%
Divorced (one parent remarried)	2.3%
Divorced (neither parent remarried)	3.1%
Separated	1.3%
Never married	2.4%
Divorced (both remarried)	0.5%

Table 3

Annual Household Income

Less than $20,000	3.2%
$21,000-35,000	5.0%
$36,000-50,000	8.4%
$51,000-65,000	10.7%
$66,000-80,000	13.7%
$81,000-100,000	17.2%
$101,000-130,000	15.5%
$131,000-160,000	9.5%
Over $161,000	16.7%

Table 4

BAPQ Means and Cut-off Scores

	Mothers/Females		Fathers/Males	
	SSC Sample	Cut-off scores[a]	SSC Sample	Cut-off scores[a]
Aloof	2.38	3.00	2.82	3.25
Rigid	2.67	3.25	2.88	3.65
Pragmatic Language	2.10	2.70	2.34	2.95
Overall	2.38	3.25	2.68	3.35

[a] Cut-off scores from Hurley, R., Losh, M., Parlier, M., Reznick, J., & Piven, J. (2007). The broad autism phenotype questionnaire. *Journal of Autism and Developmental Disorders*, *37*(9), 1679-1690, doi: 10.1007/s10803-006-0299-3

Table 5

Chi-squared Values for Covariates

	Fathers			Mothers		
Type of Substance	Alcohol	Tobacco	Illegal	Alcohol	Tobacco	Illegal
Income	113.75*	89.58*	41.35*	127.35*	89.35*	20.59
Race	83.03*	12.324	9.69	117.84*	6.35	7.51
Ethnicity	14.32*	1.86	2.63	17.67*	.134	.115

*$p < .001$

Table 6

Substance Use Prevalence Rates for SSC Sample and National Estimates

	SSC Sample		National Estimates	
	Mothers	Fathers	Males	Females
Alcohol	62.5%	70.6%	60%	44%
Tobacco	8.9%	17.8%	21%	17%
Illegal Substances	2.6%	4.9%	8.9%	8.9%

Model Invariance across Genders of the Broad Autism Phenotype Questionnaire

Neill Broderick, Jordan L. Wade, J. Patrick Meyer, Michael Hull, & Ronald E. Reeve

University of Virginia

Published in the *Journal of Autism and Developmental Disorders* (2015)

Broderick, N., Wade, J.L., Meyer, J.P., Hull, M., & Reeve, R.E. (2015). Model invariance across genders of the Broad Autism Phenotype Questionnaire. *Journal of Autism and Developmental Disorders*, *45*(10), 3133-3147.

Abstract

ASD is one of the most heritable neuropsychiatric disorders, though comprehensive genetic liability remains elusive. To facilitate genetic research, researchers employ the concept of the Broad Autism Phenotype (BAP), a milder presentation of traits in undiagnosed relatives. Research suggests that the BAP Questionnaire (BAPQ) demonstrates psychometric properties superior to other self-report measures. To examine evidence regarding validity of the BAPQ, the current study used Confirmatory Factor Analysis (CFA) to test the assumption of model invariance across genders. Results of the current study upheld model invariance at each level of parameter constraint; however, model fit indices suggested limited goodness-of-fit between the proposed model and the sample. Exploratory analyses investigated alternate factor structure models but ultimately supported the proposed three-factor structure model.

Model Invariance across Genders of the Broad Autism Phenotype Questionnaire

Autism Spectrum Disorder (ASD) refers to related neurodevelopmental conditions characterized by qualitative impairments in social communication as well as the presence of restricted, repetitive patterns of behavior and/or areas of interest (American Psychiatric Association, 2013). A diagnosis of ASD is typically determined by parent report regarding child development in addition to clinician administered standardized tests and qualitative observations across disciplines and settings; therefore, the diagnostic determination of ASD rests on clinician interpretation of findings, as there are no medical tests for diagnosis and/or underlying etiology (Geschwind, 2011; Klin, Saulnier, Tsatsanis, & Volkmar, 2005). The recent release of DSM-5 diagnostic criteria amplifies the importance of careful diagnostic discernment as clinicians transition from the DSM-IV-TR criteria.

Diagnostic practices directly influence our understanding of the prevalence of ASD, with present estimates indicating that 1 in 68 children in the United States are currently identified as having an ASD (Centers for Disease Control and Prevention, 2014). While the CDC (2014) finds this rate to best capture the current prevalence of ASD, their study revealed a broad range of prevalence rates, from 1 in 45 to 1 in 175 across data collection sites, due to varying diagnostic practices. The alarmingly high prevalence rate, which reflects a 30% increase since 2012, considered in conjunction with regional diagnostic variance, further stresses the necessity of appropriate diagnosis of ASD, including an understanding of the etiology of the disorder (CDC, 2012; Sucksmith, Roth, & Hoekstra, 2011).

The many complex presentations of ASD reflect etiological heterogeneity (Geschwind, 2011; State, 2010). While an exhaustive understanding of the origins of ASD remains elusive, several twin and family studies of individuals with ASD demonstrate high heritability and recurrence rates. These studies provide strong evidence for the genetic contributions to the disorder (Bailey et al., 1995; Folstein & Rutter, 1977; Ingersoll, Hopwood, Wainer, & Donnellan, 2011; Ozonoff et al., 2011), making ASD one of the most heritable neuropsychiatric conditions (Davidson, Goin-Kochel, Green-Synder, Hundley, Warren, & Peters, 2012; State, 2012).

More recent family studies have employed the concept of the Broad Autism Phenotype (BAP), which refers to the presentation of conceptually similar but milder traits associated with ASD in the undiagnosed relatives of individuals who warrant the diagnosis (Sucksmith, Roth, & Hoekstra, 2011). The application of the BAP allows researchers to identify which specific traits aggregate in biological family members, both diagnosed and undiagnosed (Dawson et al., 2007; Losh et al., 2008b; Sucksmith, Roth, & Hoekstra, 2011). The isolation of the traits associated with the BAP facilitates genetic research by increasing the number of individuals available for analysis as well as allowing for more precise identification of each phenotypic component, rather than a collection of traits in the full condition of ASD (Losh et al., 2008b; Piven et al., 2001; Sucksmith, Roth, & Hoekstra, 2011). Therefore, a clear delineation of the BAP could prove invaluable in determining the genetic factors underlying ASD.

In a comprehensive literature review, Sucksmith, Roth, and Hoekstra (2011) noted an increase in research efforts regarding the study of the BAP over the last 15 years. Research regarding the BAP, as evidenced in parents, suggests that it is characterized by deficits in pragmatic language skills and social reciprocity; social cognitive difficulties; heightened aloof, rigid, and hypersensitive personality traits; and increased rates of internalizing psychiatric conditions (Sucksmith, Roth, & Hoekstra, 2011). Parallel to the proliferation of research regarding the BAP phenomenon, several groups have developed instruments to quantify the BAP, conceptualizing it as a “set of continuous, quantitative traits” based on the mounting empirical support referenced above. Though research provides some evidence regarding validity of each instrument in identifying the BAP, the Broad Autism Phenotype Questionnaire (BAPQ; Hurley et al., 2007) has emerged as the foremost self-report measure of the BAP (Davidson et al., 2012; Ingersoll et al., 2011).

Hurley et al. (2007) report evidence that the BAPQ demonstrates strong psychometric properties; however, the evidence regarding its validity, while promising, is not sufficient for the emerging uses of the BAPQ. The assumption of model invariance, which refers to the stability of the measurement and structural properties of an instrument across conditions (e.g., groups), should be upheld in order to draw meaningful conclusions about the results from an instrument (Vandenberg & Lance, 2000). Recent comparisons of mean differences across gender on the BAPQ assume model invariance without explicitly testing it (Hurley et. al., 2007; Seidman, et. al., 2012). The present study sought to assess model invariance of the BAPQ explicitly.

Purpose of the Study

The current study investigated model invariance across genders on the BAPQ. More specifically, this research sought to determine if the components of the measurement and structural model of the BAPQ were invariant, or roughly equivalent, across gender groups. The following questions comprised the foci of the research: 1) Do males and females interpret the BAPQ similarly, leading them to respond to it in similar ways?; 2) Does the factorial structure of the BAPQ, as captured by measurement invariance as well as structural invariance, remain equivalent across genders?; and 3) Will the 36 items of the BAPQ combine to form three factors (aloof, rigid, and pragmatic language) as proposed by the test developers for both gender groups? The study also examined alternative factorial structures, including a one- and two-factor model based on existing literature.

Methods

Participants

Participants for the present study included families who contributed to the Simons Simplex Collection (SSC). A primary project of the Simons Foundation Autism Research Initiative (SFARI), the SSC is a large, multi-site study in the United States that employs comprehensive and consistent clinic-based assessment of “simplex families” in order to support a permanent repository of genetic and phenotypic information (SFARI, 2012). Simplex families include one child diagnosed with an ASD, who is referred to as a proband, and no other first-to third-degree relatives diagnosed with an ASD.

Subjects in the SSC are families comprised of individuals with a diagnosis of ASD (probands), ages 4-18 years who present with a nonverbal mental age greater than 18 months, as well as their parents and, where applicable, one sibling without ASD (Fischbach & Lord, 2010). A majority of probands in the SSC present with moderate to severe symptomatology associated with ASD and fairly minimal intellectual disability (Fishbach & Lord, 2010).

The current study obtained participants from the SSC dataset entitled “Simons Simplex Collection Version 14, Simons Ancillary Collection Version 2, and Simons Twin Study,” which included 2,760 families. The current study included those participants from the Simons Simplex Collection (SSC) and the Simons Twin Study (STC). The study only retained those mothers and fathers who completed a BAPQ, as some parents did not provide any BAPQ data. After refining the dataset, the sample for the present study included 4,374 parents, prior to data screening.

The probands for the current sample were comprised of 86.9% males (*n* = 1,851) and 13.1% females (*n* = 278), or a ratio of 6:1, which is commensurate with the CDC (2012) prevalence estimates’ range. SSC collected data on the race and ethnicity of the probands, where ethnicity is defined as non-Hispanic or Hispanic. The majority of probands identified as White and non-Hispanic (see Tables 1 and 2). In regard to family and parent demographics, 97.4% of the parents included in the sample were married, which is substantially more than the general US population (approximately 40%) (Goodwin, Mosher, & Chandra, 2010). In the current sample, 95.9% of parents earned a high school diploma or higher and 61% earned a Bachelor’s degree or higher, compared to the 85.4% of individuals 25 years and older in the general population of the US who obtained a high school diploma or higher, and the 28.2% who earned a Bachelor’s degree or higher. The sample included in the study was more affluent than the general population in the US, as 71% reported a family income greater than $65,000 (US Census Bureau, 2013).

Measures

Broad Autism Phenotype Questionnaire (BAPQ): Self-Report. Hurley et al. (2007) specifically designed the Broad Autism Phenotype Questionnaire (BAPQ) to efficiently, validly, and reliably screen for the personality and language characteristics identified as comprising the BAP in parents of individuals with ASD (Ingersoll et al., 2011). The BAPQ design now aligns with the empirically-supported conceptualization of the BAP traits: aloof personality, rigid personality, and pragmatic language difficulties (Hurley et al., 2007). To measure the personality and language traits associated with the BAP, the BAPQ includes 36 items that yield an overall score comprised of three subscales: Aloof, Rigid, and Pragmatic Language (Hurley et al., 2007). Each of the subscales is comprised of 12 of the 36 items on the instrument (Hurley et al., 2007). The BAPQ items instruct participants to rate how frequently each statement (e.g., "I would rather talk to people to get information than to socialize") applies to them with anchors of one (statement very rarely applies) and six (statement applies very often), which offers a range of possible responses without a neutral option (Hurley et al., 2007). Selected items (15 of the 36) include wording for reverse-scoring to minimize response set bias (Hurley et al., 2007). After reverse scoring, the administrator sums the scores on the items in each subscale and then generates an average for that subscale, whereas the total BAPQ score reflects the average of all items (Hurley et al., 2007). Therefore, all summary scores range from one to six, where higher scores are more indicative of the BAP (Hurley et al. al., 2007).

There are two versions of the BAPQ, a self-report version and an informant-report version; the versions are identical except for the use of pronouns on the informant version

(Hurley, et al., 2007). Hurley et al. (2007) administered both versions of the BAPQ to parents of children with ASD, which resulted in three scores for each parent – a self-report, informant-report, and "best-estimate" score based on an average of the former two. The same sample of parents also participated in a face-to-face clinical interview. Hurley et al. (2007) examined the internal consistency of the BAPQ subscales, yielding Cronbach's α coefficients of .94 for the Aloof subscale, .91 for the Rigid subscale, .85 for the Pragmatic Language subscale, and .95 across all 36 items. The inter-item reliability did not differ between positively and negatively valenced items, self- and informant-report ratings, male and female participants, nor parents of children with ASD and parents of children without ASD (Hurley et al., 2007). For each individual item, item-total correlations were greater than or equal to .39 relative to the other items in each subscale (Hurley et al., 2007). All subscales are significantly correlated with each other for parents of children with ASD as well as controls, and these correlations are comparable for male and female subjects as well as self- and informant-report versions. Of note, only self-report data were utilized for this study, as that was the version collected by the SSC.

Hurley et al. (2007) validated the BAPQ against direct clinical assessment of the BAP and revealed high sensitivity and specificity in a sample of parents with and without the BAP. Optimal cutoffs for the BAPQ maximized sensitivity and specificity according to Receiver Operator Curves (ROC), where the best-estimate BAPQ cutoff for Aloof is 3.25, Rigid is 3.50, Pragmatic Language is 2.75, and Total score is 3.15 (Hurley et al., 2007). Hurley et al. (2007) also used the ROC to determine cutoffs for the self- and informant-report versions as well as for genders. For the total BAPQ score, sensitivity and specificity were approximately 80% (Hurley et al., 2007). Sensitivity and specificity reached 70% for all subscales, and specificity for the Aloof and Rigid subscales exceeded 80% (Hurley et al. al., 2007). The results from Hurley et al. (2007) suggest that the BAPQ successfully distinguishes parents with a clinically defined BAP from parents without evidence of the BAP in parents of children with ASD as well as controls.

Analyses

Data Screening. In preparation for data analyses, the researchers screened the initial data (*n* = 4,374 parents) to check the relevant assumptions for multivariate statistical analysis. The authors eliminated all those parents with any missing data on the BAPQ as well as univariate and multivariate outliers. After eliminating all outliers, the data screening focused on univariate and multivariate normality. The current study assumed univariate normality according to skewness and kurtosis of the data (Curran, West, & Finch, 1996). The current study also assumed the large sample size to be robust to multivariate non-normality (Amemiya and Anderson, 1990; Anderson and Amemiya, 1988).

It is notable that the current study conducted the analyses with a sample that included the outliers, as well as a sample that excluded the above-described outliers. The analyses of both samples did not yield substantive differences in our findings or conclusions drawn from those findings. Therefore, the present study utilized the sample that included the outliers (*n* = 4,258) to present a more comprehensive perspective of how parents of children with ASD respond on the BAPQ.

Model Invariance Testing Strategy. The present study sought to determine the model invariance of the BAPQ across genders by testing equivalency of its factorial measurement as well as its underlying latent structure across genders. In order to test hypotheses regarding the BAPQ invariance across genders, the study used multi-group Confirmatory Factor Analysis (CFA) as it has emerged as the method of choice (Cheung & Rensvold, 2002; Yin & Fan, 2003). The flexibility of CFA, in addition to its ability to quantify degrees of model fit, rendered it ideal to test invariance propositions (Thompson, 2004). All models were estimated using the Analysis of Moment Structures (AMOS; Arbuckle, 2011) program operating on Maximum Likelihood (ML) estimation of covariance of structures (COVS). CFA typically uses ML estimation, which assumes that data are multivariate normally distributed, though our large sample size is robust to multivariate non-normality (Thompson, 2004).

The current study followed the Byrne (2010) hierarchy of “logically ordered and increasingly restrictive” tests of sets of model parameters to investigate the equivalences of the BAPQ across genders (p. 199). The process required constraining sets of parameters to be equivalent across genders; as each new set of parameters was tested, those previously determined to be group invariant were cumulatively constrained (Byrne, 2010). The hierarchy of tests (analyses) of model invariance for the current study commenced with a test of configural invariance then shifted to a test of metric invariance before concluding with a test of structural invariance, as proposed by Byrne (2010).

Configural invariance requires that the number of factors as well as the pattern of factor weights remains similar for each group without equality constraints imposed (Byrne, 2010; Vandenberg & Lance, 2000). A rejection of the null hypothesis of configural invariance would indicate that the BAPQ factor-structure varies across groups. Configural invariance must be upheld for subsequent tests of invariance to yield meaningful results (Vandenberg & Lance, 2000); it establishes the baseline for subsequent tests, as ML estimation goodness-of-fit change indices are additive (Byrne, 2010).

The test of metric invariance ($\Lambda^{g}_{x} = \Lambda^{g'}_{x}$) examines invariance after constraining the factor weights (also known as loadings) to be equal across gender groups. Findings that uphold metric invariance would support the pursuit of additional, more restrictive tests of invariance, such as structural invariance. Indications of noninvariance, however, may prompt the examination of partial invariance to determine the specific point where invariance breaks down.

Upon establishing equivalence of the measurement model across groups with tests of configural and metric invariance, the hierarchy shifts to a test of the structural components of the model with structural invariance, more specifically, the factor variance-covariance structure. Byrne (2010) encourages researchers to test the invariance related to factor covariances ($\phi^{g}_{jj} = \phi^{g'}_{jj}$) across groups, rather than factor variances, as covariance invariance captures the degree to which the theoretical structure underlying an instrument remains the same across groups (Byrne, 2010). Following the work of Byrne (2010) the present study constrained the covariances (or correlations) between the factors to be equal across genders (Byrne, 2010; Vandenberg & Lance, 2000), rather than factor variances.

Baseline Model. Prior to analyses, the process of testing model invariance requires an established baseline model for each group under investigation — male and female groups for the current study (Byrne, 2010). The developers of the BAPQ proposed a three-factor structure model for their instrument to capture the Broad Autism Phenotype

(BAP), which served as the baseline model for both groups (Hurley et al., 2007). Within the baseline model, there were 3 factors – Aloof Personality, Rigid Personality, and Pragmatic Language – and 36 items on the BAPQ; each item represented an observed variable that served as an indicator for only one of the three latent factors (Hurley et al., 2007). Therefore, each of the factors presumed that 12 observed variables were hypothesized to capture the underlying dimension. The three BAP factors are assumed to be correlated based on the initial psychometric study of the BAPQ (Hurley et al., 2007).

Each variable was associated with an error of measurement, and the errors were assumed to be uncorrelated. The baseline model set factor variances to unity (using 1) for scaling (Thompson, 2004). (See Figure 3 for the baseline model.) It is notable that the analysis of the baseline model did not require between-group constraints because it did not require constraints of any parameters (Byrne, 2010); therefore, the present study analyzed the baseline, unconstrained model on the ungrouped model (all individual parents included without grouping by gender for analysis) initially, then analyzed the grouped model (mothers and fathers grouped by gender for simultaneous analysis) to assess configural invariance. Subsequent tests of model invariance analyzed the groups simultaneously (Byrne, 2010).

As noted above, the baseline model for each group (and the ungrouped model) included 36 variables; therefore, the degrees of freedom for each group (and the ungrouped model) equaled 666 [36(36+1)/2] (Thompson, 2004). With two groups, the total degrees of freedom equaled 1,332. Each group (and the ungrouped model) included the following parameters: 36 error terms, 36 weights (i.e., loadings), and 3 factor correlations (36+36+3), which equaled 75 parameters. There were also 3 factor variances, however, these variances were held to one for unity, as previously noted. The model for each group was appropriately overidentified with 75<666 degrees of freedom; there are 591 unspent degrees of freedom for each group or 1,182 unspent degrees of freedom total.

Model Assessment. A model is considered acceptable if the measurement and structure implied by the hypothesized model is similar to the measurement and structure of the sample data, as indicated by specified values of goodness-of-fit indices (Byrne, 2010; Cheung & Rensvold, 2002). The present study applied the inferential chi-square (χ^2) goodness-of-fit change statistic, the inferential Comparative Fit Index (CFI) goodness-of-fit change statistic and a group of descriptive goodness-of-fit indices to assess model fit at each step of the model invariance hierarchy of tests (Byrne, 2010). To uphold invariance, the difference in χ^2 between the more constrained model and the previously tested, less-constrained model must be smaller than the χ^2 critical value. A statistically significant difference in χ^2 value reflects noninvariance, whereas a non-significant value of χ^2 supports a failure to reject the null hypothesis of invariance (Byrne, 2010; Cheung & Rensvold, 2002).

Of the goodness-of-fit change statistics, the χ^2 change statistic is the most commonly used to measure the differences between two nested models (Byrne, 2010). When using the χ^2 change statistic, however, the model is often rejected in a large sample size despite a small, even negligible, difference between the models (Byrne, 2010; Cheung & Rensvold, 2002; Thompson, 2004). Therefore, the analyses for the current study also considered the Comparative Fit Index (CFI) change statistic (Byrne, 2010; Cheung & Rensvold, 2002). Using the CFI change statistic, the absolute CFI difference between the more constrained model and the less constrained model must be less than .01 to uphold invariance (Byrne, 2010; Cheung & Rensvold, 2002). The current literature supports the use of the CFI change statistic, in addition to the χ^2 change statistic, as it is independent of model characteristics based on a rigorous Monte Carlo study (Byrne, 2010; Cheung & Rensvold, 2002).

To further assess the goodness-of-fit between the hypothesized model and the sample, the current study also considered at each test of model invariance several common model fit indices, including the goodness of fit index (GFI), adjusted goodness of fit index (AGFI), Tucker-Lewis index (TLI), comparative fit index (CFI), standardized root mean square residual (RMR), and root mean square error of approximation (RMSEA). Using the ML method, a cutoff value larger than .90-.95 for GFI, AGFI, TLI and CFI, and a cutoff value below .06 (with .08 as upper limit) for RMSEA and .08 (with .10 as upper limit) for RMR are needed in order to conclude that there is a relatively good fit between the hypothesized model and the observed data (Byrne, 2010; Hu & Bentler, 1999; Thompson, 2004; Vandenberg & Lance, 2000).

Results

Configural Invariance

The first step of analyses associated with the process of testing model invariance involved the test of configural invariance (Byrne, 2010). The configural model for the present study essentially tested the invariance of the 3-factor structure baseline model of the BAPQ in an ungrouped model (all individual parents included without grouping by gender for analysis) using a CFA as well as a grouped model (mothers and fathers grouped by gender for simultaneous analysis) using a multi-group CFA (Byrne, 2010).

The results of the CFA for the ungrouped model are detailed in Table 1. The RMR and RMSEA model fit indices support the goodness-of-fit between the sample and the model. The results of the multi-group CFA for the grouped dataset offered the opportunity to examine the χ^2 and CFI change statistics by comparing the ungrouped and grouped models. The χ^2 change upheld invariance. The CFI change statistic also upheld invariance. Despite the support from the change statistics, it is notable that only two of the model fit indices, RMR and RMSEA, indicated a good fit between the sample and the model. Though the findings upheld configural invariance, and provided evidence to proceed with the test of metric invariance, there are limitations in the fit between the sample and the model.

Metric Invariance

Given that configural invariance was upheld across the gender groups on the BAPQ, the current study proceeded with the test of metric invariance ($\Lambda^{g}_{x} = \Lambda^{g'}_{x}$), which constrained the factor weights (also known as loadings) to be equal across gender groups. Findings that upheld metric invariance supported the study to pursue additional, more restrictive tests of invariance (see Table 2).

The χ^2 change statistic did not uphold invariance. Conversely, the CFI change statistic upheld invariance. Only the RMSEA provided evidence of a good fit between the sample and the model. In conclusion, the CFI change for the test of metric invariance provided sufficient evidence to proceed with the test of structural invariance, though the χ^2 change statistic did not uphold invariance. The model fit indices again indicated limited goodness-of-fit between the sample and the model proposed by the BAPQ developers.

Structural Invariance

Upon establishing equivalence of the measurement model across groups with tests of configural and metric invariance, the study proceeded with the test of structural invariance. Following the work of Byrne (2010), the present study constrained the covariances (or correlations) between the factors to be equal across groups (Byrne, 2010; Vandenberg & Lance, 2000). (See Table 3.)

The χ^2 change statistic did not uphold invariance. Conversely, the CFI change statistic upheld invariance. Only the RMSEA provided evidence of a good fit between the sample and the model. Thus, the test of structural invariance generated a CFI change that provided sufficient evidence to conclude structural invariance. The model fit indices indicated limited goodness-of-fit between the sample and the structural invariance model proposed by the BAPQ developers.

Model Invariance Conclusions

Overall, the results of CFA and multi-group CFAs across genders on the BAPQ provided sufficient evidence to confirm model invariance, as each level of parameter constraint upheld invariance. The tests indicated that mothers and fathers of children with ASD respond similarly to the BAPQ, according to the CFI change statistic for each test of model invariance and the χ^2 statistic for configural invariance. The model fit indices at each level of parameter constraint, however, suggest limited fit between the three-factor structure model proposed by the BAPQ developers and the observed data from SSC. The minimal goodness-of-fit evidence prompted further exploration of the factor structure model and the sample.

Exploratory Analyses

Alternate Factor Structure Models According to Extant Literature. As noted earlier, the current study plan was to investigate alternate factor structure models if limited model fit indices emerged from the three-factor structure model proposed by the BAPQ developers. One alternate model included a one-factor structure based on the research of Constantino et al. (2004); in this model, each of the 36 items on the BAPQ reflect a single latent factor, Social Responsiveness. The other alternative model described earlier involved a two-factor structure, consistent with the proposed criteria for ASD in the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* (APA, 2013) and recent literature (Frazier et al., 2008; Norris et al., 2012; Snow et al., 2009). This model assigned 18 of the items of the BAPQ to the Social Communication factor and 18 items to the Restricted, Repetitive Behaviors factor, according to item content. The one-factor and two-factor structure models did not fit the sample better than the three-factor structure model proposed by the BAPQ developers.

Alternate Models According to Modification Indices Process. The literature regarding the original conceptualization of the BAP as well as the development of the BAPQ identified an Anxiety factor (Piven et al., 1997). The BAPQ developers, however, eliminated this factor from the BAPQ, given the lack of empirical support for an association between the BAP and psychiatric disorders (Bolton et al., 1998; Piven & Palmer, 1999). In the current study, a four-factor structure model was analyzed, including the three factors proposed by the BAPQ developers as well as an Anxiety factor. The four-factor structure model cross-loaded those items suggested by the modification indices on the Anxiety factor while allowing the items to load on the original factors as proposed by the BAPQ developers.

The results of the CFAs indicated that the model fit indices improved somewhat; however, no indices other than RMR and RMSEA provided good evidence of model fit. The model led to the same conclusions regarding configural, metric, and structural invariance as the three-factor structure model. Recognizing the improved, but less than ideal, model fit, the current study examined the modification indices of the four-factor structure model. The two largest modification indices after the four-factor structure model were the same as the two largest modification indices after the three-factor model, where one seemed related to Restricted Interests and the other to Theory of Mind Deficits.

Lastly, a six-factor structure model was analyzed, including latent factors of Aloof Personality, Rigid Personality, Pragmatic Language, Anxiety, Theory of Mind deficits, and Restricted Interests. The six-factor structure model cross-loaded those items on the new factors suggested by the modification indices while allowing the items to load on the original factors as proposed by the BAPQ developers. According to the results of the CFAs, the model fit indices were clearly better than the three- and four-factor models; however, only RMR and RMSEA yielded evidence of goodness-of-fit. The GFI exceeded .90, though no other relevant indices exceeded that threshold. The six-factor structure model yielded the same conclusions regarding configural, metric, and structural invariance as the three- and four-factor structure models; however, the χ^2 change upheld configural *and* metric invariance in the six-factor structure model, rather than only the configural invariance as in the three- and four-factor structure models. As in prior models, the CFI change statistic upheld invariance at all levels of parameter constraint.

Exploratory Analyses Conclusions

While alternate models with more than three factors, as supported by the modification indices process, generated better evidence of goodness-of-fit between the model and the sample, the model fit indices remained below the customary threshold of .95 (Byrne, 2010; Hu & Bentler, 1999; Thompson, 2004; Vandenberg & Lance, 2000). Furthermore, the analyses for the four- and six-factor structure models yielded the same conclusions regarding model invariance as the three-factor structure model; therefore, the current study supported the three-factor structure model, as it is the most parsimonious model.

Discussion

The current study yielded results that upheld model invariance at each level of parameter constraint – configural invariance, metric invariance, and structural invariance. It is notable that the CFI change statistic supported model invariance at each test of model invariance; however, the χ^2 change, which is sensitive to sample size, only provided evidence to uphold configural invariance. While the current study generated evidence that upheld model invariance across genders on the BAPQ, the model fit indices at each level of parameter constraint suggested limited goodness-of-fit between the proposed model and the observed data from SSC.

The limited goodness-of-fit between the model proposed by the BAPQ developers and the observed data from SSC raises some concerns regarding the BAPQ. While the three-factor structure model remained the most parsimonious to uphold model invariance, the instrument items appear to capture latent factors beyond the three-factor structure.

Limitations

The Simons Simplex Collection (SSC) dataset provided a large sample to investigate, which offers more power, or higher probability of correctly rejecting a false null hypothesis. Several aspects of the dataset, however, introduce limitations into the study. First, the demographics of the SSC dataset reflect limited diversity. The majority of SSC parents are married, well-educated, and relatively affluent. Furthermore, African-Americans, Native Americans, and Hispanics were underrepresented. The current sample was not necessarily characteristic of the population of simplex families affected by ASD, and certainly not representative of the general population in the United States (US Census Bureau, 2013). Secondly, exclusively investigating simplex families also limits the generalizability of our findings. Indeed, Losh et al. (2008a) determined a linear expression of the BAP, with the most pronounced BAP in parents of multiplex families, then simplex families, then control families; however, they did not discuss testing assumptions of model invariance.

It is also notable that the SSC does not use DSM-IV-TR diagnostic criteria to determine diagnostic labels for the probands; the labels are most consistent with the DSM-5. The SSC employs the contemporary "gold standard" of diagnostic assessment – the Autism Diagnostic Observation Schedule (ADOS) and the Autism Diagnostic Interview-Revised (ADI-R) – which contributes to the quality of the SSC database (Klin et al., 2005); however, the ADOS and ADI-R also closely aligned with the DSM-IV-TR.

The SSC sample may be more representative of the spectrum of the future population of individuals with ASD.

Additionally, the present study only accessed data from the self-report version of the BAPQ, which was limiting; the study would have been strengthened by the addition of data from the informant-report version, as it would allow for comparisons between the self- and informant-versions as well as a calculation of the “best estimate” score for additional analyses.

Future Directions

Tests of model invariance regarding the BAPQ should be applied to samples more representative of the population of individuals impacted by ASD and their families as well as different subgroups within the population of parents of individuals with ASD. For example, the current study could be replicated with multiplex families, in which parents are raising more than one child with ASD. Finally, model invariance of the BAPQ should be examined using informant-report as well as a best estimate score based on the average of self- and informant-report responses, and perhaps could be analyzed using the informant-report versus best estimate score.

References

Amemiya, Y. & Anderson, T.W. (1990). Asymptotic chi-square tests for a large class of factor analysis models. *The Annals of Statistics, 18* (3), 1453-1463.

Anderson, T.W. & Amemiya, Y. (1988). The asymptotic normal distribution of estimators in factor analysis under general conditions. *The Annals of Statistics, 16* (2), 759-771.

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders Fifth Edition*. Washington, DC: American Psychiatric Association.

American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders Fourth Edition Text Revision*. Washington, DC: American Psychiatric Association.

Arbuckle, J. L. (2011). IBM SPSS Amos 20.0 user's guide. IBM.

Bailey, A., Le Couteur, A., Gottesman, I., Bolton, P., Simonff, E., Yuzda, E., & Rutter, M. (1995). Autism as a strongly genetic disorder: evidence from a British twin study. *Psychological Medicine, 25*, (1), 5-17.

Baron-Cohen, S. (2002). The extreme male brain theory of autism. *Trends in Cognitive Sciences, 6*, 248–254. doi:10.1016/S13646613(02)01904-6.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Autism, 2*, 296 – 301.

Bernier, R., Gerdts, J., Munson, J., Dawson, G., & Estes, A. (2012). Evidence for broader autism phenotype characteristics. *Autism Research, 5*, 13-20.

Bishop, D.V.M., Maybery, M., Maley, Al., Wong, D., Hill, W., & Hallmayer, J. (2004). Using self-report to identify the broad phenotype in in parents of children with autistic spectrum disorders: a study using the autism-spectrum quotient. *Journal of Child Psychology and Psychiatry, 45*(8), 1431-1436.

Bolton, P., Macdonald, H., Pickles, A., Rios, P., Goode, S., Crowson, M., Bailey, A., & Rutter, M. (1994). A case-control family history study of autism. *Journal of Child Psychology and Psychiatry, 35*(5)*,* 877-900.

Byrne, B. (2010). Structural equation modeling with AMOS: Basic concepts, applications, and programming. 2nd Edition. New York: Routledge.

Byrne, B.M., Shavelson, R.J., & Muthen, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin, 105* (3), 456-466.

Centers for Disease Control and Prevention. (2014). CDC estimates 1 in 68 children has been identified with autism spectrum disorder. [Press Release]. Retrieved from www.cdc/mmwr.

Centers for Disease Control and Prevention. (2012). Prevalence of Autism Spectrum Disorders-Autism and Developmental Disabilities Monitoring Network, 14 Sites, United States, 2008, *Surveillance Summaries, 61* (SS03), 1-19.

Cheung, G.W. & Rensvold, R.B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 9*(2), 235-255.

Constantino, J. (2002). *The Social Responsiveness Scale*. Los Angeles, CA: Western Psychological Services.

Constantino, J.N., Davis, S.A., Todd, R.D., Schindler, M.K., Gross, M.M., Brophy, S.L., et al. (2003). Validation of a brief measure of autistic traits in the general population. *Biological Psychiatry, 57*(6), 655-660.

Constantino, J., Gruber, C.P., Davis, S., Hayes, S., Passamante, N., Przybeck, T., et al. (2004). The factor structure of autistic traits. *Journal of Child Psychology and Psychiatry, 45,* 719-726.

Curran, P.J., West, S.G., & Finch, J.F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods, 1*(1), 16-29.

Davidson, J., Goin-Kochel, R.P., Green-Snyder, L.A., Hundley, R.J., Warren, Z., Peters, S.U. (2012). Expression of the broad autism phenotype in simplex autism families from the Simons Simplex Collection. *Journal of Autism and Developmental Disorders*. Online First Publication. doi: 10.1007/s10802-012-1492-1

Dawson, G., Estes, G., Munson, J., Schellenberg, G, Bernier, R., and Abbott, R. (2007). Quantitative assessment of autism symptom-related traits in probands and parents: broader phenotype autism symptom scale. *Journal of Autism and Developmental Disorders, 37*, 523-536.

Fidell, L.S. & Tabachnick, B.G. (2003). Preparatory data analysis. In Weiner, I.B., Schinka, J.A., & Velicer, W.F. (Eds.), *Handbook of Psychology* (pp. 115-141). Hoboken, NJ: John Wiley & Sons.

Fischbach, G. & Lord, C. (2010). The Simons Simplex Collection: A resource for identification of autism genetic risk factors. *Neuron, 68*(2), 192-195.

Folstein, S. & Rutter, M. (1977). Infantile autism: A genetic study of 21 twin pairs. *Journal of Child Psychology and Psychiatry, 18*(4), 297-321.

Frazier, T.W., Youngstrom, E.A., Kubu, C.S., Sinclair, L., & Rezai, A. (2008). Exploratory and confirmatory factor analysis of the Autism Diagnostic Interview-Revised. *Journal of Autism and Developmental Disorders, 38* (3), 474-480.

Geschwind, D. (2011). Genetics of autism spectrum disorders. *Trends in Cognitive Sciences, 15* (9), 409-416.

Goodwin, P.Y., Mosher W.D., & Chandra, A. (February 2010). Marriage and cohabitation in the United States: A statistical portrait based on Cycle 6 (2002) of the National Survey of Family Growth. National Center for Health Statistics. *Vital and Health Statistics, 23*(28).

Hallmayer, J., Clevelan, S., Torres, A., Phillips, J., Cohen, B., Torigoe, T. et al. (2011). Genetic heritability and shared environmental factors among twin pairs with autism. *Archives of General Psychiatry, 68* (11), 1095-1102.

Hu, L. & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6* (1), 1-55. doi: 10.1080/10705519909540118

Hurley, E., Losh, M., Parlier, M., Reznick, J.S., & Piven, J. (2007). The broad autism phenotype questionnaire. *Journal of Autism and Developmental Disorders, 37*(9), 1679-1690.

Ingersoll, B., Hopwood, C.J., Wainer, A., & Donellan, M.B. (2011). A comparison of three self-report measures of the broader autism phenotype in a non-clinical sample. *Journal of Autism and Developmental Disorders, 41*, 1646-1657.

Klin, A., Saulnier, C., Tsatsanis, K., & Volkmar, F.R. (2005). Clinical evaluation in Autism Spectrum Disorders: psychological assessment within a transdisciplinary framework. In F.R. Volkmar, R. Paul & A. Klin (Eds.). *The handbook of autism and pervasive developmental disorders, volume 2: assessment, interventions, and policy, 3rd Edition* (pp. 772-798). New York, NY: Wiley Publishing.

Landa, R., Piven, J., Wzorek, M., Gayle, J., Cloud, D., Chase, G. & Folstein, J. (1992). Social language use in the parents of autistic individuals. *Psychological Medicine, 22*, 245-254.

Lord, C., Rutter, M., DiLavore, P.C., Risi, S. (2007). *Autism Diagnostic Observation Schedule.* WPS Edition Manual. Los Angeles, CA: Western Psychological Services.

Losh, M., Childress, D., Lam, K., & Piven, J. (2008a). Defining key features of the broad autism phenotype: a comparison of multiple- and single-incidence families. *American Journal of Medical Genetics, Part B (Neuropsychiatric Genetics)*, *147B*, 424-433.

Losh, M., Sullivan, P., Trembath, D., Piven, J. (2008b). Current developments in the genetics of autism: from phenome to genome. *Journal of Neuropathology and Experimental Neurology, 67* (9), 829-837.

Murphy, M., Bolton, P.F., Pickles, A., Fombonne, E., Piven, J., and Rutter, M. (2000). Personality traits of the relatives of autistic probands. *Psychological Medicine, 30*, 1411-1424.

Norris, M., Lecavalier, L., Edwards, M.C. (2012). The structure of autism symptoms as measured by the autism diagnostic observation schedule. *Journal of Autism and Developmental Disorders, 42* (6), 1075-1086.

Ozonoff, S., Young, G.A., Carter, A., Messinger, D., Yirmiya, N., Zwaigenbaum, L., et al. (2011). Recurrence risk for autism spectrum disorders: A baby siblings research consortium study. *Pediatrics, 128*, e488-e495.

Piven, J. (2001). The broad autism phenotype: a complementary strategy for molecular genetic studies of autism. *American Journal of Medical Genetics (Neuropsychiatric Genetics), 105* (1), 34-35.

Piven, J. & Palmer, P. (1999). Psychiatric disorder and the broad autism phenotype: evidence from a family study of multiple-incidence autism families. *American Journal of Psychiatry, 156* (4), 557-564.

Piven, J., Palmer, P., Jacobi, D., Childress, D., & Arndt, S. (1997a). Broader autism phenotype: evidence from a family history study of multiple-incidence families. *American Journal of Psychiatry, 154* (2), 185-190.

Piven, J., Palmer, P., Landa, R., Santangelo, S., Jacobi, D., & Childress, D. (1997b). Personality and language characteristics in parents from multiple-incidence

autism families. *American Journal of Medical Genetics (Neuropsychiatric Genetics), 74,* 398-411.

Piven, J., Wzorek, M., Landa, R., Lanhart, J., Bolton, P. Chase, G. & Folstein, S. (1994). Personality characteristics and language characteristics in parents from multiple incidence autism families. *Psychological Medicine, 24*, 783-795.

Rogers, S.J. (2009). What are infant siblings teaching us about autism in infancy? *Autism Research 2* (3), 125-137.

Rosenberg, R.E., Law, J.K., Yenokyan, G., McGready, J., Kaufman, W.E., & Law, P.A. (2009). Characteristics and concordance of autism spectrum disorders among 277 twin pairs. *Archives of Pediatric and Adolescent Medicine, 163* (10), 907–914.

Ruser, T.F., Arin, D., Dowd, M., Putnam, S., Winklosky, B., Rosen-Sheidley, B., Piven, J., Tomblin, B., Tager-Flusberg, H., Folstein, S. (2007). Communicative competence in parents of children with autism and parents of children with specific language impairment. *Journal of Autism and Developmental Disorders, 37*, 1323-1336.

Rutter, M., Le Couteur, A., & Lord, C. (2008). *Autism Diagnostic Interview–Revised:* WPS Edition Manual. Los Angeles, CA: Western Psychological Services.

Sasson, N.J., Lam, K.S., Childress, D., Parlier, M., Daniels, J.L., & Piven, J. (2013). The broad autism phenotype questionnaire: prevalence and diagnostic classification. *Autism Research,* Epublication ahead of print. doi: 10.1002/aur.1272.

Seidman, I., Yirmiya, N., Milshtein, S., Ebstein, R.P., & Levi, S. (2012). The broad autism phenotype questionnaire: Mothers versus fathers of children with an autism spectrum disorder. *Journal of Autism and Developmental Disorders, 42,* 837-846.

Simons Foundation Autism Research Initiative (SFARI) (September 15, 2012). SFARI Initiatives. Retrieved from (http://sfari.org/sfari-initiatives).

Snow, A.V., Lecavalier, L., & Houts, C. (2009). The structure of the Autism Diagnostic Interview-Revised: Diagnostic and phenotypic implications. *Journal of Child Psychology and Psychiatry*, *50*(6), 734-742.

Sucksmith, E., Roth, I., Hoekstra, R.A. (2011). Autistic traits below the clinical threshold: Re-examining the Broader Autism Phenotype in the 21st century. *Neuropsychology Review, 21*, 360-389.

State, M. (Presenter). (September 10, 2010). *Genetics of Autism.* [Itunes U Video Seminar]. Retrieved from the Itunes U Yale Autism and Related Disorders Seminar series.

Thompson, B. (2004). Exploratory and confirmatory factor analysis: Understanding concepts and applications. Washington, DC: American Psychological Association.

United States Census Bureau. (March 14, 2013). State and County QuickFacts: USA QuickFacts. [Population Estimates from American Community Survey]. Retrieved from http://quickfacts.census.gov/qfd/states/00000.html.

Vandenberg, R.J. & Lance, C.E. (2000). A review and synthesis of the measurement invariance literature: suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3* (4), 4-70.

Yin, P. & Fan, X. (2003). Assessing the factor structure invariance of self-concept measure across ethnic and gender groups: Findings from a national sample. *Educational and Psychological Measurement, 63* (2), 296-318.

Appendix

Table 1

Model Assessment of Configural Invariance

	CMIN	CMIN df	GFI	AGFI	CFI	TLI	RMSEA	RMR
Ungrouped Unconstrained Model	11613.390	591	.836	.815	.805	.792	.066*	.093*
Grouped Unconstrained Model	12117.620	1182	.833	.811	.799	.786	.047*	.094*
Difference	504.23	591			.006			

*Indicative of goodness-of-fit between model and observed data

Table 2

Model Assessment of Metric Invariance

	CMIN	CMIN df	GFI	AGFI	CFI	TLI	RMSEA	RMR
Unconstrained Model	12117.620	1182	.833	.811	.799	.786	.047*	.094*
Factor Loadings Constrained Model	12280.544	1218	.831	.815	.797.	.790	.046*	.102
Difference	162.924	36			.002			

*Indicative of goodness-of-fit between model and observed data

Table 3

Model Assessment of Structural Invariance

	CMIN	CMIN df	GFI	AGFI	CFI	TLI	RMSEA	RMR
Factor Loadings Constrained Model	12280.544	1218	.831	.815	.797.	.790	.046*	.102
Factor Convariances Constrained Model	12291.914	1221	.831	.816	.797	.790	.046*	.101
Difference	11.37	3			.002			

*Indicative of goodness-of-fit between model and observed data

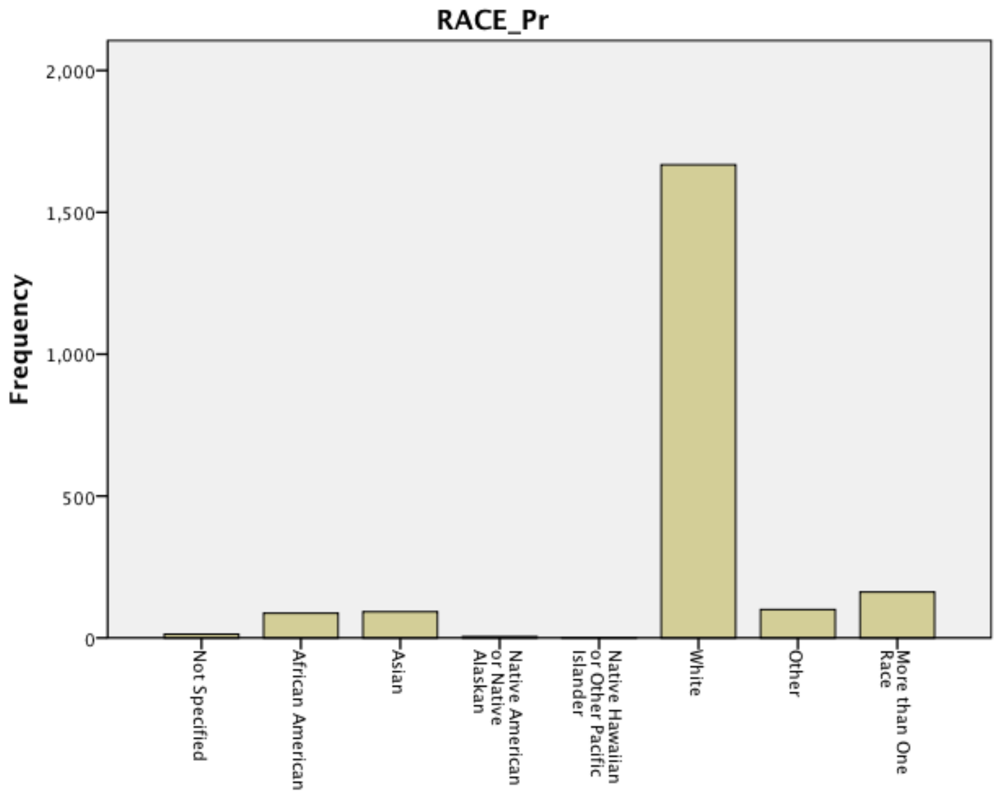

Figure 1. *Racial breakdown of probands.*

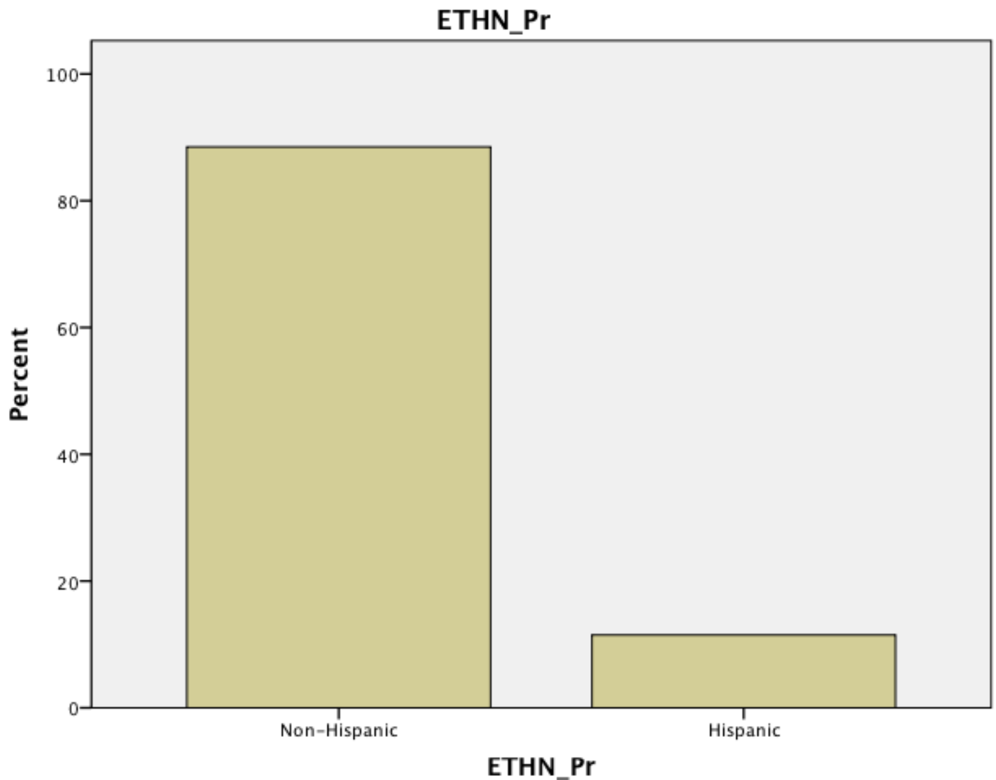

Figure 2. *Ethnic breakdown of probands.*

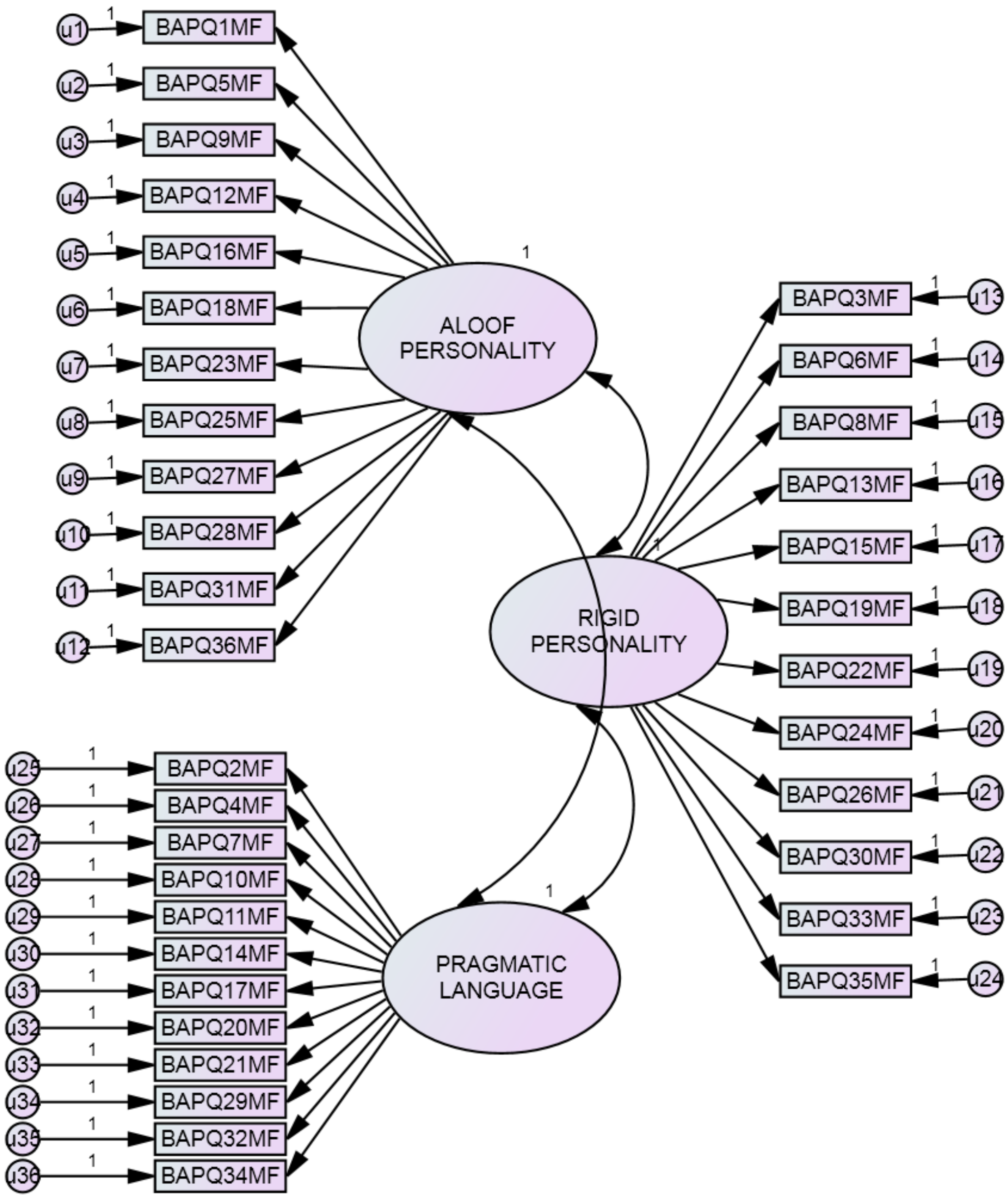

Figure 3. Broad Autism Phenotype Questionnaire baseline model.

Examination of the Factor Structure of the Autism Diagnostic Interview-Revised (ADI-R) within a Simplex Population

Jordan L. Wade & Ronald E. Reeve

University of Virginia

Abstract

Utilizing the Simons Simplex Collection (SSC), this study sought to explore the factor structure of the ADI-R within a verbal simplex population (n = 2,286). Several different models were compared to determine the best fit. Results provided support for an adjusted three-factor model in which Impaired Play constituted the third domain. Multiple-group analyses using ADOS modules 2 and 3 revealed support for configural invariance, but violations of both metric and structural invariance. Given these findings, partial invariance was explored, which ultimately yielded a model in which almost all of the subdomains were freely estimated, suggesting that the majority of the subdomains within the ADI-R are not operating equivalently across ADOS modules 2 and 3. Implications suggest a need for revision of the ADI-R factor structure, as well as additional research exploring the complexities of ASD symptoms as they present across differing expressive language and developmental levels.

Examination of the Factor Structure of the ADI-R within a Simplex Population

Autism spectrum disorder (ASD) is a neurodevelopmental disorder that is estimated to affect one in 68 individuals (Centers for Disease Control and Prevention [CDC], 2014). Despite its routine use within the literature, as well as colloquially, ASD was not an official diagnosis until the publication of the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* (DSM-5; American Psychiatric Association [APA], 2013). The previous version, the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition – Revised* (DSM-IV-TR; APA, 2000), included five diagnoses that were subsumed within the Pervasive Developmental Disorders (PDD) category: Autistic Disorder, Asperger's Disorder, Pervasive Developmental Disorder – Not Otherwise Specified (PDD-NOS), Childhood Disintegrative Disorder (CDD), and Rett's Disorder (APA, 2000).

Research suggesting that clinician judgment is less reliable in making these diagnostic distinctions, even when combined with standardized assessment measures (Fischbach & Lord, 2010; Lord et al., 2012a), led to four of the five PDD diagnoses[1] being merged into one condition, Autism Spectrum Disorder, in the DSM-5 (APA, 2013). Additionally, diagnostic criteria were collapsed into two domains (as opposed to the three in the DSM-IV-TR): 1) impairments in social communication, and 2) presence of restricted interests and/or repetitive behaviors.

There are multiple measures that have been developed to evaluate the presence of ASD, however, the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2000; Lord, Rutter, DiLavore, & Risi, 2001) and the Autism Diagnostic Interview-Revised

---

[1] Due to its clear genetic etiology, Rett's Disorder was not included in the DSM-5.

(ADI-R; Lord, Rutter, & Le Couteur, 1994; Rutter, Le Couteur, & Lord, 2008) are considered the “gold standard” of these tools (Ozonoff, Goodlin-Jones, & Solomon, 2005).

Autism Diagnostic Interview – Revised (ADI-R)

The ADI-R (Lord et al., 1994; Rutter et al., 2008) is a structured diagnostic interview that is administered to caregivers of individuals suspected of having ASD. The ADI-R can be administered to obtain information about an individual of any age; however, the person’s mental abilities must be at least the equivalent of two years, zero months (Rutter et al., 2008). According to the manual, the interview takes 1.5 to 2.5 hours to administer and score if given by a trained clinician.

The measure is based on DSM-IV (APA, 1994) and ICD-10 (World Health Organization [WHO], 1990) diagnostic criteria and includes a full developmental history; it has been extensively researched and validated (Lord et al., 1994). It consists of 93 items, which yield three core diagnostic domains – Qualitative Abnormalities in Reciprocal Social Interaction (“A”), Qualitative Abnormalities in Communication (“B”), and Restricted, Repetitive, and Stereotyped Patterns of Behavior (“C”) – in addition to a fourth domain, Abnormality of Development Evident at or Before 36 Months (“D”). The ADI-R produces two scores, one from the Diagnostic Algorithm and the second from the Current Behavior Algorithm. The diagnostic score is based on the individual’s behavior at ages 4-5; this particular age period was chosen because the behaviors evaluated by the ADI-R are typically displayed by all children by this time period, even those with intellectual disabilities who are functioning significantly below their chronological age (Hus & Lord, 2013). To obtain a diagnosis of autism, as conceptualized by the measure, the individual must meet cut-off criteria in all four diagnostic domains. In contrast, the Current Behavior Algorithm is based on an individual's functioning over the most recent months and is a less formal score. The measure's authors suggest that the current score be utilized for treatment or educational planning as it has less research validating its use (Rutter et al., 2008).

The ADI-R yields two diagnostic classifications, autism or no autism, as it was developed to provide a categorical diagnosis rather than supply information regarding symptomology severity (Tadevosyan-Leyfer et al., 2003). However, some have argued that meeting the cut-off score on the Restricted, Repetitive, and Stereotyped Patterns of Behavior domain, particularly for young children, is too stringent and results in decreased sensitivity of the measure (Ventola et al., 2006). To address this concern, many researchers have added an ASD descriptor to the ADI-R, which is operationalized as meeting cut-off criteria on two of the three domains (Bishop & Norbury, 2002; Le Couteur, Haden, Hammal, & McConachie, 2008).

The ADI-R has many strengths, among these perhaps most importantly a substantial body of research substantiating its validity, however one of its main limitations is parental bias, as underreporting may be a factor in lower scores (Chawarska, Klin, Paul, & Volkmar, 2007; Mildenberger, Sitter, Noterdaeme, & Amorosa, 2001; Noterdaeme, Mildenberger, Sitter, & Amorosa, 2002). Additionally, the measure is quite lengthy, which can make it difficult to utilize within clinical settings given that there are often time constraints (Ozonoff et al., 2005). However, despite these few limitations, the ADI-R remains one of the most metrically researched measures for the assessment of ASD and is frequently used for clinical and research purposes.

Autism Diagnostic Observation Schedule (ADOS)

The ADOS (Lord et al., 2000; Lord et al., 2001) is a semi-structured play assessment that evaluates social reciprocity and communication skills. The first version was developed for children ages five to 12 with a minimum expressive language level of a neurotypical three-year-old (Lord et al., 1989). Given the need to assess younger children and those with limited to no verbal skills, the Pre-Linguistic Autism Diagnostic Observation Scale (PL-ADOS) was developed (DiLavore, Lord, & Rutter, 1995). Additionally, a revised version of the ADOS, the ADOS-Generic (ADOS-G; Lord et al., 2000) was created with added activities for older adolescents and adults. This version of the ADOS consists of four modules based on expressive language level; Module 1 is administered to nonverbal individuals, Module 2 to those with phrase speech, Module 3 to children and adolescents with fluent speech, and Module 4 to adults with fluent speech. A trained clinician provides "social presses" (Lord et al., 1989) to evoke social interactions and communicative behaviors. Although restricted interests and repetitive behaviors are documented and scored, they are not factored into the Diagnostic Algorithm because the administration time is so short, decreasing the likelihood that these behaviors will be observed (Lord et al., 2001). However, studies have found that the presence of repetitive behaviors and/or restricted interests during the ADOS administration is diagnostically significant (LeCouteur et al., 2008).

Studies have demonstrated that the ADOS is quite good at distinguishing between children who have autism and those who do not, but is less effective at categorizing individuals who exhibit less severe symptomatology (Bishop & Norbury, 2002; de Bildt et al, 2004; Lord et al., 2000). Revised algorithms, however, have yielded improved sensitivity and specificity of the measure (Gotham et al., 2007). Although the ADOS-2 (Lord, Luyster, Gotham, & Guthrie, 2012b; Lord et al., 2012c) is the most recently published edition of the measure, this was not the version utilized in the SSC dataset as it had been developed at the time of data collection. However, it is important to note that the structure and the items of the ADOS-2 are fairly similar to those found on the ADOS; the main difference between the two versions is that the ADOS-2 has a revised Diagnostic Algorithm that is based upon DSM-5 rather than DSM-IV-TR diagnostic criteria.

Combined Use of the ADI-R and the ADOS

Both the ADOS and the ADI-R were developed to be used in tandem, and are based on a three-factor symptom structure, which was derived primarily from clinical observations rather than empirical evidence (Norris et al., 2012; Szatmari et al., 2006). Of the diagnostic tools used to assess ASD, the ADOS and the ADI-R have the strongest empirical support, with high sensitivity and specificity across studies, particularly in the identification of more classic autism (Le Couteur, et al., 2008; Lord et al., 2001; Lord et al., 1994). Research has demonstrated somewhat lower sensitivity and specificity for both measures in detection of the broader spectrum (Bishop & Norbury, 2002; de Bildt et al., 2004; Lord et al., 2000). While there is a substantial body of literature evaluating the validity and reliability of the ADOS and ADI-R individually, research directly comparing the two measures is scant (Bishop & Norbury, 2002; Le Couteur et al., 2008).

Much of the research examining the diagnostic utility of the ADOS and ADI-R in combination pertains to the instruments' use with young children (Chawarska et al., 2007; Kim & Lord, 2012a; Kim & Lord, 2012b; Le Couteur et al., 2008; Gray et al., 2008; Ventola et al., 2006; Wiggins & Robins, 2008). Ventola and colleagues (2006) compared four diagnostic measures (ADI-R, ADOS-G, Childhood Autism Rating Scale [CARS], and clinical judgment) with toddlers (16-30 months) to examine their level of agreement. The ADOS-G, CARS, and clinical judgment all demonstrated high agreement with one another, whereas the ADI-R did not demonstrate significant agreement with any of the measures. The ADI-R displayed low sensitivity in diagnosing Autistic Disorder (.556) compared to the ADOS (.899) and the CARS (.963) (Ventola et al., 2006). This finding was attributed to the stringent criteria regarding the presence of restricted interests and repetitive behaviors, which is not present on the other measures (Ventola et al., 2006). Indeed, a study by Wiggins and Robins (2008) demonstrated that, among toddlers aged 16-37 months, agreement between the ADOS and ADI-R improved from poor to fair when the behavioral domain was excluded (.521 to .754).

Other studies that have explored the level of agreement between the ADI-R and ADOS have found modest to good agreement, ranging from 63.6% to 80% (Bishop & Norbury, 2002; de Bildt et al., 2004; Gray, Tonge, & Sweeney, 2008; Papanikolaou et al., 2009; Ventola et al., 2006). The variation in level of agreement is often attributed to the differing sources of information – the ADOS is a play assessment that is scored based on the clinician's observations, whereas the ADI-R is a semi-structured interview administered to a parent or caregiver (de Bildt et al., 2004). Additionally, the two measures evaluate different time periods; the ADI-R Diagnostic Algorithm is based on the individual's functioning during the ages of 4-5, whereas the ADOS captures current functioning (de Bildt et al., 2004). Others have found that the level of agreement increases with the addition of an adaptive functioning measure, suggesting that there are behaviors not directly captured by the ADOS or the ADI-R that are diagnostically important (Tomanik, Pearson, Loveland, Lane, & Shaw, 2007).

Studies exploring the individual domains of each instrument have found weak correlations between the repetitive behaviors domains of the ADOS and ADI-R, but strong associations between the social interaction and communication domains (de Bildt et al., 2004; Le Couteur et al., 2008). In a study of young children aged 24-49 months, Le Couteur and colleagues (2008) found strong correlations between the social interaction and communication domains of the ADOS and ADI-R (r = .71 and .64, respectively), but a significantly weaker association between the repetitive behavior scores (r = .51). Similar correlations between the ADOS and ADI-R were found in a group of older children and adolescents, ages 5 to 20, with pervasive developmental disorders and intellectual disability (de Bildt et al., 2004). Weak correlations between the ADOS and the ADI-R behavioral domain may be due to the limited observational period of the ADOS, thus decreasing the likelihood of witnessing these behaviors (Le Couteur et al., 2008). Alternatively, they may indicate that the algorithm cut-off score is too stringent on the ADI-R (Ventola et al., 2006). Lastly, there is some research suggesting that the behavior domain may actually be comprised of two domains rather than one (insistence on sameness, and repetitive sensory and motor behaviors) suggesting greater complexity in their presentation (Szatmari et al., 2006).

Factor Structure of ADI-R

In addition to its diagnostic utility, the ADI-R has been used within the literature to better understand the nosology of ASD (Frazier, Youngstrom, Kabu, Sinclair, & Rezai, 2008; Hus, Pickles, Cook, Jr., Risi, & Lord, 2007; Snow, Lecavalier, & Houts, 2009; Van Lang et al., 2006). Given that some consider it to be the most comprehensive ASD measure, it is an ideal candidate for factor analysis (Snow et al., 2009). The ADOS has also been utilized for this purpose, though to a lesser degree (Gotham et al., 2007; Norris et al., 2012; Robertson et al., 1999).

Although several studies have examined the factor structure of the ADI-R (Boomsma et al., 2008; Frazier et al., 2008; Georgiades et al., 2007; Kamp-Becker, Ghahreman, Smidt, & Remschmidt, 2009; Lecavalier et al., 2006; Snow et al., 2009; Tadevosyan-Leyfer et al., 2003; Van Lang et al., 2006), the methods in which this has been done have varied considerably, as well as the demographics of the population studied. Researchers have utilized a subset of individual items (Constantino et al., 2004; Snow et al., 2009; Tadevosyan-Leyfer et al., 2003), just the algorithm items (Lecavalier et al., 2006), and the subdomains (Frazier et al., 2008; Georgiades et al., 2007; Van Lang et al., 2006). The decision regarding which items and/or subdomains to include in analyses is driven by the purpose of the study. Factor analysis of the subdomains provides information particularly pertinent to understanding the psychometric properties of the ADI-R, whereas analysis of the individual items is useful in attempts to better understand the broader autism phenotype (Snow et al., 2009). Within the literature, the most frequently studied models are one-, two-, and three-factor structures due to theoretical support for each, although the strength of this support varies.

Several studies that have examined the factor structure of the ADI-R have found support for a symptom structure different from that proposed by the DSM-IV/DSM-IV-TR (Boomsma et al., 2008; Frazier et al., 2008; Van Lang et al., 2006). Van Lang and colleagues (2006) found that, among 255 individuals with autism with a range of verbal skills and symptomatology severity, the following three factors emerged as the strongest fit for the data: impaired social communication skills, stereotyped language behaviors, and impaired make-believe and play skills. The same group of authors were able to replicate these findings with a different sample, providing further support for the model (Boomsma et al., 2008). Similarly, Georgiades and colleagues (2006) found evidence for social communication and inflexible language and behavior domains using principal components analysis (PCA) in a sample of 209. However, in contrast to the findings by Van Lang et al. (2006) and Boomsma et al. (2008), those of Georgiades et al. (2006) yielded a third domain comprising repetitive sensory and motor behaviors as opposed to impaired make-believe and play skills. Using a sample of 292 individuals, Tadevosyan-Leyfer et al. (2003) conducted a PCA of 98 ADI-R items, which yielded six factors: spoken language, social intent, compulsions, milestones, savant skills, and sensory aversions. Although six factors emerged, many of these included items from both the social interaction and communication domains, leading the researchers to suggest that a two-factor diagnostic model may be more appropriate than the three-factor model given the significant overlap between social interaction and communicative behaviors (Tadevosyan-Leyfer et al., 2003). Support for a two-factor model as opposed to a three-factor has been proposed by others, as well (Frazier et al., 2008; Snow et al., 2009).

Other studies utilizing a much larger sample (*n* = 1,170) have also found that stereotyped language has a stronger loading on restricted interests and repetitive behaviors compared to communication, in addition to providing some support for a separate factor of impairments in peer relationships and imaginative play (Frazier et al., 2008). Of note, despite the requirement of large sample sizes for structural equation modeling (Kline, 2011), few studies examining the factor structure of the ADI-R have utilized very large samples (Frazier et al., 2008). Additionally, studies have demonstrated that factor analysis of the ADI-R based on the DSM-IV-TR model is prone to estimation problems, as evidenced by the high correlation between the social interaction and communication domains (Boomsma et al., 2008; van Lang et al., 2006). A notable exception to these findings includes results reported by Lecavalier et al. (2006); using a sample of 226 children, their analyses demonstrated support for a factor model similar to DSM-IV-TR diagnostic criteria, excluding the non-verbal communication items, which loaded more strongly onto the social interaction domain. This study has yielded the most support for the original algorithm.

Taken together, these results indicate that there are mixed findings regarding the most parsimonious model, but that DSM models (based on DSM-IV/DSM-IV-TR diagnostic criteria) often yield poorer fit than other models when examining the factor structure of the ADI-R. Lastly, some studies suggest that the utility of a two-factor model is similar to that of a three-factor model, signifying the need to reevaluate the ADI-R subdomains and items (Frazier et al., 2008; Snow et al., 2009).

In addition to studying the factor structure of the ADI-R in its entirety (i.e., using all three domains), other research has examined specific domains in isolation (Robertson et al., 1999; Szatmari et al., 2006; Tanguay, Robertson, & Derrick, 1998). Analysis of the factor structure of 28 items on the social and communication domains of the ADI-R yielded three factors: affective reciprocity, joint attention, and theory of mind (Robertson et al., 1999; Tanguay et al., 1998). Szatmari and colleagues (2006) examined the factor structure of the restricted, repetitive behaviors and interests domain and found that it yielded two dimensions: insistence on sameness and repetitive sensory and motor behaviors. These findings suggest heterogeneity within the three core domains of the ADI-R, indicating that the DSM-IV/DSM-IV-TR symptom structure may be overly simplistic.

Purpose of Study

The overarching aim of this study was twofold: 1) to add to the current literature regarding the psychometric properties of the ADI-R, and 2) to evaluate invariance of the ADI-R factor structure across ADOS modules. Additionally, there is limited research examining the factor structure of the ADI-R within larger samples (Frazier et al., 2008). As such, there is a clear need to study the factor structure of the ADI-R among larger, more heterogeneous samples in order to better understand the nosology of ASD to improve the utility of this widely used diagnostic tool (Frazier et al., 2008; Snow et al., 2009). Furthermore, given the limited research examining the ADI-R and ADOS simultaneously (Falkmer et al., 2013), there is a need for further exploration of this topic. Despite being classified as the “gold standard” for ASD assessment (Ozonoff et al., 2005), research investigating their combined use is surprisingly limited (Bishop & Norbury, 2002; Falkmer et al., 2013). A review of the literature since 2000 examining the empirical support and utility of various ASD diagnostic measures found 11 studies that directly compared the ADOS and the ADI-R (Falkmer et al., 2013). However, many of these studies had small sample sizes (Bishop & Norbury, 2002; Chawarska et al., 2007; Ventola et al., 2006), and several focused specifically on the use of these instruments among very young children (Kim & Lord, 2012; Le Couteur et al., 2008; Ventola et al., 2006), thus reducing the generalizability of the results.

Given these gaps within the literature, this study sought to further understanding of the psychometric properties of the ADI-R using a simplex population. The phenotypic and genotypic presentations of individuals from simplex and multiplex families are thought to differ, highlighting the importance of studying them separately (Fischbach & Lord, 2010). The factor structure of the ADI-R has previously been studied among a multiplex population using data from the Autism Genetic Resource Exchange (AGRE; Frazier et al., 2008; Snow et al., 2009). In contrast, this study analyzed data from simplex families using the Simons Simplex Collection (SSC); to the author's knowledge, this study is the first to examine the factor structure of the ADI-R utilizing a large sample with this particular population.

Based on gaps within the current literature, this study proposed a series of questions. First, what is the level of agreement between the ADOS and ADI-R using data from simplex families? Consistent with existing literature, it was hypothesized that the ADOS and ADI-R would demonstrate modest to good agreement. Second, within a simplex population, what factor structure of the ADI-R best fits the data? Based on previous findings, it was hypothesized that the ADI-R would yield two- and three-factor models with similar goodness-of-fit. Lastly, does the factor structure of the ADI-R operate equivalently across ADOS modules 2 and 3? No hypothesis regarding this question was generated given the lack of existing research related to this area.

Methods

Participants

Data for this study were utilized from the Simons Simplex Collection (SSC), a project of the Simons Foundation Autism Research Initiative (SFARI). SSC data were collected from clinics at twelve sites, all major universities, across the country. Children between the ages of 4 and 18 years with a diagnosis of Autistic Disorder, Asperger's Disorder, or PDD-NOS were recruited. Individuals with a nonverbal mental age below 18 months were excluded, as were those with significant perinatal incidents, neurological deficits, or birth trauma. Those diagnosed with fragile X or Down syndrome were also omitted (Fischbach & Lord, 2010). Exclusionary criteria included first-, second-, or third-degree relatives with an ASD diagnosis.

The dataset utilized for this study consisted of 2,594 children and adolescents after screening for missing data, as well as univariate and multivariate outliers. Specific demographic data including age and gender for each sub-sample can be found in Table 1, as well as diagnostic and cognitive profiles. Of note, the majority of parents reported a marital status of married (90.1%), and both mothers and fathers were well educated, with the majority reporting an associate degree or higher level of education (90.2% of mothers and 84.8% of fathers). Lastly, parents reported relatively high incomes compared to the national average, with 68.8% indicating a yearly household income of $66,000 or higher.

Measures

Autism Diagnostic Interview-Revised (ADI-R). The ADI-R (Lord et al., 1994; Rutter et al., 2008) is a 93-item semi-structured interview that is administered by a trained clinician to a parent or caregiver. Of the 93 items, 34 contribute to the Diagnostic Algorithm, which is broken down into four domains based on DSM-IV (APA, 1994) and ICD-10 (WHO, 1990) diagnostic criteria for autism. These domains are: Qualitative Abnormalities in Reciprocal Social Interaction ("A"), Qualitative Abnormalities in Communication ("B"), Restricted, Repetitive, and Stereotyped Patterns of Behavior ("C"), and Abnormality of Development Evident at or Before 36 Months ("D"). Of note, cluster B is further divided into two groups – verbal and nonverbal subjects.

Cluster A consists of 15 items, which are grouped to create four behaviors: A1: Failure to use nonverbal behaviors to regulate social interaction; A2: Failure to develop peer relationships; A3: Lack of shared enjoyment; and A4: Lack of socioemotional reciprocity. Cluster B consists of 13 items for verbal subjects and seven items for nonverbal subjects. These items are grouped to create four behaviors: B1: Lack of, or delay in, spoken language and failure to compensate through gesture; B4: Lack of varied spontaneous make-believe or social imitative play; B2: Relative failure to initiate or sustain conversational interchange; and B3: Stereotyped, repetitive, or idiosyncratic speech. B1 and B4 are used for nonverbal subjects, whereas all four behaviors are used for verbal subjects. Lastly, Cluster C consists of eight items, which are grouped to create four behaviors: C1: Encompassing preoccupation or circumscribed pattern of interest; C2: Apparently compulsive adherence to nonfunctional routine or rituals; C3: Stereotyped and repetitive motor mannerisms; and C4: Preoccupation with parts of objects or nonfunctional elements of material. Item codes include 0, 1, 2, 3, 7, 8 and 9;

these are then converted to algorithm scores ranging from 0 to 2. Zero indicates no abnormality; 1 indicates some abnormality, but not necessarily consistent with ASD; 2 indicates definite abnormality, consistent with ASD; and 3 indicates markedly abnormal behavior. Scores of 7, 8, and 9 are converted to zeroes; 7 indicates abnormality not relevant to the behavior being evaluated, 8 designates no applicability, and 9 is used when the response is unknown, or the item is not asked. The mean scores and standard deviations for the ADI-R domains and subdomains by sub-sample can be viewed in Table 2.

Autism Diagnostic Observation Schedule (ADOS). The ADOS (Lord et al., 2000; Lord et al., 2001) is a standardized play assessment that is administered by a trained clinician. There are four modules, which are based on expressive language level and developmental functioning. Module 1 is administered to children who are nonverbal or have single words, Module 2 is given to children with phrase speech, Module 3 is administered to children and adolescents with fluent speech, and Module 4 is given to adolescents and adults with fluent speech. The algorithm is broken down into four domains based on DSM-IV (APA, 1994) and ICD-10 (WHO, 1990) diagnostic criteria for autism; however, the diagnostic threshold is based on only two domains, "Social Interaction" and "Communication," as well as a combination of these domains ("Communication-Social Interaction"). Item codes include 0, 1, 2, 3, 7, and 8; these are then converted to algorithm scores ranging from 0 to 2. Zero indicates no abnormality; 1 indicates some abnormality, but not necessarily consistent with ASD; 2 indicates definite abnormality consistent with ASD; and 3 indicates markedly abnormal behavior. Scores of 7 and 8 are converted to zeroes; 7 indicates abnormality not relevant to the behavior being evaluated, and 8 is used when the rating is not applicable, or the behavior was not observed.

Analyses

Data screening. The original dataset included 2,643 children and adolescents. Inspection of data yielded two cases that were missing ADI-R values; these cases were deleted, bringing the sample to 2,641. As outlined in the proposal, the data were then divided based on verbal abilities (verbal and nonverbal). Additionally, separate datasets based on ADOS modules 1-3 were created in preparation for multiple-group analyses. However, upon further inspection of the data it became apparent that including Module 1 would require alteration of the model, resulting in the analysis of 10 subdomains rather than the full 12. Although Module 1 does include some verbal children (single words), the nature of subdomains B2 and B3 made their inclusion relatively meaningless as they measure conversational abilities and stereotyped speech. Additionally, inspection of the Module 1 data (including the verbal subset, nonverbal subset, and the combined group) demonstrated extreme violations of multivariate normality, which remained even after transforming the variables. Given these various factors, Module 1 was not included in the multiple-group analyses. Although the inclusion of Module 4 made substantive sense, it was not used due to a limited number of subjects (n = 73), with the sample size falling substantially below the minimum 200 recommended for structural equation modeling (Kline, 2011).

Additionally, analyses utilizing the nonverbal sample were not pursued. Inspection of the data yielded a significant number of univariate and multivariate outliers exceeding 5% of the data. Even after the data were transformed, Mardia's statistic still indicated violations of multivariate normality. Additionally, it was not possible to run multiple-group analyses with this sample given that modules 2 and 3 are only administered to verbal children. Given these issues, no further examination of the nonverbal sample was undertaken.

Before conducting factor analyses, assumptions were checked within each dataset. Univariate and multivariate outliers were identified, followed by a check of univariate and multivariate normality. Univariate outliers were detected by standardizing the ADI-R subdomain scores and identifying values less than or greater than 3.29 (Field, 2009). Within the ADI-R verbal dataset (n = 2,327), 41 outliers across subdomains B2 and A2 were identified. Given that these only constituted 1.8% of the data, they were deleted. Analyses yielded no multivariate outliers. As such, the final ADI-R verbal dataset included 2,286 children and adolescents. In regard to univariate normality, skewness and kurtosis both fell within acceptable limits. Multivariate normality was also upheld (Mardia's statistic = 4.13).

Within the Module 2 dataset (n = 566), 21 univariate outliers were deleted, all of which were identified on the B2 subdomain. A check for multivariate outliers yielded 5 cases exceeding the critical chi-square value (χ^2 = 32.909, df = 12, $p < .001$). A total of 26 cases were deleted due to univariate and multivariate outliers, which accounted for 4.6% of the data. This brought the sample to 540. The data were then checked for univariate normality; results revealed acceptable kurtosis but slightly negative skewness for indicator B2 (SK = -2.24). Mardia's statistic (4.85) supported multivariate normality.

Analysis of the Module 3 dataset (n = 1,502) indicated 34 univariate outliers, all within the B2 subdomain. These outliers accounted for 2.3% of the data and were subsequently deleted. Two multivariate outliers were identified ($\chi^2 = 32.909$, df = 12, $p < .001$), both of which were deleted. A total of 36 univariate and multivariate outliers were deleted, accounting for 2.4% of the data and resulting in a final sample of 1,466. In checking for univariate normality, skewness and kurtosis both fell within acceptable ranges. Multivariate normality was upheld through Mardia's statistic (-.799).

Descriptive statistics. The level of agreement between diagnostic classifications on the ADOS and ADI-R within the verbal sample was 82.1%; when calculated within each group, 89.8% for Module 2, and 78.6% for Module 3. The internal consistencies of the ADI-R subdomains were evaluated using Cronbach's alpha for each sub-sample. Results indicated fair to good consistency (see Table 3).

Confirmatory factor analysis. Analysis of Moment Structures (AMOS; Arbuckle, 2006) was used to run confirmatory factor analysis (CFA) models employing the Maximum Likelihood (ML) estimation procedure. There is a substantial body of literature that has studied the underlying factors of the ADI-R; thus, it is considered to have a sound theoretical framework established, justifying the utilization of CFA in the current study (Byrne, 2010). Although the ADI-R consists of 93 items, item-level analysis was not used. Given that the purpose of this study was to further knowledge of the psychometric properties of the ADI-R, specifically as applied within a simplex population, analysis of the 12 Diagnostic Algorithm subdomains, rather than all 93 items, was warranted. Additionally, literature examining the factor structure of the ADI-R based on the diagnostic items is limited, signifying a need for such analyses to better understand the psychometric properties of the ADI-R as they relate to diagnostic implications (Frazier et al., 2008).

Two sets of CFA analyses were conducted; the first included testing several first-order CFA models with the verbal sample, and the second involved multiple-group analyses using ADOS modules 2 and 3. Five models were evaluated using the verbal sample: 1) a one-factor model (based on symptom severity) with all 12 subdomains loading on one factor; 2a) a two-factor model (based on DSM-5 diagnostic criteria) with all eight Social Interaction and Communication subdomains loading on one factor and the four RRB subdomains on a second; 2b) an adjusted two-factor model (based on research by Frazier et al., 2008) with all of the Social Interaction and Communication subdomains loading on one factor, with the exception of B3, which loaded on a second factor with the four RRB subdomains; 3a) a three-factor model (based on DSM-IV and DSM-IV-TR diagnostic criteria); and 3b) an adjusted three-factor model (based on research by van Lang et al, 2006) in which subdomains A2 and B4 loaded onto a third factor, Impaired Play. For each model, multiple measures of fit were considered including chi-square (χ^2), goodness of fit index (GFI), adjusted goodness of fit index (AGFI), Tucker-Lewis index (TLI), comparative fit index (CFI), and root mean square error of approximation (RMSEA). Adequacy of goodness-of-fit statistics were evaluated utilizing criteria proposed by Hu and Bentler (1999), with values between .90 and .95 considered historically acceptable, but with greater fit supported by values exceeding. 95.

Based on findings from the first-order CFA analyses, model 3b was used as the baseline for multiple-group analyses. As recommended within the literature, a logically organized series of increasingly stringent parameters were placed upon the model to test configural, metric, and structural invariance (Byrne, 2010). To determine statistically significant improvements in fit, differences between CFI and χ^2 values were calculated.

Indicators of invariance included a change in CFI of .01 or less and/or a non-significant change in the χ^2 value (Byrne, 2010).

Results

Confirmatory factor analysis of ADI-R

Verbal sample. Comparison of goodness-of-fit statistics across the five models demonstrated good absolute fit for models 2b (two-factor adjusted) and 3b (three-factor adjusted). Model 2a (two-factor based on DSM-5) demonstrated modest fit, with all statistics indicating acceptable fit with the exception of the TLI. Of the other two models tested, model 1 (one-factor) demonstrated poor fit and model 3a (three-factor based on DSM-IV/DSM-IV-TR) exhibited an estimation error as evidenced by the very high correlation between the Social Interaction and Communication factors ($r = 1.01$), signifying a Heywood case. This finding suggests that the subdomains within these two factors are measuring similar constructs. Of note, within model 3b, the correlation between the Social Communication and Impaired Play factors was also quite high ($r^2 = .90$), suggesting high convergent validity. This may be due, in part, to the inclusion of only two subdomains within this factor. The path diagrams for each of the five models with the correlations and variances can be viewed in Figures 1-5.

Given that the models all included the same variables, they were considered nested. The difference in χ^2 and CFI values demonstrated a statistically significant improvement in fit when model 3b (three-factor adjusted) was compared to models 1, 2a, and 2b (see Table 4). Given this, coupled with good absolute fit as evidenced by high goodness-of-fit statistics, 3b was deemed the best fitting model for the verbal sample.

Multiple-group analyses

Multiple-group analyses were undertaken to assess whether the 12 subdomains on the ADI-R operate equivalently across ADOS modules 2 and 3 within a simplex population. Given that the three-factor adjusted model demonstrated the best fit as evidenced by a statistically significant improvement in fit compared to the other four models, coupled with good absolute fit, it was used as the baseline model for multiple-group analyses.

As suggested by Byrne (2010), a set of hierarchical steps were taken, which involved placing increasingly stringent parameters on the model. This resulted in testing for three types of invariance: configural, metric, and structural. Noninvariance resulted in the testing of partial invariance, which involved releasing constraints individually by subdomains and covariances to determine the source of inequality. The configural model demonstrated good fit (CFI = .946), permitting analysis of metric invariance in which constraints were placed on all factor loadings. Analyses demonstrated measurement noninvariance, as established by both the difference in chi square and CFI values ($\Delta\chi^2_{(9)} = 84.45$, $p < .01$; $\Delta CFI = .015$). To test for the source of the noninvariance, each factor loading was then constrained separately. Results indicated noninvariance when the Social Communication factor was constrained in isolation ($\Delta\chi^2_{(4)} = 65.83$, $p < .01$; $\Delta CFI = .012$), as well as the Stereotyped Communication and Behavior factor ($\Delta\chi^2_{(4)} = 19.16$, $p < .01$; $\Delta CFI = .003$). Of note, although the difference in the CFI value indicated invariance, the difference in the chi-square value did not, prompting testing of partial invariance. The third factor, Impaired Play, supported invariance ($\Delta\chi^2_{(1)} = 0$, NS; $\Delta CFI = .000$) when constrained in isolation, signifying that the two subdomains within this factor operate equivalently across modules 2 and 3.

Next, partial invariance was tested to determine the source of invariance within the Social Communication and Stereotyped Communication and Behaviors factors. Because the Impaired Play factor demonstrated invariance, it remained constrained in testing for partial invariance. In testing for partial invariance, only items B3 and C1 demonstrated invariance within the metric model (see Table 5 for a comparison of the summary of goodness-of-fit statistics across models). As such, after testing for partial invariance the final measurement model used to test structural invariance included constraints on subdomains B3, C1, and the Impaired Play factor (comprised of subdomains A2 and B4).

To test for structural invariance, equality constraints were placed on each of the covariances between the three factors. Results indicated noninvariance ($\Delta\chi^2_{(6)} = 39.75$, $p < .01$; $\Delta CFI = .007$). Again, although the difference in CFI indicated invariance, the difference in chi-square did not, prompting testing of partial invariance. Thus, covariances were constrained in isolation. Constraining the covariance between Stereotyped Communication and Behaviors and Impaired Play demonstrated support for invariance ($\Delta\chi^2_{(4)} = 7.85$, NS; $\Delta CFI = .001$). The final model based on results of configural, metric, and structural invariance can be viewed in Figure 6.

Discussion

Results demonstrated that the two- and three-factor adjusted models produced the best fit, which differs from the current published factor structure of the ADI-R subdomains. Both these adjusted models provided support for stereotyped language loading onto the RRB domain rather than Social Communication. However, although both the two- and three-factor adjusted models displayed good fit, the improvement in fit displayed by the three-factor adjusted model was statistically significant, suggesting that it provided the best absolute fit. This finding differs from other studies that have reported better fit with the two-factor adjusted model (Frazier et al., 2008). Lastly, consistent with existing literature, the one-factor model was found to have poor fit, suggesting that one factor capturing autism symptom severity is overly simplistic (Boomsma et al., 2008; Frazier et al., 2008; Snow et al., 2009; Van Lang et al., 2006).

As has been found by others, both the two- and three-factor models based on DSM diagnostic criteria displayed poorer fit when compared to the adjusted models (Boomsma et al., 2008; Frazier et al., 2008; Snow et al., 2009; Van Lang et al., 2006). Additionally, the three-factor model demonstrated an extremely high correlation between the Social Interaction and Communication factors ($r = 1.01$), signifying substantial overlap between the subdomains. This provides support for combining these factors into one (Social Communication) as proposed by the DSM-5. However, as noted earlier, when compared to the three-factor adjusted model, the two-factor model based on the DSM-5 demonstrated statistically significant weaker fit. Thus, findings from the first-order CFA analyses suggest that neither the DSM-IV-TR nor DSM-5 model accurately capture ASD symptomology as evidenced by the stronger fit displayed by the adjusted models, although there is support for merging some of the subdomains comprising the Social Interaction and Communication factors on the ADI-R. Findings argue for an additional third factor constituting Impaired Play, for which others have also found support (Boomsma et al., 2008; Frazier et al., 2008; Van Lang et al., 2006).

Results of multiple-group analyses supported configural invariance, but metric and structural invariance were violated. After testing for partial invariance, results demonstrated that only subdomains B3 (stereotyped language) and C1 (circumscribed pattern of interest), as well as the Impaired Play factor, comprised of subdomains A2 (failure to develop peer relationships) and B4 (lack of spontaneous make-believe play), were operating equivalently across modules 2 and 3. Of note, all of the subdomains comprising the Social Communication factor (subdomains A1, A3, A4, B1, and B2) demonstrated noninvariance, suggesting that these concepts, as measured by the ADI-R, are operating differently among groups of children and adolescents with ASD based on their expressive language and developmental level (as conceptualized by the ADOS). Previous research has demonstrated support for factorial invariance of the ADI-R based on age utilizing a younger and older group (Frazier et al., 2008). Although the mean age of the module 2 group was younger than module 3 (7.25 versus 9.77), the two modules are differentiated by expressive language level, with module 2 administered to children with phrase speech and module 3 given to children who are verbally fluent. The findings of the current study suggest that the majority of the subdomains comprising the ADI-R are operating differently across modules 2 and 3, which may be influenced by expressive language level. It is also possible that the ADI-R is capturing behaviors indicative of ASD that are not correlated with current functioning as measured by the ADOS. These findings imply that although Social Communication appears to have support as a major factor, there are unaccounted for complexities in how these symptoms manifest across expressive language levels. This suggests that grouping children as either having ASD or not may not be particularly meaningful given that nuances of clinical presentations are likely overlooked; it may be more fruitful to adopt a dimensional approach, as opposed to a categorical one, in order to better capture these variances. These differences in presentations of symptoms are critical for clinicians in crafting appropriate recommendations and providing targeted interventions.

Lastly, although support for a third factor, Impaired Play, was obtained, there was a high correlation between this factor and Social Communication ($r = .90$), suggesting high convergent validity. Given that the Impaired Play factor is only comprised of two subdomains, additional research into other behaviors that may capture this concept is warranted. The measure may be improved by additional research examining other items that capture play and peer relationships in order to make this third factor more independent from Social Communication. However, it is important to note that metric invariance was upheld for this factor, suggesting that these items, unlike those on Social Communication, are operating equivalently across modules 2 and 3. The subdomains conceptualizing play, as they currently stand on the ADI-R, appear to measure the presence of these behaviors equivalently across expressive language levels when they load onto a separate factor, Impaired Play.

Taken together, this findings indicate that re-evaluation of the factor structure of the ADI-R is warranted given support for a three-factor adjusted model, as well as findings violating metric and structural invariance across modules 2 and 3. These results suggest complexities in the presentation of symptoms across expressive language levels, which warrant further study.

Limitations

Several limitations of this study should be acknowledged. First, participants in the SSC were recruited using convenience sampling, limiting the generalizability of the results. Additionally, the demographics of the SSC population are not particularly diverse in regard to race, ethnicity, marital status, or household income. As such, these results are specific to a rather homogenous group. Furthermore, the vast majority of the SSC sample met diagnostic criteria for autism. As such, factor analysis of the ADI-R within a more heterogeneous group is warranted. For the purposes of this study, only modules 2 and 3 were utilized for multiple-group analyses. However, future research should include modules 1 and 4 to test for invariance across a broader range of expressive language levels. Lastly, in regard to the ADI-R, given that it is an interview administered to caregivers rather than a direct observation of behavior, parental bias may make it particularly difficult to capture repetitive behaviors and restricted interests (Szatmari et al., 2006).

Despite these limitations, to the researcher's knowledge, this is the first study to examine the factor structure of the ADI-R within a simplex population, as well as the first to test for invariance using ADOS modules as the grouping variable. Given that the ADI-R is considered to be one of the "gold standard" instruments in the evaluation of ASD (Ozonoff et al., 2005), additional research establishing its psychometric properties across populations is imperative (Frazier et al., 2008; Snow et al., 2009), as is a better understanding of how the ADOS and ADI-R operate in combination (Falkmer et al., 2013). Ultimately, results of this study support the suggestion that the factor structure of the ADI-R should be re-evaluated and strongly encourage further study of invariance across ADOS modules.

References

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders (4th ed.)*. Washington, DC: Author.

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

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (5th ed.)*. Washington, DC: Author.

Arbuckle, J. L. (2011). IBM Amos (Version 20.0) [Computer Program]. Chicago: SPSS.

Bishop, D.V.M. & Norbury, C.F. (2002). Exploring the borderlands of autistic disorder and specific language impairment: a study using standardised diagnostic instruments. *Journal of Child Psychology and Psychiatry*, *43*(7), 917-929.

Boomsma, A., van Lang, N.D.J., de Jonge, M.V., de Bildt, A.A., van Engeland, H., & Minderaa, R.B. (2008). A new symptom model for autism cross validated in an independent sample. *Journal of Child Psychology and Psychiatry*, *49*, 809-816.

Byrne, B. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming* (2nd ed.). New York: Routledge.

Centers for Disease Control and Prevention. (2014). Prevalence of autism spectrum disorder among children aged 8 years – Autism and Developmental Disabilities Monitoring Network, 11 sites, United States, 2010. *Surveillance Summaries, 63*(SS02), 1-21. Retrieved from http://www.cdc.gov/mmwr/preview/mmwrhtml/ss6302a1.htm?s_cid=ss6302a1_w

Chawarska, K., Klin, A., Paul, R., & Volkmar, F. (2007). Autism spectrum disorder in the second year: Stability and change in syndrome expression.*Journal of Child Psychology and Psychiatry*, *48*(2), 128-138.

Constantino, J.N., Gruber, C.P., Davis, S., Hayes, S., Passanante, N., & Przybeck, T. (2004). The factor structure of autistic traits. Journal of Child Psychology and Psychiatry, 45, 719-726.

de Bildt, A., Sytema, S., Ketelaars, C., Kraijer, D., Mulder, E., Volkmar, F., & Minderaa, R. (2004). Interrelationship between Autism Diagnostic Observation Schedule – Generic (ADOS-G), Autism Diagnostic Interview – Revised (ADI-R), and the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR) classification in children and adolescents with mental retardation. *Journal of Autism and Developmental Disorders*, *34*(2), 129-137.

DiLavore, P., Lord, C., & Rutter, M. (1995). Pre-Linguistic Autism Diagnostic Observation Schedule (PL-ADOS). *Journal of Autism and Developmental Disorders*, *25*, 355-379.

Falkmer, T., Anderson, K., Falkmer, M., & Horlin, C. (2013). Diagnostic procedures in autism spectrum disorders: A systematic literature review. *European Child and Adolescent Psychiatry*, *22*, 329-340.

Field, A. (2009). Discovering statistics using SPSS (3rd ed.). SAGE Publications: London.

Fischbach, G.D. & Lord, C. (2010). The Simons Simplex Collection: A resource for identification of autism genetic factors. *Neuron*, *68*, 192-195.

Frazier, T. W., Youngstrom, E. A., Kubu, C. S., Sinclair, L., & Rezai, A. (2008). Exploratory and confirmatory factor analysis of the Autism Diagnostic Interview-Revised. *Journal of Autism and Developmental Disorders*, *38*(3), 474-480.

Gray, K.M., Tonge, B.J., & Sweeney, D.J. (2008). Using the Autism Diagnostic Interview-Revised and the Autism Diagnostic Observation Schedule with young children with developmental delay: Evaluating diagnostic validity. *Journal of Autism and Developmental Disorders*, *38*, 657-667.

Georgiades, S. Szatmari, P., Zwaigenbaum, L., Duku, E., Bryson, S., Roberts, W…. Mahoney, W. (2007). Structure of the autism symptom phenotype: A proposed multidimensional model. *Journal of the American Academy of Child and Adolescent Psychiatry*, *46*, 188-196.

Gotham, K., Risi, S., Pickles, A., & Lord, C. (2007). The Autism Diagnostic Observation Schedule: Revised algorithms for improved diagnostic validity. *Journal of Autism and Developmental Disorders*, *37*(4), 613-627.

Hu, L. & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6* (1), 1-55. doi: 10.1080/10705519909540118

Hus, V. & Lord, C. (2013). Effects of child characteristics on the Autism Diagnostic Interview-Revised: Implications for Use of Scores as a Measure of ASD Severity. *Journal of Autism and Developmental Disorders*, *43*, 371-318.

Hus, V., Pickles, A., Cook, Jr., E. H., Risi, S., & Lord, C. (2007). Using the Autism Diagnostic Interview-Revised to increase phenotypic homogeneity in genetic studies of autism. *Biological Psychiatry*, *61*(4), 438-448.

Kamp-Becker, I., Ghahreman, M., Smidt, J., & Remschmidt, H. (2009). Dimensional structure of the autism phenotype: Relations between early development and current presentation. *Journal of Autism and Developmental Disorders*, *39*, 557-571.

Kim, S. H., & Lord, C. (2012a). New Autism Diagnostic Interview-Revised algorithms for toddlers and young preschoolers from 12 to 47 months of age. *Journal of Autism and Developmental Disorders*, *42*(1), 82-93.

Kim, S. H., & Lord, C. (2012b). Combining information from multiple sources for the diagnosis of autism spectrum disorders for toddlers and young preschoolers from 12 to 47 months of age. *Journal of Child Psychology and Psychiatry*, *53*(2), 143-151.

Kline, R.B. (2011). Principles and practice of structural equation modeling (3rd ed.). The Guilford Press: New York.

Lecavalier, L., Aman, M.G., Scahill, L., McDougle, C.J., McCracken, J.T., Vitiello, B…. Bodfish, J. (2006). Validity of the autism diagnostic interview – revised. *American Journal on Mental Retardation*, *111*, 199-215.

Le Couteur, A., Haden, G., Hammal, D., & McConachie, H. (2008). Diagnosing autism spectrum disorders in pre-school children using two standardised assessment instruments: The ADI-R and the ADOS. *Journal of Autism and Developmental Disorders*, *38*, 362-372.

Le Couteur, A., Lord, C., & Rutter, M. (2003). *The Autism Diagnostic Interview-Revised* (ADI-R). Los Angeles, CA: Western Psychological Services.

Lord, C., Petkova, E., Hus, V., Gan, W., Lu, F., Martin, D. M.,…Risi, S. (2012a). A multisite study of the clinical diagnosis of different autism spectrum disorders. *Archives of General Psychiatry, 69(3)*, 306-313.

Lord C., Luyster R. J., Gotham K., & Guthrie W. (2012b). *Autism Diagnostic Observation Schedule, Second Edition (ADOS-2) manual (Part II): Toddler module*. Torrance, CA:Western Psychological Services.

Lord, C., Risi, S., Lambrecht, L., Cook, H.E., Leventhal, B.L., DiLavore, P.C. … Rutter, M. (2000). The Autism Diagnostic Observation Schedule – Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*, *30*, 205-223.

Lord, C., Rutter, M., DiLavore, P.D., & Risi, S. (2001). *Autism Diagnostic Observation Schedule* (ADOS). Los Angeles, CA: Western Psychological Services.

Lord C., Rutter M., DiLavore P. C., Risi S., Gotham K., & Bishop S. (2012c). *Autism Diagnostic Observation Schedule, Second Edition*. Torrance, CA: Western Psychological Services.

Lord, C., Rutter, M., Goode, S., Heemsbergen, J., Jordan, H., Mawhood, L., & Schopler, E. (1989). Autism Diagnostic Observation Schedule: A standardized observation of communicative and social behavior. *Journal of Autism and Developmental Disorders*, *19*, 185-212.

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

Mildenberger, K., Sitter, S., Noterdaeme, M., & Amorosa, H. (2001). The use of the ADI-R as a diagnostic tool in the differential diagnosis of children with infantile autism and children with a receptive language disorder. *European Child & Adolescent Psychiatry*, *10*, 248-255.

Norris, M., Lecavalier, L., & Edwards, M.C. (2012). The structure of autism symptoms as measured by the Autism Diagnostic Observation Schedule. *Journal of Autism and Developmental Disorders*, *42*, 1075-1086.

Noterdaeme, M., Mildenberger, K., Sitter, S., & Amorosa, H. (2002). Parent information and direct observation in the diagnosis of pervasive and specific developmental disorders. *Autism*, *6*(2), 159-168.

Ozonoff, S., Goodlin-Jones, B.L., & Solomon, M. (2005). Evidence-based assessment of autism spectrum disorders in children and adolescents. *Journal of Clinical Child and Adolescent Psychology*, *34*(3), 523-540.

Papanikolaou, K., Paliokosta, E., Houliaras, G., Vgenopoulou, S., Giouroukou, E., Pehlivanidis, A., … Tsiantis, I. (2009). Using the Autism Diagnostic Interview – Revised and the Autism Diagnostic Observation Schedule – Generic for the diagnosis of autism spectrum disorders in a Greek sample with a wide range of intellectual abilities. *Journal of Autism and Developmental Disorders*, *39*, 414-420.

Robertson, J., Tanguay, P., L'Ecuyer, S., Sims, A., & Waltrip, C. (1999). Domains of social communication handicap in autism spectrum disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, *38*(6), 738-745.

Rutter, M., Le Couteur, A., & Lord, C. (2008). *Autism Diagnostic Interview – Revised WPS edition manual*. Los Angeles, CA: Western Psychological Services.

Snow, A.V., Lecavalier, L., & Houts, C. (2009). The structure of the Autism Diagnostic Interview – Revised: Diagnostic and phenotypic implications. *Journal of Child Psychology and Psychiatry*, *50*(6), 734-742.

Szatmari, P., Georgiades, S., Bryson, S., Zwaigenbaum, L., Roberts, W., Mahoney, W…. Tuff, Lawrence. (2006). Investigating the structure of the restricted, repetitive behaviours and interests domain of autism. *Journal of Child Psychology and Psychiatry*, *47*(6), 582-590.

Tanguay, P.E., Robertson, J., Derrick, A. (1998). A dimensional classification of autism spectrum disorder by social communication domains. *Journal of the American Academy of Child and Adolescent Psychiatry*, *47*, 582-590.

Tadevosyan-Leyfer, O., Dowd, M., Mankoski, R., Winklosky, B., Putnam, S., McGrath, L…. Folstein, S. (2003). A principal components analysis of the Autism Diagnostic Interview-Revised. *Journal of the American Academy of Child and Adolescent Psychiatry*, *42*, 864-872.

Tomanik, S.S., Pearson, D.A., Loveland, K.A., Lane, D.M., & Shaw, J.B. (2007). Improving the reliability of autism diagnoses: Examining the utility of adaptive behavior. *Journal of Autism and Developmental Disorders*, *37*, 921-928.

Van Lang, N. D., Boomsma, A., Sytema, S., de Bildt, A. A., Kraijer, D. W., Ketelaars, C., & Minderaa, R. B. (2006). Structural equation analysis of a hypothesised symptom model in the autism spectrum. *Journal of Child Psychology and Psychiatry*, *47*(1), 37-44.

Ventola, P.E., Kleinman, J., Pandey, J., Barton, M., Allen, S., Green, J., … Fein, D. (2006). Agreement among four diagnostic instruments for autism spectrum disorders in toddlers. *Journal of Autism and Developmental Disorders*, *36*, 839-847.

Wiggins, L. D., & Robins, D. L. (2008). Brief report: Excluding the ADI-R behavioral domain improves diagnostic agreement in toddlers. *Journal of Autism and Developmental Disorders, 38*(5), 972-976.

World Health Organization. (1990). *International classification of diseases (10th revision)*. Geneva: World Health Organization.

Appendix

Table 1

Demographics, Diagnostic, and Cognitive Profiles by Sub-sample

	Verbal M (SD)	**Nonverbal M (SD)**	**Module 2 M (SD)**	**Module 3 M (SD)**
N	2286	308	540	1466
Age	9.18 (3.57)	8.12 (3.60)	7.25 (3.19)	9.77 (3.17)
Gender (% male)	86.6	86.4	83	88.5
Race (%)				
White	79.9	67.2	73.7	83.7
African-American	3.6	6.2	5.7	2.3
Asian	3.5	9.4	5.4	2.6
Native American	0.2	0.3	0.2	0.2
Native Hawaiian	0.1	-	-	0.1
Not specified	0.8	0.3	1.1	0.7
More than one race	7.9	7.8	7.8	7.3
Other	4.0	8.8	6.1	3.2
Hispanic or Latino	10.9	13.3	14.7	9.2
ADI-R diagnosis (%)				
Autism	90.5	95.8	92.0	89.3
ASD	9.5	4.2	8.0	10.7
ADOS diagnosis (%)				
Autism	88.1	94.2	95.6	85.2
ASD	11.9	5.8	4.4	14.8
Overall Verbal IQ	84.42 (26.24)	28.67 (16.78)	69.27 (21.93)	93.63 (20.76)
Overall Nonverbal IQ	89.59 (22.47)	46.24 (19.10)	79.98 (21.22)	96.15 (18.79)

Table 2

Score Means and SD for ADI-R Domains and Sub-domains by Sub-sample

	Verbal	**Nonverbal**	**Module 1**	**Module 2**	**Module 3**
Soc Int (A) Total	19.82 (5.59)	24.66 (4.17)	24.20 (4.32)	20.10 (6.05)	19.32 (5.38)
A1	3.85 (1.61)	4.25 (1.62)	4.22 (1.57)	3.67 (1.72)	3.87 (1.56)
A2	5.89 (1.72)	7.23 (1.10)	7.15 (1.12)	6.13 (1.70)	5.69 (1.71)
A3	4.18 (1.73)	5.27 (1.03)	5.15 (1.19)	4.26 (1.75)	4.08 (1.74)
A4	5.90 (2.14)	7.91 (1.73)	7.69 (1.81)	6.04 (2.27)	5.69 (2.05)
Comm (B) Total					
Verbal	16.62 (4.21)	-	-	17.55 (4.00)	16.12 (4.26)
Nonverbal	8.90 (3.40)	12.11 (2.19)	11.73 (2.48)	9.45 (3.41)	8.46 (3.32)
B1	4.37 (2.57)	6.80 (1.65)	6.57 (1.85)	4.92 (2.52)	3.96 (2.51)
B2	3.68 (.595)	-	-	3.87 (.334)	3.61 (.642)
B3	4.03 (1.77)	-	-	4.23 (1.55)	4.04 (1.85)
B4	4.53 (1.40)	5.31 (.918)	5.16 (1.02)	4.53 (1.45)	4.50 (1.38)
RRB (C) Total	6.64 (2.55)	5.87 (1.87)	6.22 (2.07)	7.00 (2.53)	6.55 (2.58)
C1	1.96 (1.16)	1.36 (1.28)	1.47 (1.26)	1.81 (1.19)	2.05 (1.13)
C2	1.64 (1.45)	.84 (.921)	1.12 (1.17)	1.89 (1.50)	1.58 (1.44)
C3	1.41 (.814)	1.77 (.561)	1.74 (.596)	1.53 (.744)	1.34 (.839)
C4	1.63 (.609)	1.90 (.306)	1.89 (.318)	1.76 (.505)	1.58 (.639)

Table 3

Internal Consistency of ADI-R Domains by Sub-sample

	Verbal	**Nonverbal**	**Module 1**	**Module 2**	**Module 3**
Social Interaction	.774	.731	.733	.822	.752
Communication	.427	.508	.551	.421	.450
RRB	.415	.237	.268	.404	.441

*Only includes nonverbal sub-domains (B1 and B4)

Table 4

Fit Indices for Verbal Sample

Model	**χ^2**	**DF**	**p**	**χ^2/df**	**RMSEA**	**CFI**	**TLI**	**GFI**	**AGFI**
One factor (1)	834.62	54	<.0001	15.46	.080*	.861	.830	.931*	.901*
Two factor (2a)	604.85	53	<.0001	11.41	.068*	.901*	.877	.955**	.934*
Two factor adjusted (2b)	398.42	53	<.0001	7.52	.053*	.938*	.923*	.971**	.957**
Three factor (3a)	601.30	51	<.0001	11.80	.069*	.902*	.873	.955**	.931*
Three factor adjusted (3b)	363.99	51	<.0001	7.14	.052*	.944*	.928*	.973**	.959**

*Indicates historically acceptable fit
**Indicates current standards for acceptable fit

Table 5

Summary of Goodness-of-fit Statistics for Multiple-group Analyses

Model description	**Comparative model**	χ^2	**df**	$\Delta\chi^2$	**Δdf**	**Statistical significance**	**CFI**	**ΔCFI**
Configural model (1)	-	365.23	102	-	-	-	.946	-
Measurement model (2a)	2a to 1	449.68	111	84.45	9	$p < .01$	.931	.015
Only SocComm constrained equal (2b)	2b to 1	431.06	106	65.83	4	$p < .01$	.934	.012
Only Stereotyped CommBeh constrained equal (2c)	2c to 1	384.39	106	19.16	4	$p < .01$	.943	.003
Only Impaired Play constrained equal (2d)	2d to 1	365.23	103	0	1	NS	.946	.000
2d + A1 (2e)	2e to 1	423.83	104	58.6	2	$p < .01$	.935	.011
2d + A3 (2f)	2f to 1	421.77	104	56.54	2	$p < .01$	.935	.011
2d + A4 (2g)	2g to 1	429.89	104	64.66	2	$p < .01$	.933	.013
2d + B1 (2h)	2h to 1	413.67	104	48.44	2	$p < .01$	.937	.009
2d + B3 (2i)	2i to 1	367.90	104	2.67	2	NS	.946	.000
2d, B3, and C1 (2j)	2j to 1	371.33	105	6.1	3	NS	.946	.000
2d, B3, C1, and C2 (2k)	2k to 1	382.49	106	17.26	4	$p < .01$	.944	.002
2d, B3, C1, and C3 (2l)	2l to 1	374.78	106	9.55	4	$p < .025$	.945	.001
Structural model 2j + covariances among three factors equal (3)	3 to 1	404.98	108	39.75	6	$p < .01$	.939	.007
Only covariance between SocComm and Impaired Play	3a to 1	398.90	106	33.67	4	$p < .01$	.940	.006

(3a)								
Only covariance between SocComm and Stereotyped CommBeh (3b)	3b to 1	398.57	106	33.34	4	p < .01	.940	.006
Only covariance between Stereotyped CommBeh and Impaired Play (3c)	3c to 1	373.08	106	7.85	4	NS	.945	.001

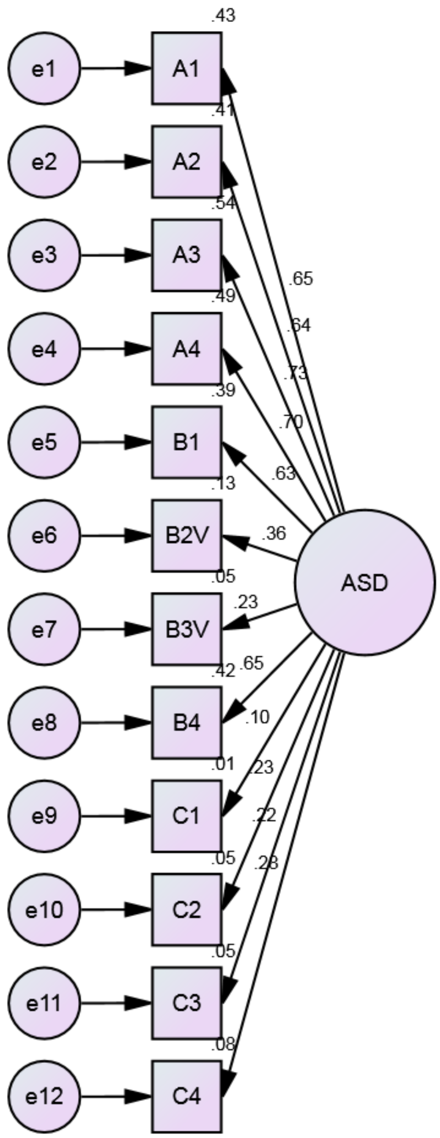

Figure 1. *Path diagram with output for one-factor model.*

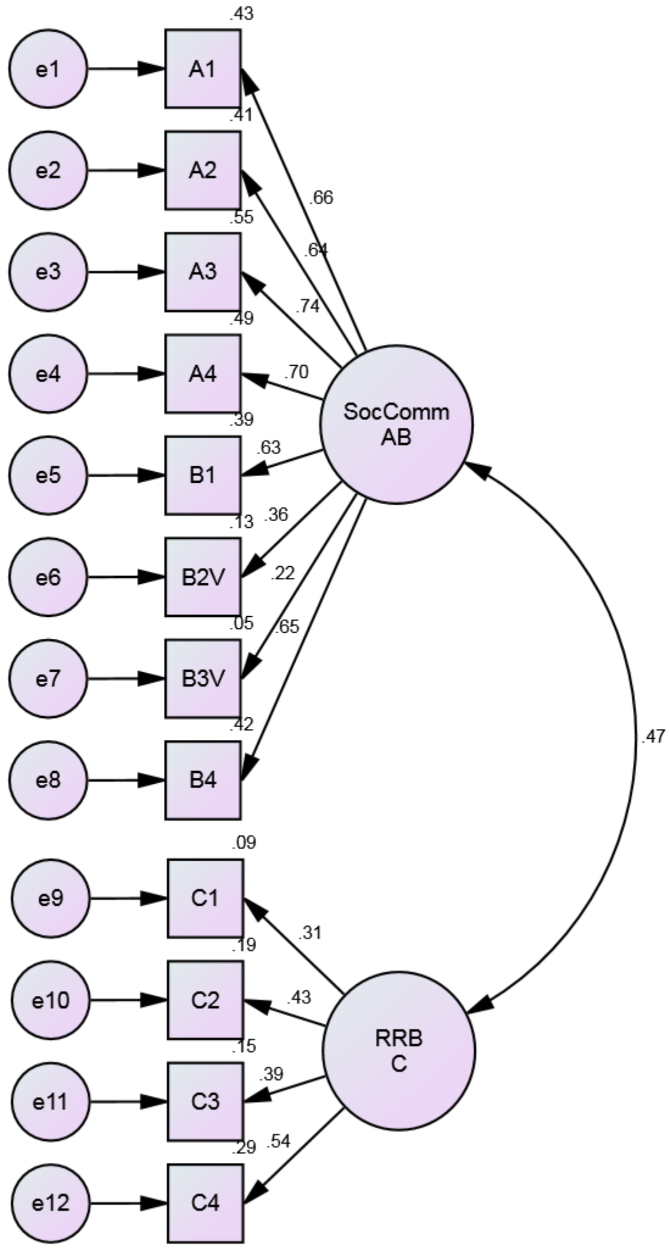

Figure 2. *Path diagram with output for two-factor model, based on DSM-5 diagnostic criteria.*

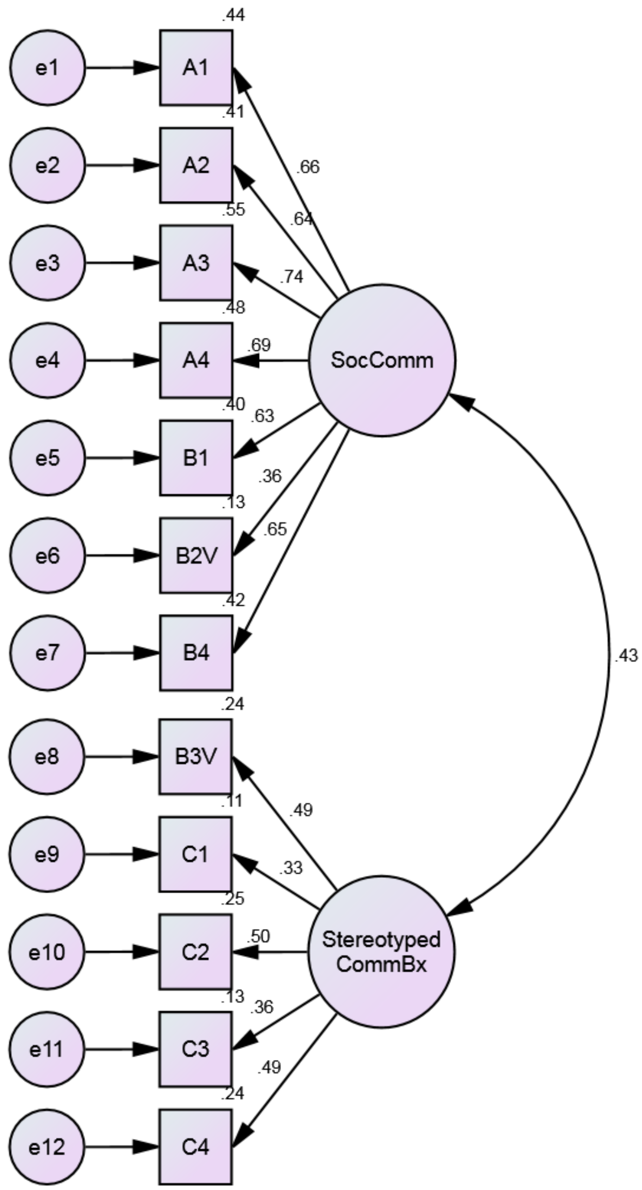

Figure 3. *Path diagram with output for two-factor adjusted model.*

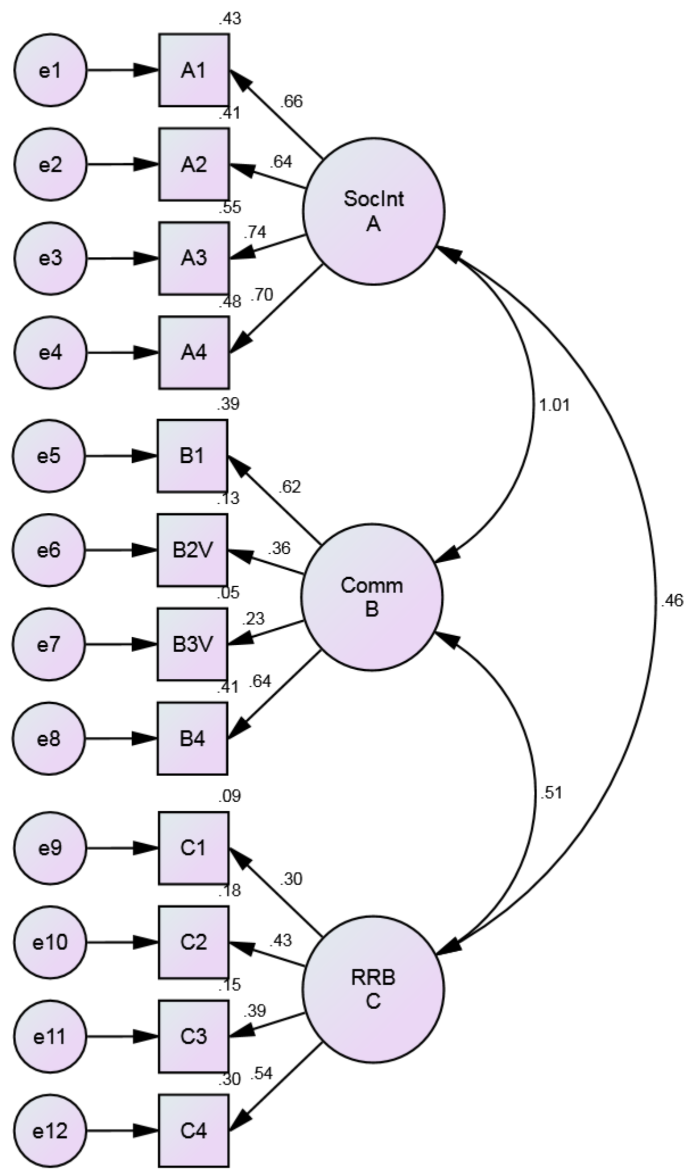

Figure 4. Path diagram with output for three-factor model, based on DSM-IV/DSM-IV-TR diagnostic criteria.

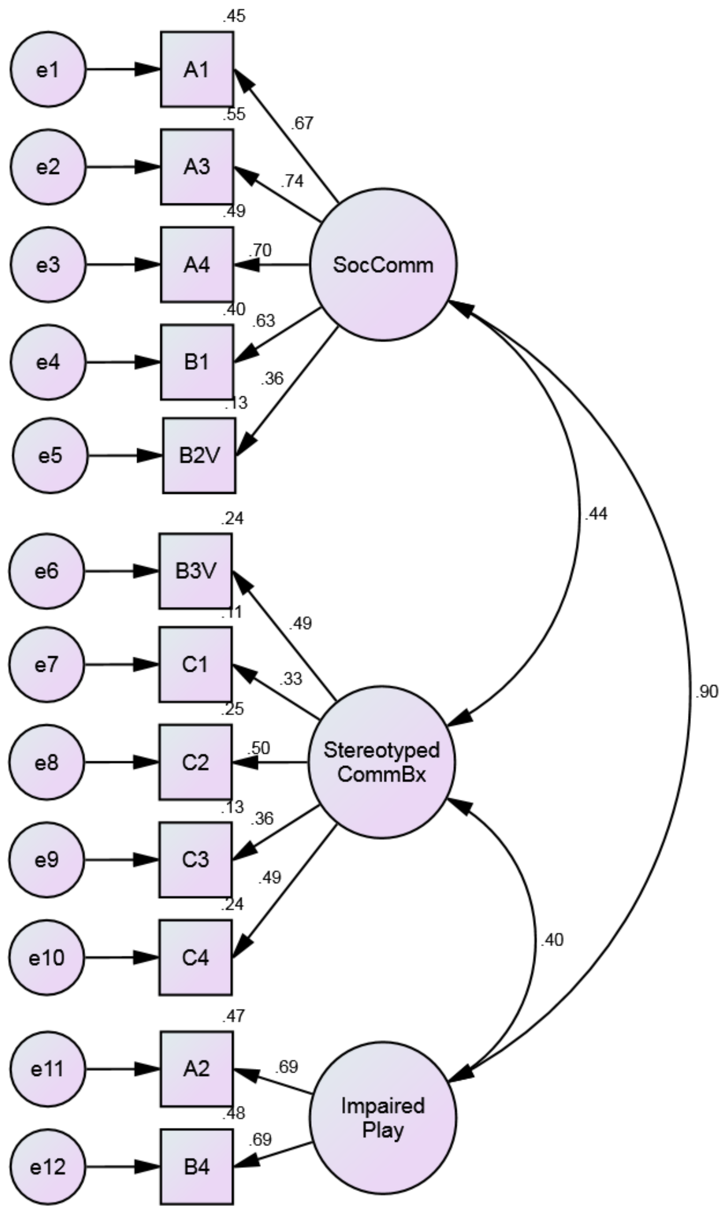

Figure 5. *Path diagram with output for adjusted three-factor model.*

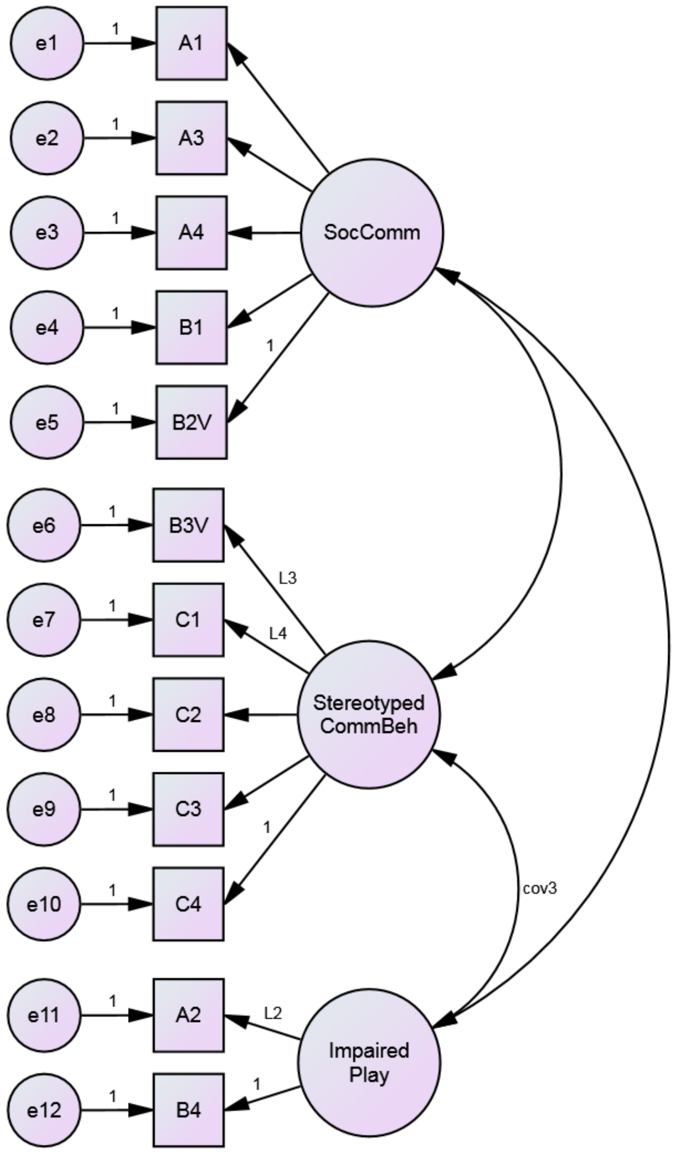

Figure 6. Final structural model for multiple-group analyses

Neill Broderick, Ph.D.
Curry School of Education
University of Virginia
Curry Programs in Clinical and School Psychology
PO Box 400270
Charlottesville, VA 22904

May 22, 2015

Dear Dissertation Committee:

This letter is to confirm that Jordan Wade was an integral second author on the following paper: *Model Invariance across Genders of the Broad Autism Phenotype Questionnaire*.

Mrs. Wade participated in the initial conceptualization of the study, contributed to data analyses preparation, and edited the literature review. As such, Mrs. Wade made central contributions to the article and fulfilled her responsibilities as second author.

The *Journal of Autism and Developmental Disorders* recently accepted the manuscript for publication, which is the leading peer-reviewed journal for autism research.

If you have any questions, please contact me via email at neillbroderick@gmail.com.

Sincerely,

Neill Broderick, Ph.D.
Postdoctoral Fellow
Treatment and Research Institute for Autism Spectrum Disorders (TRIAD)
Vanderbilt Kennedy Center
Vanderbilt University