AUTISM SPECTRUM DISORDER IN GIRLS: SEX DIFFERENCES, SUBTYPES, AND SYMPTOM STRUCTURE

Melissa Rourke

A DISSERTATION SUBMITTED TO
THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN PSYCHOLOGY
YORK UNIVERSITY
TORONTO, ONTARIO

October 2024

ABSTRACT

The present study investigated how autism spectrum disorder manifests in female children through three separate studies. Two separate datasets were used. The Surrey Place (SP) dataset (N=1725) was composed of children screened to participate in an intensive treatment program. The SFARI dataset (N=2757), is a collection of data representing autistic children who participated in the Simons Simplex Collection (SSC; Fischbach & Lord 2010). Each study boasts a large sample of autistic females (n=301, and n=375, respectively). The datasets differed from one another in many important ways, such as the characteristics of the sample, method of data collection, and the measures used.

Study 1 comprised an investigation of sex differences between female and male autistic children across a wide range of developmental and diagnostic characteristics. Groups of male and female children were compared using frequency data, correlations, and Multivariate Analysis of Covariance. Results of this study showed significant differences based on sex. That is, boys tended to have higher FSIQ, Non-Verbal IQ and adaptive abilities than girls. No differences were found in terms of total autism symptom scores. Boys were found to have higher levels of RRBIs and Self Injurious Behaviour.

Study 2 used Latent Profile Analysis to explore subtypes within the sample of girls; that is, autistic girls with specific characteristics that cluster together, and likewise within the samples of boys. Results indicated three distinct profiles (or classes) for the SP dataset and these were similar in boys and girls. Class 1 consisted of children who were slightly younger than the sample mean, with adaptive scores close to the sample mean, and autism scores that showed mild to moderate symptoms of autism. Class 2 again consisted of younger children with better adaptive skills, and mild to no autism symptoms. Class 3 consisted of older children, with lower adaptive skills, and more severe autism symptoms. The SFARI sample included more developmental

variables and allowed for a more nuanced analysis. As a result, many possible combinations of profiles emerged. Using LPA fit statistics and clinical judgement, a five-profile solution was deemed to best fit the data. Visual comparison across the profiles for girls and boys found that four of the five profiles were largely similar across boys and girls. One of the five profiles showed unique and distinct profiles for girls and boys.

Study 3 used confirmatory factor analysis to investigate the symptom structure of autism between females and males, based on factor structures reported in previous research. Overall, the factor structure was found to be largely consistent for males and females.

These results suggest that sex differences in the current study were minimal, although some subtle differences were noted. More research is needed to better understand the unique qualities of autistic females. Neuro-affirming diagnostic measures and more representative research samples may allow for the discovery of more sex differences.

ACKNOWLEDGMENTS

This dissertation is dedicated to my loving family. To my parents, Laurie and Richard, I would not have gotten this far without your support and example. To my husband, Mathew, thank you for always supporting and encouraging me, even when this season in our lives stretched on for longer than we planned. To my beautiful children, Bennett, Claire, and our new little one who is still on the way. You are my guiding light, and my commitment to your futures is what has kept me going. I love you endlessly.

I am incredibly grateful to my supervisor and mentor, Dr. Adrienne Perry. This dissertation would never have been possible without her unwavering support, encouragement, wisdom, and guidance. I admire her strong, steady presence, as well as her contributions to the field, her community, the lives of the people around her, and the environment. Having her guidance throughout my graduate school career has shaped me into the person I have become today. I would like to communicate my deepfelt thanks and gratitude.

Thank you to Dr. Monique Herbert and Dr. Nancy Freeman for agreeing to be on my committee. Thank you for your patience, wisdom, and feedback. Your expertise has been a tremendous help, and it has been a privilege to work with you both. I would like to make special mention of the many hours Dr. Herbert provided, coaching me through the answers to many statistical questions and helping to make the work meaningful and understandable.

Many people at York University and beyond have been hugely influential and supportive as I have navigated graduate school. I would like to thank my Perry Lab family and my cohort, who continue to provide me with friendship, perspective, guidance, and community. Your support was hugely influential, and I am grateful to have each and every one of you in my lives.

I would like to make special mention of Dr. Rebecca Shine who spent countless hours meticulously collecting the Surrey Place data. She then graciously showed me the ropes and

allowed me to be a part of such a meaningful piece of work. I am forever grateful for her support and friendship. Also, thank you to Dr. Azin Taheri who put in countless hours to create and organize the SFARI dataset, helped me to navigate and understand the many variables, spent her time answering my questions, and whose incredible work was a cornerstone of Study 3. Another special mention goes to Dr. Nidhi Luthra, whose work was also replicated in Study 3.

I am especially grateful to all of the families and children who participated in this research and the people who helped to collect the data. Without the families and staff at Surrey Place, as well as the families and researchers at the participating Simons Simplex Collection (SSC) sites, none of this would have been possible. As well, thank you to the principal investigators of the SFARI research project (A. Beaudet, R. Bernier, J. Constantino, E. Cook, E. Fombonne, D. Geschwind, R. Goin-Kochel, E. Hanson, D. Grice, A. Klin, D. Ledbetter, C. Lord, C. Martin, D. Martin, R. Maxim, J. Miles, O. Ousley, K. Pelphrey, B. Peterson, J. Piggot, C. Saulnier, M. State, W. Stone, J. Sutcliffe, C. Walsh, Z. Warren, E. Wijsman).

I am grateful to have obtained access to phenotypic data on SFARI Base.

Approved researchers can obtain the SSC population dataset described in this study (https://www.sfari.org/resource/simons-simplex-collection/) by applying at https://base.sfari.org.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVATIONS

Abbreviation	Definition
ASD	Autism Spectrum Disorder
DSM-IV-TR	Diagnostic and Statistical Manual of Mental Disorders – Fourth Edition, Text Revision
DSM-5	Diagnostic and Statistical Manual of Mental Disorders – Fifth Edition
FSIQ	Full-Scale Intelligence Quotient
IBI	Intensive Behavioural Intervention
ID	Intellectual Disability
IQ	Intelligence Quotient
RRB	Restricted and Reptitive Behaviour
RRBI	Restricted and Repetitive Behaviours and Interests
SES	Socio-Economic Status
SFARI	Simons Foundation Autism Research Initiative
SP	Surrey Place
SSC	Simons Simplex Collection
PDD-NOS	Pervasive Developmental Disorder – Not Otherwise Specified
TPAS	Toronto Partnership for Autism Services
ADI-R	Autism Diagnostic Interview – Revised
ADOS-2	Autism Diagnostic Observation Schedule, Second Edition
ASQ	Autism Spectrum Quotient – Short Form
CARS	Childhood Autism Rating Scale
CBCL	The Child Behavior Checklist
DAS-II	Differential Ability Scale, Second Edition
MCHAT-R	Modified Checklist for Autism in Toddlers – Revised
MSEL	Mullen Scales of Early Learning
Q-ASC	Questionnaire for Autism Spectrum Conditions
RBS-R	Repetitive Behavior Scale-Revised (RBS-R)
SCQ	Social Communication Questionnaire
SRS	Social Responsiveness Scale
TRF	Teacher Report Form
VABS	Vineland Adaptive Behavior Scales, Second Edition
WASI	Wechsler Abbreviated Scale of Intelligence, Second Edition
WISC-IV	Wechsler Intelligence Scale for Children, Fourth Edition

Abbreviation (cont)	Definition
AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
CSS	Calibrated Severity Score
CFA	Confirmatory Factor Analysis
CFI	Comparative Fit Index
LPA	Latent Profile Analysis
M	Mean
MANOVA	Multivariate Analysis of Variance
MANCOVA	Multivariate Analysis of Covariance
SD	Standard Deviation
SME	Standardized Mean Estimates
SS	Standard Scores
SRMR	Standardized Root Mean Square Error
TLI	Tucker-Lewis Index
RMSEA	Root Mean Square Error of Approximation
VABS ABC	Vineland Adaptive Behavior Composite

Introduction

Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by social communication differences, along with restricted and repetitive behaviours and interests (RRBI; APA, 2013). It is widely understood that the abilities of autistic individuals vary widely across developmental variables, such as, autism severity, cognitive abilities, adaptive abilities, and social and emotional health. The cognitive ability of an autistic child can range from above average intellectual ability to a cognitive level that falls within the range of intellectual disability (ID). Many autistic children have multiple exceptionalities, as autism is also often comorbid with other developmental disorders or mental health challenges (such as, attention-deficit hyperactivity disorder and/or anxiety). To capture these differences, autistic children are described as having differing support needs that vary over time and according to their environment (i.e., Level 1, Level 2, Level 3). Autistic children may require a range of supports throughout their development and into adulthood, for example, they may require accommodations at school, and/or targeted and individualized interventions (provided by many different types of professionals at school or in the community). It is commonly mentioned that if you have met one autistic child, then you have met one autistic child, as the autism spectrum is wide, and no two children are exactly alike.

ASD occurs in approximately 2% of the Canadian population (Canadian Health Survery of Canadian Youth, 2019), and it has long been established that more males are diagnosed with ASD than females (~4:1; Baird et al., 2006). The proportion of girls, or children assigned female at birth, included in autism research varies dramatically across studies, and males are generally overrepresented (Watkins et al., 2014; Young et al., 2018). Disparities between autistic girls and boys have long been discussed, but how these sex differences manifest, and how they influence

diagnosis and treatment is unclear. This dissertation will involve an investigation into the nature of sex differences, including delineating symptom profiles in girls and boys separately, determining whether there are similar or different subtypes of autism in girls and boys, and investigating the factor structure of ASD symptoms in girls and boys separately.

Diagnosis of Girls with Autism

Currently, in research focusing on autistic children, gender is often defined by the sex a child is assigned at birth, i.e., male or female, and is reported by parents. It is important to acknowledge that categorizing children in this way may not accurately reflect their gender identity. For readability, we use the terms 'girls' and 'boys' or 'female' and 'male' to describe children who are assigned female or male at birth.

As mentioned above, the diagnostic ratio of girls to boys varies dramatically across research, epidemiological, and intervention studies but usually over-represents males (Young et al., 2018). This sex imbalance makes research on girls more challenging and contributes to small sample sizes, which in turn influences the power of a study to find differences, and increases the likelihood of Type 2 error.

As such, ASD research is very male-centered, and the studies upon which we create and standardize our diagnostic measures are predominately male. This means that diagnostic measures are more likely to identify males, further biasing research samples (Haney, 2016). On average four males are diagnosed to every one female, however, the 4:1 ratio of males to females varies with Intelligence Quotient (IQ). The ratio can be as low as 2:1 when ASD is co-morbid with intellectual disability, and as high as 6-8:1 in higher IQ groups (Fombonne, 2005). This suggests that ASD may be easier to diagnose in female children with lower IQs (more equivalent

diagnostic rates to males), but more challenging to diagnose in females in higher IQ groups. Diagnostic measures may be less sensitive at detecting the female autism phenotype for certain girls. In part, due to this lower sensitivity in our diagnostic measures, autistic girls may be more likely to receive a later diagnosis (despite similar ages of first concern), or be misdiagnosed altogether (Rutherford et al., 2016). Evidence also suggests that girls without an intellectual impairment or language delay may be misdiagnosed with other conditions, such as personality disorders, schizophrenia, eating disorders, borderline personality disorder, selective mutism, separation anxiety disorder, depression, or specific phobias (Christensen, et al., 2016; Rynkiewicz & Lucka, 2015).

Clinically, selecting appropriate diagnostic tools when assessing females for ASD is critical and comprehensive assessments should include measures that are sensitive to differences between sexes. Two of the most well-known measures used both clinically and in research studies when diagnosing autism are the Autism Diagnostic Observation Schedule, Second Edition (ADOS-2) and the Autism Diagnostic Interview – Revised (ADI-R). These measures are considered to be the "gold-standard" for diagnosing autism. However, these measures are not adequately sensitive to differences based on gender. The ADOS-2 was not adequately standardized on females, and the normative sample did not include individuals with a more mild ASD presentation. Additionally, it does not measure social masking or imitation, which are important distinctions to be considered in the presentation of girls without intellectual disability or language delay (Cheslack-Postava & Jordan-Young, 2012; Kamp-Becker et al., 2018; Lai et al., 2015; Rynkiewicz et al., 2016). Masking or camouflaging refers to an autistic person's ability to conciously or unconciously: 1) suppress or reduce their autistic traits in order to appear more neurotypical, or 2) their ability to mimic or imitate reciprocal interactions after carefully evaluating the emotions or actions of others, or typical social norms. Emerging research suggests

that autistic females have shown an increased ability to camouflage (Cook et al., 2021; Goddard et al., 2014; Hull et al., 2019; Simone, 2010). For example, compared to autistic males, autistic females have been found to apologize more often to appease others during social situations. Research in this area is emerging and a consensus on how to measure camouflaging has not been reached. Currently, mainstream diagnostic measures do not have the ability to detect the presence or degree of camouflaging, nor the impact, (i.e, whether or not a female child who is 'high-masking' is accurately diagnosed).

On the ADI-R, Beggiato et al. (2017) investigated which items differentiated between males and females, and they evaluated the weighting of these items in the final diagnosis of autism. The researchers found that several items that are part of the diagnostic alogrithm, significantly differentiated between males and females. This bias may contribute to lower rates of diagnosis in females and the proportions of these differences in diagnosis likely vary depending on sample characteristics. Despite many screening measures having separate norms for girls and boys, the algorithm of the ADI-R has not been reformulated over time to address gender differences.

There are some screening instruments that include items designed to assess female-specific ASD traits. A study by Simcoe et al., (2022) used the Questionnaire for Autism Spectrum Conditions (Q-ASC, Attwood et al., 2011; Simcoe et al., 2022) to explore differences between genders. They found that autistic females had higher scores on gendered behaviour (i.e., how a child is expected to behave given their assigned sex), sensory sensitivity, social masking, and imitation than autistic males.

In addition to less sensitivity in most diagnostic measures of autism across sexes, there may also be differences in how parents and professionals perceive and report symptoms of ASD in girls and boys. Research has emerged investigating the effects of society and social gender

stereotypes on how we perceive and interpret ASD in boys and girls. How gender is conceptualized in our society influences the expectations we have for girls and boys, and this in turn influences behaviour in both children and adults. Whitlock et al. (2020) found that educators, who often are the first to notice ASD symptoms and begin the referral process of a child for an assessment, fall prey to gender stereotyping, leading to missed symptoms in girls. The researchers propose that this is due to a "female autism phenotype" (a term first proposed by Lai and colleagues) that doesn't fit with the male-centric view of the disorder (Lai, et al., 2017). As a part of this study, educators were given four vignettes to read, each depicting different fictional children with four different diagnostic profiles, (i.e., male autism phenotype, female autism phenotype, separation anxiety, and attention-deficit hyperactivity disorder). The child's sex/gender was randomized. Core characteristics of the female autism phenotype presented in the vignette were based on the literature and featured core characteristics, including difficulty socializing, restricted interests with a social or animal focus, camouflage/mimicking, autism-related emotional/behavioural challenge, an autistic characteristic that impacts relationships with a teacher, mild eczema, and anxiety. The male phenotype was quite similar, however restricted interests were more male-focused (e.g., Harry Potter), the vignette depicted challenge with transitions and behavioural rigidity, as well as peer conflict. Results indicated that the educators in the study were significantly more likely to identify cases describing the male phenotype. They were also more likely to identify boys as having autism than girls; to rate boys as more severe; and they were more likely to recommend supportive services for boys.

Similarly, a review by Tomlinson et al. (2020) noted that school staff and parents often had contrasting perspectives when describing girls' experiences, suggesting that sex and gender influence how autism symptoms are perceived. A review by Young et al. (2018) discussed how sociocultural and familial influences are important to consider when thinking about how autism is

perceived in girls. They proposed that parents may be more likely to expect girls to engage in pretend play with dolls, and thus may take time to teach these skills to girls, while allowing boys to engage in more cause and effect types of games and activities. Both educators and parents can fall prey to biases that affect how they perceive the female autism phenotype. Overall, many factors, including gender bias, may impact when or even whether girls are referred for a diagnosis of ASD, as well as the diagnostic process itself.

Symptoms of ASD in Girls

Understanding the female autism phenotype and how it differs from the male autism phenotype is critical to accurate diagnosis and support. Even within samples of diagnosed children (not capturing girls who do not meet our current criteria), there is evidence that symptom profiles may be different for autistic boys and girls. Differences have been found in areas involving pretend play, looking at faces, gesturing, communicating internal states, camouflauging/mimicking or masking, social motivation, quality of friendships, frequency and type of restricted and repetitive interests, and internalizing symptoms.

Autistic girls seem to engage in more pretend play, although the quality of this pretend play may be different from that of a typically developing child and may be repetitive in nature. It is important to keep in mind that, on a diagnostic measure such as the ADOS, the mere presence of pretend play may mask a child's social differences. In support of this finding, Harrop et al. (2019) found that girls were more likely than boys to look at faces, making it more challenging to notice social differences. Rynkiewic et al. (2016) found that autistic girls tended to use gestures more vividly than boys. The term "vivid" meaning girls are more likely to use longer gestures, presented in a short amount of time, that are more likely to be noticed by an observer. Kauschke

et al. (2016) found that, when communicating a narrative from a wordless picture book, autistic girls verbalized internal states more often than boys. Song et al. (2021) found that autistic girls talk more about social groups indicating that they may be more aware of social exclusion than boys. Sedgewick et al. (2016) found that autistic girls showed similar social motivation and friendship quality to non-autistic girls, while autistic boys reported having both qualitatively different friendships and less motivation for social contact relative to boys without autism and to girls with and without autism. Also notable, autistic girls reported high levels of relational aggression within their friendships, suggesting that girls on the autism spectrum may struggle to identify and deal with conflict in their social relationships.

In stark contrast, two recent studies found evidence that young autistic girls show more severe deficits in social interaction and communication compared to autistic boys. Evans et al. (2019) found that, specifically at 4 to 5 years of age, autistic girls were reported by parents to have significantly more severe social interaction and communication deficits than boys on the Social Communication Questionnaire (SCQ; Rutter, 2003). In typically developing children, the pattern was reversed with boys showing more social interaction and communication deficits than girls. Ros-Demarize et al. (2020) found that female toddlers (age 18–35 months) showed higher levels of clinician-reported social interaction and communication deficits, while preschool girls (age 36–72 months) showed higher levels of parent-reported social interaction and communication deficits than boys. The authors suggested that the two most commonly used screening tools for ASD, the Modified Checklist for Autism in Toddlers-Revised (MCHAT-R; Robins et al., 2014) and the SCQ may underidentify RRBIs in toddler and preschool girls as screening scores seemed to be only influenced by social interaction and communication deficits. However, other studies have found no evidence of social differences between autistic boys and girls (Wilson et al., 2016).

There are mixed findings regarding the occurrence of RRBIs across genders with some researchers finding no signficant differences in frequency or intensity of RRBIs (Hull et al., 2017; Ros-Demarize et al., 2020), while others have found consistently lower levels of RRBIs, particularly restricted interests, in autistic females (Frazier & Harden, 2016; Frazier et al., 2014; Mandy et al., 2012; Szatmari et al., 2012). Kaat et al. (2020) found that, in a large sample, there were minimal differences between girls and boys on standard measures of autism symptoms when accounting for age, IQ, and language level. However, they did find that boys received more severe scores on restrictive and repetitive behaviour on both the ADI-R and the ADOS-2, while girls received more severe scores on the Social Responsiveness Scale (SRS), particularly in adolescence. While the frequency and intensity of RRBIs in girls is similar to or less than that of boys, there is evidence to support the idea that girls may have restricted interests more typical of their gender (e.g., jewellery; Sutherland et al., 2017), making them more difficult to identify. Girls may also present with RRBIs that are more socially-oriented than boys (McFayden et al., 2019), for example, they may have a special interest in learning the social conventions of a friendship. Overall, researchers continue to find subtle but inconsistent differences between boys and girls in terms of symptoms of ASD, and further research to clarify these conflicting findings is necessary. For example, a 2019 study by Allely systematically reviewed 19 articles investigating restricted and repetitive behaviours and interests (RRBIs) in autistic females or the differences between autistic males and females. They found that five studies found no significant evidence to support the notion of sex differences in RRBIs in ASD, one study did not report any differences in RRBIs between males and females, 12 studies found evidence that autistic males had significantly more RRBIs compared to females, and one study using the RBS-R found that the items that were found to best-discriminate gender were greater stereotyped behaviours and restricted interests in the boys and compulsive, sameness, restricted, and self-injurious behaviour

items in the girls (Antezana et al., 2018). It is likely that the inconsistencies across studies are a result of many factors; some likely factors such as: (1) sample characteristics (e.g., is the sample younger or older on average? What is the average IQ or autism severity of the children?), (2) the sample size and methodology of the studies (e.g., larger samples may allow more sophisticated analyses; are the measures direct observational measures or parent report?; do they have multiple time points or a control group?), and (3) the variables available to investigate (e.g., did they covary for IQ?).

Differences between autistic boys and girls have been found that are relevant for planning intervention. In studies comparing typically developing children to autistic children, autistic boys were more likely to show externalizing behaviours, such as hyperactivity and inattention (when compared to typically developing children and autistic girls), while girls were more likely to internalize and mask their symptoms when compared to autistic boys (May et al., 2016). These differences seem to persist over time; a study by Solomon et al. (2012) showed that girls with ASD showed higher rates of internalizing symptoms and depression in adolescence than did boys with ASD and typically developing girls. Moreover, and as mentioned above, girls may present with more internalizing symptoms and engage in less externalizing and/or problematic behaviours, making them less likely to be noticed or identified by teachers. It is unclear whether current social interventions are more geared towards the needs of boys, (e.g., initiating social contact, learning coping tools to manage emotions and challenging behaviours), rather than the possible needs of girls, (e.g., understanding and engaging in meaningful and safe social relationships, advocating for oneself, learning emotion regulation skills, supporting sensory needs, and support for symptoms of depression and anxiety). Waizbard-Bartov et al. (2020) evaluated a group of autistic children, 89 of whom were male, and 36 of whom were female.

Children were assessed for severity of ASD at three time points, one year apart, beginning at age 3. The authors classified the children into groups based on whether their autism severity increased, remained stable, or decreased (a two-point decrease on the autism calibrated severity score on the ADOS-2 was used to calculate a change score). They recorded whether the children were in treatment or had received treatment, as well as the intensity of said treatment, all based on parent report (the researchers did not have control over this variable). Results showed that girls tended to decrease in autism severity more often and increase in severity less often than boys, even when controlling for IQ and adaptive behaviour, suggesting that girls may respond differently than boys to treatment. It is important to note that the intensity and amount of intervention received did not significantly differ across the three groups, although it is not indicated if the amount of treatment received was different for autistic girls and boys. Waizbard-Bartov et al. (2020) also note that, as in many studies of this nature, the numbers of boys and girls were not equal, and the small sample of girls may have limited the results. Moran et al. (2019) found that, in adolescence, autistic girls showed less social inclusion than boys on the ASD-KidsLifeScale (Gomez et al., 2020). Even when controlling for level of intellectual disability (ID), support needs, and age, these differences remained. This finding is surprising, given that the evidence indicating that girls are more motivated to be included in social interactions. The authors suggest that supports and programs for ASD may be more geared towards the needs of boys.

Subtypes

Wing coined the term Autism Spectrum in 1981. Since then, autism has generally been considered an umbrella, consisting of different disorders. According to the Diagnostic and

Statistical Manual of Mental Disorders – Fourth Edition (DSM-IV) the umbrella consisted of Pervasive Developmental Disorders, including: autistic disorder, Asperger's disorder, Rett's disorder and childhood disintegrative disorder, and pervasive developmental disorder – not otherwise specified (PDD-NOS; American Psychiatric Association, 1994). The DSM-5 (2013) reconceptualized autism as ASD, a single condition on a spectrum, although (like all DSM-5 disorders), they allowed for delineation of three levels of severity: Level 1 "Requiring support", Level 2 "Requiring Substantial Support", and Level 3 "Requiring very substantial support". This change in how the disorder was conceptualized arose because of a lack of research evidence clearly differentiating the distinct disorders. Verte et al. (2006) performed a cluster analysis based on the ADI-R subscales and found that the three groups differed primarily in terms of severity level, rather than exhibiting unique profiles. More recent research has found support for distinct profiles within the larger umbrella of autism, however these differences are not consistent across measures and vary depending on the characteristics of the sample being studied (Stevens et al. 2000; Klopper, Testa, Pantelis, and Skafidas, 2017). Despite these unclear findings, some researchers still believe that it would be helpful to think of autism as consisting of different subtypes, rather than one condition (Cholemkery et al., 2016). Given that children with autism show heterogeneity in their clinical presentation and variability in their response to treatment, understanding different profiles may be helpful in order to tailor treatment programs for groups of children with differing profiles. Given the literature supporting sex differences in children with autism across developmental variables, autism symptoms, RRBIs, and social and emotional functioning, it stands to reason that the existence of different subtypes of girls may help to explain conflicting findings.

Some researchers have found that autistic females exhibit a more severe form of the disorder, (i.e., lower IQ, more severe autism symptoms, and deficits in adaptive functioning;

Volkmar et al., 1993; Tsai & Beisler, 1983), while others describe a group of girls without intellectual impairment or language delays, who are more likely to be undiagnosed or misdiagnosed and not included in research (e.g., Baldwin & Costley, 2016; Rynkiewicz et al., 2016). Typically, epidemiological studies have found more males with autism, as we see in many other developmental disorders (about 4 to 1; Fombonne, 2003). However, several researchers have found that the proportion of girls to boys is more equal in samples of children who have ASD and severe intellectual disability, while boys are over represented in average- to high-IQ groups (Banach et al., 2009; Bryson et al., 1988; Wing, 1981). In support, other researchers have also found that a larger proportion of girls with ASD have concurrent ID, and they are more likely to have epilepsy than boys (Amiet et al., 2008). There are several competing genetic theories that attempt to explain why we see this pattern (Kirkovski, et al., 2013; Banach et al., 2009), however results are mixed and contradictory. It is well known that ASD is a polygenic disorder, that is, no one chromosome or genetic profile has been identified to cause the disorder, so it is not surprising that there are multiple genetic theories and seemingly contradictory findings. Researchers have qualitatively described or suggested different subtypes of girls, however small sample sizes have prevented more formal statistical analysis of subtypes in all-girl samples. Much of the research on subtypes to date has been conducted on predominately male samples and is discussed below.

Individuals with ASD are known to fall along a spectrum of functioning, however, some researchers have investigated the possibility of subtypes of individuals who cluster together with particular characteristics. Research on different subtypes in samples of both boys and girls has shown mixed results. Munson et al. (2008) conducted a latent class analysis in a sample of 456 children with ASD to investigate the possibility of more than one distinct subtype of autism based on IQ. They found evidence for multiple IQ-based subgroups, indicating that subtypes of

autism differ in level of intellectual functioning, patterns of verbal and non-verbal behaviour, cognitive strengths and weaknesses, and severity of autism symptoms. Miles et al. (2005) defined two subtypes of autism that he labelled, “complex” autism, and “essential autism”. The “complex” subtype is described as having a biological/genetic cause (e.g., fragile X syndrome), and individuals are more likely to have more seizures, dysmorphic physical features, microcephaly, lower IQs, and a lower male to female ratio. The “essential” subtype is described as having higher sibling reccurence, family history of ASD, higher male to female ratio, higher likelihood of regression and macrocephaly, and overall higher IQs. Wing and Gould (1979), found evidence for three subtypes based on behavioural patterns and social interaction. These subgroups were termed aloof, passive, and active-but-odd. More recent research by Dekker et al. (2020) used latent class analysis to identify subgroups of autistic children who benefit from social skills training. They found four subgroups based on social-communicative skills before, and in response to training. Overall, the literature suggests that IQ is the most significant contributor to discriminating between clusters and subgroups (Miller & Ozonoff, 2000). In many studies, IQ strongly predicts social functioning, adaptive behaviour, severity of symptoms and prognosis (Howlin et al., 2004). Despite these patterns, it is unclear whether distinct subgroups exist as cluster analysis seems to yield different results depending on the sample and study methodology.

No studies using formal latent class analysis have been conducted looking specifically at subtypes of girls. This is because samples of girls are small and studies often lack the power to perform these analyses. Still, some researchers have suggested that girls with particular symptoms may cluster together, creating particular subtypes. Unsurprisingly, cognitive ability may be a factor mediating subgroup differences, and females with concurrent ID may be a diagnostic subgroup with a specific symptom profile different from higher IQ females (Amiet et al., 2008). In support of this theory, Lundstrom et al. (2019) found that girls in their sample

represented the more extreme end of the spectrum in terms of ASD symptoms. Girls in this sample were more severely autistic and had more other concerns, such as attention-deficit/hyperactivity disorder, learning disabilities, oppositional defiant disorder and ID. The authors suggest that future research should look into sex-specific cutoffs on diagnostic measures for ASD. Haney (2016) suggests that females who show average intelligence, and milder ASD are less likely to be recognized by clinicians and may form a specific subtype. Evaluating this hypothesis may be difficult as we must find ways to research females who have not been diagnosed. Also, investigating these subtypes proves to be a very challenging task without an adequate sample of girls.

All of the studies described above lacked a large enough sample size of girls (not enough power) to perform formal latent class analysis. The current study has a much larger sample of girls than seen before and is well suited to find further evidence of symptom subtypes in girls with ASD if such subtypes do exist. Performing latent class analysis on such a large sample has not been done and this research study is a novel and important contribution to the field, furthering our understanding of girls with autism. Research suggests that we may find evidence supporting two main subtypes of girls with ASD, one group characterised by low IQ, concurrent ID, other medical problems, and more severe ASD symptoms; and another group characterised by higher IQ, and less severe ASD symptoms, however we may also find evidence for other more novel subtypes.

Factor Structure

Factor analysis is a statistical technique that may be used to represent the structure of a disorder by identifying how symptoms are related to one another and how symptoms cluster

together (e.g., Sorge et al., 2015). Understanding the structure and relationships among ASD symptoms has important clinical and theoretical implications. Clinically, accurate and evidence-based diagnostic criteria are essential for diagnosis. The structure of ASD symptoms informs the instruments we use to measure and diagnose ASD. Children who do not meet the criteria for diagnosis are not included in our research samples and may not gain access to critical supports and resources. The Diagnostic and Statistical Manual of Mental Disorders – Fifth Edition (*DSM-5*; American Psychological Association, 2013) presents a diagnosis of ASD based on research demonstrating a two-factor structure, consisting of a) Social-Communication and b) Restricted and Repetitive Behaviour and Interests (RRBI). It is mentioned as a note, that gender should be considered when making a diagnosis of autism using the *DSM*.

Although the *DSM* is typically the gold standard for diagnosis, the literature indicates that this conceptualization of autism is not the only model proposed by researchers to accurately diagnose autism. The diagnosis of ASD in the previous version of the *DSM* (the *DSM-IV-TR*) was based on a triad of impairment (i.e., domains included reciprocal social interaction, communication, and restricted, repetitive and stereotyped behaviour). Current research suggests that the *DSM-5* and *DSM-IV-TR* are similarly sensitive when being used as diagnostic tools, however both may be missing cases, especially girls (Haney, 2016).

Factor analytic studies differ in terms of the specific diagnostic measures and symptoms being examined (Shuster et al., 2014). A large number of studies are based on the Autism Diagnostic Interview-Revised (ADI-R; Lord et al., 1994), a parent-interview measure of autism symptoms. Others are based on direct observation of children's symptoms using the ADOS-2 (Lord et al., 2000) and some studies have used combined data from both the ADOS-2 and the ADI-R. A small number of studies have used the Childhood Autism Rating Scale (CARS; Schopler et al., 1988), an observational scale that has been widely used clinically and in research

(Filipek et al., 1999; Perry et al., 2005). Researchers have also conducted factor analytic studies on measures of a particular dimension of autism symptoms, such as the Repetitive Behavior Scale-Revised (RBS-R) which looks specifically at RRBIs.

Shuster and colleagues conducted a review of 36 studies published between 1980 and 2011, that used factor analysis to report on the structure of autism symptoms. Factor analysis was conducted on 13 different measures across studies, including the ADOS and ADI-R (Shuster et al., 2014). These studies were heterogeneous in terms of the sample size, the age of the children, and the IQ range. All studies reported significantly more males than females, with the percentage of males ranging from 50-95%. None of these studies specifically investigated a factor structure based on sex. The authors found evidence for different factor structures of autism symptoms. Data in some studies fit best the *DSM-IV-TR* triad structure, others supported the two-factor *DSM-5* criteria (Social-Communication and RRBI), some supported both, and some supported different models altogether (Boosma et al., 2008; Frazier et al., 2008; Van Lang et al., 2006). Shuster et al. (2014) discussed possible reasons for the differences in findings and emphasized that it is important for "future studies to examine the role of additional variables, such as age, cognitive ability, and language ability, in order to better understand the relationship between these variables and ASD symptom structure". The authors suggest using these variables to define different samples or subgroups in which to examine factor structure.

An unpublished study by Taheri (2020) did just this by testing four different factor structures of autism, including those based on factor structures found in the literature (i.e., the DSM-5 2-factor structure; the DSM-IV-TR 3-factor structure; a 3-factor structure reported in some literature) and a novel 4-factor structure. Taheri (2020) examined these four different models in a large sample, separately by IQ level (average, borderline, mild/moderate ID, severe/profound ID). Results demonstrated that the 4-factor model fit the data best, for each IQ

level, using scores from the ADI-R, and findings were replicated with ADOS-2 scores. The four factors were labelled: Social, Communication, Play/Peer, and RRBI. The present study builds on Taheri's research by investigating boys and girls separately.

Few studies have looked at the factor analytic structure of the CARS. Evidence for 3-, 4-, and 5-factor structures has been found although the power of these studies may have been limited as a result of small sample size (DiLalla & Rogers, 1994; Magyar & Pandolfi, 2007; Stella et al., 1998). Luthra (2013) investigated the factor structure of the CARS in a large and heterogenous (in terms of age and IQ) sample of children with autism. She found that a 3-factor structure somewhat resembling the triad of impairment found in the DSM-IV-TR best fit the data. The three factors found by Luthra were termed: Social Interaction, Communication, and Emotion Regulation. The present study will examine this further for girls and boys separately.

Investigating the nature of RRBIs specifically is important because it has allowed for a better understanding of how these behaviours and interests affect a child's ability to interact with others (Harrop et al., 2014). Traditionally, RRBIs have been characterized as either "higher order" or "lower order or sensorimotor" (Mosconi et al., 2009; Turner, 1999). The higher order category consists of preoccupations/circumscribed interests, compulsive routines, and ritualistic behaviour, while the Sensorimotor category includes stereotyped and repetitive body movements and sensory preoccupations. Recent research has addressed how the individual behaviours and interests may group together using factor analysis. Several studies have found evidence supporting the presence of multiple distinct subtypes of RRBIs (Bishop et al., 2013; Lam & Aman, 2007; Lam et al., 2011). Factor analysis studies have been conducted on several different measures in order to investigate the structure of RRBIs, specifically the RRBI domains of the ADI-R (Rutter et al., 2003) and the ADOS (Lord et al., 2002), and the Repetitive Behaviors Scale-Revised (RBS-R; Bodfish et al., 2000).

A study by Lam and colleagures (2008) used exploratory factor analysis to investigate the factor structure of RRBIs by looking at relevant items taken from the ADI-R. They found three distinct factors: Repetitive Motor Behaviors, Insistence on Sameness, and Circumscribed Interests. Repetitive motor behaviours were found to be associated with a variety of subject characteristics such as IQ, age, social and communication impairments, and the presence of regression. Insistence on sameness was associated with social and communication impairments. The review by Shuster et al. (2014) found evidence that a two-factor model of RRBI symptoms is supported in the literature. These studies looked specifically at the RRBI domain found in the ADI-R. The two-factor model found consisted of a Repetitive Sensory Motor factor and a Resistence to Change factor (also referred to as Insistence on Sameness; Cuccaro et al., 2003). Another study by Warsof (2014) used exploratory and confirmatory factor analysis for the ADOS and ADI-R and found evidence for a two-factor structure, with a Cogntive-Restrictive factor and a Motor-Sensory factor. They found that overall, RRBI items correlated poorly with one another across the two measures, that RRBI scores were highly negatively correlated with IQ, and that adolescents showed significantly lower RRBI scores than younger children. Overall, the factor structure found in the ADI-R and ADOS-2 (in items related to RRBIs) is limited by the small number of RRBI items, possibly narrowing the number of factors found.

In contrast, the RBS-R is a detailed measure aimed at capturing the nature and severity of RRBIs in individuals with autism more specifically. Bodfish and colleagues (2000) developed this parent questionnaire that is used to specifically measure repetitive behaviour in autistic children and adolescents. It was designed with six rationally-derived subscales: Stereotyped Behaviour, Self-Injurious Behaviour, Compulsive Behaviour, Routine Behaviour, Sameness Behaviour, and Restricted Behaviour. The authors investigated this model on a sample of 32 autistic individuals (8 female) and 34 individuals who did not have ASD, but who did have ID,

and found that autistic adults scored significantly higher (showed higher levels of repetitive behaviour) than adults with ID but not autism. Another study done by Lam and Aman (2007) used factor-analytic procedures to examine the items on the RBS-R and found that a 5-factor structure fit their data best, based on a sample of 307 caregivers of autistic children. The factors they described were: Ritualistic/Sameness Behavior, Stereotypic Behavior, Self-Injurious Behavior, Compulsive Behavior, and Restricted Interests. Similarly, Bishop et al. (2013) also found a 5-factor structure in a sample of 1825 autistic individuals aged 4-18 years of age. Similar factors as those presented by Lam and Aman (2007) were found, however there was some variability in terms of the items that loaded on to each factor. Mirenda et al. (2010) was the first study to use confirmatory factor analysis and found a 5-factor model in a sample of 287 preschool children aged 24 to 64 months of age. The factors comprised a Stereotypic Behavior factor (items 1–6), Self-Injurious Behavior factor (items 7–14), Restricted Interests factor (items 40–43), a Compulsive Behavior factor (items 15–22) and a Ritualistic/Sameness Behavior factor (items 23–39). A more recent study by Hooker et al. (2019) investigated the factor structure of RRBIs in the RBS-R in a sample of 350 autistic children ages 2 to 9 years. The authors used confirmatory factor analysis to examine six proposed structural models, including the 5-factor model found by Mirenda and colleagues (2010). Again, this 5-factor model demonstrated the best fit, although 4- and 6-factor models also fit adequately-to-well. All of the subtypes demonstrated signficant associations with IQ and adaptive behaviour. Conversly, another study by Georgiades et al. (2010) found a 2-factor structure best fit their data (205 autistic individuals aged 2 to 48 years), labelled as a Compulsive-Ritualistic-Sameness-Restricted Behaviors factor and a Stereotyped Self-Injurious Behavior factor.

Research on RRBIs has been limited by a lack of clarity in terms of which subtypes are stable across measures. This variability in measurement, likely contributes to mixed findings in

terms of how RRBIs are related to other areas of functioning (e.g., IQ), or how they may differ by sex. Many factors such as age, diagnostic characterstics, and sample size may also explain some of these inconsistencies. No researchers have investigated the factor structure of RRBIs in autistic children separately by sex, despite evidence showing differences between boys and girls in this area.

A recent systematic review was conducted by Bourson and Prevost (2021). They reviewed characteristics of restricted interests in autistic girls compared to boys. Their final analysis comprised 15 separate studies. They found that girls demonstrated fewer restricted interests than boys, and that the interests of autistic girls were more similar to neurotypical girls than those of autistic boys. They also found that age and IQ were two factors that were associated with variations in restricted interests according to gender. One of the goals of the present study is to take a more nuanced look at this core domain of ASD symptomatolgy and its role in differences between girls and boys overall.

Recently, two groups of researchers have looked specifically at whether the same factor structure for autism is appropriate for autistic boys and girls. Frazier and Harden (2017) used the Simons Simplex Collection to investigate the factor structure of ASD in 2643 children (352 females). They mapped items from the Social Responsiveness Scale (SRS) and the ADI-R onto nine a priori symptom dimensions. The SRS is a parent- and teacher-report tool used to distinguish autism spectrum conditions from other child psychiatric conditions by identifying the presence and extent of social impairment characteristic of autism. The ADI-R , as noted above, is a semi-structured and standardized caregiver interview examining symptoms of autism in three domains: Language/Communication, Reciprocal Social Interactions, and Restricted, Repetitive, and Stereotyped Behaviours. The authors used multi-group confirmatory factor models to examine whether males and females demonstrated measurement or structural differences in

autism symptom constructs. The results of the analysis showed that males and females did differ in terms of restricted interests; specifically, females with high verbal IQ had substantially lower levels of repetitive interests. Otherwise, the autism symptom factor structure for males and females was highly similar. A strength of this study included its large sample, allowing for sophisticated analysis.

Grove et al. (2017) looked at the factor structure of the Autism Spectrum Quotient Short Form (ASQ-SF) in a sample of autistic adults (265 males, 285 females). The ASQ is a diagnostic questionnaire designed to measure the expression of autism-spectrum traits in an individual, by his or her own subjective self-assessment. The adults included had received a formal diagnosis of ASD from a qualified clinician; diagnoses were done separately from this study. Participants were recruited from a large online database and were over 16 years of age with an IQ of above 70. Factor analysis was done on the ASQ. The authors found that the overall factor structure supported a two-factor structure that was composed of two higher-order factors assessing 'Social Behavioral Difficulties' and 'a Fascination for Numbers/Patterns'. This factor structure was consistent across groups indicating that it captured both men and women adaquately. Grove et al. (2017) found that the ASQ was not biased towards the male phenotype, at least for higher IQ autistic adults.

Recent studies seem to indicate that the factor structure for boys and girls is highly similar, however the authors all identify that some subtle differences do exist and that more research is required. It is important to acknowledge that factor analysis research is based on symptoms measured using available diagnostic instruments. It remains a limitation that we cannot study girls who are not diagnosed with these tools and as such they cannot be included in this research. Overall, few studies have investigated the factor structure of autism based on sex differences. From the few studies that have, we can conclude that there is evidence for some

subtle differences between the sexes (e.g., repetitive behaviour), but in contrast, it appears that the factor structure for boys and girls is statistically similar. Still, research in this area is new and studies replicating these findings and delving further into the topic are needed. Additionally, the factor structure of autism in girls may be impacted by the presence of distinct subtypes of girls.

Current study

It appears that there are several reasons why studying sex differences can be challenging (e.g., biased samples and diagnostic measures, sociocultural influences on how males and females are perceived, subtypes and factor structure specific to autistic girls not clear, and small sample sizes with insufficient power to make comparisons). In the current study I hope to address some of the concerns identified in the literature on sex differences in ASD. To do this sex differences were investigated in two different large samples, the Surrey Place (SP) and the Simons Foundation Autism Research Initiative (SFARI) datasets, described below. See Table 1 for characteristics of the two datasets. Important gaps in the literature are addressed and contradictory findings are clarified by conducting three separate studies.

Study 1 addressed differences between boys and girls, using both the SP and SFARI datasets on a wide range of variables. Boys and girls were compared on direct observational measures, parent report measures, and teacher report measures. The measures themselves include diagnostic measures of autism symptoms, adaptive behaviour, intelligence quotient (IQ), socio-emotional functioning, social responsiveness, and repetitive and restrictive behaviour; as evidence has been found for sex differences in all these areas of functioning (see Table 2).

It was hypothesized that there may be cognitive differences between girls and boys in our samples, that is, that girls may be more likely to have lower cognitive abilities (and/or comorbid ID) as the literature suggests. How boys and girls differ socio-emotionally, specifically in terms of internalizing and externalizing behaviours or tendencies, was investigated. Based on the literature, it was hypothesized that girls would show higher levels of internalizing behaviours, while boys would show higher levels of externalizing behaviours. Differences and similarities in terms of social responsiveness were investigated. It was suspected that girls would show lower scores on the Social Responsiveness Scale, compared to boys, which corresponds to less

difficulty with social responding. Finally, repetitive and restricted behaviours and interests were investigated. It was suspected that boys and girls would show similar levels of RRBIs (though the content of their RRBIs may differ).

Study 2 focused on exploring subtypes in girls versus boys, based on variables from both samples (combinations of diagnostic, developmental, and metal health characteristics). I wanted to answer the following research question: Are there different subtypes of girls and of boys and are the subtypes similar in the two groups or are there subtypes that are unique to one sex or the other? In order to investigate for these subtypes (pinpoint children with similar characteristics who cluster together), I conducted latent class analysis in two large datasets. Performing these investigations in two samples with different measures and different sample characteristics helped to strengthen and confirm our findings.

Study 3 was designed to investigate the factor structure of ASD based on sex. I conducted confirmatory factor analysis for the factor structure of ASD in several different diagnostic measures in both SP and SFARI datasets. I built on the findings of Taheri (2020) to see if the factor structures investigated for different IQ groups, in the same SFARI dataset, fit well in girls and boys separately. Taheri tested four different factor structures, as described above, using items from the ADI-R. I used confirmatory factor analysis to see how well these different models fit the data separately for boys and girls using the ADI-R from the SFARI sample.

I also conducted confirmatory factor analysis to build on the findings of Luthra (2013) using the 15 individual items of the CARS, in the SP sample. As mentioned above, she found evidence supporting a three-factor structure *somewhat* resembling the triad of impairment found in the DSM-IV-TR (Luthra's three factors being, Social Interaction, Communication, and Emotion Regulation).

For the RRBI domain, I conducted this analysis based on the original six domains of the measure (Bodfish et al., 2000) and a 5-factor structure found by several researchers showed some stability across several studies (Bishop et al. 2013; Hooker et al. 2019; Lam & Aman, 2007; Mirenda et al. 2010). Based on the literature, we expected that the factor structures would be similar for boys and girls in the different measures, although we believed that there may be subtle differences in terms of social communication and RRBIs. We hoped to clarify the contradictory findings other researchers have discovered in these areas and provide comparisons across measures.

Overall, the findings of this study will add important information to the literature on sex differences in ASD, especially given that these studies investigate many important questions, using different types of measures, heterogenous populations, and large samples. The results are crucial to help clinicians, teachers and parents enhance their understanding of how ASD may differ in girls, possibly influencing more accurate ASD diagnoses and helping them to consider the important variable of “sex” when creating and implementing individualized treatment plans.

Method

This research project was initially approved by the York University Human Research Ethics Board from January 31st, 2019 to January 31st 2020, and subsequently renewed from February 7th 2020, to February 7th 2021 (certificate number: e2019-039). The study was approved by Surrey Place on March 1st 2019 through to March 1st 2022 (REB ID#190104AP), and it was also approved by the Simons Foundation SFARI from January 29th 2019 through to completion.

Participants: Surrey Place Dataset

The Surrey Place (SP) dataset includes the eligibility screening data for 1725 children referred to the Toronto Partnership for Autism Services (TPAS), conducted between the years 2000 and 2016 when agencies providing intervention to children with autism were publicly funded. The data were collected as a part of an archival file review study (Shine, 2021). The purpose of the eligibility screening was to determine if children already diagnosed with ASD, met the Ministry requirement (at the time), of falling toward the severe end of the autism spectrum, and if intensive behavioural intervention (IBI) was an appropriate intervention. Additionally, the initial autism program had an age cut-off of 5 years, 11 months, as is reflected in the age range of the SP participants.

Initially, 1859 screenings were conducted and entered into the database. The database was examined carefully with regard to missing data, and cases were removed if they were missing either the Vineland Adaptive Behavior Scales (VABS; 2nd ed.; Sparrow et al., 2005) or the CARS; 108 cases (5.7% of the sample). In cases where a particular domain standard score (VABS) or item score (CARS) was missing, the mean of the other scores available for that individual was used as an estimate of that score (this only happened in three cases). A visual

inspection of the histograms and stem-and-leaf box plots indicated that the data were positively skewed (more younger children). As a result of the non-normality of the data, children were included in the study if they were aged 87 months (7.25 years) or younger; 66 cases or 3.5% of the sample were deleted based on this cut-off. Following this correction, the distribution of scores improved, although it remained skewed. A visual inspection of the histograms and stem-and-leaf box plots indicated that scores across all other measures were normally distributed.

Of the remaining 1725 children, 301 (17.4%) were female. The children's ages ranged from 1 year and 2 months to 7 years and 3 months (M = 4 years, 1 month; SD = 1 year, 8 months).

Measures: Surrey Place Dataset

The SP dataset includes the following measures: the Childhood Autism Rating Scale (CARS; Schopler et al., 1988; 2010), and the Vineland Adaptive Behavior Scales (VABS; 2nd ed.; Sparrow et al., 2005).

The Childhood Autism Rating Scale, first (CARS; Schopler et al., 1988) or second edition (CARS-2; Schopler et al., 2010)

The CARS is a clinical evaluation of a child's behaviour based on direct observation (supplemented by interaction and parent report), used as part of an assessment battery to diagnose ASD. The CARS has 15 items scored on a 7-point Likert-type scale ranging from 1 to 4 with half points. Higher scores represent more severe autism symptomatology. Items include: (1) Relating to people, (2) Imitation, (3) Emotional response, (4) Body use, (5) Object use, (6) Adaptation to change, (7) Visual response, (8) Listening response, (9) Taste, smell, and touch response and use, (10) Fear or nervousness, (11) Verbal communication, (12) Nonverbal communication, (13)

Activity level, (14) Level and consistency of intellectual response, and (15) General impressions. A CARS total score is generated and used to evaluate the level of symptoms of autism a child demonstrates. A score of 15-29.5 (for children under 13 years of age) is indicative of No Minimal to Symptoms of ASD; a score of 30-36.5 is indicative of Mild to Moderate Symptoms of autism spectrum disorder; a score of 37 and higher is indicative of Severe Symptoms of autism spectrum disorder.

The Vineland Adaptive Behavior Scales - Survey Interview, either the first (Vineland; Sparrow et al., 1989) or second edition (Vineland-II; Sparrow et al., 2005)

The Vineland is a parent interview used to measure children's adaptive behaviour, i.e., how a child functions in everyday life, in four domains: Communication, Daily Living Skills, Socialization, and Motor Skills (only for young children), with an overall score termed the Adaptive Behavior Composite (ABC). Standard scores (M = 100; SD = 15) and age equivalent scores are available for this measure. For consistency, the VABS ABC score was derived by calculating the average across the first three domain standard scores for participants of all ages. This was done in order to eliminate undue weighting of the Motor domain standard score which is only used to calculate the ABC score for younger children and tends to inflate the ABC total score for those children. Domain and ABC Standard Scores are qualitatively described according to the following adaptive levels, that is, scores from 20 to 24 are considered to be indicative of profoundly low adaptive abilities; scores between 25 and 39 are considered severely low; scores between 40 and 54 are considered moderately low; scores between 55 and 69 are considered mildly low; scores between 70 and 84 are considered borderline low; and scores above 85 are considered average. <u>This means that low scores indicate poorer adaptive abilities, while higher scores indicate better adaptive functioning.</u>

Participants: SFARI Dataset

The second dataset was taken from the Simons Foundation for Autism Research Initiative (SFARI), a project that aims to improve the understanding of causes, genetics, and diagnosis of ASD. Recruitment and data collection were completed in collaboration with 12 university affiliated research clinics (11 U.S. and 1 Canadian). The Simons Simplex Collection (SSC; Fischbach & Lord, 2010) is a core project whose researchers collected genetic samples from simplex families (i.e., where only one child has ASD). In addition to this genetic information, extensive phenotypic data was also collected for each child. All participants included in the study received a clinical-best-estimate diagnosis of autistic disorder, Asperger's disorder, or pervasive developmental disorder – not otherwise specified (PDD-NOS) according to the *DSM-IV-TR* (APA, 2000). For the original data collection, SFARI exclusion criteria included: child age younger than 4 or older than 18 years, nonverbal mental age below 18 months, severe neurological deficits, birth trauma, perinatal complications, or evidence of a known genetic disorder (e.g., fragile X or Down syndrome).

Cases were examined carefully for missing data. Children were included in the study if their sex (male or female) was indicated. Cases were removed if they were missing a significant amount of data across multiple variables. There were 98 cases or 3.4% of the sample that were removed as a result. In cases where a particular total score (e.g., FSIQ), domain standard score (e.g., VABS) or item score (e.g., ADI-R, RBS-R) was missing, the mean of the other scores was used as an estimate of that score. Otherwise, missing data was dealt with using a pair-wise technique, or an imputation technique in the case of the Latent Profile Analysis.

Following an investigation of histograms and stem-and-leaf box plots, it was determined that age was slightly negatively skewed towards younger ages. Scores across all other measures

were normally distributed, except for the RBS-R and its subscales, where scores were highly positively skewed due to a high number of 0 scores, which is typical of the measure.

A dataset of 2757 autistic children, between 4 and 18 years of age (M = 9 years; SD = 3 years, 7 months) was compiled, including 375 autistic females (13.6%). The SFARI dataset includes a large variety of standardized diagnostic and behavioural measures from this research sample that were not included in the brief clinical screenings that led to the SP dataset. The SFARI dataset is also more heterogenous in terms of age, adaptive level, IQ, and ASD severity of the children.

Table 7 shows the demographic characteristics of the SFARI sample. A majority of parent respondents were married (90.1%). Annual household income and education levels of the parents indicates a sample skewed towards higher socio-economic status.

Children in the sample were diagnosed based on DSM-IV-TR diagnostic criteria, with a majority of children falling in the category of Autistic Disorder and fewer in the categories of PDD-NOS and Asperger's disorder. The majority of children were assessed using ADOS-2 modules 1 through 3, with a small percentage assessed using module 4.

Measures: SFARI Dataset

Children in the SFARI dataset participated in comprehensive assessments that measured their functioning in a variety of different areas, including using direct observational measures, parent-report measures, and teacher-report measures. The specific areas of functioning assessed include, adaptive behaviour, IQ, autism symptoms, restricted and repetitive behaviour and interests, socio-emotional and behavioural functioning (parent and teacher), and social

responsiveness difficulties (see the lower section of Table 2 for the constructs, measures, and variables). The following measures relevant to this analysis will be described in detail below.

The Autism Diagnostic Observation Schedule, Second Edition (ADOS-2; Lord et al., 1999)

The ADOS is an instrument for diagnosing and assessing autism. It is a direct-observational measure that involves a series of structured and semi-structured tasks that require social interaction between the examiner and the child. There are four original modules, as well as a Toddler module, and the appropriate module is selected based on the individual's expressive language and chronological age. Each module consists of a number of tasks that take about 30 to 60 minutes to administer and are designed to elicit specific behaviours related to language and communication, reciprocal social interaction, play and imagination, and restricted and repetitive behaviours and interests. Children are then classified as having "autism", "autism spectrum", or "non-spectrum" based on cut-off scores, with higher cut-off scores required for a diagnosis of "autism". In the current study, the original ADOS-2 (Lord et al., 1999) was administered to study participants, while the newest ADOS-2 algorithms were used to classify children. Several scores are generated through the ADOS-2 algorithm and used to describe a child's symptoms of autism in different domains. The scores used herein include: the Communication and Socialization Total score, Social Affect Total score, Restricted and Repetitive Behaviour Total score, and the ADOS-2 Calibrated Severity Score (CSS). The Calibrated Severity Score is a score used to determine the clinical severity of a child's autism symptoms regardless of which specific module is administered. A score of 1 or 2 is indicative of minimal to no evidence of autism; a score of 3 or 4 is indicative of a low-level of symptoms of autism; scores of 5, 6, and 7 are indicative of

moderate autism symptoms, and scores of 8, 9, and 10 are indicative of high levels of autism symptoms. This means that higher scores indicate more severe symptoms of autism.

The Autism Diagnostic Interview – Revised (ADI-R; Rutter, et al., 2003)

The ADI-R is a standardized semi-structured caregiver interview examining symptoms of autism in three domains: Language/Communication, Reciprocal Social Interactions, and Restricted, Repetitive, and Stereotyped Behaviors. Each of the three domains consists of related subdomain items. The ADI-R is designed for individuals with mental ages of 18 months and above. Overall, there are a total of 93 interview questions, with up to 42 used in scoring. For those with verbal ability, some of the scoring items are collapsed together resulting in up to 34 items used as part of the final algorithm. Individuals without verbal ability do not complete verbal items, resulting in 27 algorithm items. Most items include parent report of both current and historical functioning. Classification of "autism" is based on prescribed cut-off scores for all three domains. Several scores are generated and used herein to describe a child's autism symptoms based on different domains. The scores used in the current study include the ADI-R Reciprocal Social Interaction Domain total score and the ADI-R Restricted and Repetitive Behaviour Total score. Higher scores on this measure are indicative of more severe autism symptoms.

Cognitive Abilities

As is common in autism research, several different IQ measures were used to determine a child's cognitive abilities, depending on the child's age and ability level. Within the SFARI dataset, these included the Differential Ability Scale, Second Edition (DAS-II; Elliott, 2007;

82%), the Mullen Scales of Early Learning (MSEL; Mullen, 1995; 9%), the Wechsler Intelligence Scale for Children, Fourth Edition (WISC-IV; Wechsler, 1999; 2%), and the Wechsler Abbreviated Scale of Intelligence, Second Edition (WASI-II; Wechsler, 2011; 4%), and a combination of these measures (3%). The scores on these measures were then used to determine overall IQ scores, including the Full-Scale Intelligence Quotient (FSIQ), a Verbal IQ score, and a Non-Verbal IQ score. The FSIQ variable includes both Deviation IQs (i.e., IQs derived from standard scores) and Ratio IQs (i.e., IQs derived by dividing their mental age by chronological age and multiplying by 100) for individuals who were not able to obtain a Deviation IQ. Nonverbal (or Performance) and Verbal IQ scores were also used to describe the intelligence of the current sample. Standard-deviation-based criteria were used to classify the IQ range. Scores above 85 were considered to fall with the Average range of intellectual functioning, scores between 70 and 84 were considered to fall in the Borderline range, scores between 55 and 69 were considered to fall in the Mild range of Intellectual Disability, scores between 40 and 54 were considered to fall in the Moderate range of Intellectual Disability, scores between 25 and 39 were considered to fall in the Severe range of Intellectual Disability, and scores below 25 were considered to fall in the Profound range of Intellectual Disability. This means that higher IQ scores are indicative of stronger cognitive abilities.

The Child Behavior Checklist (CBCL; Achenbach, 2011)

The CBCL is a parent report checklist used to detect emotional and behavioural problems in children and adolescents. The CBCL also has a Teacher Report Form (TRF), completed by teachers.

Both forms are used for children aged 6 to 18, or an early version for children aged 2 to 5, and consist of 113 questions, scored on a three-point Likert scale. The results group into two broad-band scores, i.e., Internalizing and Externalizing problems. There are also 8 syndrome scales, i.e., Anxious/Depressed, Withdrawn/Depressed, Somatic Complaints, Social Problems, Thought Problems, Attention Problems, Rule-Breaking Behaviour, and Aggressive Behaviour. An overall T-score is generated for the Internalizing Behavior domain and the Externalizing Behavior domain. Clinical ranges are used to qualitatively describe the level of behaviour exhibited. Children with T-scores of 65 and below demonstrate internalizing or externalizing behavior in the Average range; T-scores of 65 to 70 are considered to fall in the Borderline Clinical range; while T-scores of 70 and above are considered to fall in the clinical range. Higher scores on this measure are indicative of more social-emotional difficulties.

The Social Communication Questionnaire (SCQ; Rutter, 2003)

The SCQ is a brief ASD screening tool. It includes 40 yes/no questions and produces a total score with a specific cuttoff of 15. Scores above the cutoff suggest that a full evaluation for autism is warranted. The questionnaire collects information related to social relating, communication, and range of interests. The SCQ requires that a child have a mental age of at least 2 years and takes 10 minutes to complete. Higher scores on this measure are indicative of more autism symptoms.

The Social Responsiveness Scale (SRS; Constantino, 2002)

The SRS is a parent and teacher report tool used to distinguish autism spectrum conditions from other child psychiatric conditions by identifying the presence and extent of social

impairment characteristic of autism. It is appropriate for children 4 to 18 years old and takes 15 to 20 minutes to complete. A Total score is generated for the SRS, which is then turned into a T-Score that can be used to qualitatively describe a child's level of social impairment. T-Scores of 59 and below are considered to fall within normal limits. T-Scores of 60 to 65 are considered to fall within the Mild range. Scores in this range indicate deficiencies in reciprocal social behaviour that are clinically significant and may lead to mild to moderate interference with everyday social interactions. T-Scores of 66 to 75 are considered to fall within the Moderate range. Scores in this range indicate deficiencies in reciprocal social behaviour that are clinically significant and lead to substantial impairment with everyday social interactions. T-Scores of 76 or higher are considered to fall within the Severe range. Scores in this range indicate deficiencies in reciprocal social behaviour that are clinically significant and lead to severe interference with everyday social interactions. Higher scores are indicative of more severe social deficits.

The Repetitive Behavior Scale-Revised (RBS-R; Bodfish et al., 2000)

The RBS-R is a 44-item parent-report questionnaire that is used to measure repetitive behaviour in children, and adolescents, ages 6 to 17 years old, with ASD. It consists of six subscales including; Stereotyped Behavior, Self-Injurious Behavior, Compulsive Behavior, Routine Behavior, Sameness Behavior, and Restricted Behavior. Parents are asked to rate how problematic a particular behaviour is on a 4-point Likert scale, as well as to provide a general rating for how problematic repetitive behaviours are overall for their child on a scale from 1 to 100. In addition to the RBS-R total score based on the original Bodfish article (2000), subscales derived by Lam and Aman (2007) were calculated using the mean score on each specific combination of items: Ritualistic Sameness (*M* of items, 26, 27, 28, 30, 31, 32, 33, 34, 35, 37, 38,

and 39), Self-Injurious Behavior (*M* of items 7, 8, 9, 10, 11, 12, 13, and 14), Stereotypy (*M* of items 1, 2, 3, 4, 5, 6, 22, 42, and 43 Compulsive Behavior (*M* of items 15, 16, 17, 18, 19, and 20), and Restricted Behavior (*M* of items 36, 40, and 41). Although difficult to interpret clinically, a higher mean score is associated with higher levels of RRB.

Study 1: Comparing Girls and Boys

Results: Study 1

Objective 1: Research Questions and Analyses; Investigating Sex Differences in the SP dataset

The goal of this study was to explore the differences between girls and boys. We examined descriptive statistics (means and standard deviations) and conducted visual inspection of the distributions of boys and girls separately for all variables of interest. Bivariate correlations were conducted, and the two groups were compared statistically using, *t*-tests, chi-squares, and Multivariate Analysis of Variance (MANOVAs; with appropriate post-hoc tests if significant, and controlling for Type I error [e.g., Bonferroni]), to investigate the research questions related to the sex differences. Please see the upper section of Table 2 for the specific constructs, measures, and variables investigated.

Using the SP dataset, we explored whether boys and girls differ, based on parent report (VABS) and a direct clinical observational measure by professionals (CARS). We conducted bivariate correlations across variables to explore relationships among variables of interest. By Cohen's convention (1988), a correlation of $r < .10$, represents no relationship, a correlation of $.10 < r > .29$ represents a small or weak relationship, a correlation of $.30 < r > .49$ represents a moderate strength relationship, while a correlation of $r > .50$ indicates a strong relationship. We conducted a *t* test to see if the VABS Composite score differed between girls and boys. We also investigated the domain-specific scores, Daily Living Skills, Socialization, and Communication using a MANOVA with post-hoc tests. Cohen's *d* was also used to describe effect size. Cohen's *d* effect sizes of 0.2 are considered to be small, 0.5 are considered medium, and 0.8 are considered to be large (1988). We conducted a *t* test to see if the CARS total score differed for girls and boys. We used a stringent alpha of .001 for all *t* tests.

Objective 1: Results

Table 3 shows the descriptive results, including means and standard deviations for the Vineland Adaptive Behavior Composite (ABC) score, and the three domain scores, as well as the CARS total score.

In order to investigate the variables of interest, bivariate correlations across age, VABS ABC, VABS Communication standard score, VABS Daily Living Skills standard score, VABS Socialization standard score, and CARS total score were conducted. Correlation coefficients are presented in Table 4. Age showed a moderate negative relationship with three of four VABS scores (ABC: $r(1723)=-.36$, $p<.001$; Soc: $r(1723)=-.42$, $p<.001$; DLS: $r(1723)=-.40$, $p<.001$), and a weak negative relationship on the Communication domain score (Comm: $r(1723)=-.21$, $p<.001$). These scores demonstrate that older children show lower adaptive functioning standard scores (as would be expected). VABS scores were highly correlated with one another, ranging from $r(1723)=.70$, $p<.001$ to $r(1723)=.91$, $p<.001$, so children who scored highly in one domain on the VABS were likely to also have higher functioning in other domains, and subsequently in their total score. CARS total scores were strongly negatively correlated with VABS scores, ranging from $r(1723)=-.52$, $p<.001$ to $r(1723)=-.60$, $p<.001$, indicating that children with higher adaptive functioning were likely to have lower autism severity scores.

The mean VABS ABC score and the mean CARS total score for girls and boys were compared using *t* tests (see Table 5). A MANOVA was also conducted to investigate the differences between girls and boys on each of the three VABS standard score domains (see Table 6). Results suggest that when using a stringent alpha of .001, there is no significant effect of sex on VABS ABC total score, $t(1722)=2.178$, $p=.03$, $d=.14$, despite boys having a slightly higher mean score than girls (Girls: $M=62.36$, $SD=11.79$; Boys: $M=63.96$, $SD=11.55$). This finding

indicates that the mean scores for girls and boys are not significantly different from one another. Similarly, there is no significant effect of sex on CARS total score, $t(1722)=.474$, $p=.64$, $d=.03$, again indicating no significant mean differences (Girls: $M=30.36$, $SD=4.48$; Boys: $M=30.50$, $SD=4.74$). Table 6 shows the mean scores for the Vineland Communication Domain SS, the Vineland Daily Living Skills SS, and the Vineland Socialization SS, along with *F* values and effect sizes (see Table 6). The results indicate no significant differences between girls and boys across the communication, socialization, and daily living skills domains on the Vineland.

Objective 2: Research Questions and Analyses; Investigating Sex Differences in the SFARI dataset

Using the SFARI dataset, we explored whether girls and boys differ on: developmental variables, diagnostic variables, RRBI characteristics, and psychopathology symptoms, using specific measures and variables as shown in Table 2. The general analytical approach involved exploring bivariate correlations among variables of interest, and then comparing boys and girls on global scores first using a *t* test and Cohen's *d* to provide effect size. If significant and not trivial, subscales or domains within that construct were then examined in more detail, using MANOVAs and appropriate post-hoc tests, with Bonferroni corrections, and effect size was reported to avoid overinterpreting trivial differences. This helped us to avoid the high Type I error that is likely to manifest in a large sample when conducting many statistical tests. Because IQ is highly correlated with many of the variables of interest we also repeated our analyses while co-varying for IQ.

Objective 2: Results

Table 8 shows the descriptive statistics for the SFARI sample, including means and standard deviations for the variables of interest.

Table 9 shows the bivariate correlations for the SFARI dataset. Sex was not included in this analysis as it is a dichotomous variable and was analyzed separately as one of the main research questions. Correlations are described based on magnitude as described above. Results of the correlations indicate several relevant findings. FSIQ, Verbal IQ and Non-verbal IQ showed high positive correlations with one another, unsurprisingly (ranging from: $r(2750)=.84$, $p<.001$ to $r(2750)=.97$, $p<.001$). As a result of these high correlations, we decided to use only Verbal and Non-Verbal IQ in subsequent analyses in order to minimize multicollinearity. Scores on the VABS were also highly correlated with one another and with IQ (ranging from: $r(2750)=.53$, $p<.001$ to $r(2750)=.92$, $p<.001$). Similar to the SP dataset, age was moderately negatively correlated with adaptive functioning, such that older children tended to have lower adaptive scores (ranging from: $r(2750)=-.20$, $p<.001$ to $r(2750)=-.33$, $p<.001$). This occurs because the number of skills required for a particular score in a younger child is less than the same score for an older child (more is required of them). So as children age, the gap between their adaptive skills and that of typically developing peers widens, leading to lower scores. This finding is commonly found in the literature.

The subscales on the ADI-R, Reciprocal-Social Interaction and Communication were moderately negatively correlated with all IQ measures, and Vineland domains (ranging from: $r(2750)=-.25$, $p<.001$ to $r(2750)=-.53$, $p<.001$). This means that as children's IQ increase and scores on the Vineland increase (better adaptive skills), their scores on the ADI-R Reciprocal-Social Interaction decrease (fewer social challenges). This ADI-R Reciprocal-Social Interaction score was also moderately positively correlated with the ADI-R Communication domains (for both verbal and non-verbal children; $r(2421)=.66$, $p<.001$ & $r(331)=.61$, $p<.001$). This pattern

was also seen for the ADOS-2 Communication and Social domain and the Social Affect domain ($r(2681)=.93$, $p<.001$) and for the ADOS-2 Calibrated Severity Score ($r(2681)=.78$, $p<.001$ & $r(2681)=.79$, $p<.001$).

In order to investigate mean differences between girls and boys in the SFARI sample, *t*-tests were conducted for all of the variables of interest. As mentioned above, because of the number of comparisons a strict alpha level was chosen (i.e., .001). Results of the *t*-tests can be found in Table 11, including means, standard deviations, *t* values, degrees of freedom, *p* values, and effect sizes using Cohen's *d*. FSIQ (G: $M=75.20$, $SD=28.23$; B: $M=82.11$, $SD=27.81$) and Non-Verbal IQ (G: $M=77.99$, $SD=26.13$; B: $M=85.57$, $SD=26.02$) differed significantly for girls and boys, such that boys had significantly higher scores than girls, but the effect sizes were small (mean difference of 6.91 and 7.58 IQ points respectively; $t(2750)=-4.46$, $p<.001$, $d=.25$; $t(2754)=-5.23$, $p<.001$, $d=.29$).

On the VABS, the VABS ABC total score (G: $M=72.54$, $SD=12.34$; B: $M=75.15$, $SD=12.42$), Communication standard score (G: $M=74.63$, $SD=14.11$; B: $M=77.45$, $SD=14.61$), and the Daily Living Skills standard score (G: $M=73.72$, $SD=13.52$; B: $M=76.81$, $SD=13.88$) were significantly different for girls compared to boys (ABC: $t(2755)=-3.79$, $p<.001$, $d=-.21$; Comm: $t(2755)=-3.49$, $p<.001$, $d=.19$; DLS: $t(2755)=-4.03$, $p<.001$, $d=.22$). Again, boys showed higher adaptive skill levels than girls, with small effect sizes (mean differences of 2.61, 2.82, and 3.09 points).

No significant differences were detected across measures of autism symptoms, via the ADI-R, a parent-report measure, or the ADOS-2, a direct-observational measure. No differences were seen across the domain total scores of the ADI-R, as well as the core descriptive scores and the module-specific domain scores on the ADOS-2 (see Table 11).

When looking specifically at RRBI, a significant sex difference was found for the total score on the RRBI domain of the ADI-R (G: *M*=6.13, *SD*=2.46; B: *M*=6.58, *SD*=2.50), such that boys showed significantly higher scores than girls (increased restricted and repetitive behaviours and interests; $t(2754)=-3.23$, $p=.001$, $d=.07$). The size of this effect was small. No differences emerged on the RRBI totals on the ADOS-2 (across any of the modules). A significant sex difference was found on the Self-Injurious Behavior subscale on the RBS-R (G: *M*=.63, *SD*=.51; B: *M*=.71, *SD*=.53), such that boys showed significantly higher scores than girls ($t(2731)=-2.64$, $p=.008$, $d=.15$); the effect size was small.

A significant sex difference was also found on the CBCL specific to the Externalizing subscale as rated by teachers (G: *M*=61.22, *SD*=8.54; B: *M*=58.38, *SD*=8.15), indicating, contrary to expectations, that girls showed more externalizing symptoms than boys ($t(1238)=4.28$, $p<.001$, $d=.35$), and the effect size was moderate (2.84 T score point mean difference).

In order to consider the influence of related variables on the differences between girls and boys, nine Multivariate Analyses of Covariance (MANCOVA) were conducted (see Tables 12 through 19). The variables of interest were grouped conceptually into related domains: adaptive skills, autism symptoms, restricted and repetitive behaviour across different measures (grouping related measures together in three separate analyses), social responsiveness as rated by parents, and internalizing and externalizing behaviour as rated by parents and teachers (in separate analyses). Girls and boys significantly differed in terms of their IQ and IQ was moderately correlated with 15 of the 34 variables of interest, so it was included as a covariate in each analysis, in order to control for its influence.

Measures based on teacher report were only available for approximately half the sample and were therefore examined separately from parents' data. As before, a stringent alpha level was

used to determine significance, (i.e., .001). Assumptions were checked and satisfied. In each case, a MANCOVA was run with an interaction term of sex*IQ in order to test the heterogeneity of the slopes. If the interaction term was not significant, indicating that the slopes were sufficiently parallel to detect a significant difference between girls and boys, then the interaction term was removed from the model and the MANCOVA was run again as per convention.

The first MANCOVA was conducted to test the hypothesis that there would be one or more mean differences between girls and boys, on scores of adaptive functioning, specifically, the Communication, Socialization, and Daily Living Skills domains of the Vineland, while controlling for IQ (see Table 12). The overall multivariate test was not significant, Pillai's Trace = .001, $F(3,2751) = .980$, $p=.401$, $\eta_p^2<.001$, indicating that child sex does not account for a significant percentage of the variance in adaptive functioning.

The second MANCOVA was conducted to test the hypothesis that there would be one or more mean differences between girls and boys, on scores of autism symptoms, specifically, the ADI-R Reciprocal Social Interaction total score, the ADOS-2 Calibrated Severity Score, the ADOS-2 Communication and Social total score, and the ADOS-2 Social Affect total score, while controlling for IQ (see Table 13). The overall multivariate test was not significant, Pillai's Trace = .001, $F(4,2675) = .717$, $p=.580$, $\eta_p^2<.001$, indicating that a significant percentage of the variance in autism symptoms is not accounted for by the sex of the child.

A third MANCOVA was conducted to test the hypothesis that there would be one or more mean differences between girls and boys, on scores of restricted and repetitive behaviour across three measures, specifically, the ADI-R RRBI total score, the ADOS RRBI total score, and the RBS-R total score, while controlling for IQ (see Table 14). The multivariate test for the model was not significant for the interaction term (Pillai's Trace = .004, $F(3,2747) = 3.979$, $p=.008$), however, the test of between-subjects effects for the interaction term was significant for the ADI-

R RRBI total score [$F(3,2747) = 11.506$, $p=.001$], indicating that the slopes are not heterogenous or parallel and violating this assumption. The interaction term was then removed from the model and the MANCOVA was run again. The overall multivariate test was not significant, Pillai's Trace = .005, $F(3,2748) = 4.696$, $p=.003$. Between-subjects effects for the individual total scores show that the ADI-R RRBI total score was significantly different for girls and boys, with boys having higher levels of RRBIs according to parent report. However, the effect size for this finding was very small ($\eta_p^2=.004$).

In order to further investigate different types of RRBIs, a fourth MANCOVA was conducted to test the hypothesis that there would be one or more mean differences for girls and boys, across different domains of RRBIs on the RBS-R, specifically, the average of the following subscale scores, the Ritualistic/Sameness Behavior subscale, the Stereotypic Behavior subscale, the Self-Injurious Behavior subscale, the Compulsive Behavior subscale, and the Restricted Interests subscale, while controlling for IQ (see Table 15). The overall multivariate test was significant, Pillai's Trace = .016, $F(5,2696) = 8.883$, $p<.001$, indicating that a significant percentage of the variance in RRBI domains is accounted for by the sex of the child. Between-subjects effects for the individual subscale scores show that the Self-Injurious Behavior subscale score was significantly different for girls and boys, with boys having higher levels of Self-Injurious Behavior according to parent report on the RBS-R. The size of this effect, however, was very small ($\eta_p^2=.007$).

The SRS also measures other aspects of problematic autism symptoms as reported by parents, with higher scores indicative of more severe behaviours. A MANCOVA was conducted for SRS Parent raw scores. The MANCOVA was conducted to test the hypothesis that there would be one or more mean differences for girls and boys, on scores of autism symptoms on the Parent SRS, specifically, the Awareness raw score, the Cognition raw score, the Communication

raw score, and the Motivation raw score, while controlling for IQ (see Table 17). The overall multivariate test was not significant, Pillais' Trace = <.001, $F(4,2733) = .315$, p=.868, indicating that a significant percentage of the variance in autism symptoms, as reported by parents, is not accounted for by the sex of the child. Between-subjects effects for the individual total scores were also not significant, and effect sizes indicate no effect (η_p^2<.001).

In order to investigate externalizing and internalizing behaviour for girls and boys with autism we conducted a MANCOVA to test the hypothesis that there would be one or more mean differences for girls and boys, on scores of externalizing and internalizing behaviour on the CBCL (a parent-report measure), while controlling for IQ (see Table 18). The overall multivariate test was not significant, Pillai's Trace = .002, $F(2,2745) = 3.123$, p=.044, indicating that a significant percentage of the variance in internalizing and externalizing symptoms is not accounted for by the sex of the child. Between-subjects effects for the individual total scores were also not significant, and effect sizes indicate no effect (η_p^2=.001 & η_p^2<.001).

Another MANCOVA was run in order to investigate the same variables, but those based on teacher report on the parallel measure, the TRF (see Table 19), for the partial sample with teacher data. The overall multivariate test was significant, Pillai's' Trace = .018, $F(2,1236)$ = 11.347, p<.001, indicating that a significant percentage of the variance in internalizing and externalizing symptoms as reported by teachers is accounted for by the sex of the child. Between-subjects effects for the individual *T* scores indicated that the Externalizing *T* score was significantly different for girls and boys [$F(1,1237) = 16.251$, p<.001] indicating that girls showed significantly more externalizing behaviour, as reported by teachers, than boys, although the effect size was small with only 1.3% of the variance in externalizing behaviour being explained by the sex of the child (η_p^2=.013).

Discussion: Study 1: Comparing Girls and Boys

Notably, the results of Study 1, Objectives 1 and 2, found some significant sex differences. There were no significant differences found between autistic girls and boys in adaptive functioning or autism severity in the SP sample. In the SFARI sample, autistic boys tended to have higher FSIQ, non-verbal IQ, and adaptive skills, than girls, with small effect size. This is consistent with the literature showing evidence that, on average, autistic girls tend to demonstrate lower IQ, adaptive skills, and more severe autism symptoms.

When comparing girls and boys, no differences were found across autism symptom total scores across measures. However, boys showed significantly higher levels of RRBIs than girls, and they were more likely to engage in self injurious behaviour, to small effect. After covarying for IQ, many of these findings were no longer significant implying that IQ may subsume variance of differences in autism symptoms. For example, when covarying for IQ, no sex differences were found in terms of autism symptoms, social responsiveness, and internalizing and externalizing behaviours. A significant difference in RRBI total score was found on the ADI-R, although the effect size was very small and no other significant differences were found for the other measures (i.e., ADOS-2 or RBS-R). The difference in rates of self-injurious behaviour also remained significant, however the effect size was small. This supports the literature, as traditionally boys have been found to have higher rates of RRBIs. Overall, results from both samples suggest that sex has a small influence on the variables of interest according to the analyses conducted in Study 1. However, it is important to note these group comparisons may not fully capture sex differences, especially if subtypes are present in the sample.

Study 2: Latent Profile Analysis

For this study we investigated data from girls and boys separately to see if particular subtypes emerged. Separate analyses were conducted in the SP and SFARI datasets. We conducted a latent profile analysis to see if girls and boys with particular characteristics group together / form clusters. Our investigation included a selection of the following variables: developmental level (age, IQ, and adaptive functioning), ASD severity or specific characteristics, socio-emotional functioning, social responsiveness, and RRBI severity.

In order to identify potential subgroups within the sample, a Latent Profile Model was created in each of the SP and SFARI datasets for both girls and boys. Latent Profile Analysis (LPA) is a multivariate-latent-modeling technique that identifies latent or unobserved subgroups of individuals within a population, in this case autistic children, based on a set of continuous indicators (Wang & Hanges, 2011). This is a person-focused statistical technique, used to identify groups of children with similar profiles and to determine the ideal number of unique profiles in a given sample. Profiles are interpreted and named by the researcher based on the patterns of standardized variable means. Clinically, this technique helps to identify groups of children who have distinct profiles across many areas of functioning, which may improve how we define each group's diagnostic and therapeutic needs.

Results: Study 2

LPA in the SP dataset: Results

In order to investigate profiles of autistic girls in the SP dataset an LPA analysis was conducted using the child variables that were available in that dataset. Specifically, these included child age; adaptive functioning, as measured by the Vineland domain standard scores, Communication, Daily Living Skills, and Socialization; along with autism severity, as measured using the CARS total score.

The best fitting LPA models were determined by running multiple models, beginning with two classes and adding one additional class in each subsequent iteration up to five classes, in order to explore different options. Model selection was guided by the Akaike Information Criterion (AIC; Akaike, 1974), Bayesian Information Criterion (BIC; Schwarz, 1978), and Entropy (Celeux & Soromenho, 1996) values. The AIC and BIC values provide information on how well the model fits the data, with lower values indicating better model fit. Entropy is used to measure how well the model defines or distinguishes between classes (accuracy of classification), with higher values of entropy indicative of better classification and class separation (entropy values approach a maximum of 1 and should be no lower than .8). The BIC, AIC, and Entropy values were acceptable for each model and continued to improve with each iteration (see Tables 20 and 21). Model 4 showed 9/301 children in its smallest class, or 2.9% of the sample, at which point the number of participants went below 5% of the total sample (see Tables 22 and 23). This means that the smallest class in Model 4 held too few children for a meaningful interpretation. As a result, models four and five were ruled out. The 3-Class solution showed good fit for both girls and boys, while also demonstrating distinct and clinically informative profiles, so it was therefore selected as the model best representing the data. Participants were not equally distributed amongst the 3 classes, with almost twice as many children in class 1 as in the other two classes.

The class with the fewest participants consisted of 66/301 (21.9%) of the sample for girls, and 270/1424 (21.7%) for boys. A plot comparing models can be found in Figures 1 and 2 respectively.

The 3-Class model is defined by three unique classes or profiles. Based on the limited number of variables available in this dataset, these profiles are largely dictated by a child's functioning level. They represent how much a particular profile deviates from the sample mean, for a particular measure, for female children. Children in Class 1 are defined as having lower age at the time of the screening assessment, higher adaptive skills on all three Vineland domains (i.e., Communication, Daily Living Skills, and Socialization), and lower (i.e., milder) autism severity scores on the CARS. Children in Class 2 are characterized by lower age at the time of assessment, moderate scores on all three Vineland domains, and moderate autism severity scores on the CARS. Children in Class 3 are characterized by older age at assessment, lower adaptive functioning across all three domains on the Vineland, and higher autism severity or CARS total score. This profile pattern was very similar for girls and boys. Figure 3 depicts a plot of the Standardized Mean Estimates (SME) for Class 1 for girls and Class 3 for boys in isolation from the other profiles, allowing for visual and qualitative comparison between profiles with similar characteristics. Figure 3 demonstrates that Class 1 for girls and Class 3 for boys are highly similar in the pattern of scores for girls and boys, as well as the number of boys and girls in each class. Similarly, Figure 4 shows a highly similar profile of scores for Class 2 for girls and Class 2 for boys. Lastly, Figure 5 compares Class 3 for girls and Class 1 for boys. Again, there are no discernable differences between the profiles of scores across the variables of interest for girls and boys.

Figure 1

Latent Profile Model – 3-Class Solution for Girls in the SP Dataset

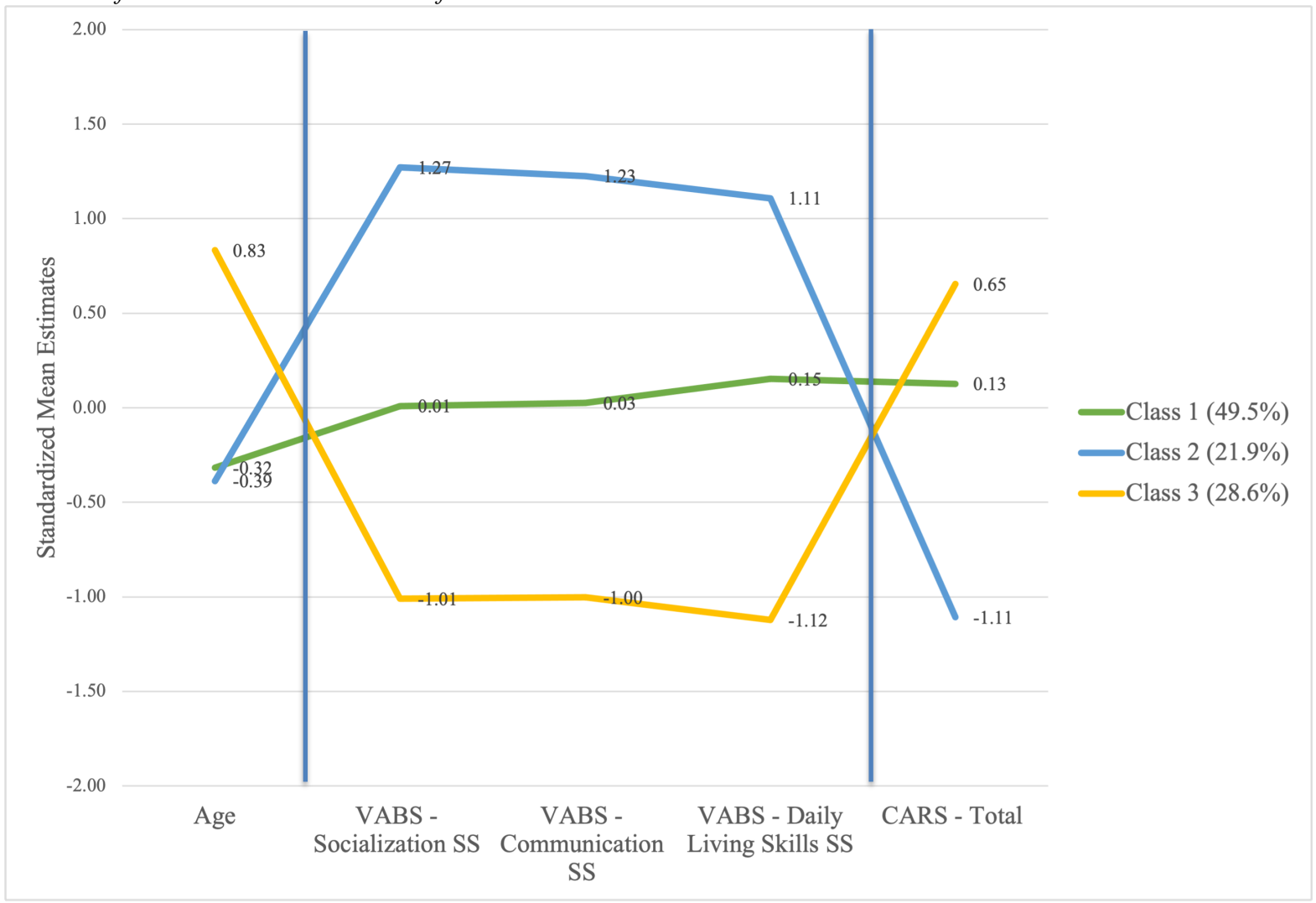
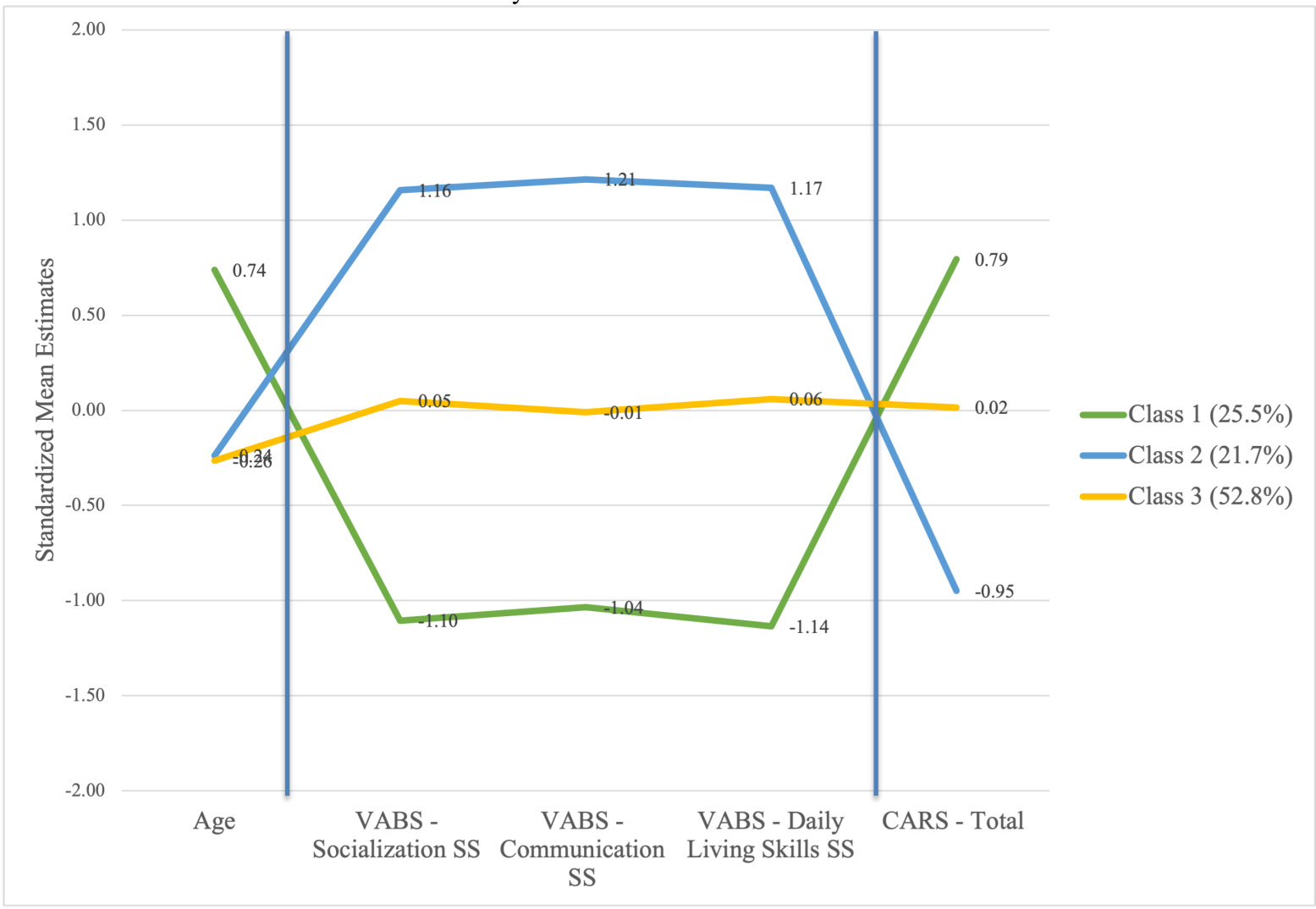


Figure 2

Latent Profile Model – 3-Class Solution for Boys in the SP Dataset



Explaining the Classes for the 3-Class Model

Each class is defined as having a distinct profile of scores across the selected variables in relation to the overall standardized means in the SP sample. This means that differences can be understood in terms of their standard deviations for the standardized mean, allowing comparison across variables with differing scales. In order to bring clinical meaning to the profiles, we also generated the unstandardized means for each variable, within each separate class. The unstandardized means for those variables with clinically meaningful values or qualitative descriptors will be described. The overall means for the total sample, for the variables selected, can be found in Table 26.

Class 1. The children who would be most likely to fall within this profile (based on probabilities generated through the Latent Profile Analysis) are younger than the overall average of the SP sample (*SD*=-.32). Their adaptive abilities are close to the overall sample average (*SD*=.01, .03, .15, respectively). Their autism symptoms are also close to the sample average (*SD*=.13).

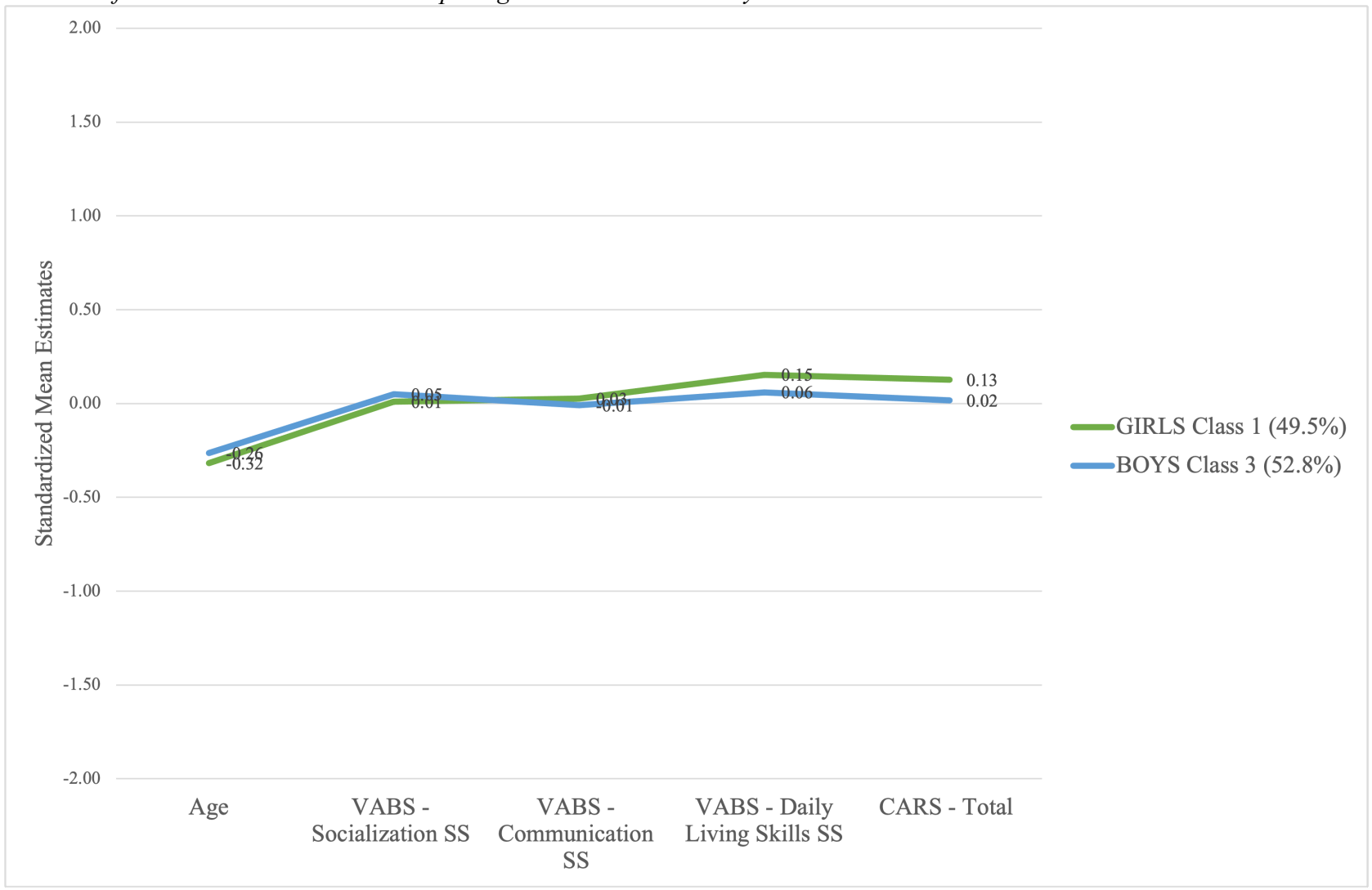
The unstandardized means indicate that the children likely to fall within this profile are approximately 3-1/2 years of age (*M*=42.22; *SD*=11.88). This translates to a difference of ~5 months from the overall sample mean for girls of 4 years and 1 month of age (*M*=4 years 1 month; *SD*=1 year, 8 months). Their scores on the Vineland indices of Socialization, Communication, and Daily Living Skills, were 62.78 (*SD*=5.22), 61.51 (*SD*=8.94), and 65.81 (*SD*=6.91), respectively. These scores are minimally different from the overall sample means for girls (*M*=62.64, *SD*=9.71, *M*=61.08, *SD*=14.74; *M*=63.46, *SD*=13.87). Based on the Vineland Qualitative Descriptors, these scores all fall within the Mildly Low range (standard scores 55 to 69). The mean CARS Total score for the children likely to fall within Class 1, was 30.96

(*SD*=3.38), a score that is similar to the overall sample mean for girls (*M*=30.36, *SD*=4.48) and falling within the CARS Severity Group - Mild to Moderate Symptoms of Autism Spectrum Disorder (Total score range: 30-36).

As evident in Figure 3, a comparison between the profiles of girls and boys across girls in Class 1 and Boys in Class 3 found minimal differences.

Figure 3

Latent Profile Model – SP Dataset – Comparing Girls in Class 1 to Boys in Class 3



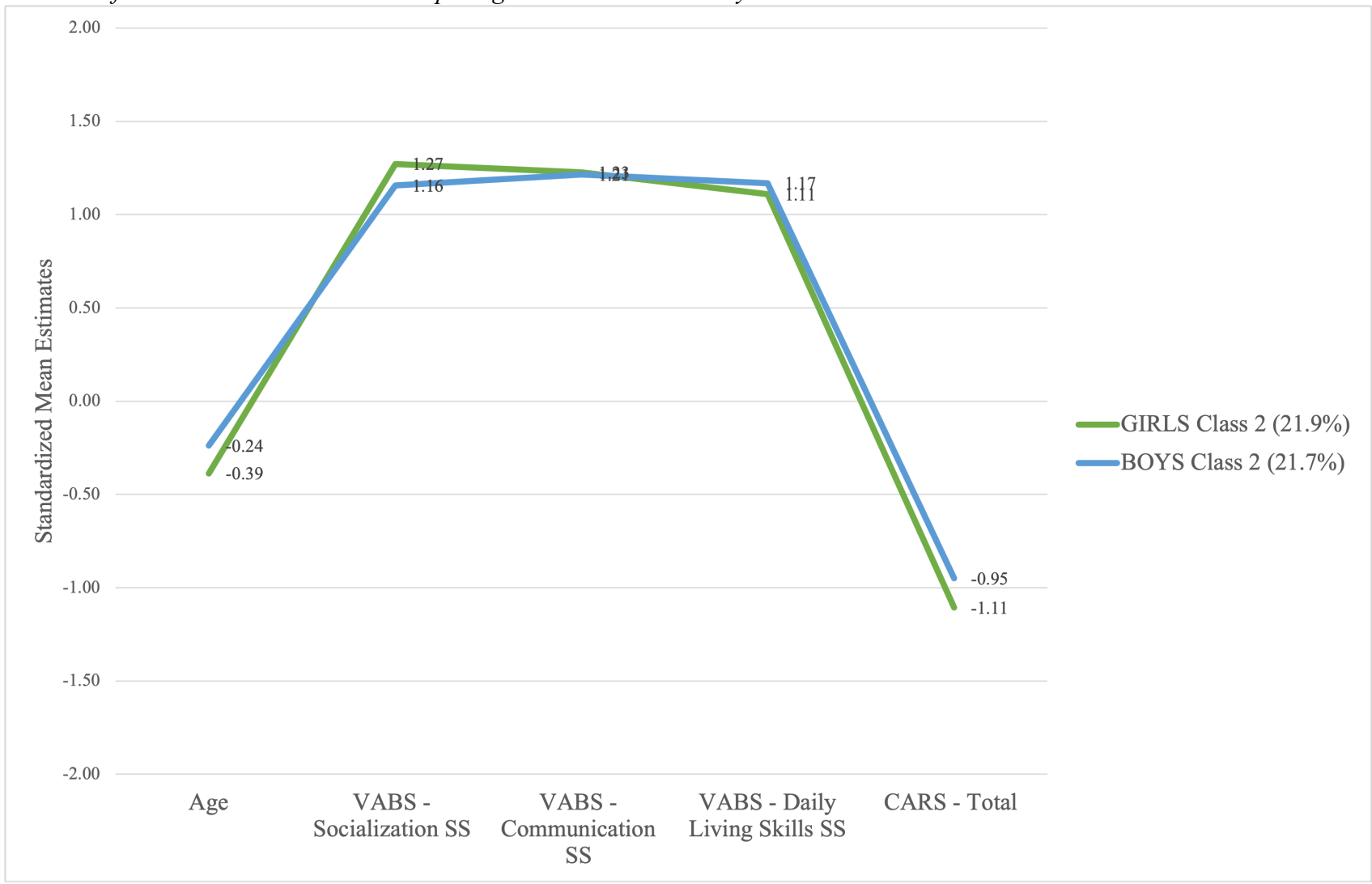
Class 2. The children who would be most likely to fall within this profile were younger than the overall sample average (*SD*=-.39). Their adaptive abilities were higher than the sample average (*SD*=1.27, 1.23, 1.11, respectively), while their autism symptoms were lower than the overall sample average (*SD*=-1.11).

The unstandardized means indicate that the children likely to fall within this profile were approximately 3-1/2 years of age (*M*=41.71, *SD*=12.86), a difference of approximately 5 months less than the overall sample mean age and very similar to Class 1. Their scores on the Vineland indices of Socialization, Communication, and Daily Living Skills, were 79.59 (*SD*=10.81), 78.81(*SD*=9.61), and 75.41 (*SD*=7.07), respectively, translating to 19, 16, and 13 points higher than the overall sample means for girls (*M*=62.64, *SD*=9.71, *M*=61.08, *SD*=14.74; *M*=63.46, *SD*=13.87) and falling within the Borderline Low range (standard scores 70 to 84). The mean CARS Total score for the children likely to fall within Class 2 was 25.11 (*SD*=2.92), translating to a 5.3-point difference from the overall sample mean for girls (*M*=30.36, *SD*=4.48) and falling within a different Severity Group than the overall sample and the other two classes, (i.e., Minimal-to-No Symptoms of Autism Spectrum Disorder).

As evident in Figure 4, a comparison between the profiles of girls and boys across girls in Class 2 and Boys in Class 2 found minimal differences.

Figure 4

Latent Profile Model – SP Dataset – Comparing Girls in Class 2 to Boys in Class 2



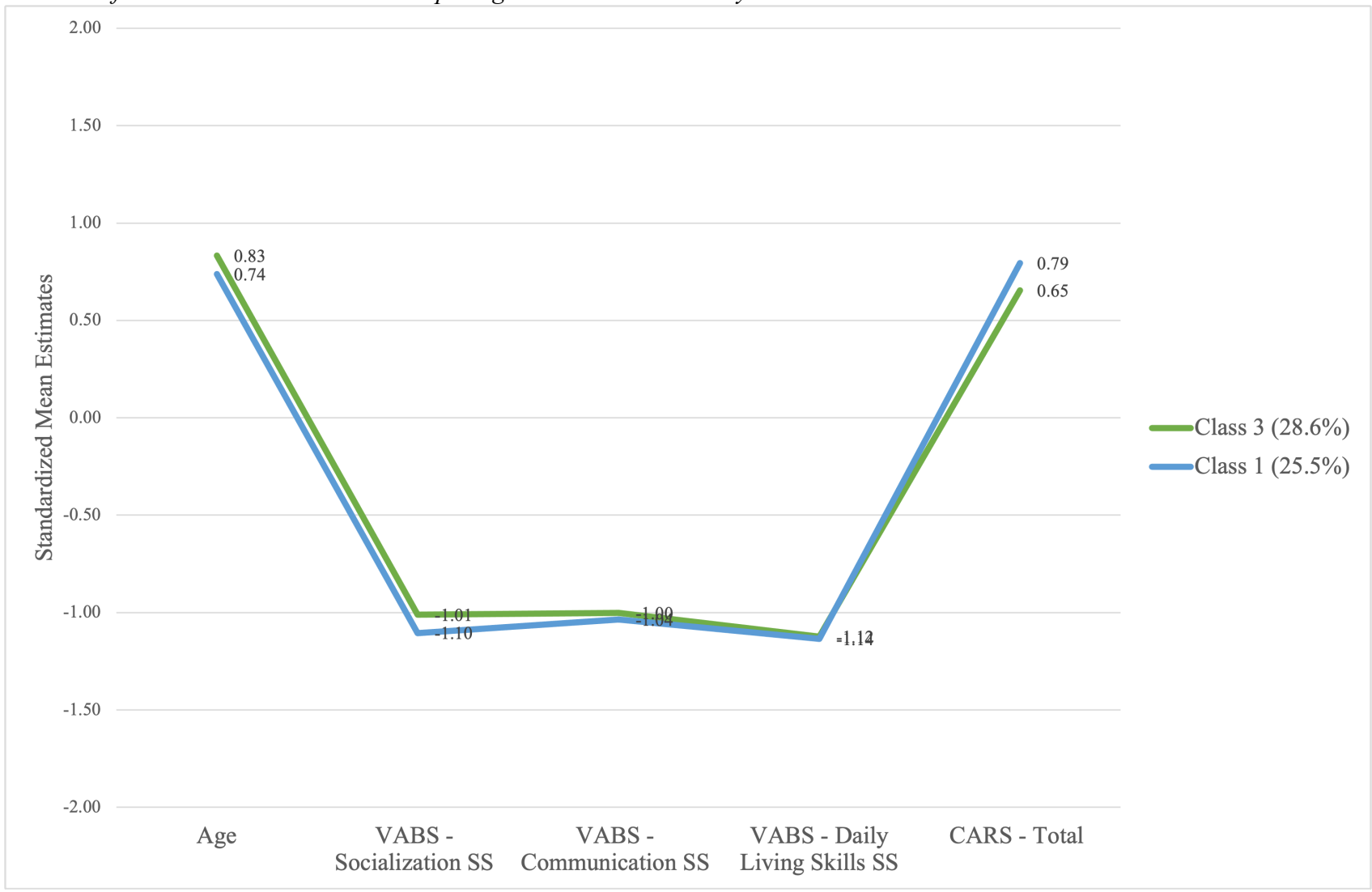
Class 3. The children who would be most likely to fall within this profile were older than the overall sample average (*SD*=.83). They demonstrated adaptive abilities lower than the sample average (*SD*=-1.01, -1.00, -1.12, respectively), and their autism symptoms were rated to be higher (i.e., more severe) than the overall sample average (*SD*=.65).

The unstandardized means indicate that the children likely to fall within this profile were approximately 4 years, 11 months of age (*M*=59.47, *SD*=11.19); a difference of approximately 10 months older than the overall sample mean age. Their scores on the Vineland indices of Socialization, Communication, and Daily Living Skills, were 46.14 (*SD*=6.81), 47.59 (*SD*=8.84), and 52.62 (*SD*=4.37), respectively. This translates to 17, 13, and 11 points lower than the overall sample means for girls (*M*=62.64, *SD*=9.71, *M*=61.08, *SD*=14.74; *M*=63.46, *SD*=13.87). These scores fall within the Moderately Low range (standard scores of 40 to 54). The mean CARS Total score for the children likely to fall within Class 3 was 33.35 (*SD*=3.63), which is approximately a 3-point difference from the mean of the overall sample for girls (*M*=30.36, *SD*=4.48); and falls within the CARS Severity Group - Mild to Moderate Symptoms of Autism Spectrum Disorder (score range: 30-36).

As evident in Figure 5, a comparison between the profiles of girls and boys across girls in Class 3 and Boys in Class 1 found minimal differences.

Figure 5

Latent Profile Model – SP Dataset – Comparing Girls in Class 3 to Boys in Class 1



LPA in the SFARI dataset: Results

As mentioned in the methods section, there are many variables in the SFARI dataset. Based on clinical judgement and a thorough investigation of the literature, variables most relevant to addressing the research question were selected for use in the LPA. The variables selected represent a range of clinically significant behavioural constructs related to how children with ASD function. They reflect direct observational measures, as well as parent-report measures, and are sampled from clinicians and parents. Variables from the Vineland, ADI-R, ADOS-2, RBS-R, SRS, CBCL were selected. Variables included, age, Verbal and Non-Verbal IQ; Vineland Socialization, Communication, and Daily Living Skills, Standard scores; Socialization Total score from the ADI-R; Calibrated Severity Score, Communication/Social Total score, Social Affect Total score from the ADOS; Repetitive and Restrictive Behaviour Total scores across the ADI-R and ADOS-2, along with Average Domain scores for the RBS-R (i.e., Ritual/Sameness Average, Self-Injurious Behaviour, Stereotypy, Compulsive, and Restricted); SRS Total scores for parents, and CBCL Internalizing and Externalizing. Teacher data was available for only half of the sample and was therefore not used for this analysis. The profile analysis was conducted separately for girls and boys, using similar procedures to those described above.

As in the previous LPA, the best fitting LPA models were determined by running multiple LPA models, beginning with two classes, and adding one additional class in each subsequent iteration until the number of participants in each group went below the 5% threshold. The BIC, AIC, and Entropy values were acceptable for each model and continued to improve with each iteration (see Tables 27 & 28). As a result, clinical judgement was needed to examine the clinical relevance, and interpretability of the patterns in the profiles. After running multiple LPA models

with the 20 indicator variables mentioned above, the 5-Class solution showed good fit for both girls and boys, while also demonstrating distinct and clinically informative profiles. Participants were relatively equally distributed amongst the 5 classes (see Tables 29 & 30). The class with the fewest participants consisted of 49/375 (13.1%) of the sample for girls, and 270/2382 (11.3%) for boys. Based on the fit statistics, as well as clinical interpretation, it was decided that the 5-Class model should be retained for both girls and boys, although models 3, 4, and 6 also showed adequate fit. The 5 and 6-Class models showed differences between the profiles of girls and boys that were not seen in the 2-, 3- and 4-Class models. The 6-Class model (and models with more classes) provided solutions that were not well defined or meaningfully different from one another and therefore lacked clinical relevance. Given the goal to investigate sex differences we decided to investigate model 5. A plot depicting the Standardized Mean Estimates (SME) for the selected Model with 5 Classes is shown in Figures 5 and 6, for girls and boys respectively.

Figure 6

Latent Profile Model – 5-Class Solution for Girls in the SFARI Dataset

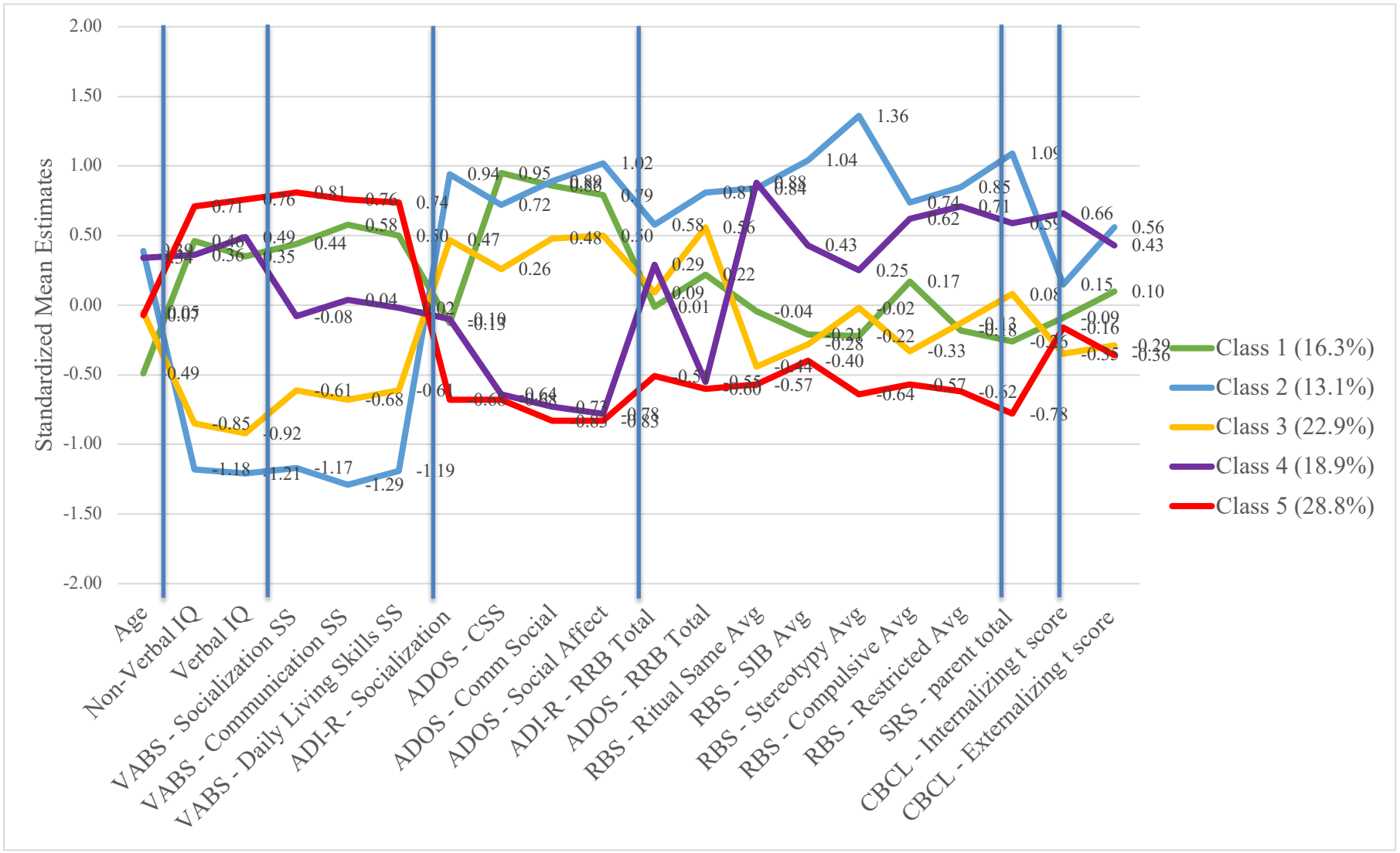
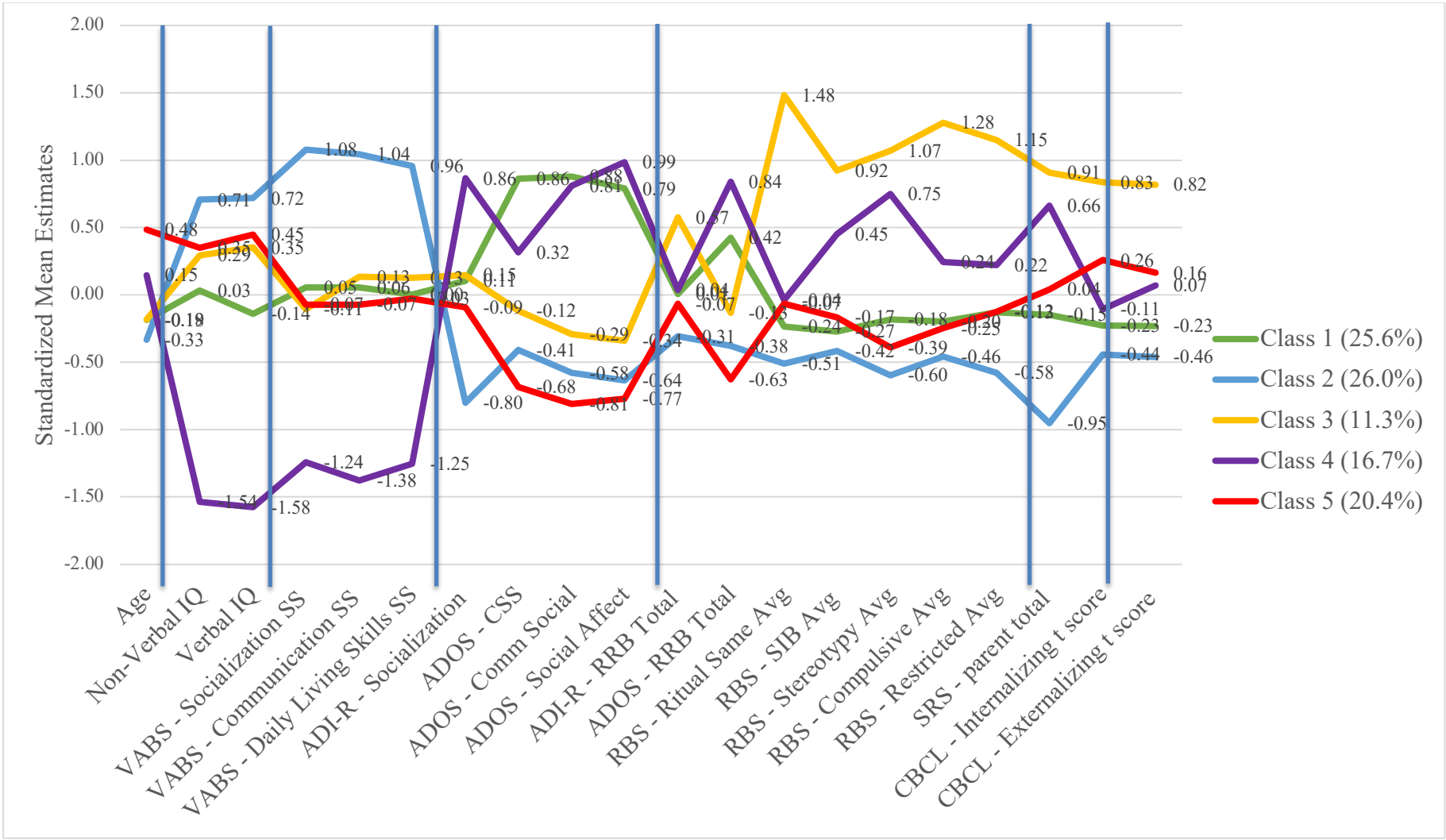


Figure 7

Latent Profile Model – 5-Class Solution for Boys in the SFARI Dataset



Explaining the Classes for the 5-Class Model. Below is an explanation of the Classes for Girls in the SFARI sample. Each class is defined as having a distinct profile of scores across the selected variables in relation to the overall standardized means in the SFARI sample. This means that differences can be understood in terms of their standard deviations for the standardized mean, allowing comparison across variables with differing scales (These values can be found in Tables 31 and 32, for girls and boys respectively). In order to bring clinical meaning to the profiles, we also generated the unstandardized means for each variable, within each separate class. The unstandardized means for those variables with clinically meaningful values or qualitative descriptors will be described and can be found in Table 33. All overall sample means reported in Table 33 for reference are the overall means for girls only. The overall means for the total sample, for the variables selected, can be found in Table 8 for reference.

Class 1. Children who are likely to fall within this profile (see Figure 8 below) are younger than the average of the total sample of female children (*SD*=-0.49). They were on average 7 years and 5 months of age, a difference of 1 year and 8 months from the overall total sample mean for girls of 9 years and 2 months. Their non-verbal and verbal IQ scores are higher than the overall sample average (*SD*=0.46, & 0.35, respectively). Their Nonverbal IQ was considered to fall in the Average range of intellectual functioning (*M*=90.17, *SD*=17.67), representing a 12.18-point difference from the overall sample mean of 77.99 (*SD*=26.13). Their Verbal IQ fell in the Average range of intellectual functioning (*M*=84.98, *SD*=18.70), an 11.03-point difference from the overall sample mean of 73.95 (*SD*=32.38). Similarly, their adaptive functioning on the Vineland indices of Socialization, Communication, and Daily Living Skills were above the overall sample average (*SD*=0.44, 0.58, 0.50, respectively). More specifically,

their Socialization score fell within the Borderline Low range (*M*=75.11, *SD*=9.01). A 5.85-point difference from the overall sample mean of 69.26 (*SD*=12.71). As did their Communication score, (*M*=82.87, *SD*=9.45), an 8.24-point difference from the overall sample mean of 74.63 (*SD*=14.11); and their Daily Living Skills score (*M*=80.82, *SD*=9.10), which was a 7.1-point difference from the overall sample mean of 73.72 (*SD*=13.52).

The autism symptoms of girls in Class 1, as measured by the ADI-R Social Interaction Domain total score, were largely similar to the overall sample mean (*SD*=-0.13), while their scores on the ADOS-2, including the Calibrated Severity Score, the Communication and Socialization Total score, and the Social Affect Total score were above the overall sample means (*SD*= 0.95, 0.86, 0.79, respectively). More specifically, the Social Interaction Domain Total score on the ADI-R (*M*=19.69, *SD*=5.63) represented a 0.86-point difference from the overall sample mean of 20.55 (*SD*=6.00). On the ADOS-2, the CSS was classified as High (*M*=9.13, *SD*=0.85), and represented a 1.67-point difference from the overall sample mean of 7.46 (*SD*=1.72). The ADOS-2 Communication and Social Total score (*M*=17.74, *SD*=2.70), represented a 4.41-point difference from the overall sample mean of 13.77 (*SD*=4.45), and the ADOS-2 Social Affect Total score (*M*=15.00, *SD*=2.48) represented a 3.40-point difference from the overall sample mean of 11.60 (*SD*=4.20).

Similarly, the children likely to fall in Class 1 have restricted and repetitive behaviour scores that differ based on the measure used. Restricted and Repetitive Behaviour measures are best understood using standardized means as their unstandardized means have been transformed and do not have clinical descriptors, meaning that they are not easily interpretable clinically. The children in Class 1 show RRB scores close to the overall sample mean on the ADI-R Restricted and Repetitive Behaviour Total score (*SD*=-0.01). They show RRB scores similar to the overall sample mean when measured using the ADOS-2 Restricted and Repetitive Behaviour Total score

(*SD*=0.22). On the RBS-R, the children's scores vary in terms of effect size across the different types of restrictive and repetitive behaviour. The average score for the Ritualistic Sameness score is close to the overall average (*SD*=-0.04), moderately below the overall average for the Self-Injurious Behaviour score (*SD*=-0.21), the Stereotypy score (*SD*=-0.22), and the Restricted Interests score (*SD*=-0.18), and above the overall average for the Compulsive Behaviour score (SD=0.17).

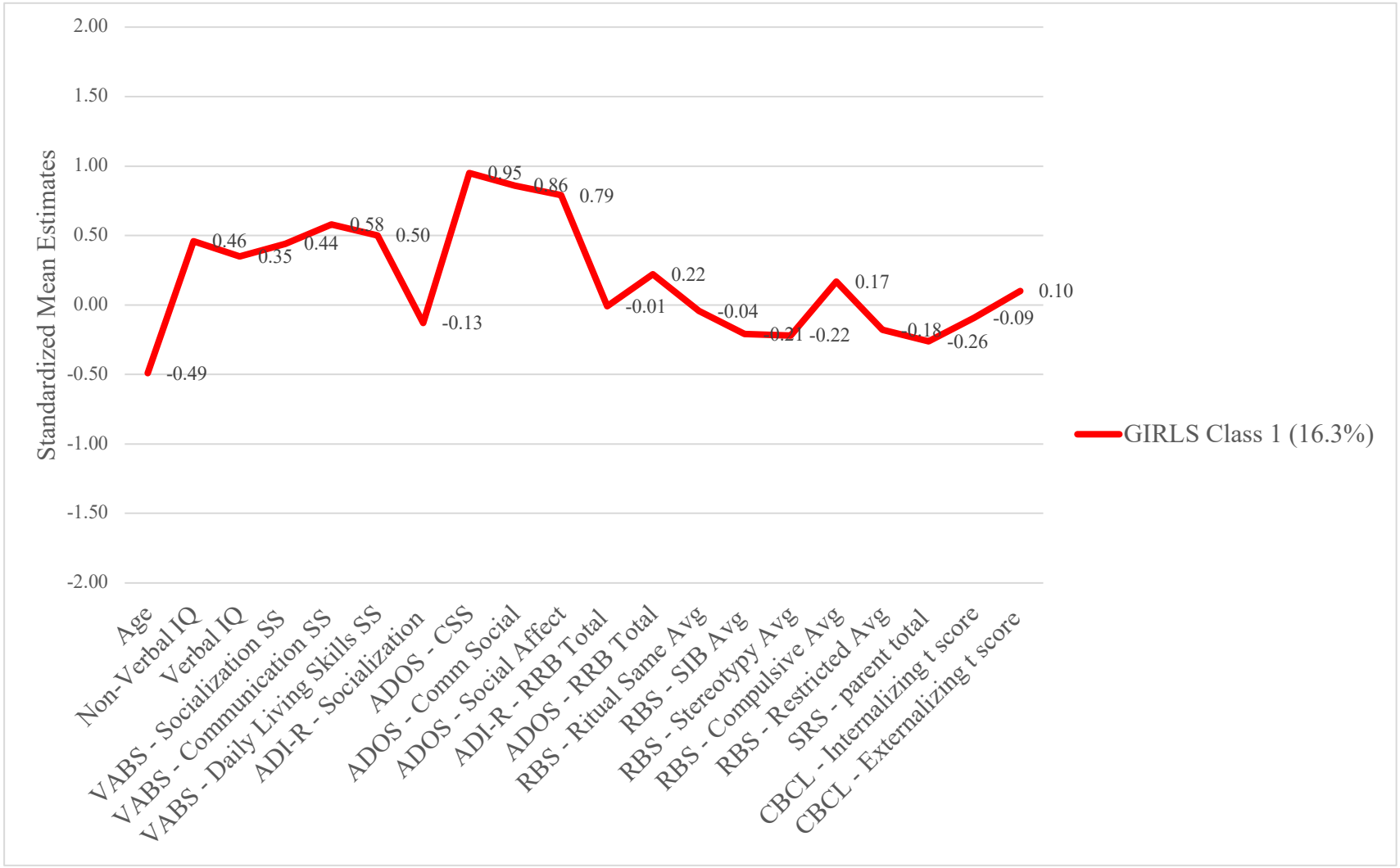
Class 1 children's social responsiveness was less impaired (lower score) than the overall sample mean (*SD*=-0.26), as measured by the SRS Total score, although the effect size was moderate. Children's social responsiveness as reported by parents, corresponded to a T-score of 73 and fell within the Moderate range (SRS Parent total: *M*=92.48, *SD*=22.67). This corresponds to a 3-point T-score difference from the overall sample mean of 99.95 (*SD*=27.30; T-score of approximately 76). Scores in this range indicate deficiencies in reciprocal social behaviour that are clinically significant and lead to substantial interference with everyday social interactions. Such scores are typical for children with autism spectrum disorders of moderate severity.

Children in Class 1 had internalizing and externalizing behaviours that were close to the overall sample average (*SD*=-0.09, 0.10, respectively). The CBCL Internalizing T-score Parent (*M*=58.65, *SD*=9.96) fell within the Acceptable range. A 1.15-point difference from the overall sample mean of 59.80 (*SD*=10.04). Externalizing behaviour also fell within the Acceptable clinical range (CBCL Externalizing T-score Parent: *M*=58.23, *SD*=9.60), a 0.71-point difference from the overall sample mean of 57.52 (*SD*=10.02).

In summary, relative to the whole sample of girls, Class 1 girls tended to be younger, with higher cognitive and adaptive functioning (Average and Borderline range). They showed higher autism symptom severity but slightly lower social responsiveness difficulties. However, they tended to be in the normal range for internalizing and externalizing behaviours.

Figure 8

Latent Profile Model –SFARI Dataset –Girls in Class 1



Class 2. Children who are likely to fall within this profile were older than the average age of the total sample of girls (*SD*=0.39; see Figure 9). They were approximately 10 years and 8 months of age, i.e., 1 year, 6 months older than the overall total sample mean for girls of 9 years, 2 months. Their non-verbal and verbal IQ were considerably below the overall average (*SD*=-1.18, & -1.21). Specifically, their Nonverbal IQ fell in the range of Moderate Intellectual Disability (*M*=46.78, *SD*=17.36), a 31.21-point difference from the overall sample mean of 77.99 (*SD*=26.13); while their Verbal IQ fell in the range of Severe Intellectual Disability (*M*=34.71, *SD*=20.66), a 39.24-point difference from the overall sample mean of 73.95 (*SD*=32.38). Similarly, their adaptive functioning on the Vineland indices of Socialization, Communication, and Daily Living Skills were below the overall sample average (*SD*=-1.17, -1.29, -1.19, respectively). On the measure itself, their scores on the Socialization domain fell within the Moderately Low range (VABS Soc: *M*=54.12, *SD*=8.81, a 15.14-point difference from the overall sample mean of 69.26 (*SD*=12.71). Their Communication domain score fell within the Mildly Low range (*M*=56.20, *SD*=9.89), an 18.43-point difference from the overall sample mean of 74.63 (*SD*=14.11), as did their Daily Living Skills score (*M*=57.47, *SD*=10.20), representing a 16.25-point difference from the overall sample mean of 73.72 (*SD*=13.52).

Their autism symptoms, as measured by the ADI-R Social Interaction Domain total score, as well as the ADOS-2 Calibrated Severity Score, the ADOS-2 Communication and Socialization Total score, and the ADOS-2 Social Affect Total score were higher for girls in Class 2 than the overall sample means (*SD*=0.98, 0.72, 0.89, 1.02, respectively). More specifically, their ADI-R Social Interaction Domain Total score (*M*=26.49, *SD*=3.44), represented a 5.94-point difference from the overall sample mean of 20.55 (*SD*=6.00). On the ADOS-2 their CSS was classified as High (*M*=8.67, *SD*=1.42), a 1.21-point difference from the overall sample mean of 7.46

(*SD*=1.72). Their ADOS-2 Communication and Social Total score (*M*=17.76, *SD*=3.33) represented a 3.99-point difference from the overall sample mean of 13.77 (*SD*=4.45). Their ADOS-2 Social Affect Total score (*M*=15.82, *SD*=2.64), represented a 4.22-point difference from the overall sample mean of 11.60 (*SD*=4.20).

The children likely to fall in Class 2 had restricted and repetitive behaviour scores higher than the overall mean on the ADI-R Restricted and Repetitive Behaviour Total score and the ADOS-2 Restricted and Repetitive Behaviour Total score (*SD*=0.58, & 0.81). Children likely to fall in Class 2, showed scores higher than the sample average across the different types of RRBs measured on the RBS-R; including the Ritualistic Sameness score (*SD*=0.84), the Self-Injurious Behaviour score (*SD*=1.04), the Stereotypy score (*SD*=1.36), the Compulsive Behaviour score (*SD*=0.74), and the Restricted Interests score (*SD*=0.85).

In Class 2, the children's social responsiveness was more impaired (higher score) than the overall sample mean (*SD*=1.09), as measured by the SRS Total score. Children's social responsiveness as measured by the SRS Parent total corresponded to a T-score of 87 and fell within the Severe range (*M*=130.64, *SD*=17.06). Scores in this range indicate deficiencies in reciprocal social behaviour that are clinically significant and lead to severe interference with everyday social interactions. Such scores are strongly associated with clinical diagnosis of an autism spectrum disorder. This corresponds to an 11-point T-score difference from the overall sample mean of 99.95 (*SD*=27.30; T-score of approximately 76).

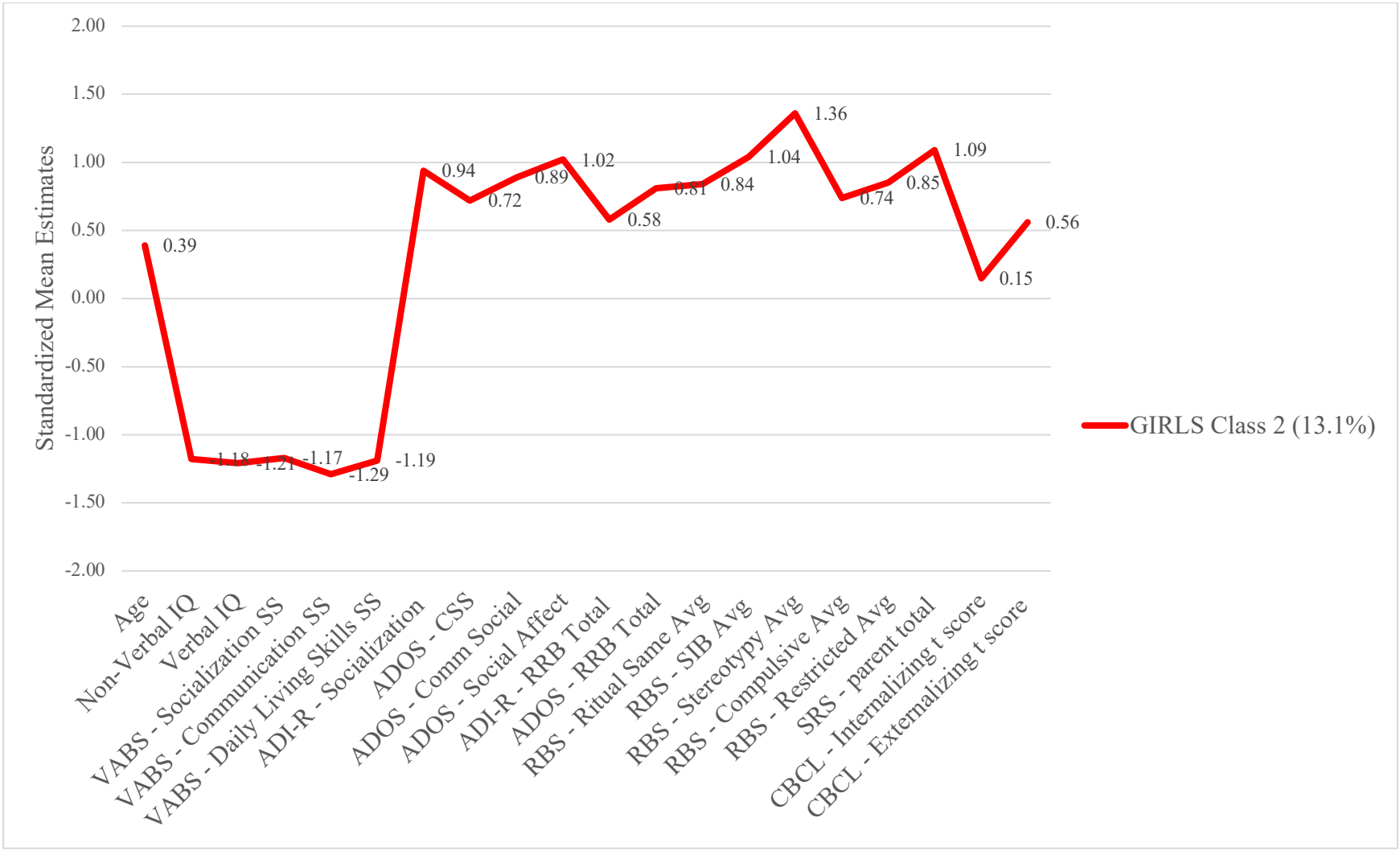
Internalizing behaviours for children in Class 2 were largely similar to the overall sample mean however, their externalizing behaviours were moderately higher than the overall sample average (*SD*=0.15, 0.56, respectively). Internalizing and externalizing behaviour fell within the Borderline Clinical range [CBCL Internalizing T-score Parent: *M*=61.35, *SD*=9.58, a 1.55-point difference from the overall sample mean of 59.80 (*SD*=10.04); CBCL Externalizing T-score

Parent: *M*=63.21, *SD*=9.62, a 5.69-point difference from the overall sample mean of 57.52 (*SD*=10.02)].

In summary, relative to the whole sample of girls, Class 2 girls tended to be older, with lower cognitive functioning (in the Moderate to Severe range). In terms of their adaptive functioning, their socialization skills were moderately low while their socialization and daily living skills were mildly low. They showed higher autism symptom severity, higher levels of restricted and repetitive behaviour, and more impaired social responsiveness (Severe range). They showed internalizing behaviours close to the sample average and within the Borderline range, while they showed high externalizing behaviours.

Figure 9

Latent Profile Model –SFARI Dataset –Girls in Class 2



Class 3. The children who are likely to fall within this profile are similar in age to the average of the total sample of girls (*SD*=-0.05). They were on average 8 years, 11 months of age, a 3-month difference from the overall total sample mean for girls of 9 years, 2 months. Their non-verbal and verbal IQ were below the overall average (*SD*=-0.85, & -0.92, respectively). Their Nonverbal IQ fell in the range of Mild Intellectual Disability (*M*=55.90, *SD*=16.03), a 22.09-point difference from the overall sample mean of 77.99 (*SD*=26.13). Their Verbal IQ fell in the range of Moderate Intellectual Disability (*M*=44.44, *SD*=18.05), a 29.51-point difference from the overall sample mean of 73.95 (*SD*=32.38). Similarly, their adaptive functioning on the Vineland indices of Socialization, Communication, and Daily Living Skills were below the overall sample average (*SD*=-0.61, -0.68, -0.61, respectively). Their Communication, Socialization, and Daily Living Skills scores fell within the Mildly Low range, [VABS Soc: *M*=61.56, *SD*=8.72; a 7.7-point difference from the overall sample mean of 69.26 (*SD*=12.71); VABS Comm: *M*=65.14, *SD*=7.83, a 9.49-point difference from the overall sample mean of 74.63 (*SD*=14.11); and VABS DLS: *M*=65.41, *SD*=7.49, an 8.31-point difference from the overall sample mean of 73.72 (*SD*=13.52)].

For girls in Class 3, autism symptoms, as measured by the ADI-R Social Interaction Domain total score, as well as the ADOS-2 Calibrated Severity Score, the ADOS-2 Communication and Socialization Total score, and the ADOS-2 Social Affect Total score were moderately higher (i.e., more severe) than the overall sample means (*SD*=0.47, 0.26, 0.48, 0.50, respectively). In terms of their autism symptoms, their ADI-R Social Interaction Domain Total score (*M*=23.38, *SD*=4.44), represented a 2.83-point difference from the overall sample mean of 20.55 (*SD*=6.00). On the ADOS-2 their CSS was classified as Moderate (*M*=7.92, *SD*=1.41), a 0.46-point difference from the overall sample mean of 7.46 (*SD*=1.72). Their ADOS-2

Communication and Social Total score (*M*=16.01, *SD*=3.47), represented a 2.24-point difference from the overall sample mean of 13.77 (*SD*=4.45). Their ADOS-2 Social Affect Total score (*M*=13.67, *SD*=3.18) represented a 2.07-point difference from the overall sample mean of 11.60 (*SD*=4.20).

The children likely to fall in Class 3 have restricted and repetitive behaviour scores that differ based on the measure used. The children show RRB scores similar to the overall mean on the ADI-R Restricted and Repetitive Behaviour Total score (*SD*=0.09). They show RRB scores moderately higher than the overall sample mean when measured using the ADOS-2 Restricted and Repetitive Behaviour Total score (*SD*=0.56). On the RBS, the children's scores vary across the different types of restrictive and repetitive behaviour. Children likely to fall in Class 3, show scores lower than the sample average on the Ritualistic Sameness score (*SD*=-.044), scores lower than the overall sample average on the Self-Injurious Behaviour score (*SD*=-0.28), scores close to the overall sample mean on the Stereotypy score (SD=-0.02), scores lower than the overall sample average on the Compulsive Behaviour score (*SD*=-0.33), and scores lower than the overall sample mean on the Restricted Interests score (*SD*=-0.13).

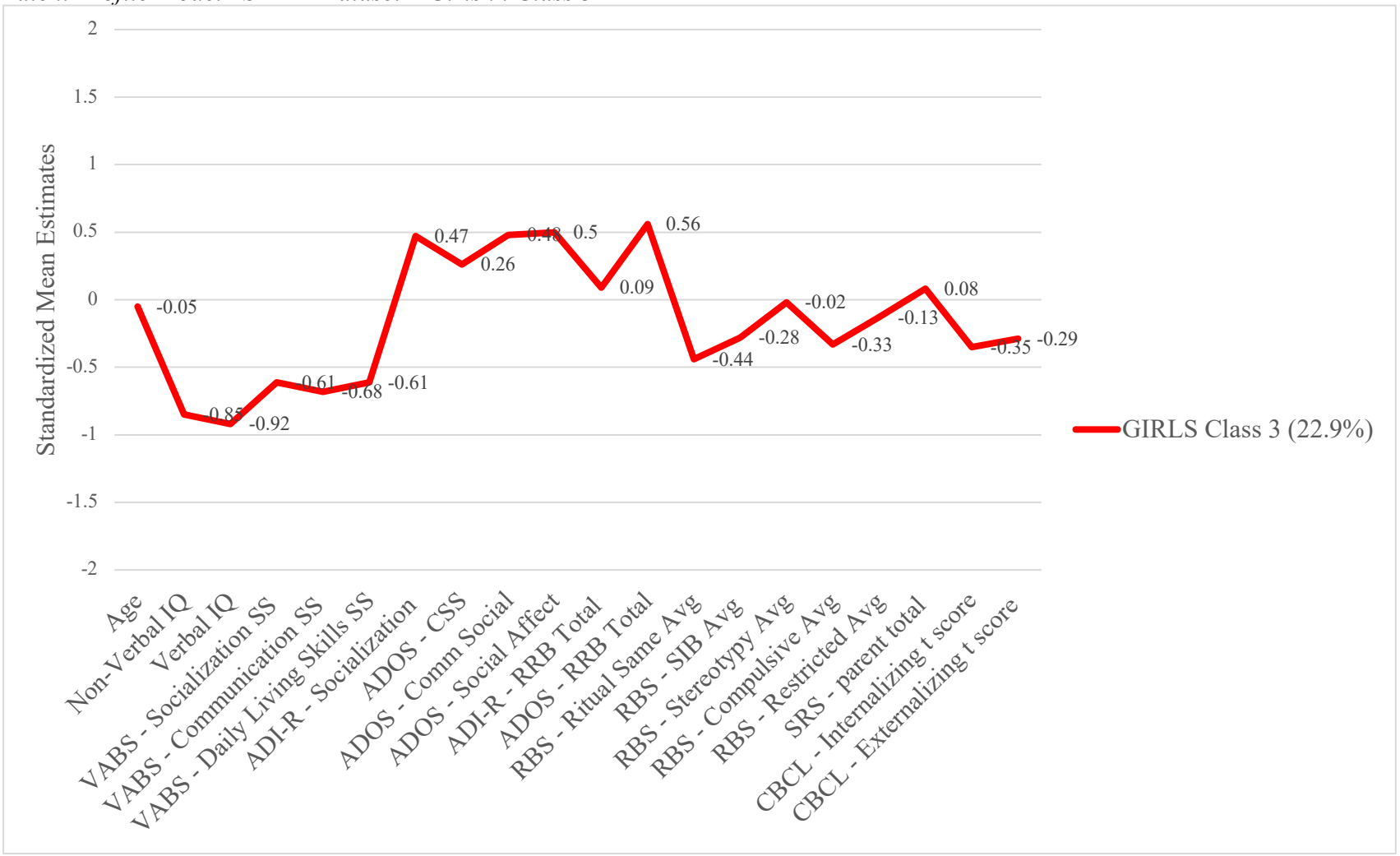
Children's social responsiveness was similar to the overall sample mean (*SD*=0.08), as measured by the SRS Total score. The SRS Parent total (*M*=102.86, *SD*=20.13) corresponded to a T-score of 77 and fell within the Severe range. Scores in this range indicate deficiencies in reciprocal social behaviour that are clinically significant and lead to severe interference with everyday social interactions. Such scores are strongly associated with clinical diagnosis of an autism spectrum disorder. The overall sample mean for the SRS is 99.95 (*SD*=27.30; T-score of approximately 76), which falls within the Moderate range. This corresponds to a 1-point T-score difference from the overall sample mean.

Their internalizing and externalizing behaviours were lower than the overall sample average (*SD*=-0.35, -0.29, respectively). More specifically, the CBCL Internalizing T-score Parent (*M*=56.56, *SD*=9.47), fell within the Average/Acceptable range. A 3.24-point difference from the overall sample mean of 59.80 (*SD*=10.04). As did the CBCL Externalizing T-score Parent (*M*=54.78, *SD*=8.79), a 2.74-point difference from the overall sample mean of 57.52 (*SD*=10.02).

In summary, relative to the whole sample of girls, Class 3 girls tended to be similar age close to the average of the overall sample, with lower cognitive functioning (in the Mild to Moderate range). Their adaptive skills were all in the Mildly Low range. They showed moderate autism symptom severity. They showed close to average levels of restrictive and repetitive behaviour on the ADI-R, and higher levels of restricted and repetitive behaviour on the ADOS-2. On the RBS-R their scores were generally lower than the average. Their social responsiveness was similar to the sample average and fell in the Moderate/Severe range, indicating more severe social deficits. They showed internalizing and externalizing behaviours lower than the sample average and within the Average/Acceptable range.

Figure 10

Latent Profile Model –SFARI Dataset – Girls in Class 3



Class 4. Children who were likely to fall within this profile are older than average of the total sample of girls (*SD*= 0.34). They were on average approximately 10 years, 4 months of age, a difference of 1 years and 2 months from the overall sample mean for girls of 9 years, 2 months. They have moderately higher non-verbal and verbal IQ (*SD*=0.36 & 0.49). Their Nonverbal IQ fell within the Average range of intellectual ability (*M*=87.58, *SD*=20.57), a 9.59-point difference from the overall sample mean of 77.99 (*SD*=26.13). Their Verbal IQ also fell within the Average range of intellectual ability (*M*=90.18, *SD*=23.14), a 16.23-point difference from the overall sample mean of 73.95 (*SD*=32.38). They have adaptive skills on the Vineland indices of Socialization, Communication, and Daily Living Skills that are close to the overall sample average (*SD*=-0.08, 0.04, -0.02, respectively). On the Vineland their Socialization total score fell within the Mildly Low range (*M*=68.28, *SD*=9.78), a 0.98-point difference from the overall sample mean of 69.26 (*SD*=12.71). While their Communication and Daily Living Skills scores fell within the Borderline Low range [VABS Comm: *M*=75.41, *SD*=8.18, a 0.78-point difference from the overall sample mean of 74.63 (*SD*=14.11); VABS DLS: *M*=73.58, *SD*=10.93, a 0.14-point difference from the overall sample mean of 73.72 (*SD*=13.52)].

Their autism symptoms, as measured by the ADI-R Social Interaction Domain total score, are closer to the overall sample mean (*SD*=-0.10), while their autism symptoms measured by the ADOS-2 Calibrated Severity Score, the ADOS-2 Communication and Socialization Total score, and the ADOS-2 Social Affect Total score were lower than overall sample means (*SD*=-0.64, -0.73, -0.78, respectively). Their ADI-R Social Interaction Domain Total score (*M*=20.04, *SD*=4.82), represented a 0.51-point difference from the overall sample mean of 20.55 (*SD*=6.00). On the ADOS-2 their CSS was classified as Moderate (*M*=6.28, *SD*=1.31), a1.18-point difference from the overall sample mean of 7.46 (*SD*=1.72). Their ADOS-2 Communication and Social

Total score (*M*=10.54, *SD*=2.51), and represented a 3.23-point difference from the overall sample mean of 13.77 (*SD*=4.45); while their ADOS-2 Social Affect Total score (*M*=8.31, *SD*=2.35), represented a 3.29-point difference from the overall sample mean of 11.60 (*SD*=4.20).

The children likely to fall in Class 4 have restricted and repetitive behaviour scores that vary depending on the measure used. Their scores on the ADI-R Restricted and Repetitive Behaviour Total score are moderately higher than the overall sample average (*SD*=0.29), while their scores on the ADOS-2 Restricted and Repetitive Behaviour Total score were lower than the overall sample mean (*SD*=-0.55). On the RBS-R, the children show RRB moderately higher than the overall sample means on the Ritualistic Sameness score (*SD*=0.88), the Self-Injurious Behaviour score (*SD*=0.43), the Stereotypy score (*SD*=0.25), the Compulsive Behaviour score (*SD*=0.62), and the Restricted Interests score (*SD*=0.71).

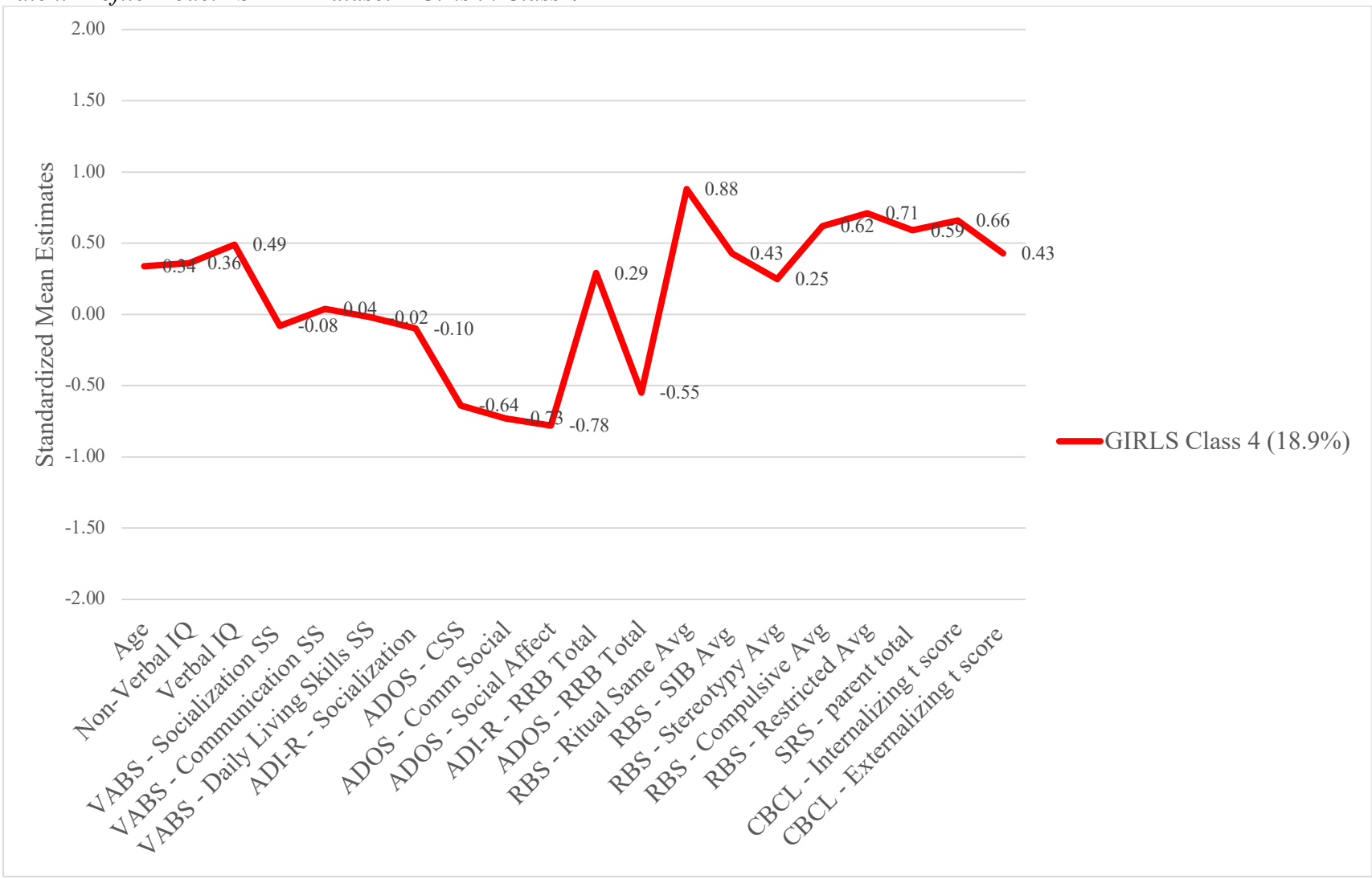
Children's social responsiveness was somewhat more impaired (higher score) than the overall sample mean (*SD*=0.59), as measured by the SRS Total score. The SRS Parent total (*M*=115.20, *SD*=20.10) corresponded to a T-score of 82 and fell within the Severe range. Scores in this range indicate deficiencies in reciprocal social behaviour that are clinically significant and lead to severe interference with everyday social interactions. This is higher than the overall sample mean of 99.95 (*SD*=27.30; T-score of approximately 76), an approximately 6-point T-score difference.

Children's internalizing and externalizing behaviours were moderately higher than the overall sample average (*SD*=0.66, 0.43, respectively). Their CBCL Internalizing T-score Parent (*M*=66.13, *SD*=8.61) represented a 6.33-point difference from the overall sample mean of 59.80 (*SD*=10.04) and their CBCL Externalizing T-score Parent (*M*=61.94, *SD*=7.74), represented a 4.42-point difference from the overall sample mean of 57.52 (*SD*=10.02).

In summary, relative to the whole sample of girls, Class 4 girls tended to be older, with higher cognitive functioning (in the Average range of intellectual ability). Their adaptive skills fell in the Mildly Low to Borderline Low ranges. They showed moderate autism symptom severity. They showed higher levels of restrictive and repetitive behaviour on the ADI-R, and lower levels of restricted and repetitive behaviour on the ADOS-2. On the RBS-R they showed higher RRBs. Their social responsiveness was somewhat more impaired than the overall sample mean and fell in the Severe range. They showed higher internalizing and externalizing behaviours.

Figure 11

Latent Profile Model –SFARI Dataset – Girls in Class 4



Class 5. Children who are likely to fall within this profile were similar in age to the average of the total sample of girls (*SD*= -0.07). They were on average 8 years and 11 months of age, a difference of 3 months from the overall sample mean for girls of 9 years 2 months. They have higher non-verbal and verbal IQ (*SD*=.71, 0.76, respectively). Their Nonverbal IQ fell in the Average range of intellectual functioning (*M*=96.68, *SD*=16.03). An 18.69-point difference from the overall sample mean of 77.99 (*SD*=26.13). Their Verbal IQ also fell within the Average range of intellectual functioning (*M*=98.45, *SD*=20.17), a 24.5-point difference from the overall sample mean of 73.95 (*SD*=32.38). Similarly, they had higher adaptive skills on the Vineland indices of Socialization, Communication, and Daily Living Skills (*SD*=0.81, 0.76, 0.74, respectively). On the Vineland, their Socialization score fell within the Borderline Low range (*M*=79.61, *SD*=8.99), a 10.35-point difference from the overall sample mean of 69.26 (*SD*=12.71). Their Communication domain score fell within the Average range (*M*=85.38, *SD*=11.20), a 10.75-point difference from the overall sample mean of 74.63 (*SD*=14.11). While their Daily Living Skills score fell within the Borderline Low range (*M*=83.87, *SD*=10.74), a 10.15-point difference from the overall sample mean of 73.72 (*SD*=13.52).

The autism symptoms for Class 5, as measured by the ADI-R Social Interaction Domain total score, as well as the ADOS-2 Calibrated Severity Score, the ADOS-2 Communication and Socialization Total score, and the ADOS-2 Social Affect Total score were lower than overall sample means (*SD*=-0.68, -0.68, -0.83, -0.83, respectively). Their ADI-R Social Interaction Domain Total score (*M*=16.43, *SD*=5.60) represented a 4.12-point difference from the overall sample mean of 20.55 (*SD*=6.00). On the ADOS-2, their CSS was classified as Moderate (*M*=6.28, *SD*=1.25), a 1.18-point difference from the overall sample mean of 7.46 (*SD*=1.72). Their ADOS-2 Communication and Social Total score (*M*=10.06, *SD*=2.31), represented a 3.71-

point difference from the overall sample mean of 13.77 (*SD*=4.45). While their ADOS-2 Social Affect Total score (*M*=7.99, *SD*=2.16), represented a 3.61-point difference from the overall sample mean of 11.60 (*SD*=4.20).

The children likely to fall in Class 5 have restricted and repetitive behaviour scores that are lower than the overall sample averages across the ADI-R Restricted and Repetitive Behaviour Total score (*SD*=-0.51), and the ADOS-2 Restricted and Repetitive Behaviour Total score (*SD*=-0.60). On the RBS, the children show less RRB than the overall sample, i.e., the Ritualistic Sameness score (*SD*=-0.57), the Self-Injurious Behaviour score (*SD*=-0.40), the Stereotypy score (*SD*=-0.64), the Compulsive Behaviour score (*SD*=-0.57), and the Restricted Interests score (*SD*=0.62).

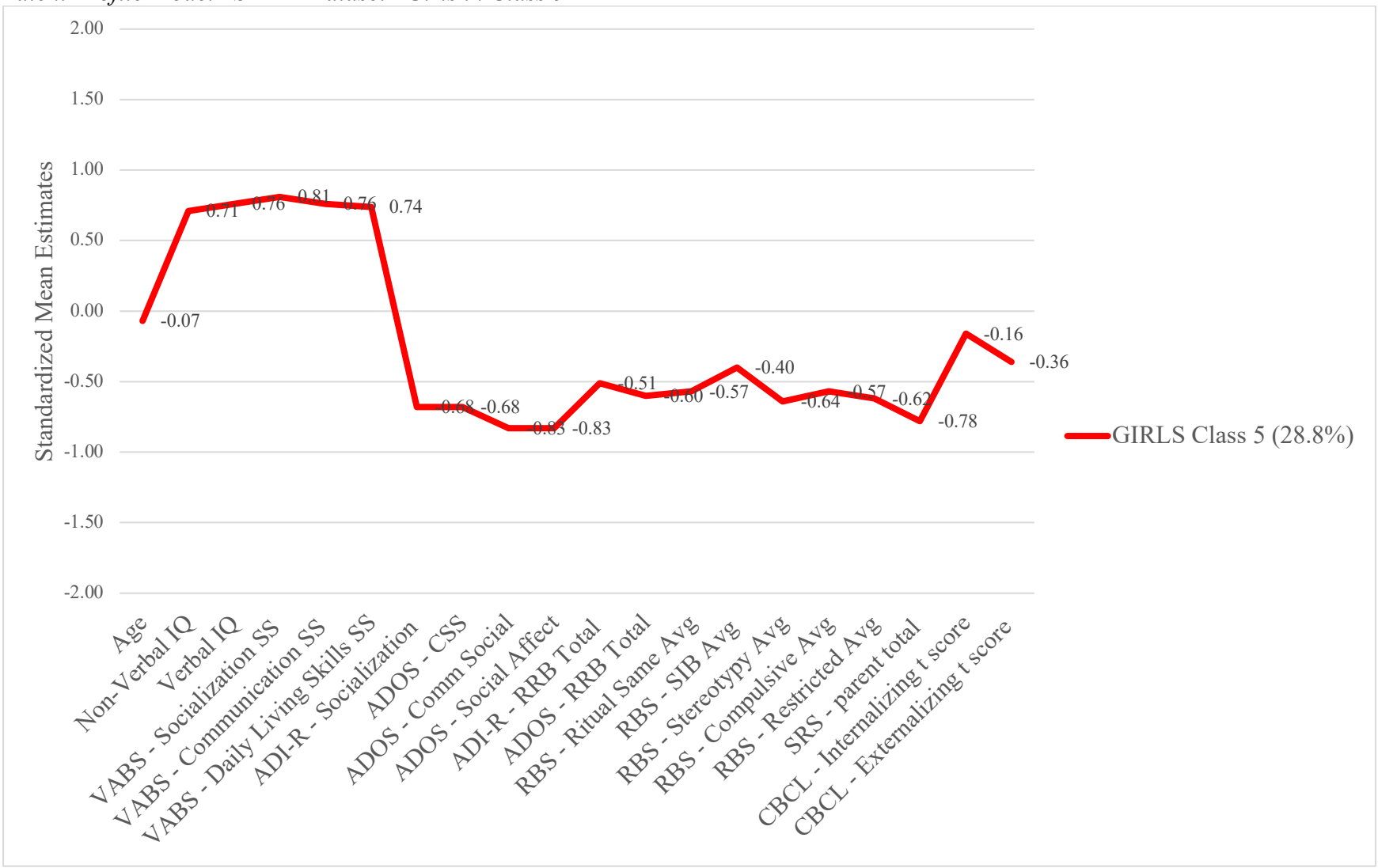
Children's social responsiveness was less impaired (lower score) than the overall sample mean (*SD*=-0.78), as measured by the SRS Total score. Their SRS Parent total (*M*=78.56, *SD*=22.65), corresponded to a T-score of 67, and fell within the Moderate range. Scores in this range indicate deficiencies in reciprocal social behaviour that are clinically significant and lead to substantial interference with everyday social interactions. Such scores are typical for children with autism spectrum disorders of moderate severity. This corresponds to a 9-point T-score difference from the overall sample mean of 99.95 (*SD*=27.30; T-score of approximately 76).

Children's internalizing and externalizing behaviours were moderately lower than the overall sample average (*SD*=-0.16, -0.36, respectively). Their internalizing and externalizing scores fell within the Average range [CBCL Internalizing T-score Parent: *M*=58.19, *SD*=9.79, a1.61-point difference from the overall sample mean of 59.80 (*SD*=10.04); CBCL Externalizing T-score Parent: *M*=53.88, *SD*=10.49, a 3.64-point difference from the overall sample mean of 57.52 (*SD*=10.02)].

In summary, relative to the whole sample of girls, Class 5 girls tended to be average age, with higher cognitive functioning (in the Average range of intellectual functioning). Their adaptive skills were all higher than average and in the Borderline Low to Average range of the measure. They showed moderate autism symptom severity, lower than the sample average. They showed lower levels of restrictive and repetitive behaviour on the ADI-R, ADOS-2, and RBS-R. Their social responsiveness was less impaired than the sample average and fell in the Moderate range. Similarly, their internalizing and externalizing behaviours were lower than the sample average and within the Average/Acceptable range.

Figure 12

Latent Profile Model –SFARI Dataset –Girls in Class 5



Examining the Classes across girls and boys for the 5-Class Model.

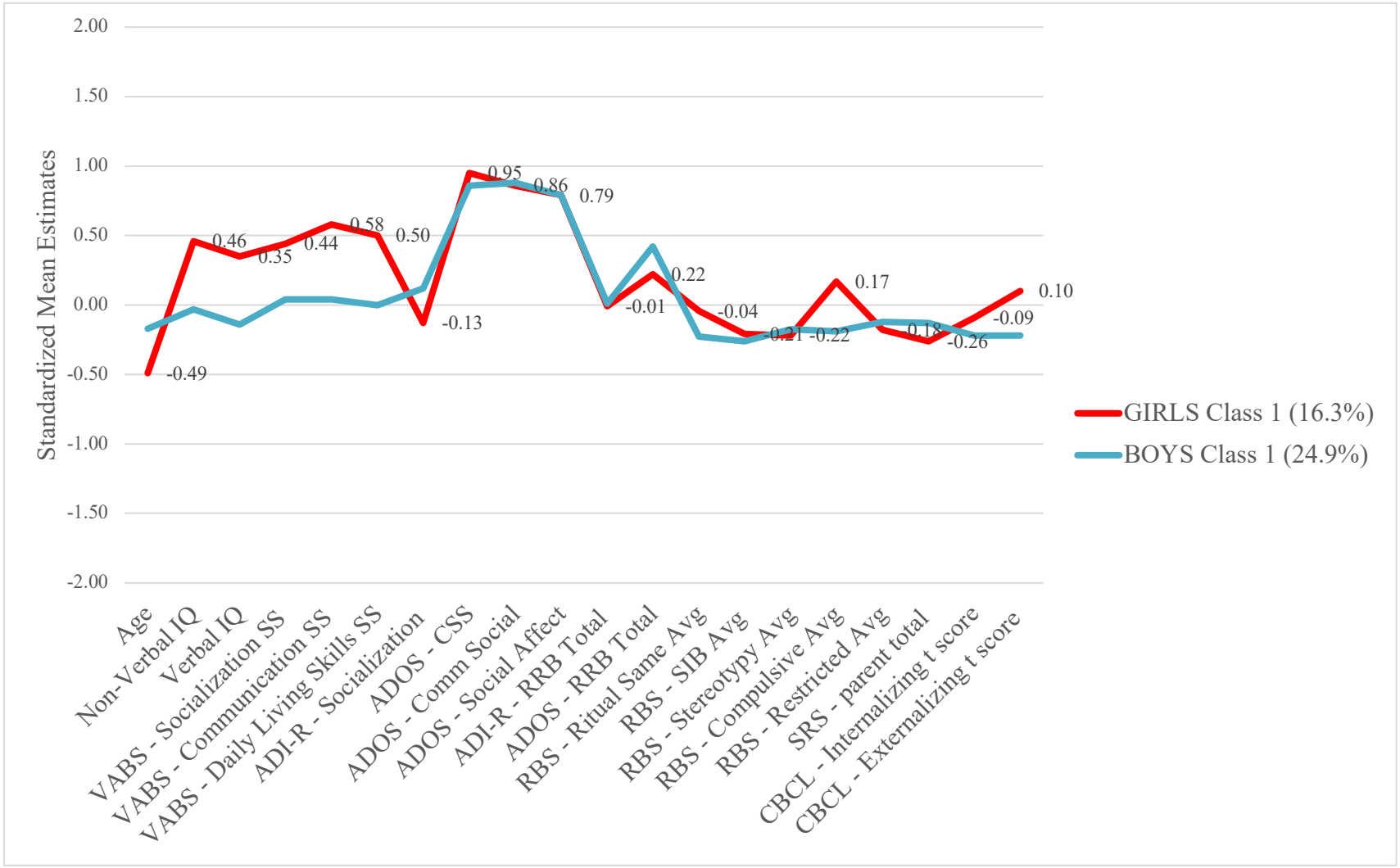
Below is an examination of similar profiles across girls and boys. These profiles were paired based primarily on a visual inspection of the the profiles, looking for patterns in the data, and also by evaluating the sample sizes of each profile. There are several limitations to this comparison that must be identified at the outset. Although these limitations also apply to the SP LPA results, they are especially relevant in the SFARI sample given this sample's diverse number and type of variables and the complexity of interpreting the results. Firstly, LPAs for girls and boys were completed using separate analyses. This means that the standardized mean differencs used to create these graphs each compare to their own respective means, that is different means for girls and boys. Since we know that some of the means for boys and girls do differ across measures, this certainly impacts how we compare each difference to its respective mean, and how we interpret the differences across sex/group. This means that although the profiles may look similar, that there are likely some differences that we are unable to capture and comment on. In addition, we have not completed a statistical analysis of comparison, or post-hoc tests, as this type of analysis (comparing boys and girls separately) has not yet been done and no such analyses was available at the time this study was conducted. Similarly, this is the first exploratory study of its kind, so no validation work, or comparison to another sample, can or has been done to see if there is stability and consistency across these findings. While these comparisons are interesting and are unique contributions that we want to highlight, it is truly expoloratory, and we are simply commenting on how the profiles are similar or different qualitatively.

Class 1 for girls; Class 1 for boys. Figure 13 presents Class 1 for girls and boys. These profiles are largely similar. Of note, girls do appear to be younger, and have higher IQs and

adaptive abilities. Girls also seemed to have slightly higher levels of compulsive behaviour than boys in Class 1. This may represent our first evidence for girls who are cognitively strong, with RRBs that are more compulsive in nature. Measuring the quality rather than type of their RRBs may be a better way to detect them based on this comparison.

Figure 13

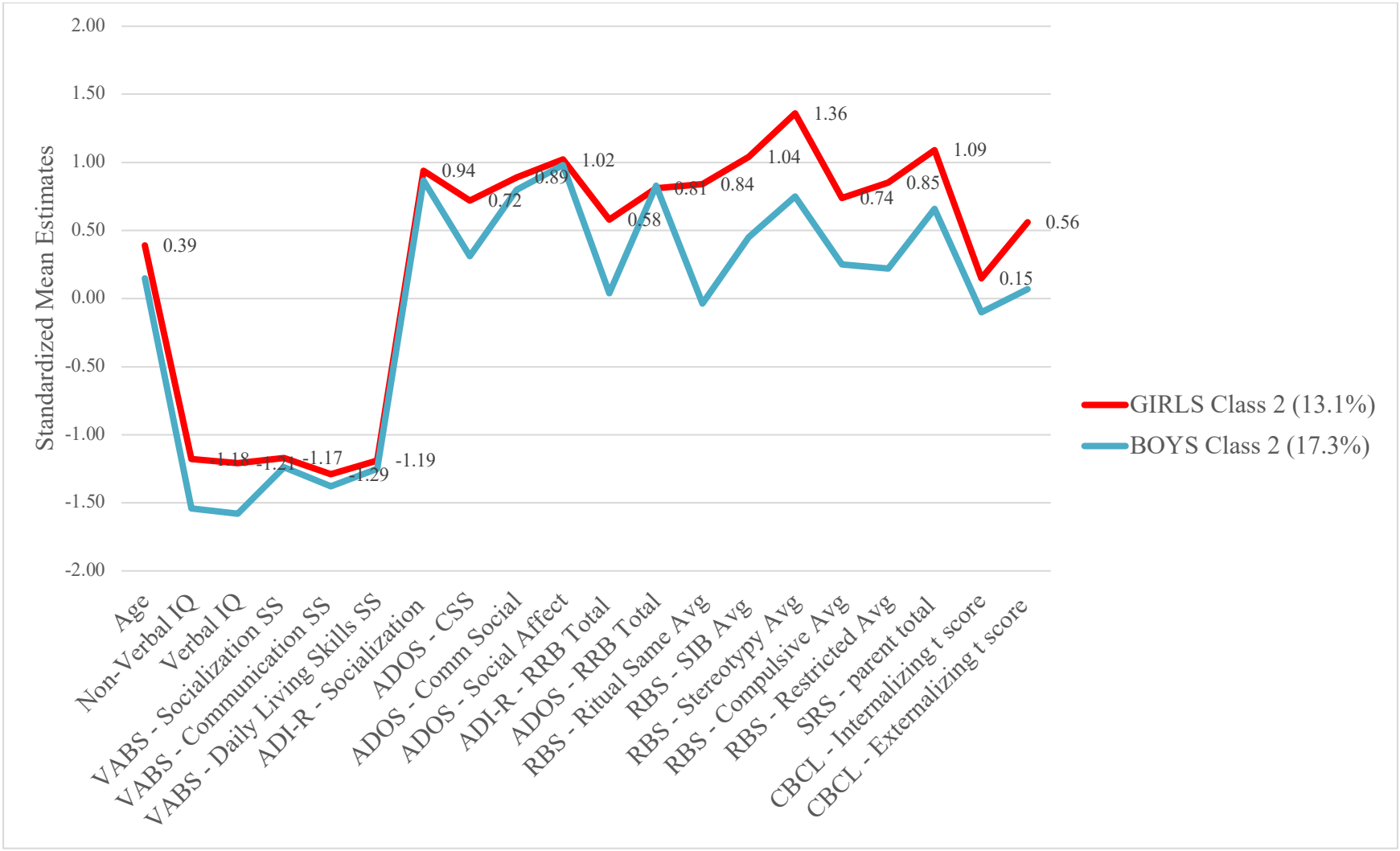
Latent Profile Model –SFARI Dataset – Comparing Girls in Class 1 to Boys in Class 1



Class 2 for girls; Class 2 for boys. Figure 14 presents Class 2 for girls and boys. Again, these profiles are highly similar. Girls and boys are similar ages, boys appear to have lower IQs (although both profiles show low IQs). Boys appear to show slightly lower autism severity on the ADOS-2 CSS, while similar autism symptoms are seen across other severity and symptom scores. Boys seems to show lower RRBs on the ADI-R and RBS than girls (parent report measures), while similar levels of RRBs are found on the ADOS-2 across sexes. Girls and boys showed similar levels of internalizing symptoms, while girls showed higher externalizing symptoms. So girls in this group show higher RRBs and externalizing behaviours than boys based on parent report, a deviation from our previous comparison, and consistent with the finding from Study 1 where girls showed more RRBs and externalizing behaviours. Although captured here, this finding is not reflected in all of our comparisons, suggesting that simply comparing groups may be missing the nuance found here.

Figure 14

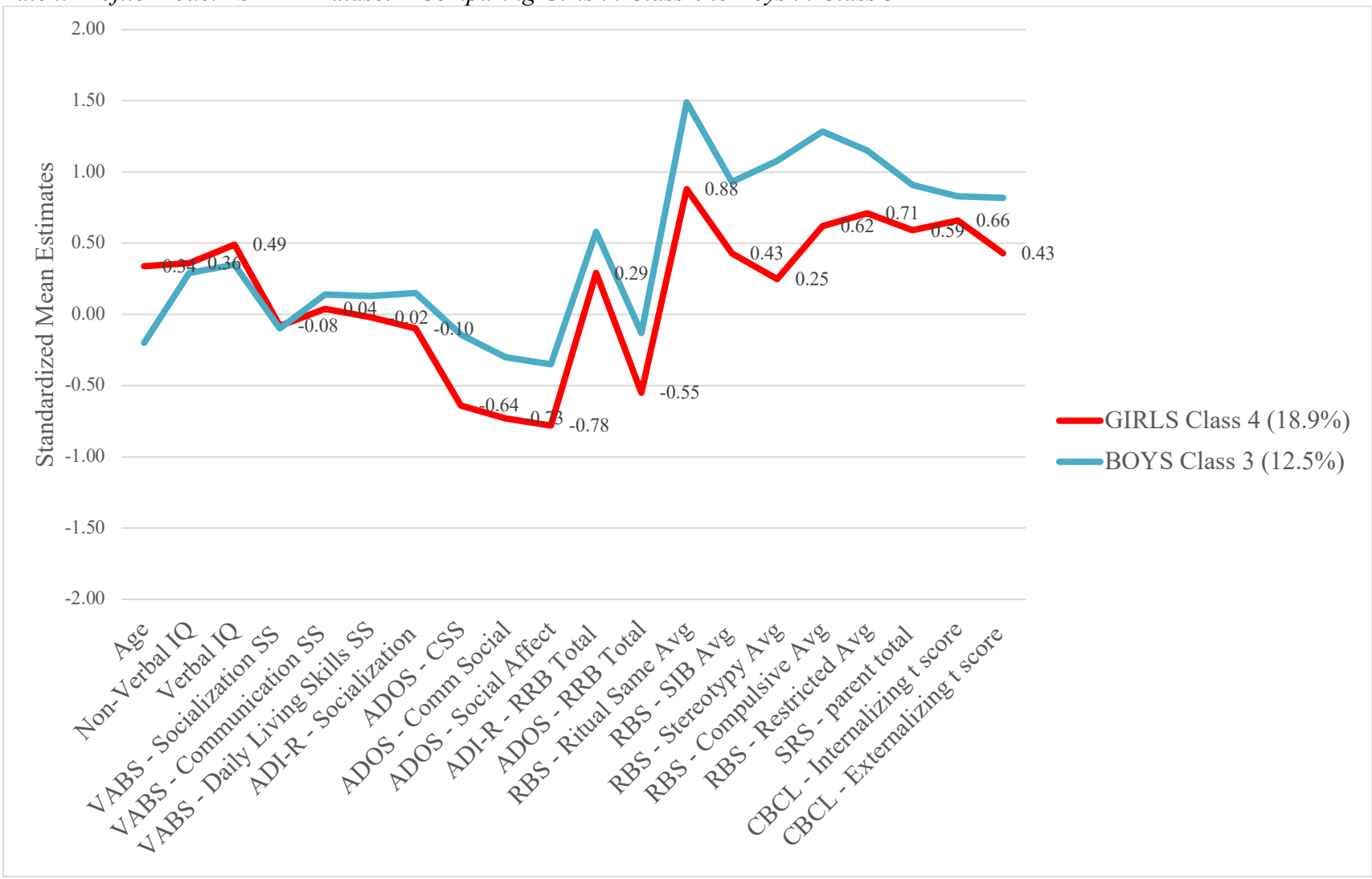
Latent Profile Model –SFARI Dataset – Comparing Girls in Class 2 to Boys in Class 2



Class 4 for girls; Class 3 for boys. Figure 15 compares the girls in Class 4 to the boys in Class 3. Girls appear to be slightly older than the boys, while both sexes appear to have similar IQs and adaptive abilities. Girls showed lower levels of autism symptoms across measures. Girls showed lower levels of RRBs across measures but a similar pattern to boys across the types of RRBs. Similarly, both groups showed less RRBs on the ADOS-2 than parent-report measures. Boys showed more impaired social responsiveness. Girls showed lower levels of internalizing and externalizing behaviours. Here we see the oppositive pattern to our overall group finding. Girls in this comparison show lower levels of internalizing and externalizing behaviours. We also see clear differences emerging in how different types of measures capture RRBs in different profiles of children. This provides support for investigating sex differences as there are already differences in the strengths/challenges of the assessment measures themselves.

Figure 15

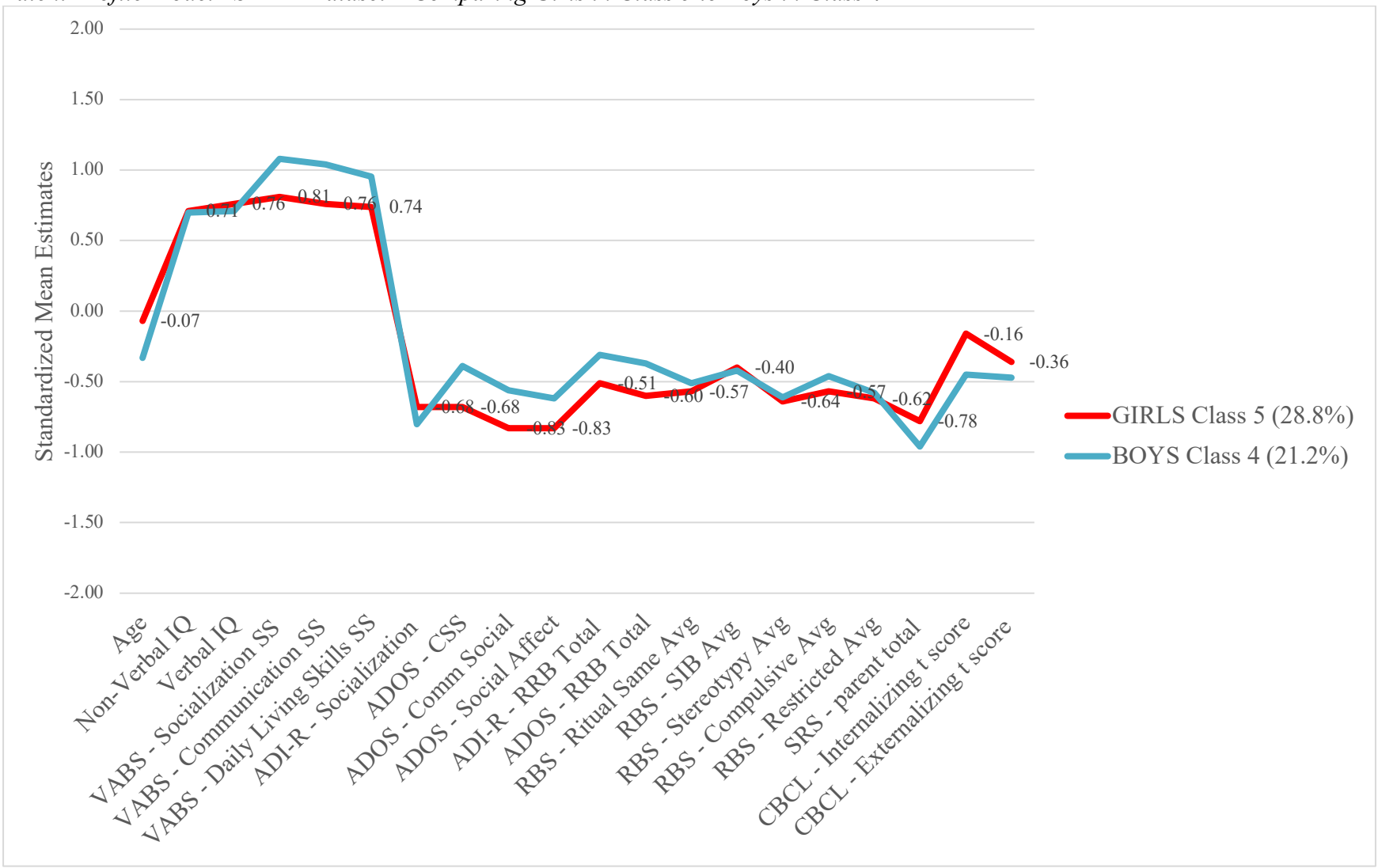
Latent Profile Model –SFARI Dataset – Comparing Girls in Class 4 to Boys in Class 3



Class 5 for girls; Class 4 for boys. Figure 16 compares girls in Class 5 to boys in Class 4. The profiles are largely similar with few notable differences. Boys appear to have slightly higher (or stronger) adaptive abilities. They also appear to have slightly higher autism symptoms. This comparison suggests similarities between girls and boys in these profiles.

Figure 16

Latent Profile Model –SFARI Dataset – Comparing Girls in Class 5 to Boys in Class 4



Class 3 for girls and Class 5 for boys were graphed individually in Figures 17 and 18, as they do seem to be quite distinct from one another.

Class 3 for girls. Girls in Class 3 are aged close to the sample average, their IQ and adaptive abilities are lower than the mean, and fall Mild and Moderate Intellectual Disability range, respectively (Non-Verbal IQ & Verbal IQ). Their autism symptoms are higher than the mean and fall within the Moderate range. They show higher RRBs on the ADOS-2, but close to the mean or lower levels of RRBs across parent-report measures, as well as lower internalizing and externalizing symptoms. These appear to be girls who have lower IQ, poorer adaptive functioning, and higher autism symptoms. Their RRBs were best captured on the ADOS, while parents reported fewer concerns in this area. Again, we see subtle evidence that these measures may vary in their ability to detect symptoms/differences in different types of children; and here we see evidence for sex differences as well. This finding is worthy of further exploration.

Class 5 for boys. Boys in Class 5 are older than the mean, have moderately higher IQs, that fall in the Average range of intellectual functioning. They show close to the mean adaptive abilities, in the Borderline Low range. They show lower autism symptoms, in the Moderate range. They demonstrate moderately fewer RRBs on the ADOS-2, but score closer to the mean on parent-report measures. Their social responsiveness is similar to the overall average for boys, they show moderately higher levels of internalizing behaviours, and externalizing behaviours similar to the mean.

Figure 17

Latent Profile Model –SFARI Dataset – Girls in Class 3

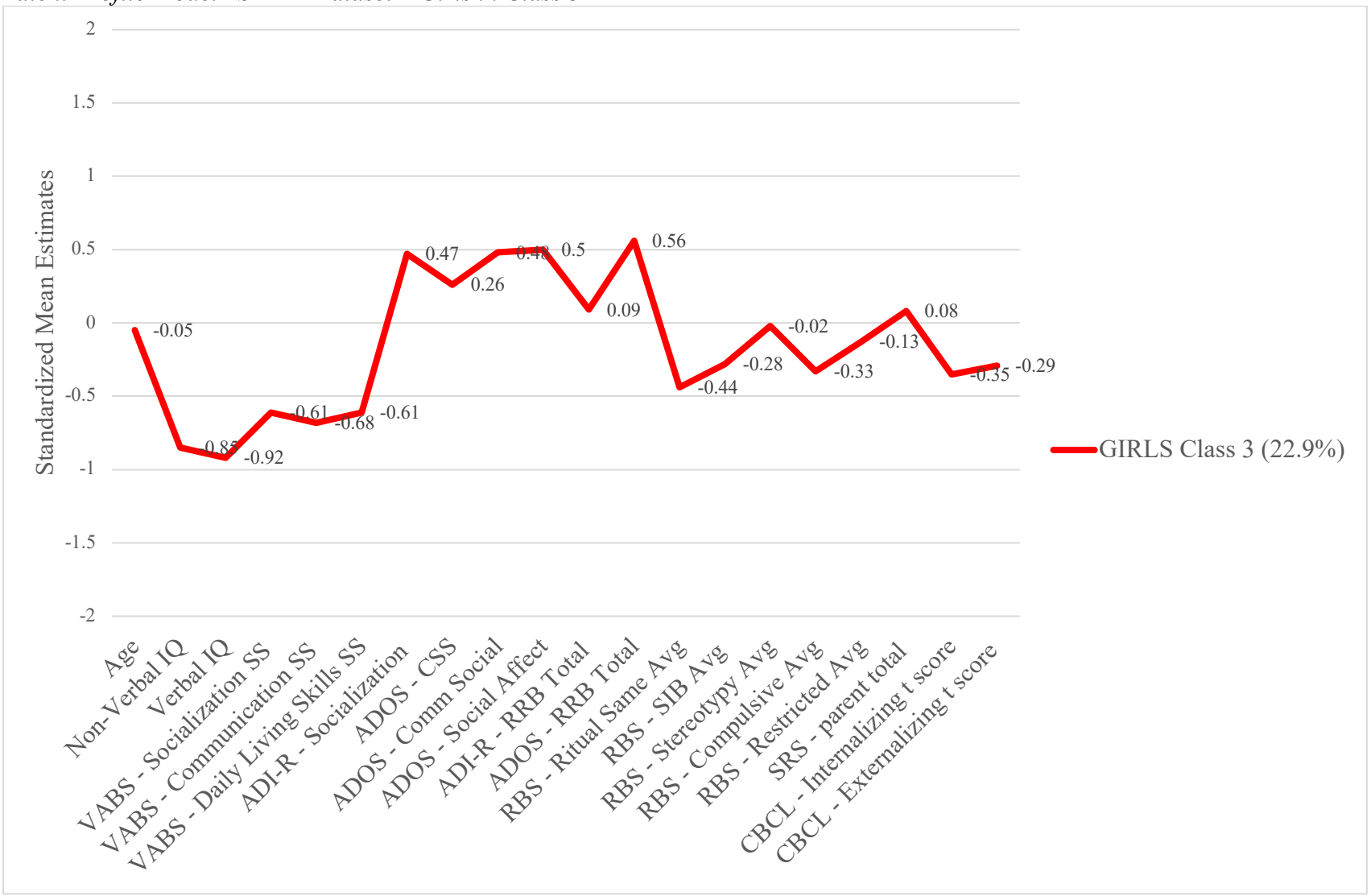
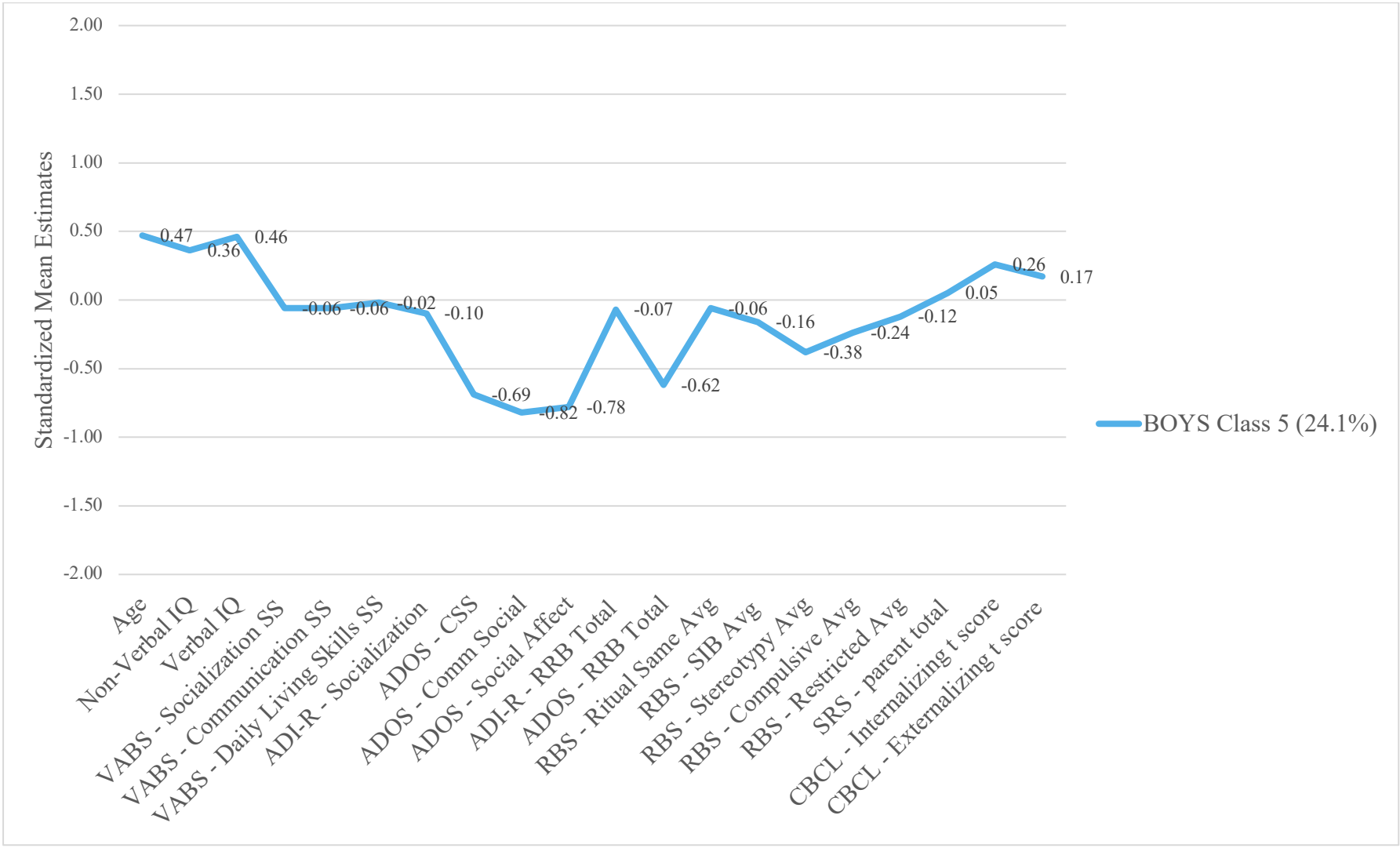


Figure 18

Latent Profile Model –SFARI Dataset – Boys in Class 5



Discussion for Study 2: Latent Profile Analysis

Given the literature supporting sex differences in children with autism across developmental variables, autism symptoms, RRBIs, and social and emotional functioning (see intro), the current study investigated to see if different profiles for girls and boys might emerge.

For the SP sample, we found evidence for three subtypes of children, varying based upon their developmental level and autism severity. The profile analysis for girls is described herein. Class 1 consisted of children who were younger than the sample average, with adaptive abilities and autism severity close to the sample average. Their adaptive scores fell in the Mildly Low range, and their CARS total score fell in the Mild to Moderate range in terms of their autism symptom severity. If creating a treatment to suit the most children, children in this group most closely represent the average of the sample and may be more likely to have their needs met. Class 2 consisted of children who were younger than the overall sample average, their adaptive abilities fell in the Borderline Low range, while their autism severity indicated minimal to no symptoms of autism. It is likely that these children would have different treatment needs and the targets of treatment might need to be adjusted. As an example, their treatment plan might be less focused upon basic adaptive skills, rather they may need support socializing with their peers, recognizing and maintaining healthy relationships, or learning organization and planning skills. Class 3 consisted of children who were older than the overall sample, their adaptive skills fell in the Moderately Low range, and their CARS severity score was in the Mild to Moderate range. These children may benefit from treatment goals focusing on their basic adaptive skills, such as toileting, safety, and interacting with others.

Putting this together, although girls and boys did not differ significantly in this sample, investigating how children score on these variables helps in terms of planning treatment. The

results here clearly point to three groups of children with different qualities, who would likely benefit from very different types of interventions, targeting very different skills. These results might be useful if a particular program were looking to present three different group-based interventions, uniquely tailored to meet the needs of as many children as possible.

Also important to consider here is how these results relate to how autism functioning has been conceptualized in the past, and how we might make changes to better support autistic children in the future. Based upon these results, children who we might have labelled in the past as being from the "high functioning" class would likely need just as much support as children we might label as "low functioning". The most important take home message would be that these groups may differ in the types of supports needed, but still may need similar amounts of support. Their interventions may be delivered by different professionals, in different amounts, at varying intensities, and in different settings.

The neurodiversity movement is moving away from labelling children as "high" or "low" functioning, and for good reason. The DSM-5 currently assigns three support levels to children at the time of diagnosis that determine a child's support needs. These levels have been highly criticized as they require a clinician to determine the level of support a child needs in a less rigorous and one-dimensional way, and at only one point in time. Investigating how an individual child functions across a variety of developmental variables is being touted as much more helpful in terms of determining support needs. Tailoring our diagnostic process to capture as much of this nuance as possible will help us to tailor treatment plans, make accurate recommendations, and adjust these treatment needs over time, as children learn and grow. It is quite clear that support needs will and should vary over time. Reconceptualizing these developmental levels may be a worthwhile practice to find better ways to support children and to move away from the stigma brought on by the labels "high and low functioning".

A visual comparison of the profiles of girls and boys explored different patterns among the profiles of boys and girls. These graphs can be found in Figures 1 through 5. In the SP dataset specifically, the profiles for girls and boys were easily matched and appeared to be highly similar across sexes.

Within the SP sample, a clear and somewhat uncomplicated picture arises. That is, girls and boys show similar profiles, and there are three groups that differ based on their age, ability level, and autism symptomatology. This pattern of results is helpful when conceptualizing the classes, although it may be overly simplistic and does not capture many additional important developmental variables that are crucial to consider when planning a treatment or predicting a child's suitability for or response to treatment. The SFARI sample allows for a more nuanced investigation of different subtypes based on a larger selection of variables.

For the SFARI sample, we found evidence for five subtypes of children, varying based upon their developmental level, autism severity, RRBs, and social and emotional functioning.

The profiles for girls will be discussed herein. Having a wider range of developmental variables allows a more in-depth and nuanced look into each of the classes. Each profile provides distinct and meaningful information.

Class 1 is composed of girls who show average IQ and adequate adaptive abilities, but higher autism symptoms, especially when measured clinically. Interestingly, while their autism symptoms were better measured directly, their RRBs were better captured through parent report. Their internalizing and externalizing behaviour fell in the normal range suggesting that they do not demonstrate high anxiety, as is sometimes assumed for higher-IQ girls. Children in Class 1 are younger than the average of the total sample. Their non-verbal and verbal IQ scores are higher than the overall sample average and fall in the Average range of intellectual functioning. Their adaptive abilities are also higher than the overall sample average and fall in the Moderately

Low range, indicating that they still require some support to develop adaptive skills across the three areas (Socialization, Communication, and Daily Living Skills). Their autism symptoms were similar to the overall sample mean on a parent-report measure, while their autism symptoms on a direct report measure were above the overall means and their Calibrated Severity score (CSS) on the ADOS-2 fell in the Severe range. Children likely to be in Class 1 have RRBs that differ slightly based on the measure used. On a parent report measure they show similar to average RRBs, while on a direct observational measure, they show slightly lower levels of RRBs. When looking at different types of RRBs, they show higher levels on the Compulsive Behaviour subscale, while they show slightly higher levels of Ritualistic Sameness, Restricted Interests, Self-Injurious Behaviour, and Stereotypy. Overall, these differences are minimal. The social responsiveness of these children was less impaired than the overall sample and their score falls in the Moderate range. Children in Class 1 did not demonstrate clinically significant difficulties with symptoms of internalizing and externalizing behaviour and their scores fell in the Normal range for the measures used. In terms of an appropriate treatment plan, a child likely to be in Class 1 may benefit from therapeutic support targeting their ability to socialize and interact with others, perhaps in the form of Cognitive Behaviour Therapy and/or Social Stories.

Class 2 is composed of children who seem similar to the more severe subtype commonly referenced in the literature, i.e., girls who have lower IQs, and more severe autism symptoms, concurrent medical disorders, etc.,. In our sample, these children were relatively older, with relatively low cognitive abilities, falling within the range of Moderate to Severe Intellectual Disability, and mild-moderate delays in adaptive abilities. Children in Class 2 have higher than average autism symptoms including. Their social responsiveness was also more impaired that the overall sample mean, falling in the Severe range. They have higher RRBs across all measures, so these behaviours are easily observable by clinicians and parents. Children in Class 2 have more

difficulties with internalizing and externalizing symptoms than the overall sample (falling within the Borderline Clinical range). When receiving a diagnosis of autism it is likely that children in this profile would be categorized as Level 3 based on the DSM-5, and require very substantial support in all areas of functioning. Interestingly, in our current sample, only one out of the five profiles selected demonstrates this subtype. Based on the literature, we might have expected a larger percentage of the sample to fit within this profile. However, only 49 children fell within Class 2, or 13.1 percent of the overall sample of girls, representing our smallest class.

Girls in Class 3 are part of the class that diverges from the contrasting male class. Compared to the contrasting male class, they do present quite differently and represent an additional lower-IQ group that the male LPA did not demonstrate. This suggests some support that there is an additional group of autistic girls who are more severe, possibly impacting the overall group means of girls. It is important to parse out this finding and recognise that it does not apply to the whole sample of girls, rather a subset. This may be why group-based research on girls is so variable as findings rely on the composition of the sample and the girls included. Class 3 is composed of children who are close in age to the overall sample average. They have lower non-verbal and verbal IQs in the range of Mild and Moderate Intellectual Disability, respectively. Their adaptive abilities fall in the Mildly Low range. Children in Class 3 have higher than average autism symptoms and more impaired social responsiveness. Children in Class 3 also have RRB scores that vary depending on the measure used to capture them. Their RRB scores are more easily captured through direct clinical testing compared to parent report. On a parent report measure their scores are similar to the overall sample mean, however on a direct observational measure their scores are higher than the overall sample mean. When looking at different types of RRB, their scores on the RBS-R are generally lower than the overall sample average in all areas, except the Stereotypy subscale, upon which they are close to the overall sample average. Their

Internalizing and Externalizing symptoms fell within the Average range, indicating that they do not struggle with clinically significant social and emotional difficulties. These children would likely benefit from moderate supports in academics as well as adaptive and life skills. They may not struggle overly with anxiety or depression, although they may require support in terms of areas of functioning commonly impacted by symptoms of autistm, such as social skills, cognitive flexibility, sensory needs.

Class 4 seems to describe the type of girls who may be harder to diagnose, especially without the use of a comprehensive assessment using multiple measures and information sources. They are older, have average cognitive abilities, and their adaptive abilities fall in the Midly Low range across the area of socialization, while they fall in the Borderline Low range across the areas of communication, and daily living skills. Children in Class 4 show autism symptoms similar to the mean on a parent-report measure, while their autism symptoms are lower/less severe, possibly harder to detect, on a direct-report measure. Their CSS score on the ADOS-2 falls within the moderate range. They also show more severe socialization deficits on the SRS, another parent-report measure. Similarly, these children show RRB scores that vary depending on the measure used to capture them. On a parent report measure their scores are higher than the overall sample mean, however on a direct observational measure their scores are lower than the overall sample mean. When looking at different types of RRB, their scores on the RBS-R are generally higher than the overall sample average in all areas. These children may have difficulty with change to their routine, and transitions. They may have higher levels of special interests and compulsive behaviour, although these behaviours are best captured using parent report and do not seem to be as easily captured through direct testing. These girls experience Clinically Significant levels of Internalizing symptoms, such as anxiety, and they experience At-Risk or Borderline levels of externalizing behaviours, indicating that they struggle with social and emotional difficulties.

Clinically speaking, these children may benefit from routine and structure. They may become more engaged when their special interests are highlighted. In addition, they may benefit from social and emotional support for symptoms of anxiety and depression, as well as support for externalizing behaviour, such as emotion dysregulation and distress tolerance.

When discussing Class 5, we see again, another Average IQ group. This group of children may be challenging to diagnose. Children who are likely to fall within Class 5 are younger than the average of the total sample. They have average intellectual functioning and adaptive abilities that are higher than the overall sample, but nonetheless fall in the Borderline Low range on the Socialization and Daily Living Skills domains, and in the Average range across the Communication domain, indicating that they still require some support to develop certain adaptive skills. Their autism symptoms are lower than the overall sample average, although their CSS score falls in the Moderate range. Children in Class 5 had similar levels of social impairment to the overall average, with their score falling in the Moderate range. Children in Class 4 have lower levels of RRBs than the overall sample average across measures. They also show Borderline levels of internalizing symptoms, suggesting that they may struggle with symptoms of anxiety. They showed Average levels of externalizing behaviour. In summary, they represent higher-IQ children with lower autism symptoms and lower levels of RRBs across measures, suggesting that their autism symptoms might be more challenging to measure during the diagnostic process. They seem to experience anxiety and internalize their challenges, rather than externalize; perhaps making it more challenging for their teachers and parents to identify their autistic symptoms and refer them for diagnosis. Supports for these children might include supporting their development of social skills based on their own socialization goals, support to develop emotion regulation skills, independence, and self care skills; as well as understanding and finding ways to cope with their anxiety.

In summary, we found three classes of females with Average IQ. These three classes varied depending on the severity of their autism symptoms, how their RRBs were best captured, and whether they showed social and emotional challenges that require support. One Average-IQ group showed higher autism symptoms, while the other two showed fewer. Interestingly, RRBs in this group with higher autism symptoms showed RRBs that were similar to or fewer than the sample average. In the other two Average-IQ classes, one showed higher RRBs that were best captured using parent report, while the other showed fewer RRBs than the sample average across measures. There were also variations in internalizing and externalizing symptoms, with two groups showing scores in the normal range, and one Average-IQ group showing clinically significant internalizing symptoms. We also found two classes of girls whose IQs fell in the mild, moderate, or severe IQ range. Both of these groups showed high autism symptoms. The lowest IQ group showed high levels of RRBS across measures, while the other showed some variability with direct testing being the best way to capture RRBs. The lowest IQ group showed social and emotional challenges in the Borderline range, while the other did not require support in terms of their internalizing and externalizing behaviours.

Figures 6, 7, and Figures 13 through 16, compare similar classes. Some subtle differences between four of the pairs of profiles of girls and boys were found, while two classes showed very discrepant patterns (Figures 17 and 18). As the LPA was done separately for girls and boys, the profiles cannot be easily compared to determine if differences are significant, as mentioned above. Each profile analysis is standardized according to the overall sample means for the sample selected (i.e., male or female), so differences are not easily comparable as girls and boys demonstrate different means and standard deviations. A complicated statistical analysis to compare differences was beyond the scope of the current project, however it may be a

worthwhile future endeavor. Still, a visual analysis of the class comparisons is worthwhile, and may generate important questions for future research.

Overall, our results suggest that the profiles of girls and boys across developmental variables are largely similar, with some nuanced and subtle differences across girls and boys emerging within the profiles. Our results clearly point to the need for future research investigating these sex differences further as we did find different and contrasting patterns based on sex in the profile analysis, that did not always match the group differences found in Study 1. This suggests that comparing means and lumping all children together, may not fully capture sex differences, as well as subtypes found within the sample as a whole. The classes found provided rich and meaningful information, describing groups of children based upon developmental variables such as IQ, adaptive functioning, autism symptom severity, RRBs, and/or social and emotional symptoms. Defining these profiles is helpful clinically to inform the appropriate supports and treatments best suited for different groups of children. A one-sized fits all treatment is not appropriate for all children, and clinicians and policy makers are encouraged to use information gathered across many variables when designing treatment programs and determining eligibility. If only severity level, IQ, age, and adaptive behaviour are considered, then other variables not measured may impact a child's ability to participate in a treatment, may limit their response to treatment and may lessen the usefulness/efficiency of the intervention.

A study by Zheng, Grove, and Eapen (2019) conducted a latent profile analysis to determine whether restrictive and repetitive behaviour on the RBS-R could be used to identify potential subtypes of autism. The researchers used a slightly different type of LPA than the current study. Specifically, they used unsupervised clustering algorithms to differentiate subgroups of individuals. They found three groups, consisting of low, medium, and high levels of restricted and repetitive behaviours. The groups also differed on a range of clinical measures

including problematic behaviour, autistic traits and adaptive behaviours. The authors concluded that subgroups could be identified, however, that RRBs may be best understood under a dimensional continuum of severity and within the context of other developmental variables.

Studies have found that children with higher IQ are more likely to struggle with anxiety (Mingins, et al., 2021). While this may be true for some children, profiles 3 and 4, show that not all autistic children with high IQ struggle with anxiety. Both are higher IQ groups; children likely to fall in profile 3 show high levels of internalizing and externalizing symptoms, while children likely to fall in profile 4 do not show clinically significant levels of anxiety or externalizing behaviour, suggesting that different treatment targets would be supported for these groups.

The current study is unique in that we studied two very large samples of girls, who varied significantly across the developmental variables sampled (e.g., age, autism severity, IQ, etc.). Our results did differ depending on the sample and measures included, supporting the varied results found by different researchers. Both the SP and SFARI LPAs provided unique and meaningful information that helps to direct future research, improve diagnostic practices, and tailor individual treatment programs.

Study 3: Factor Analysis for Girls and Boys

Study 3: Analyses

In this study we investigated the factor structure of ASD symptoms in boys and girls separately, using different measures of autism symptoms in the two samples.

(1) Using the SP dataset, we used confirmatory factor analysis (CFA) on the CARS for girls and boys. The factor structure tested was based on Luthra's (2013) 3-factor structure, composed of Social-Interaction, Communication, and Emotion-Regulation factors.

(2) In the SFARI dataset, we used CFA to build on the findings of Taheri (2020) using the ADI-R, looking specifically at sex by running CFAs separately for boys and girls. The four factor structures evaluated were as follows:

- Two-factor model based on the *DSM-5* (APA, 2013) consisting of (1) Social Communication and (2) RRBI.
- Three-factor model based on the *DSM-IV-TR* (APA, 2002) consisting of (1) Social (2) Communication, and (3) RRBI.
- Modified three-factor model consisting of (1) Social Communication, (2) Play/Peer, and (3) RRBI.
- Four-factor model consisting of (1) Social, (2) Communication, (3) Play/Peer, and (4) RRBI. This model was the one favoured by Taheri (2020) in all four IQ-level groups.

(3) We used CFA to investigate the factor structure of the RBS-R from the SFARI dataset in boys and girls separately, based on the structure proposed by Lam and Aman (2007).

Factor Analysis in the SP dataset: Results

For this analysis, model fit was thoroughly evaluated using four different fit indices: The root mean square error of approximation (RMSEA), comparative-fit index (CFI), standardized root-mean-square residual (SRMR), and Tucker-Lewis index (TLI). Good or adequate fit is indicated in bold in the tables. Hu and Bentler (1999) suggested adequate model fit is evident when RMSEA is close to .06 or less, SRMR is close to .08 or less, and CFI or TLI is close to .95 or greater. Others have proposed more flexible guidelines for acceptable fit, including values of .90 or greater for CFI and TLI (Kline, 2011) and .08 or less for the RMSEA (MacCallum, Browne, & Sugawara, 1996). Current best practice for evaluating model fit involves considering more than one type of fit index across two or more competing models, and not adhering to the Hu and Bentler guidelines too strictly (Flora, 2018). For the purposes of this analysis, models with CFI and Tucker-Lewis Index TLI values at or greater than .90 and SRMR and RMSEA values at or less than .08 were considered to have good fit, with some flexibility.

CARS

Confirmatory Factor Analysis was conducted in the SP sample separately for girls and boys, to confirm Luthra's three-factor model (2015) using the CARS. Her sample included file review data obtained for 642 children with autism, 111 of whom were female. As mentioned in the Method section, Luthra found that a 3-factor model, composed of Social-Interaction, Communication, and Emotion-Regulation factors, best fit her data.

For the current study, polychoric correlations among all CARS items were examined for girls and boys (see Tables 34 & 35). The correlations among the majority of the items were small, ranging from .00 to .26 for both girls and boys. Some of the items had no relation to one another, for example, items 2 and 3 had .00 correlation with one another.

Fit indices for girls and boys are presented in Table 36. The three-factor model for girls demonstrated good fit across all indices (X^2=186.9, CFI=.959, TLI=.948, RMSEA=.074, SRMR=.054), while the three-factor model for boys showed good fit across most fit indices, except RMSEA (X^2=833.7, CFI=.946, TLI=.930, RMSEA=.087, SRMR=.048).

The standardized factor loading estimates for the three-factor model are reported in Table 37 with the majority of loadings >.30 for both models (i.e., girls and boys). Within the Social Interaction factor, Item 3 (Emotional Response), was not a viable factor. This is different from the previous study, however it is still a viable factor under Emotional Regulation. Within the Emotion Regulation factor, Item 5 (Object Use) and Item 15 (General Impressions), were not viable factors. Again this is different from the previous study. Object Use was, however, a viable factor under the Communication factor, and General Impressions was a viable factor under the Social Interaction Factor. Generally speaking, the factor loadings were consistent with the original study and were similar for girls and boys.

Factor Analysis in the SFARI dataset: Results

For each case, fit indices, using the thresholds discussed in the SP factor analysis above, were examined to determine whether a model had good fit (indicated in bold in the tables).

ADI-R

Confirmatory Factor Analysis (CFA) was conducted in the SFARI sample separately for girls and boys, using the CFA models found by Taheri (2020).

Polychoric correlations among all included ADI-R items were examined for girls and boys (see Tables 38 & 39). The correlations among the majority of the items ranged from small (e.g., .10) to moderate (e.g., .40) for both groups. A few of the items had no relation to one

another, for example, items 33 (inappropriate questions) and 36 (stereotyped utterances) had .00 correlation with one another.

Fit indices for CFA models for girls and boys, using the ADI-R are presented in Tables 40 and 41. For both groups, four models were tested. Model 4 demonstrated the best fit across all measures, especially for boys, while other models did not meet all of the thresholds required for good fit (Girls Model 4: CFI=.94; TFI=.94; RMSEA=.05; SRMR=.08; Boys Model 4: CFI=.91; TFI=.90; RMSEA=.06; SRMR=.07).

Since Model 4 was the only model that had good fit for both groups based on an examination of fit indices, the standardized factor loading estimates for Model 4 are reported in Table 42. Model 4 was a 4-factor model consisting of Social, Communication, Play/Peer, and RRBI factors. The majority of items had significant positive loadings (> .30) on corresponding factors across both groups with very few exceptions. For both groups, item 36 (Inappropriate Questions), item 38 (Neologisms), and item 68 (Circumscribed Interests) did not significantly load onto the RRBI Factor.

RBS-R

Confirmatory Factor Analysis (CFA) was conducted in the SFARI sample separately for girls and boys, using the CFA model based upon the RBS-R, found by Lam and Aman (2007). As mentioned in the methods, Lam and Aman (2007) derived subscales using specific combinations of items. These 5 factors included Ritualistic Sameness, Self-Injurious Behavior, Stereotypy, Compulsive Behavior, and Restricted Interests.

Polychoric correlations among RBS-R items were examined for the total sample (see Table 43). The correlations among the items ranged from small (e.g., .10) to moderate (e.g., .68).

Fit indices for CFA models for girls and boys, using the RBS-R are presented in Table 44. For both girls and boys the Five-Factor model demonstrated adequate fit across all measures (Girls Five-Factor Model: CFI=.95; TFI=.94; RMSEA=.04; SRMR=.08; Boys Five-Factor Model: CFI=.92; TFI=.91; RMSEA=.05; SRMR=.06).

Standardized factor loading estimates for the Five-Factor model are reported in Table 45. All of the items had significant positive loadings (> .30) on corresponding factors across both groups (i.e., girls and boys) with no exceptions.

Discussion Study 3: Factor Analysis for Girls and Boys

Factor Analysis in the SP dataset: Discussion

CARS

Our results confirm that Luthra's 3-factor model (2013) using the CARS showed good fit. For girls, all fit indices were good, while for boys all fit indices showed good fit except for RMSEA, which was slightly above the acceptable limit. Considering our flexible understanding of model fit, we conclude that this model showed good fit. The factor loadings for the majority of the items were greater than .30, however, we might suggest dropping Emotional Response from Social Interaction factor, as well as dropping General Impressions and Object Use from the Emotion Regulation factor as they were not viable items on these factors in our data. Future research with this measure might consider exploratoring different models to investigate an RRBI factor. Overall, the findings were generally consistent with the previous study. These results were largely consistent across girls and boys, indicating that the factor structure of autism seems consistent across sexes.

Factor Analysis in the SFARI dataset

ADI-R

Our results confirm Taheri's findings (2020) that, out of the four FA models investigated, a 4-Factor model showed the best fit in our data. The majority of the items loaded positively onto their respective factors, however items 36, 38, and 68 did not significantly load into their RRBI factor indicating that they might be removed from the model. These results were consistent across girls and boys, indicating that the factor structure of autism seemed consistent across sexes. It is important to acknowledge that this is the same sample as used in the original study, just exploring different groups (i.e., sex), so our results are perhaps unsurprising. Still, the 4-Factor model

seems to be very robust, working well for all IQ groups in Taheri's study and both sexes in the present study.

RBS-R

Our results confirm that Lam and Aman's 5-Factor model demonstrated adequate fit in our data (Lam & Aman, 2007). All items loaded positively onto their respective factors. Results were consistent across girls and boys.

Overall, the factor structure of autism was not observed to differ by sex. However, it is important to remember that the symptoms assessed on our diagnostic measures are standardized and validated using predominately male samples. As a result, only children who fit the diagnostic criteria are included in research samples. It is therefore perhaps unsurprising that few differences are found across sexes. It is also important to acknowledge that while other measures, such as social-emotional assessment tools, have different norms for girls and boys, current autism measures do not.

General Discussion

The current study boasts one of the largest datasets of autistic girls researched thus far. Over 600 autistic girls and their families were included in this research. This represents a huge strength of our study and a huge contribution to the literature on autistic women and girls. Having such a large sample size allowed us to conduct more sophisticated and in-depth statistical analyses never done before.

Our in-depth examination of the differences between male and female autistic children found evidence consistent with the literature. We found evidence in support of significant overall group differences between autistic girls and boys, that is, that boys tended to have higher FSIQ, non-verbal IQ and adaptive abilities than girls. No differences were found in terms of total autism symptom scores. Boys were found to higher levels of RRBIs and Self Injurious Behaviour, although these differences were no longer significant after covarying for IQ.

Our novel exploration of sex differences using Latent Profile Analysis, provided some new evidence expanding on these findings. Results of the Latent Profile Analyses supported the idea that exploring subtypes within samples of autistic children may be helpful to group children with similar characteristics together, which could be particularly important in order to provide appropriate supports tailored to a given subtype. We found evidence for four distinct subtypes that were quite similar in boys and girls, but also one relatively unique profile for each group.

Interestingly, some differences were found depending on how RRBs were measured. In several instances differences were found across parent-report measures compared to direct-observational measures. These differences were also found across sexes, indicating that perhaps how RRBs are best measured may differ across subtypes and across sexes. This suggests that a comprehensive diagnostic assessment with multiple and varied sources of information is

warranted. Future research should explore these findings further and expand on these unique findings.

The results of the factor analysis study largely confirmed the factor structures found in previous research as adequately fitting both boys and girls when examined separately. Some evidence suggested that the boys showed slightly better model fit, however this may simply be related to the fact that the number of boys was much larger and thus they are likely to show more stable results. Overall, female and male samples both showed adequate fit for the models predicted and did not differ in terms of the factor structure investigated. This is unsurprising given that girls and boys included in the studies are diagnosed using diagnostic measures designed using these factor structures. Overall, our results confirm that sex differences are minimal, however, we did find evidence for some subtle differences that warrant further exploration (e.g., how certain items load onto more than one factor).

Limitations and Future Directions

Our study is one of the first studies to include two large samples of autistic females. On their own, each of our datasets includes some of the largest sample sizes of autistic girls in the literature to date. When considered together, our samples consisted of different populations with different characteristics that capture a broad range of autistic children and families. The SP sample consists of younger Canadian children, who were all referred for IBI services in the Toronto area. This implies that these children had high support needs and show more severe autism. A strength of this sample is its representativeness with respect to socio-cultural context. The families referred for treatment likely come from diverse cultural backgrounds and ethnicities, and their socio-economic status is varied. Of course, it is important to acknowledge the

limitations of our samples. A limitation of the SP data included the limited number of measures used and the lack of IQ data in particular.

In contrast, the SFARI dataset represents a primarily American sample of families who volunteered to participate in an intensive research study. A strength of this study included its many measures including parent report, teacher report, and direct testing, as well as the numerous variables measured. This allowed for a complex statistical analysis and nuanced understanding of the data not possible in the SP dataset. Despite this strength, the SFARI dataset includes participants with, on average, higher SES who are able and willing to participate in an intensive research study with many measures conducted at one testing session. These represent highly resourced families who may not be representative of all families of autistic children (e.g., they have access to transportation and the ability to take time off work to complete testing sessions). Families were also 90 percent married, which may not be representative of all families of autistic children. It is also important to note that the sample consisted only of simplex families (i.e., only one child with autism) and participants were required to meet specific criteria on the measures used in the study, which suggests that the sample may not include more complex or ambiguous cases.

Due to the numerous data points collected, only certain scores were selected to represent the developmental variables we wished to explore. For example, only total scores, and/or algorithm items were used on the ADI-R and ADOS-2 (for the LPA and factor analyses). This may not represent all of the behaviours associated with autism. This is highly relevant, especially since it is possible that unique characteristics of autistic girls may be better captured using other symptoms on diagnostic measures that were not included in the algorithm scores used in our study, as well as, potentially, other symptoms or characteristics entirely.

Each measure included in this study has its own limitations. The data used were collected using various measurement techniques, including direct observational methods, and through parent-report questionnaires. Parent-report data has many benefits including being time-effective, low cost, easy to obtain, and it taps into a valuable information source as parents know their children over time and in a variety of different contexts. Despite the advantages, it is important to acknowledge bias inherent in parent-report data, especially those related to sex/gender. As mentioned in the introduction, parents are subject to cultural bias and gender stereotypes. Parents also vary in their understanding of autism, their children's development, and in the accuracy of their reporting. Given that we are investigating sex differences, it is important to acknowledge that some of these biases could have a direct impact on our ability to find or miss differences in girls and boys.

Direct observational measures may also be subject to some of these biases. Individual assessors are trained to administer the various measures. As a part of the SFARI study, assessors had to become "research reliable" in their administration of the ADOS-2 and ADI-R. Still, some variability across assessors is expected. Similarly, different assessors may have different beliefs about sex and gender that may influence how accurately they notice symptoms in girls compared to boys. The current study could not measure the degree to which these biases may affect the diagnostic and assessment process.

It is important to acknowledge the statistical strengths and limitations of the study. A strength of this study includes its large samples and the possibility of conducting multiple and complex statistical analyses in two different samples. The number of analyses conducted increases the risk of Type 1 error (i.e., concluding there are differences when there actually are none). In order to minimize Type 1 error, a strict alpha level was chosen, (i.e., $p<.001$) and effect sizes were computed in order to determine the clinical significance of the findings that were

significant. Still, with the number of *t*-tests done in Study 1, the probability that some of the findings were random or by chance is high and must be acknowledged. On the other hand, the conservative p value may have resulted in us overlooking findings that may have been meaningful (although given small *d* values this is unlikely). For Study 2 (the LPA), missing data may have impacted the accuracy of the profiles and contributed to some random findings. The nature of LPA means that the statistical software predicts the likelihood that a child belongs to a certain profile. When there is missing data, each time the analysis is run it is slightly different than the previous iteration. There is inherently some unpredictability involved in this process and the profiles may be less reliable and replicable than ideal.

In terms of the FA, Study 3, results were limited to the measures and the individual items used. The FA was confirmatory and thus replicated the findings of previous researchers. An exploratory FA may have generated new or different results, considering different types of models that might fit the data. An exploratory FA may be an important next step for future research.

For the purposes of the current study, children were categorized as male or female, based on their biological sex, or their sex assigned at birth. This way of defining sex allows researchers to investigate sex differences, although it certainly limits our findings in many important ways. Autistic people are more likely than the typical population to be gender-diverse (i.e., individuals whose gender does not always correspond to the sex they were assigned at birth). Currently, 0.4-1.3% of the general population in North America is estimated to be transgender or gender diverse (Meerwijk & Sevelius, 2017; Waite & Denier, 2019; Zucker, 2017), although this is likely an underestimate. In recent years, studies have found higher rates of gender variance in autistic populations of children and adults, ranging from 4-5%, with similar rates in males and females (as assigned at birth) (May et al., 2016; Janssen et al., 2016). In children, these rates are based on

parent report. Since sex and gender may not match for all children sampled herein, this may add "noise" in the data and our ability to find clear sex differences may be diminished. Although sex differences are well established in the neurotypical population, such differences may be harder to measure in autistic populations. When gender itself is flexible and hard to define, differences based upon those imprecise categories may not be clear. Of course, it is also possible that sex or gender may be less relevant variables to consider when differentiating among autistic people. Other types of variables may be more relevant and meaningful when attempting to better understand, diagnose, and support/provide intervention to autistic children and adults.

Future research should attempt to find and support the development of diagnostic measures that are standardized on male and female samples. Measures that directly assess for and capture the female phenotype, such as the Q-ASC (Attwood et al., 2011; Simcoe, et al., 2022) should be incorporated into diagnostic batteries. As discussed in the introduction, Simcoe et al., (2022) used the Q-ASC to differentiate autistic females from neurotypical females through the measurement of gendered behaviour (e.g., is s/he interested in looking feminine?), sensory sensitivity, compliant behaviours, imagination, and imitation subscales. Incorporating these domains into our standard diagnostic tools would be worthwhile to improve diagnostic specificity for female children.

As discussed previously, it is challenging to accurately measure the potentially distinguishing characteristics of girls or females as they are more likely to camouflage and mask their symptoms (Attwood et al., 2011). Autistic girls who successfully camouflage may be more likely to be missed or diagnosed later in life (Rutherford et al., 2016). This means that research on samples of female children does not fully capture all autistic females; instead, it captures those girls who meet the predominantly male diagnostic criteria (Haney, 2016). This certainly applies to the current study.

It may be worthwhile for future research to begin to measure and consider camouflaging in autistic children. Emerging research finds that autistic females have shown an increased ability to camouflage (Cook et al., 2021; Goddard et al., 2014; Hull et al., 2019; Simone, 2010). There are many factors that influence whether an autistic person engages in camouflaging (e.g., cognitive ability and executive functioning) so it may only apply to a subset of autistic children/adults (and likely does not apply to all of the children in the current study). Current research largely samples autistic adults and relies upon self-report, although some studies have attempted to directly measure camouflaging. There is debate among the field regarding how to measure it and valid and reliable measures have not yet been created and standardized. As a result, findings are mixed and comparison across camouflaging studies is challenging.

Understanding and measuring characteristics of the female autism phenotype is critical, as failing to support autistic females (or those assigned female at birth) can lead to poor mental health outcomes. A systematic review by Cook et al., (2021) found that (1) adults with more self-reported autistic traits report greater engagement in camouflaging, (2) autistic females and girls/women appear to demonstrate more camouflaging than autistic boys/men, and (3) higher self-reported camouflaging is associated with worse mental health outcomes. Autistic individuals report that engaging in camouflaging is overwhelming, exhausting, and leads to stress, autistic burnout, anxiety, sadness, and identity confusion (Bargiela et al., 2016; Hull et al., 2017). It appears clear that investigating the links between camouflaging and mental health is important in improving well-being and quality of life for autistic adults. Reliably identifying it earlier in childhood may help to prevent mental health challenges in adolescence and adulthood. The current study focused on younger autistic children and no measure of camouflage was available. In addition, the current study did not capture/include all children who are high in camouflaging, specifically those ‘missed’ by our diagnostic measures.

Adding to this, clinicians and researchers would do well to support the development of neuro-affirming diagnostic measures that highlight both the strengths and challenges of autistic females as they navigate their lives. Research investigating symptom profiles tends to be focused on deficits rather than the strengths autistic children possess. It also presents the neurotypical child as the benchmark and ideal for social interactions. Strengths-based assessment and recommendations help to support autistic individuals to honour their unique strengths and talents, advocate for themselves, and find their way in a predominantly neurotypical society. Current research has found that autistic individuals communicate effectively with other autistic people (i.e., The Double Empathy Problem; Crompton, et al., 2020; Davis & Crompton, 2021), suggesting that they may be better able to communicate comfortably and naturally with someone with a familiar neurotype. Neuro-affirming research has also discovered that autistic individuals have strengths that neurotypical individuals do not. Evidence is converging to support the idea that neurotypical people often misread social situations when interacting with autistic people. For example, a study by Sheppard, et al. (2016) found that non-autistic people interpret facial emotions of autistic people less accurately than do autistic individuals. They may also be more likely to misread the mental state of an autistic person (Edey et al., 2016) and hold a negative bias towards the autistic person with whom they are communicating (Alkhaldi et al., 2019). In research, comparing a group of autistic children to a control group consisting of 'neurotypical' children assumes that neurotypical children communicate in the preferred way and may miss many important differences and strengths. Clinically, current diagnostic tools measure a child's social skills against those of neurotypical children, possibly missing some of the strengths of autistic individuals or unique aspects of their communication style that are not as valued in neurotypical society. Neuro-affirming research is critical to capture many strengths in autistic females that are, as of now, not prominently featured in research.

Conclusion

In summary, the current dissertation included three unique studies to investigate sex differences in two large samples of girls with varying characteristics. These samples represent some of the largest sample sizes of autistic girls studied in the literature to date. Access to and investigation of these samples allowed a unique contribution to the literature.

Study 1 used group-based comparisons to investigate sex differences between female and male autistic children and found differences based on sex, although the effect sizes of these findings were small. Traditional research has relied on some of these 'less sophisticated' statistical analyses in research on girls and women often due to low sample sizes.

Study 2 used Latent Profile Analysis to explore subtypes of girls, that is, autistic girls with specific characteristics that cluster together. This is a novel contribution to the literature as no other researchers have conducted an LPA on an all-female sample (no other studies have a sample size large enough to allow for this type of analysis). The results of Study 2 found three distinct profiles for the SP dataset. Profile 1 consisted of younger children with better adaptive skills, and lower autism symptoms. Profile 2 consisted of children with scores close to the sample mean across measures. Profile 3 consisted of older children, with lower adaptive skills, and more severe autism symptoms. These three profiles were essentially the same for boys and girls. The SFARI sample consisted of more developmental variables and allowed for a more nuanced and complex analysis. Using LPA fit statistics and clinical judgement, we decided that a five-profile solution best fit the data. Visual comparison across the profiles for girls and boys found that four of the five graphs were largely similar across boys and girls. One of the five profiles showed unique and distinct profiles for girls that did not have a similar counterpart for boy and one unique profile was found for boys that did not have a parallel in the girls' sample. Future researchers are encouraged to use post-hoc testing to test the significance of the differences

found, which was beyond the scope of the current study. Although we selected the five-profile solution, other solutions were plausible and showed good fit, despite not demonstrating clinically meaningful differences among girls and boys. Future research may want to explore these alternative profiles in different samples. Clinically speaking, fewer profiles seemed to be more clinically meaningful. Investigating these profiles further is worthwhile as it could be used to group similar children and adapt treatment programs to meet their needs.

Study 3 used confirmatory factor analysis to investigate potential differences in the symptom structure of autism between females and males, based on findings from previous research. Overall, the factor structure was found to be largely consistent for males and females. This is perhaps not surprising as the factor structures investigated were fashioned after our current and past diagnostic measures. Since these measures are used to diagnose autistic girls, we would not expect them to capture girls who fall outside of these parameters. Future research may wish to re-conceptualize female-specific features of autism, find a way to accurately measure them, and add these features to our diagnostic measures. There is much more work to be done in order to better capture the unique qualities of neurodiversity in females.

References

Achenbach, T. M., & Edelbrock, C. (1983). *Manual for the Child Behavior Checklist and Revised Child Behaviour Profile.* University of Vermont Department of Psychiatry.

Alkhaldi, R.S., Sheppard, E., & Mitchell, P. (2019). Is there a link between autistic people being perceived unfavorably and having a mind that is difficult to read? *Journal of Autism and Developmental Disorders. 49*(10), 3973-3982. 10.1007/s10803-019-04101-1

Allely, C. (2019). Exploring the female autism phenotype of repetitive behaviours and restricted interests (RBRIs): a systematic PRISMA review. *Advances in Autism, 5*(4), 171-186. https://doi.org/10/1108/AIA-09-2018-0030

American Psychiatric Association (APA). (2020). *Publication Manual of the American Psychological Association,* 7th ed.

American Psychiatric Association (APA). (2000). *Diagnostic and Statistical Manual of Mental Disorders,* 4th ed., Text Revision.

American Psychiatric Association (APA). (2013). *Diagnostic and Statistical Manual of Mental Disorders,* 5th ed.

Amiet, C., Gourfinkel-An, I., Bousamondo, A., Tordjman, S., Baulac, M., Lechat, P., Mottron, L., & Cohen, D. (2008). Epilepsy in autism is associated with intellectual disability and gender: Evidence from a meta-analysis. *Biological Psychiatry, 64,* 577-582. 10.1016/j.biopsych.2008.04.030

Antezana, L., Factor, R. S., Condy, E. E., Strege, M. V., Scarpa, A., & Richey, J. A. (2019). Gender differences in restricted and repetitive behaviors and interests in youth with autism. *Autism research : official journal of the International Society for Autism Research*, *12*(2), 274–283. https://doi.org/10.1002/aur.2049

Attwood, T., Garnett, M. S., & Rynkiewicz, A. (2011). Questionnaire for Autism Spectrum Conditions (Q-ASC) [Measurement instrument].

Autism Spectrum Disorder: Highlights from the 2019 Canadian Health Survey on Children and Youth. (2022). Health Promotion and Chronic Disease Prevention in Canada: Research, Policy and Practice, *42(3)*, 122. https://doi.org/10.24095/hpcdp.42.3.06

Baird, G., Simonoff, E., Pickles, A., Chandler, S., Loucas, T., Meldrum, D., & Charman, T. (2006). Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: the special needs and autism project (SNAP). *The Lancet 368*, 210–215. 10.1016/S0140-6736(06)69041-7

Baio, J., Wiggins, L., Christensen, D. L., Maenner, M. J., Daniels, J., Warren, Z., Kurzius-Spencer, M., Zahorodny, W., Robinson-Rosenberg, C., White, T., Durkin, M. S., Imm, P.,

Nikolaou, L., Yeargin-Allsopp, M., Lee, L. C., Harrington, R., Lopez, M., Fitzgerald, R. T., Hewitt, A., Pettygrove, S., Dowling, N. F. (2018). Prevalence of Autism Spectrum Disorder Among Children Aged 8 Years - Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2014. *Morbidity and Mortality Weekly Report. Surveillance Summaries, 67*(6), 1–23.

Baldwin, S., Costley, D. (2016). The experiences and needs of female adults with high-functioning autism spectrum disorder. *Autism Spectrum Australia, 20*(4), 483-495. https://doi.org/10.1177/1362361315590805

Banach, R., Thompson, A., Szatmari, P., Goldberg, J., Tuff, L., Zwaigenbaum, L., Mahoney, W. (2009). Brief report: Relationship between non-verbal IQ and gender in autism. *Journal of Autism and Developmental Disorders, 39*, 188-193. 10.1007/s10803-008-0612-4

Bargiela, S., Steward, R., & Mandy, W. (2016). The Experiences of late-diagnosed women with autism spectrum conditions: An investigation of the female autism phenotype. *Journal of Autism and Developmental Disorders*, *46*(10), 3281–32.

Beggiato, A., Peyre, H., Maruani, A., Scheid, I., Rastam, M., Amsellem, F., Gillberg, C.I., Leboyer, M., Bourgeron, T., Gillberg, C. and Delorme, R. (2017). Gender differences in autism spectrum disorders: Divergence among specific core symptoms. *Autism Research, 10(4),* 680-689. https://doi.org/10.1002/aur.1715

Bishop, S.L., Hus, V., Duncan, A. Huerta, M., Gotham, K., Pickles, A., Kreiger, A., Buja, A., Lund, S., Lord, C. (2013). Subcategories of restricted and repetitive behaviors in children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 43,* 1287–1297. https://doi.org/10.1007/s10803-012-1671-0

Bodfish, J. W., Symons, F. J., Parker, D. E., & Lewis, M. H. (2000). Varieties of repetitive behavior in autism: Comparisons to mental retardation. *Journal of Autism and Developmental Disorders, 30*(3), 237–243. https://doi.org/10.1023/a:1005596502855

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*(8), 809-816. 10.1111/j.1469-7610.2008.01897.

Bourson, L., & Prevost, C. (2022). Characteristics of restricted interests in girls with ASD compared to boys: A systematic review of the literature. *Journal of European Child and Adolescent Psychiatry,* 10.1007/s00787-022-01998-5. Advance online publication. https://doi.org/10.1007/s00787-022-01998-5

Bryson, S. E., Clark, B. S., & Smith, I. M. (1988). First report of a Canadian epidemiological study of autistic syndromes. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 29*(4), 433–445. doi:10.1111/j.1469-7610.1988.tb00735.x.

Cheslack-Postava, K., Jordan-Young, R.M. (2012). Autism spectrum disorders: Towards a gendered embodiment model. *Social Science and Medicine. 74*(11), 1667-1674. 10.1016/j.socscimed.2011.06.013.

Christensen, D.L., Baio, J., Van Naarden, Braun, K., Bilder, D., Charles, J., Constantino, J.N., Daniels, J., Durkin, M.S., Fitzgerald, R.T., Kurzius-Spencer, M., Lee, L.C., Pettygrove, S., Robinson, C., Schulz, E., Wells, C., Wingate, M.S., Zahorodny, W., Yeargin-Allsopp, M. (2016). Centers for Disease Control and Prevention (CDC). Prevalence and characteristics of autism spectrum disorder among children aged 8 years. *Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States*. *65*(3), 1-23. 10.15585/mmwr.ss6503a1.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Lawrence Erlbaum Associates.

Constantino, J. N. (2002). *The Social Responsiveness Scale*. Western Psychological Services.

Cook, J., Hull, L., Crane, L., Mandy, W. (2021) Camouflaging in autism: A systematic review. *Clinical Psychology Review, 89,* 102080. 10.1016/j.cpr.2021.102080.

Crompton, C.J., Ropar, D., Evans-Williams, C.V., Flynn, E.G., & Fletcher-Watson, S. (2020). Autistic peer-to-peer information transfer is highly effective. *Autism, 24*(7), 1704-1712. https://doi.org/10/1177/1362361320919286

Cuccaro, M.L., Shao, Y., Grubber, J. Slifer, M., Wolpert, C.M., Donnelly, S.L., Abramson, R.K., Ravan, S.A., Wright, H.H., DeLong, G.R., Pericak-Vance, M.A. (2003). Factor Analysis of Restricted and Repetitive Behaviors in Autism Using the Autism Diagnostic Interview-R. *Child Psychiatry and Human Development, 34*, 3–17. https://doi.org/10.1023/A:1025321707947

Davis, R., Crompton, C.J. (2021). What do new findings about social interaction in autistic adults mean for neurodevelopmental research? *Perspectives on Psychological Science. 16*(3), 649-653. 10.1177/1745691620958010

Dean, M., Harwood, R., & Kasari, C. (2017). The art of camouflage: Gender differences in the social behaviors of girls and boys with autism spectrum disorder. *Autism, 21*(6), 678-689. 10.1177/1362361316671845

Dekker, V., Nauta, M. H., Timmerman, M. E., Mulder, E. J., Hoekstra, P. J., & de Bildt, A. (2020). Application of Latent Class Analysis to Identify Subgroups of Children with Autism Spectrum Disorders who Benefit from Social Skills Training. *Journal of Autism and Developmental Disorders*, 10.1007/s10803-020-04678-y. Advance online publication. https://doi.org/10.1007/s10803-020-04678-y

DiLalla, D. L., & Rogers, S. J. (1994) Domains of the childhood autism rating scale: Relevance

for diagnosis and treatment. *Journal of Autism and Developmental Disorders, 24*, 115–128.

Edey, R., Cook, J., Brewer, R., Johnson, M.H., Bird, G., Press, C. (2016). Interaction takes two: Typical adults exhibit mind-blindness towards those with autism spectrum disorder. *Journal of Abnormal Psychology. 125*(7), 879-885. 10.1037/abn0000199

Elliott, C. (2007). *Differential Ability Scales* (2nd ed.) Harcourt Assessment.

Evans, S. C., Boan, A. D., Bradley, C., & Carpenter, L. A. (2019). Sex/gender differences in screening for autism spectrum disorder: Implications for evidence-based assessment. *Journal of Clinical Child and Adolescent Psychology, 48(6),* 840-854. http://dx.doi.org.ezproxy.library.yorku.ca/10.1080/15374416.2018.1437734

Filipek, P. A., Accardo, P. J., Baranek, G. T., Cook, E. H., Jr., Dawson, G., & Gordon, B. (1999) The screening and diagnosis of autistic spectrum disorders [published erratum appears in Journal of Autism Developmental Disorders 2000 Feb;30(1):81]. *Journal of Autism Developmental Disorders, 29*, 439–484.

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

Flora, D. B, (2018). *Statistical methods for the social and behavioural sciences: A model-based approach.* Sage Publications.

Fombonne, E. (2003). Epidemiological surveys of autism and other pervasive developmental disorders: An update. *Journal of Autism and Developmental Disorders, 33*, 365–382. doi:10.1023/ A:1025054610557.

Frazier, T. W., & Hardan, A. Y. (2017) Equivalence of symptom dimensions in females and males with autism. *Autism, 21*(6), 749-759. doi:10.1177/1362361316660066

Frazier, TW, Georgiades, S, Bishop, SL. (2014a) Behavioral and cognitive characteristics of females and males with autism in the Simons Simplex Collection. *Journal of the American Academy of Child and Adolescent Psychiatry 53*(3): 329–340.e321–e323.

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*, 474-480.

Georgiades, S., Papageorgiou, V. & Anagnostou, E. (2010) Brief Report: Repetitive Behaviours in Greek Individuals with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders, 40,* 903–906. https://doi.org/10.1007/s10803-009-0927-9

Goddard, L., Dritschel, B., & Howlin, P. (2014). A preliminary study of gender differences in autobiographical memory in children with an autism spectrum disorder. *Journal of*

Autism and Developmental Disorders, *44*(9), 2087–2095. doi: https://doi.org/10.1007/s10803-014-2109-7

Gómez, L. E., Morán, L., Alcedo, M. A., Arias, V. B., Verdugo, M. A. (2020). Addressing Quality of Life of Children with Autism Spectrum Disorder and Intellectual Disability. *Journal of Intellectual and Developmental Disabilities, 58*(5), 393-408.

Grove, R., Hoekstra, R. A., Wierda, M., & Begeer, S. (2017) Exploring sex differences in autistic traits: A factor analytic study of adults with autism. *Autism, 21*(6), 760-768. 10.1177/1362361316667283

Haney, J.L. (2016). Autism, females, and the DSM-5: Gender bias in autism diagnosis. *Social Work in Mental Health, 14*(4), 396-407.

Harrop, C., McConachie, H., Emsley, R., Leadbitter, K., & Green, J. (2014). Restricted and repetitive behaviors in autism spectrum disorders and typical development: Cross-sectional and longitudinal comparisons. *Journal of Autism and Developmental Disorders, 44*(5), 1207–1219. https://doi.org/10. 1007/s10803-013-1986-5

Harrop, C., Jones, D. R., Sasson, N. J., Zheng, S., Nowell, S. W., & Parish-Morris, J. (2019). Social and object attention is influenced by biological sex and toy gender congruence in children with and without autism. *Autism Research.* http://dx.doi.org.ezproxy.library.yorku.ca/10.1002/aur.2245

Hooker, J. (2019). Psychometric analysis of the Repetitive Behavior Scale-Revised using confirmatory factor analysis in children with autism. *Autism Research., 12*(9). https://doi.org/10.1002/aur.2159

Howlin, P., Goode, S., Hutton, J., & Rutter, M. (2004). Adult outcome for children with autism. *Journal of Child Psychology and Psychiatry, 45*, 212–229.

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

Hull, L., Petrides, K. V., Allison, C., Smith, P., Baron-Cohen, S., Lai, M. C., & Mandy, W. (2017). "Putting on my best normal": Social camouflaging in adults with autism spectrum conditions. *Journal of Autism and Developmental Disorders, 47*(8), 2519–2534. https://doi. org/10.1007/s10803-017-3166-5.

Hull, L., Mandy, W., Lai, M. C., Baron-Cohen, S., Allison, C., Smith, P. (2019). Development and validation of the camouflaging autistic traits questionnaire (CAT-Q). *Journal of Autism and Developmental Disorders, 49*(3), 819–833. https://doi.org/10.1007/s1080 3-018-3792-6.

Hull, L., Lai, M. C., Baron-Cohen, S., Allison, C., Smith, P., Petrides, K. V., & Mandy, W. (2019). Gender differences in self-reported camouflaging in autistic and non-autistic

adults. *Autism*, *24*, 352–363.

Hull, L., Levy, L., Lai, M. *et al.* (2021). Is social camouflaging associated with anxiety and depression in autistic adults? *Molecular Autism, 12*, 13. https://doi.org/10.1186/s13229-021-00421-1

Janssen, A., Huang, H., Duncan, C. (2016). Gender variance among youth with autism spectrum disorders: A retrospective chart review. *Transgender Health. 1*(1), 63-68. 10.1089/trgh.2015.0007

Jorgenson, C., Lewis, T., Rose, C., Kanne, S. (2020). Social camouflaging in autistic and neurotypical adolescents: A pilot study of differences by sex and diagnosis. *Journal of Autism and Developmental Disorders, 50*, 4344–4355. https://doi-org.ezproxy.library.yorku.ca/10.1007/s10803-020-04491-7

Kaat, A. J., Shui, A. M., Ghods, S. S., Farmer, C. A., Esler, A. N., Thurm, A., Georgiades, S., Kanne, S.M., Lord, C., Kim, Y., Bishop, S. L. (2020). Sex differences in scores on standardized measures of autism symptoms: A multisite integrative data analysis. *Journal of Child Psychology and Psychiatry.* http://dx.doi.org.ezproxy.library.yorku.ca/10.1111/jcpp.13242

Kamp-Becker, I., Albertowski, K., Becker, J., Ghahreman, M., Langmann, A., Mingebach, T., Poustka, L., Weber, L., Schmidt, H., Smidt, J., Stehr, T., Roessner, V., Kucharczyk, K., Wolff, N., Stroth, S. (2018). Diagnostic accuracy of the ADOS and ADOS-2 in clinical practice. *European Child and Adolescent Psychiatry. 27*(9), 1193-1207. 10.1007/s00787-018-1143-y.

Kauschke, C., van der Beek, B., & Kamp-Becker, I. (2016). Narratives of girls and boys with autism spectrum disorders: Gender differences in narrative competence and internal state language. *Journal of Autism and Developmental Disorders, 46,* 840-852. 10.1007/s10803-015-2620-5

Kirkovski, M., Enticott, P.G., Fitzgerald, P.B. (2013). A review of the role of female gender in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 43*, 2584-2603.

Kline, R. B. (2011). *Methodology in the Social Sciences. Principles and practice of structural equation modeling (3rd ed.)*. Guilford Press.

Kreiser, N. L., White, S. W. (2014). ASD in females: Are we overstating the gender difference in diagnosis? *Clinical Child Family Psychological Review, 17*, 67-84.

Lai, M., Baron-Cohen, S., Buzbaum, J.D. (2015). Understanding autism in the light of sex/gender. *Molecular Autism. 6*(24), https://doi.org/10.1186/s13229-015-0021-4

Lai, M., Lombardo, M.V., Ruigrok, A.N.V., Chakrabarti, B., Auyeung, B., Szatmari, P., Happe, F., & Baron-Cohen, S. (2017). Quantifying and exploring camouflaging in men and

women with autism. *Autism, 21*(6), 690-702. 10.1177/1362361316671012

Lam, K. S., Aman, M. G. (2007). The Repetitive Behavior Scale-Revised: independent validation in individuals with autism spectrum disorders. *Journal of Autism and Developmental Disorders. 37*(5), 855-866. DOI: 10.1007/s10803-006-0213-z.

Lam, K. S. L., Bodfish, J. W., & Piven, J. (2008). Evidence for three subtypes of repetitive behavior in autism that differ in familiarity and association with other symptoms. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 49*(11), 1193–1200. https://doi.org/10.1111/j.1469-7610.2008.01944.x

Livingston, L. A., Shah, P., & Happé, F. (2019). Compensatory strategies below the behavioural surface in autism: A qualitative study. *The Lancet Psychiatry, 6*(9), 766–777. https://doi.org/10.1016/S22150366(19)30224-X.

Lord, C., Rutter, M., DiLavore P., & Risi, S. (1999) *Autism diagnostic observation schedule: Manual*. 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 Developmental Disorders, 24,* 659–685.

Lundström, S., Mårland, C., Kuja-Halkola, R., Anckarsäter, H., Lichtenstein, P., Gillberg, C., & Nilsson, T. (2019). Assessing autism in females: The importance of a sex-specific comparison. *Psychiatry Research, 282,* 4. http://dx.doi.org.ezproxy.library.yorku.ca/10.1016/j.psychres.2019.112566

Luthra, N., & Perry, A. (2013, August). *Developmental and symptom profiles of children with autism at different ages and levels of Intellectual Disability.* Poster presentation at the American Psychological Association, Honolulu, HI.

Magyar, C. I., & Pandolfi, V. (2007) Factor structure evaluation of the childhood autism rating scale. *Journal of Autism and Developmental Disorders, 37*, 1787–1794.

Mandy, W, Chilvers, R, Chowdhury, U. (2012) Sex differences in autism spectrum disorder: Evidence from a large sample of children and adolescents. *Journal of Autism and Developmental Disorders 42*(7): 1304–1313.

May, T., Cornish, K., & Rinehart, N.J. (2016). Gender profiles of behavioral attention in children with autism spectrum disorders. *Journal of Attention Disorders, 20*(7), 627-635. 10.1177/1087054712455502

May, T., Pang, K., Williams, K.J. (2016). Gender variance in children and adolescents with autism spectrum disorder from the national database for autism research. *International Journal of Transgenderism. 18*(1) 7-15. https://doi.org/10.1080/15532739.2016.1241976

McFayden, T. C., Albright, J., Muskett, A. E., & Scarpa, A. (2019). Brief report: Sex differences

in ASD diagnosis—A brief report on restricted interests and repetitive behaviors. *Journal of Autism and Developmental Disorders, 49*(4), 1693–1699. https://doi. org/10.1007/s10803-018-3838-9.

Meerwijk, E.L., Sevelius, J.M. (2017). Transgender population size in the United States: A meta-regression of population-based probability samples. *American Journal of Public Health. 107*(2), e1-e8. 10.2105/AJPH.2016.303578.

Miles, J. H., Takahashi, T. N., Bagby, S., Sahota, P. K., Vaslow, D. F., Wang, C. H., Hillman, R. E., & Farmer, J. E. (2005). Essential versus complex autism: Definition of fundamental prognostic subtypes. *American Journal of Medical Genetics, 135A*, 171–180.

Miller, J. N., & Ozonoff, S. (2000). The external validity of Asperger disorder: Lack of evidence from the domain of neuropsychology. *Journal of Abnormal Psychology, 109*, 227–238.

Mingins, J.E., Tarver, J., Waite, J., Jones, C., & Surtees, A.DR. (2021). Anxiety of intellectual functioning in autistic children: A systematic review and meta-analysis. *Autism, 25*(1) 18-32. Doi: 10.1177/1362361320953253

Mirenda, P., Smith, I.M., Vaillancourt, T., Georgiades, S., Duku, E., Szatmari, P., Bryson, S., Fombonne, E., Roberts, W., Volden, J., Waddell, C., Zwaigenbaum, L. (2010). Validating the Repetitive Behavior Scale-revised in young children with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 40*(12), 1521-30. doi: 10.1007/s10803-010-1012-0.

Morán, M. L., Gómez, L. E., Alcedo, M. Á., & Pedrosa, I. (2019). Gender differences in social inclusion of youth with autism and intellectual disability. *Journal of Autism and Developmental Disorders, 49*(7), 2980-2989. http://dx.doi.org.ezproxy.library.yorku.ca/10.1007/s10803-019-04030-z

Mosconi, M.W., Kay, M., D'Cruz, A.M., Seidenfeld, A., Guter, S., Stanford, L.D., and Sweeney, J.A. (2009). Impaired inhibitory control is associated with higher-order repetitive behaviors in autism spectrum disorders. *Psychological Medicine 39*(9), 1559–1566.

Mullen, E. M. (1995) *Mullen scales of early learning*. American Guidance Service.

Munson, J., Dawson, G., Sterling, L., Beauchaine, T., Zhou, A., Elizabeth, K., Lord, C., Rogers, S., Sigman, M., Estes, A., & Abbott, R. (2008). Evidence for latent classes of IQ in young children with autism spectrum disorder. *American Journal of Mental Retardation, 113*(6), 439–452. https://doi.org/10.1352/2008.113:439-452

Perry, A., Condillac, R. A., Freeman, N. L., Dunn Geier, J. & Belair, J. (2005) Multi-site study of the Childhood Autism Rating Scale (CARS) in five clinical groups of young children. *Journal of Autism and Developmental Disorders, 35,* 625-634.

Robins, D. L., Casagrande, K., Barton, M., Chen, C. M. A., Dumont-Mathieu, T., & Fein, D. (2014). Validation of the modified checklist for autism in toddlers, revised with follow-up

(M-CHAT-R/F). *Pediatrics, 133*(1), 37–45.

Ros-Demarize, R., Bradley, C., Kanne, S. M., Warren, Z., Boan, A., Lajonchere, C., Park, J., & Carpenter, L. A. (2019). ASD symptoms in toddlers and preschoolers: An examination of sex differences. *Autism Research,* http://dx.doi.org.ezproxy.library.yorku.ca/10.1002/aur.2241

Rutherford, M., McKenzie, K., Johnson, T., Catchpole, C., O'Hare, A., McClure, I., Forsyth, K., McCartney, D., & Murray, A. (2016). Gender ratio in a clinical population sample, age of diagnosis and duration of assessment in children and adults with autism spectrum disorder. *Autism, 20(5),* 628–634. 10.1177/1362361315617879

Rutter, M., Bailey, A., & Lord, C. (2003). *The Social Communication Questionnaire*. Los Angeles: Western Psychological Services.

Rynkiewicz, A., Schuller, B., Marchi, E., Piana, S., Camurri, A., Lassalle, A., & Baron-Cohen, S. (2016). An investigation of the 'female camouflage effect' in autism using a computerized ADOS-2 and a test of sex/gender differences. *Molecular Autism, 7*(10). 10.1186/s13229-016-0073-0

Rynkiewicz, A., Lucka, I. (2018) Autism spectrum disorder (ASD) in girls. Co-occurring psychopathology. Sex differences in clinical manifestation. *Psychiatry Polska*. *52*(4), 629-639. 10.12740/PP/OnlineFirst/58837

Schopler, E., Reichler, R. J., & Renner, B. R. (1988) *The Childhood Autism Rating Scale (CARS)*. Western Psychological Services.

Sedgewick, F., Hill, V., Yates, R., Pickering, L., Pellicano, E. (2016) Gender differences in the social motivation and friendship experiences of autistic and non-autistic adolescents. *Journal of Autism and Developmental Disorders, 46*, 1297–1306. https://doi.org/10.1007/s10803-015-2669-1

Sheppard, E., Pillai, D., Wong, T-L, G., Ropar, D., Mitchell, P. (2016). How easy is it to read the minds of people with autism spectrum disorder? *Journal of Autism and Developmental Disorders. 46*(4), 1247-1254. 10.1007/s10803-015-2662-8

Shine, R. (2021). Examining stress and coping in parents of children with autism spectrum disorder. Unpublished Doctoral Dissertation, York University, Toronto, ON.

Shuster, J., Perry, A., Bebko, J., Toplak, M. E., (2014). Review of factor analytic studies examining symptoms of autism spectrum disorders. *Journal of Autism and Developmental Disorders, 44*, 90–110. https://doi.org/10.1007/s10803-013-1854-3

Simcoe, S., Brownlow, C., Garnett, M.S., Rynkiewicz, A., Attwood, T. (2018) Profiling autism symptomatology: An exploration of the Q-ASC parental report scale in capturing sex differences in autism. *Journal of Autism and Developmental Disorders*. *48*(2), 389-403. 10.1007/s10803-017-3324-9

Simcoe, S.M., Gilmour, J., Garnett, M.S. (2022). Are there gender-based variations in the presentation of Autism amongst female and male children? *Journal of Autism and Developmental Disorders. 53*, 3627–3635. https://doi.org/10.1007/s10803-022-05552-9

Solomon, M., Miller, M., Taylor, S. L., Hinshaw, S. P., & Carter, C. S. (2012). Autism symptoms and internalizing psychopathology in girls and boys with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 42*(1), 48-59. http://dx.doi.org.ezproxy.library.yorku.ca/10.1007/s10803-011-1215-z

Song, A., Cola, M., Plate, S., Petrulla, V., Yankowitz, L., Paney, J., Schultz, R.T., Parish-Morris, J. (2021). Natural language markers of social phenotype in girls with autism. *Journal of Child Psychology and Psychiatry, 62(8),* 949-960. doi:10.1111/jcpp.13348

Sorge, G.B., Flora, D.B., van der Maas, M., Vingilis, E., Erickson, P., Kolla, N.J., Mann, R.E., Seeley, J., Toplak, M.E. (2015) Using factor analytic models to examine the association between attention-deficit/hyperactivity disorder symptoms and health-related outcomes in a representative population survey. *Attention Deficit Hyperactivity Disorders. 7*(3), 225-35. 10.1007/s12402-015-0167-0

Sparrow, S. S., & Cicchetti, D. V. (1989). The Vineland Adaptive Behavior Scales. In C. S. Newmark (Ed.), *Major psychological assessment instruments, 2,* 199–231. Allyn & Bacon.

Sparrow, S. S, Cicchetti, D. V., & Balla, D.A. (2005). *Vineland Adaptive Behavior Scales (2nd ed)*. AGS Publishing.

Stella, J., Mundy, P. & Tuchman, R. (1999). Social and nonsocial factors in the childhood autism rating scale. *Journal of Autism and Developmental Disabilities, 29,* 307–317.

Sutherland, R., Hodge, A., Bruck, S., Costley, D., & Klieve, H. (2017). Parent-reported differences between school-aged girls and boys on the autism spectrum. *Autism, 21*(6), 785–794. https://doi. org/10.1177/1362361316668653

Szatmari, P, Liu, XQ, Goldberg, J. (2012) Sex differences in repetitive stereotyped behaviors in autism: Implications for genetic liability. *American Journal of Medical Genetics Part B Neuropsychiatric Genetics 159B*(1), 5–12.

Taheri, A. (2020). Examining autism symptom severity and structure in IQ groups [Unpublished doctoral dissertation]. York University.

Tomlinson, C., Bond, C., & Hebron, J. (2020). The school experiences of autistic girls and adolescents: A systematic review. *European Journal of Special Needs Education, 35(2),* 203-219. http://dx.doi.org.ezproxy.library.yorku.ca/10.1080/08856257.2019.1643154

Tsai, L., & Beisler, J. (1983). The development of sex differences in infantile autism. *British*

Journal of Psychiatry, 142(4), 373-378. doi:10.1192/bjp.142.4.373

Turner, M. (1999). Repetitive behaviour in autism: A review of psychological research. *Journal of Child Psychology and Psychiatry and Allied Disciplines., 40*(6). https://doi.org/10.1111/1469-7610.00502

Van Lang, N.D.J., Boomsma, A., Sytema, S., De Bildt, A.A., Kraijer, D.W., Ketelaars, C., & Minderaa, R.B. (2006). Structural equation analysis of a hypothesized symptom model in the autism spectrum. *Journal of Child Psychology and Psychiatry, 47*(1), 37-44. 10.1111/j.1469-7610.2005.01434.x

Volkmar, F. R., Szatmari, P., & Sparrow, S. S. (1993). Sex differences in pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 23*, 579–591. doi:10.1007/ BF01046103.

Waite, S., Denier, N. (2019). A research note on Canada's LGBT Data Landscape: Where we are and what the future holds. *Canadian Review of Sociology. 56*(1), 93-117. 10.1111/cars.12232.

Waizbard-Bartov, E., Ferrer, E., Young, G. S., Heath, B., Rogers, S., Wu Nordahl, C., Solomon, M., Amaral, D. G. (2020). Trajectories of autism symptom severity change during early childhood. *Journal of Autism and Developmental Disorders,* http://dx.doi.org.ezproxy.library.yorku.ca/10.1007/s10803-020-04526-z

Wang, M., Hanges, P.J. (2011) Latent class procedures: Applications to organizational research. *Organizational Research Methods. 14*(1)*,* 24-31. https://doi.org/10.1177/1094428110383988

Warsof, B. D. (2014). *Repetitive and restricted behaviors and interests in children and adolescents with autism spectrum disorder* (Order No. AAI3592009) [Doctoral dissertation]. Available from APA PsycInfo®. (1543435747; 2014-99120-441). Retrieved from http://ezproxy.library.yorku.ca/login?url=https://www-proquest-com.ezproxy.library.yorku.ca/dissertations-theses/repetitive-restricted-behaviors-interests/docview/1543435747/se-2?accountid=15182

Watkins, E.E., Zimmermann, Z.J., & Poling, A. (2014). The gender of participants in published research involving people with autism spectrum disorders. *Research in Autism Spectrum Disorders, 8,* 143-146.

Wechsler D. (1999) *The Wechsler Intelligence Scale for Children,* 4th ed. Psychological Corporation.

Wechsler D. (1999). *Wechsler Abbreviated Scale of Intelligence.* Psychological Corporation.

Whitlock, A., Fulton, K., Lai, M., Pellicano, E., & Mandy, W. (2020). Recognition of girls on the autism spectrum by primary school educators: An experimental study. *Autism Research*. http://dx.doi.org.ezproxy.library.yorku.ca/10.1002/aur.2316

Wilson, C. E., Murphy, C. M., McAlonan, G., Robertson, D. M., Spain, D., Hayward, H., Woodhouse, E., Deeley, P. Q., Gillan, N., Ohlsen, J. C., Zinkstok, J., Stoencheva, V., Faulkner, J., Yildiran, H., Bell, V., Hammond, N., Craig, M. C., & Murphy, D. G. (2016). Does sex influence the diagnostic evaluation of autism spectrum disorder in adults?. *Autism, 20*(7), 808–819. https://doi.org/10.1177/1362361315611381

Wing, L., & Gould, J. (1979). Severe impairments of social interaction and associated abnormalities in children: Epidemiology and classification. *Journal of Autism and Childhood Schizophrenia, 9*, 11-29.

Wing, L. (1981). Sex ratios in early childhood autism and related conditions. *Psychiatry Research, 5,* 129–137. doi:10.1016/0165- 1781(81)90043-3.

Young, H., Oreve, M.-J., & Speranza, M. (2018). Clinical characteristics and problems diagnosing autism spectrum disorder in girls. *Archives de Pediatrie, 25,* 399-403.

Zucker, K.J. (2017). Epidemiology of gender dysphoria and transgender identity. *Sex Health. 14*(5), 404-411. 10.1071/SH17067.

Table 1

Description of SP and SFARI Datasets

	SP	SFARI
N	1725	2757
N (%) girls	301 (17.4%)	375 (13.6%)
Age range	1 year 2 months - 7 years 3 months (*M* = 4 years 1 month, S*D* = 1 year 8 months)	4 - 18 years of age (*M* = 9 years 2 months; *SD* = 3 years 7 months)
Sample description	- Screening for IBI	- Full diagnostic assessment for research
	- Not selected (file review) so no volunteer bias	- Volunteer sample
	- Younger, more severe	- Wider range of age and IQ
	- Demographically diverse	- Demographically less diverse (likely)
	- Canadian dataset	- Mostly American dataset

Table 2

Constructs, Measures, and Variables Investigated

SP dataset		
Construct	Measure	Score
Adaptive Behaviour	Vineland Adaptive Behavior Scales	Adaptive behavior composite for 3 domains Socialization domain standard score Communication domain standard score Daily living skills domain standard score
Autism Severity	CARS	CARS total score CARS item scores 1 – 15
SFARI dataset		
Construct	Measure	Score
Cognitive Abilities	IQ	Full scale IQ (ratio and deviation IQ scores)
	Non-Verbal IQ	Non-Verbal IQ score
	Verbal IQ	Verbal IQ score
Adaptive Behaviour	Vineland Adaptive Behavior Scales	Adaptive Behavior Composite standard score Communication domain standard score Daily Living Skills domain standard score Socialization domain standard score
Autism Symptoms	ADI-R	Reciprocal Social Interaction domain total score Verbal Communication domain total Non-Verbal Communication domain total Age first noticed not right in lang, relation, or behaviour (age in months) Onset as perceived with hindsight (scale 1-9)

	ADOS-2	Calibrated Severity Score (CSS; Gotham, Pickles, & Lord, 2009) Communication + Social Total score calculated for all individuals recommended for use with module 4 only Social Affect Total score for modules 1,2,&3 (Gotham, Risi, Pickles, & Lord, 2007)
Restricted and repetitive behaviour and interests	ADI-R	Total score for the RRBI domain
	ADOS-2	RRB Totals from modules 1-3
	RBS-R	Total score on the RBS-R
		RBS-R Stereotyped Behavior RBS-R Self Injurious RBS-R Compulsive Behavior RBS-R Ritualistic Behavior RBS-R Sameness Behavior RBS-R Restricted Behavior
	SRS	Mannerism subscale score Parent
Socio-emotional and behavioural functioning – as reported by parent	Preschool CBCL	Internalizing Total score for 2-5 yr olds Externalizing Total score for 2-5 yr olds
	CBCL	Internalizing Total score for 6-18 yr olds Externalizing Total score for 6-18 yr olds
Socio-emotional and behavioural functioning – as reported by teacher	C-TRF	Internalizing Total score for 2-5 yr olds Externalizing Total score for 2-5 yr olds
	TRF	Internalizing Total score for 6-18 yr olds Externalizing Total score for 6-18 yr olds

Social responsiveness difficulties – as reported by the parent	SRS	Parent T score
	SRS subscale scores	Awareness subscale total - parent Cognition subscale total - parent Communication subscale total - parent Motivation subscale total - parent

Table 3

Sample Descriptives for SP dataset (n=1725)

	Minimum	Maximum	*M (SD)*
Vineland Adaptive Behavior Composite	24.67	110.67	63.67 (11.61)
Vineland Communication Domain standard score	<20	129	62.61 (15.28)
Vineland Daily Living Skills standard score	21	117	65.37 (13.58)
Vineland Socialization standard score	<20	102	63.05 (9.60)
CARS Total Score	<20	49	30.48 (4.70)

Table 4

Bivariate Correlations for SP dataset

	1.	2.	3.	4.	5.	6.
1. Age	1					
2. Vineland ABC 3	-.360*	1				
3. Vineland Comm SS	-.207*	.909*	1			
4. Vineland DLS SS	-.397*	.913*	.709*	1		
5. Vineland Soc SS	-.416*	.882*	.700*	.766*	1	
6. CARS total	.060	-.604*	-.581*	-.519*	-.533*	1

Note. A correlation of $r < .10$, represents no relationship, a correlation of $.10 < r > .29$ represents a small or weak relationship, a correlation of $.30 < r > .49$ represents a moderate strength relationship, while a correlation of $r > .50$ indicates a strong relationship.

**$p<.001$

Table 5

Means, Standard Deviations, and independent samples t-tests between boys and girls, comparing adaptive behaviour and autism severity for SP dataset

Construct	Girls (*n*=301)		Boys (*n*=1421)		*t*(1722)	*p*	Cohen's *d*
	M	*SD*	*M*	*SD*			
Vineland ABC Score 3 Domains	62.36	11.79	63.96	11.55	2.178	.03	.138
CARS total score	30.36	4.48	30.50	4.74	.474	.64	.030

Table 6

Means, Standard Deviations, and Multivariate One-Way Analysis of Variance between boys and girls, comparing adaptive behaviour for SP dataset

Construct	Girls (*n*=301)		Boys (*n*=1421)		*F*(1, 1722)	*p*	η_p^2
	M	*SD*	*M*	*SD*			
Vineland Socialization Domain SS	62.64	9.71	63.15	9.58	.676	.41	.000
Vineland Communication Domain SS	61.08	14.74	62.95	15.37	3.72	.05	.002
Vineland Daily Living Skills SS	63.46	13.87	65.79	13.46	7.420	.006	.004

Table 7

Demographics for SFARI sample (n=2757)

	n	*%*
Marital Status		
Married	2484	90.1
Child lives with		
Both parents	2536	92.0
Annual household income (USD)		
<$20,000	241	9.3
$21,000-$35,000	223	8.6
$36,000-$50,000	281	10.8
$51,000-$65,000	361	13.9
$66,000-$80,000	450	17.3
$81,000-$100,000	399	15.3
$101,000-$130,000	78	3.0
$131,000-$160,000	135	5.2
> $161,000	433	16.6
Total	2601	100
Mother's education		
Less than high school	6	0.22
Some high school	22	0.80
High-school diploma/GED	224	8.17
Some post-secondary	597	21.76
Associate degree	219	7.98
Bachelor's degree	978	35.65
Graduate degree	697	25.41
Total	2743	100
Father's education		
Less than high school	13	0.48
Some high school	52	1.91
High-school diploma/GED	326	11.96
Some post-secondary	527	19.34
Associate degree	183	6.72
Bachelor's degree	855	31.38
Graduate degree	769	28.22
Total	2725	100
Diagnosis		
Autism	2496	90.5
ASD	204	7.4
Aspergers	56	2.0

Table 8

Sample Descriptives for SFARI dataset (n=2757)

	Minimum	Maximum	*M (SD)*
IQ			
FSIQ	<20	167	81.17 (27.96)
Verbal IQ	<20	167	78.05 (31.27)
Non-Verbal IQ	<20	161	84.54 (26.16)
Adaptive Behaviour			
Vineland			
Adaptive Behavior Composite	27	115	73.14 (12.07)
Socialization Standard Score	34	117	70.93 (12.59)
Daily Living Skills Standard Score	25	127	76.39 (13.87)
Communication Standard Score	30	132	77.06 (14.57)

	Minimum	Maximum	*M (SD)*
Autism Symptoms			
ADI-R			
Reciprocal Social Interaction Domain Total score	8	30	20.34 (5.71)
Communication Domain: Verbal subjects Total score *(n=2423)*	6	26	16.50 (4.26)
Communication Domain: Non-verbal subjects Total score *(n=333)*	5	14	12.08 (2.23)
ADOS Core Descriptive			
Calibrated Severity Score (CSS)	4	10	7.44 (1.68)
Communication and Social Total score	4	24	13.33 (4.15)
Social Affect Total score	2	20	11.17 (4.00)
ADOS Module 1 (n=515)			
CSS for Mod 1	4	10	7.48 (1.44)
Communication and Social Total for Mod 1	6	22	15.18 (3.45)
Social Affect Total for Mod 1	5	20	14.38 (3.09)

	Minimum	Maximum	*M (SD)*
ADOS Module 2 (n=603)			
CSS for Mod 2	4	10	7.56 (1.50)
Communication and Social Total for Mod 2	6	24	15.83 (4.35)
Social Affect Total for Mod 2	4	20	12.14 (3.86)
ADOS Module 3 (n=1565)			
CSS for Mod 3	4	10	7.38 (1.80)
Communication and Social Total for Mod 3	4	22	11.84 (3.58)
Social Affect Total for Mod 3	2	20	9.73 (3.57)
ADOS Module 4 (n=74)			
Communication and Social Int Total for Mod 4	7	18	11.45 (2.92)
Communication Total for Mod 4	2	6	3.45 (1.29)
Social Interaction Total for Mod 4	4	13	8.00 (2.14)

	Minimum	Maximum	*M (SD)*
Restrictive and Repetitive Behaviour and Interests			
ADI-R			
Total Score for the RRBI domain	0	12	6.52 (2.50)
ADOS			
RRB Totals from Modules 1-3	0	8	3.96 (2.06)
Restricted and Repetitive Behavior Total for Mod 1	0	8	5.32 (1.84)
Restricted and Repetitive Behavior Total for Mod 2	0	8	4.91 (1.85)
Restricted and Repetitive Behavior Total for Mod 3	0	8	3.26 (1.81)
Restricted and Repetitive Behavior Total for Mod 4	0	6	1.66 (1.30)
RBS-R			
Total Score on the RBS-R (Bodfish, 2000)	0	105	27.13 (17.39)
Stereotyped Behavior subscale (Bodfish, 2000)	0	18	4.46 (3.36)
Self-Injurious Behavior subscale (Bodfish, 2000)	0	21	2.09 (2.87)
Compulsive Behavior subscale (Bodfish, 2000)	0	24	4.11 (3.93)

	Minimum	Maximum	*M (SD)*
Ritualistic Behavior subscale (Bodfish, 2000)	0	18	5.04 (3.96)
Sameness Behavior subscale (Bodfish, 2000)	0	33	7.76 (5.99)
Restricted Behavior subscale (Bodfish, 2000)	0	12	3.67 (2.83)
Avg Ritualistic/Sameness Beh subscale (Lam & Aman, 2007)	0	3.00	.70 (.54)
Avg Stereotypic Beh subscale (Lam & Aman, 2007)	0	2.63	.26 (.36)
Avg Self-Injurious Beh subscale (Lam & Aman, 2007)	0	2.89	.70 (.53)
Avg Compulsive Beh subscale (Lam & Aman, 2007)	0	3.00	.50 (.50)
Avg Restricted Interests subscale (Lam & Aman, 2007)	0	3.00	1.24 (.82)
SRS (parent; n=2746)			
Mannerism subscale score	0	36	18.68 (6.84)

	Minimum	Maximum	*M (SD)*
Socio-emotional and behavioural functioning			
CBCL (parent)			
Internalizing Total *T* score for all ages *(n=2750)*	33	90	60.32 (9.56)
Externalizing Total *T* score for all ages *(n=2750)*	32	97	56.58 (10.61)
TRF (teacher)			
Internalizing Total *T* score for all ages *(n=1241)*	34	91	60.60 (8.49)
Externalizing Total *T* score for all ages *(n=1240)*	36	94	58.79 (8.27)
Social responsiveness difficulties			
SRS (parent; n=2746)			
Parent Total raw score	11	177	98.02 (27.00)
Awareness subscale total	1	24	12.57 (3.67)
Cognition subscale total	1	34	18.55 (5.63)
Communication subscale total	2	60	33.45 (9.92)
Motivation subscale total	0	32	14.78 (5.75)

Table 9

Bivariate Correlations for SFARI dataset – Total Sample

	1	2	3	4	5	6	7	8	9	10	11	12
1. Age	1											
2. FSIQ	-.007	1										
3. Non-Verbal IQ	-.028	.966*	1									
4. Verbal IQ	-.001	.940*	.838*	1								
5. Vineland ABC 3	-.303*	.703*	.680*	.686*	1							
6. Socialization SS	-.302*	.546*	.532*	.541*	.896*	1						
7. Communication SS	-.329*	.723*	.690*	.713*	.921*	.741*	1					
8. Daily Living Skills SS	-.195*	.636*	.623*	.607*	.911*	.726*	.756*	1				
9. ADI-R Comm Domain - Verbal	.099*	-.284*	-.246*	-.308*	-.377*	-.367*	-.326*	-.311*	1			
10. ADI-R Comm Domain – Non-Verbal	.329*	-.384*	-.367*	-.372*	-.466*	-.463*	-.441*	-.351*	_	1		
11. ADI-R Rec Social Interact	.211*	-.388*	-.357*	-.409*	-.509*	-.531*	-.468*	-.397*	.658*	.611*	1	
12. ADOS CSS	-.010	-.185*	-.159*	-.196*	-.144*	-.154*	-.116*	-.125*	.202*	.156*	.194*	1
13. ADOS Comm + Social	-.131*	-.437*	-.383*	-.465*	-.326*	-.297*	-.293*	-.300*	.250*	.265*	.288*	.791*
14. ADOS Social Affect	-.050	-.492*	-.442*	-.517*	-.411*	-.381*	-.382*	-.360*	.252*	.262*	.346*	.777*
15. ADI-R RRBI total	.041	-.029	-.019	-.033	-.051	-.099*	-.015	-.032	.302*	.155	.221*	.097*
16. ADOS RRBI total	-.224*	-.440*	-.400*	-.448*	-.276*	-.245*	-.240*	-.269*	.223*	.154	.228*	.532*
17. RBS-R total	-.044	-.157*	-.155*	-.150*	-.238*	-.269*	-.192*	-.193*	.232*	.060	.230*	.069*
18. Avg Rit/Sameness	.028	.027	.013	.037	-.095*	-.158*	-.049	-.062*	.156*	.078	.132*	.029
19. Avg Stereotypic	.028	-.170*	-.176*	-.160*	-.244*	-.237*	-.225*	-.205*	.117*	.060	.154*	.011
20. Avg Self-Injurious	-.171*	-.324*	-.317*	-.312*	-.320*	-.309*	-.284*	-.284*	.240*	.071	.269*	.112*
21. Avg Compulsive	-.012	-.109*	-.095*	-.117*	-.141*	-.171*	-.122*	-.095*	.160*	-.032	.161*	.029
22. Avg Restricted	-.006	-.121*	-.116*	-.110*	-.182*	-.200*	-.144*	-.156*	.185*	.008	.175*	.073*
23. Mannerisms Parent	.125*	-.190*	-.208*	-.150*	-.331*	-.371*	-.270*	-.270*	-.271*	.213*	.293*	.127*
24. Mannerisms Teacher	.041	-.329*	-.319*	-.331*	-.319*	-.296*	-.306*	-.274*	.182*	.101	.247*	.279*
25. SRS Parent total	.104*	-.251*	-.259*	-.228*	-.444*	-.498*	-.375*	-.347*	.304*	.294*	.408*	.112*

26. Awareness parent	-.025	-.197*	-.202*	-.185*	-.372*	-.418*	-.297*	-.309*	.264*	.257*	.352*	.094*
27. Cognition parent	.060	-.282*	-.276*	-.270*	-.399*	-.416*	-.360*	-.318*	.309*	.288*	.350*	.080*
28. Comm parent	.097*	-.234*	-.242*	-.216*	-.428*	-.493*	-.361*	-.324*	.277*	.261*	.388*	.103*
29. Motivation parent	.133*	-.149*	-.155*	-.141*	-.326*	-.377*	-.277*	-.244*	.156*	.210*	.334*	.061
30. Internalizing CBCL	.050	.148*	.114*	.166*	-.032	-.096*	-.002	.002	-.005	-.038	.043	-.049
31. Externalizing CBCL	-137*	-.012	-.043	.029	-.113*	-.160*	-.057	-.099*	.014	-.067	.023	-.021
32. Internalizing TRF	.014	-.014	-.013	-.023	-.070	-.090*	-.062	-.042	-.041	-.045	.036	.076
33. Externalizing TRF	-.087	-.153*	-.157*	-.149*	-.180*	-.183*	-.150*	-.166*	.022	-.080	.073	.125*

	13	14	15	16	17	18	19	20	21	22	23	24	25
13. ADOS Comm + Social	1												
14. ADOS Social Affect	.929*	1											
15. ADI-R RRBI total	.060	.041	1										
16. ADOS RRBI total	.486*	.402*	.129*	1									
17. RBS-R total	.091*	.099*	.370*	.122*	1								
18. Avg Rit/Sameness	-.011	-.014	.305*	-.019	.865*	1							
19. Avg Stereotypic	.049	.075*	.143*	.072*	.559*	.343*	1						
20. Avg Self-Injurious	.191*	.216*	.273*	.292*	.739*	.443*	.428*	1					
21. Avg Compulsive	.067*	.079*	.301*	.047	.744*	554*	.309*	.471*	1				
22. Avg Restricted	.067*	.063*	.294*	.106*	.719*	.596*	.288*	.456*	.470*	1			
23. Mannerisms Parent	.110*	.133*	.327*	.167*	.604*	.521*	.341*	.530*	.360*	.494*	1		
24. Mannerisms Teacher	.279*	.299*	.115*	.335*	.197*	.104*	.150*	.282*	.091	.131*	.276*	1	
25. SRS Parent total	.171*	.216*	.246*	.150*	.579*	.493*	.347*	.503*	.371*	.442*	.850*	.266*	1
26. Awareness parent	.174*	.219*	.169*	.158*	.403*	.315*	.267*	.396*	.248*	.300*	.573*	.229*	.755*
27. Cognition parent	.144*	.167*	.230*	.146*	.523*	.436*	.291*	.445*	.359*	.404*	.671*	.213*	.846*
28. Comm parent	.175*	.222*	.169*	.133*	.465*	.399*	.299*	.411*	.296*	.338*	.713*	.247*	.934*
29. Motivation parent	.119*	.175*	.142*	.033	.430*	.382*	.252*	.339*	.295*	.321*	.554*	.150*	.768*
30. Internalizing CBCL	-.091*	-.070*	.144*	-.124*	.401*	.409*	.263*	.244*	.247*	.271*	.424*	-.007	.477*
31. Externalizing CBCL	-.020	-.009	.128*	.001	.436*	.424*	.353*	.299*	.254*	.280*	.411*	.072	.424*
32. Internalizing TRF	.076	.086*	.040	.035	.084	.082	.051	.074	.062*	.023	.077	.406*	.104*
33. Externalizing TRF	.168*	.182*	.022	.160*	.128*	.089	.193*	.113*	.078*	.061	.133*	.512*	.139*

	26	27	28	29	30	31	32	33
26. Awareness parent	1							
27. Cognition parent	.605*	1						
28. Comm parent	.694*	.733*	1					
29. Motivation parent	.439*	.547*	.657*	1				
30. Internalizing CBCL	.268*	.372*	.400*	.509*	1			
31. Externalizing CBCL	.352*	.364*	.394*	.243*	.522*	1		
32. Internalizing TRF	.022	.067	.081	.178*	.258*	.076	1	
33. Externalizing TRF	.156*	.117*	.142*	.036	.015	.342*	.458*	1

Note. A correlation of $r < .10$, represents no relationship, a correlation of $.10 < r > .30$ represents a small or weak relationship, a correlation of $.30 < r > .50$ represents a moderate strength relationship, while a correlation of $r > .50$ indicates a strong relationship.

*$p < .001$

Table 10

Bivariate Correlations for SFARI dataset – Girls

	1	2	3	4	5	6	7	8	9	10	11	12
1. Age	1											
2. FSIQ	-.027	1										
3. Non-Verbal IQ	-.052	.968*	1									
4. Verbal IQ	-.021	.951*	.857*	1								
5. Vineland ABC 3	-.341*	.720*	.700*	.703*	1							
6. Socialization SS	-.349*	.583*	.566*	.584*	.912*	1						
7. Communication SS	-.332*	.750*	.713*	.744*	.919*	.752*	1					
8. Daily Living Skills SS	-.259*	.641*	.640*	.599*	.922*	.772*	.766*	1				
9. ADI-R Comm Domain - Verbal	.083	-.421*	-.389*	-.420*	-.428*	-.410*	-.394*	-.357*	1			
10. ADI-R Comm Domain – Non-Verbal	.165	-.357*	-.317*	-.460*	-.436*	-.331*	-.435*	-.384*	_	1		
11. ADI-R Rec Social Interact	.122	-.456*	-.433*	-.470*	-.518*	-.539*	-.464*	-.427*	.736*	.504*	1	
12. ADOS CSS	-.072	-.234*	-.200*	-.243*	-.190*	-.189*	-.171*	-.165*	.260*	.000	.268*	1
13. ADOS Comm + Social	-.189*	-.440*	-.376*	-.473*	-.338*	-.315*	-.319*	-.298*	.306*	.161*	.379*	.813*
14. ADOS Social Affect	-.095	-.477*	-.416*	-.513*	-.429*	-.421*	-.406*	-.354*	.288*	.160	.419*	.806*
15. ADI-R RRBI total	.104	-.198*	-.170*	-.212*	-.224	-.236*	-.197*	-.185*	.350*	.204	.288*	.127*
16. ADOS RRBI total	-.239*	-.456*	-.423*	-.452*	-.297*	-.257*	-.279*	-.282*	.343*	.102	.314*	.553*
17. RBS-R total	.062	-.194*	-.174*	-.207*	-.299*	-.308*	-.275*	-.242*	.200*	.081	.288*	.138*
18. Avg Rit/Sameness	.125	-.018	-.012	-.031	-.161*	-.199*	-.130	-.116*	.087	.157	.147*	.067
19. Avg Stereotypic	.111	-.209*	-.198*	-.208*	-.259*	-.215*	-.277*	-.218*	.063	.098	.166*	-.005
20. Avg Self-Injurious	-.118	-.351*	-.341*	-.334*	-.360*	-.327*	-.334*	-.328*	.289*	.026	.354*	.226*
21. Avg Compulsive	.078	-.112	-.087*	-.139*	-.167*	-.181*	-.171*	-.108	.117	-.136	.204*	.098
22. Avg Restricted	.108	-.128	-.116	-.133	-.244*	-.281*	-.202*	-.191*	.178*	.148	.247*	.103
23. SRS Parent total	.237*	-.251*	-.246*	-.255*	-.508*	-.519*	-.444*	-.438*	.291*	.291	.405*	.187*
24. Awareness Parent	-.008	-.167*	-.160*	-.172*	-.369*	-.404*	-.281*	-.337*	.273*	.203	.342*	.192*
25. Cognition Parent	.192*	-.269*	-.256*	-.273*	-.436*	-.416*	-.360*	-.318*	.309*	.288*	.350*	.080*

26. Comm Parent	.188*	-.248*	-.238*	-.261*	-.494*	-.493*	-.361*	-.324*	.277*	.261*	.388*	.103*
27. Motivation Parent	.261*	-.132*	-.134*	-.143*	-.373*	-.377*	-.277*	-.244*	.156*	.210*	.334*	.061
28. Mannerisms Parent	.291*	-.218*	-.227*	-.204*	-.429*	-.371*	-.270*	-.270*	-.271*	.213*	.293*	.127*
29. Internalizing CBCL	.114	.186*	.153*	.207*	-.025	-.096*	-.002	.002	-.005	-.038	.043	-.049
30. Externalizing CBCL	-.073*	-.037	-.047	-.028	-.105*	-.160*	-.057	-.099*	.014	-.067	.023	-.021
31. Internalizing TRF	.083	-.018	-.026	-.026	-.165*	-.090*	-.062	-.042	-.041	-.045	.036	.076
32. Externalizing TRF	-.107	-.269*	-.218*	-.320*	-.293*	-.183*	-.150*	-.166*	.022	-.080	.073	.125*

	13	14	15	16	17	18	19	20	21	22	23	24	25
13. ADOS Comm + Social	1												
14. ADOS Social Affect	.929*	1											
15. ADI-R RRBI total	.060	.041	1										
16. ADOS RRBI total	.486*	.402*	.129*	1									
17. RBS-R total	.091*	.099*	.370*	.122*	1								
18. Avg Rit/Sameness	-.011	-.014	.305*	-.019	.865*	1							
19. Avg Stereotypic	.049	.075*	.143*	.072*	.559*	.343*	1						
20. Avg Self-Injurious	.191*	.216*	.273*	.292*	.739*	.443*	.428*	1					
21. Avg Compulsive	.067*	.079*	.301*	.047	.744*	554*	.309*	.471*	1				
22. Avg Restricted	.067*	.063*	.294*	.106*	.719*	.596*	.288*	.456*	.470*	1			
23. SRS Parent total	.171*	.216*	.246*	.150*	.579*	.493*	.347*	.503*	.371*	.442*	1		
24. Awareness parent	.174*	.219*	.169*	.158*	.403*	.315*	.267*	.396*	.248*	.300*	.755*	1	
25. Cognition parent	.144*	.167*	.230*	.146*	.523*	.436*	.291*	.445*	.359*	.404*	.846*	.605*	1
26. Comm parent	.175*	.222*	.169*	.133*	.465*	.399*	.299*	.411*	.296*	.338*	.934*	.694*	.733*
27. Motivation parent	.119*	.175*	.142*	.033	.430*	.382*	.252*	.339*	.295*	.321*	.768*	.439*	.547*
28. Mannerisms Parent	.110*												
29. Internalizing CBCL	-.091*	-.070*	.144*	-.124*	.401*	.409*	.263*	.244*	.247*	.271*	.477*	.268*	.372*
30. Externalizing CBCL	-.020	-.009	.128*	.001	.436*	.424*	.353*	.299*	.254*	.280*	.424*	.352*	.364*
31. Internalizing TRF	.076*	.086*	.040	.035	.084*	.082*	.051	.074*	.062	.023*	.104*	.022	.067*
32. Externalizing TRF	.168*	.182*	.022	.160*	.128*	.089*	.193*	.113*	.078*	.061	.139*	.156*	.117*

	26	27	28	29	30	31	32
26. Comm parent	1						
27. Motivation parent	.657*	1					
28. Mannerisms Parent	.713*	.554*	1				
29. Internalizing CBCL	.400*	.509*	.424*	1			
30. Externalizing CBCL	.394*	.243*	.411*	.522*	1		
31. Internalizing TRF	.081*	.178*	.077*	.258*	.076	1	
32. Externalizing TRF	.142*	.036	.133*	.015	.342*	.458*	1

Note. A correlation of $r < .10$, represents no relationship, a correlation of $.10 < r > .30$ represents a small or weak relationship, a correlation of $.30 < r > .50$ represents a moderate strength relationship, while a correlation of $r > .50$ indicates a strong relationship.

*$p < .001$

Table 11

Means, Standard Deviations, and Independent t-tests for SFARI dataset (n=2757)

	Girls		Boys		*t (2756)*	*p*	Cohen's *d*
	M	*SD*	*M*	*SD*			
Age							
Age at ADOS	9.17	3.70	9.00	3.55	1.02 (2755)	.310	.056
IQ							
FSIQ	75.20	28.23	82.11	27.81	-4.46 (2750)	<.001	.248
Verbal IQ	73.95	32.38	78.70	31.06	-2.73 (2754)	.006	-.152
Non-Verbal IQ	77.99	26.13	85.57	26.02	-5.23 (2754)	<.001	.291
Adaptive Behaviour							
Vineland							
Adaptive Behavior Composite based on 3 domains	72.54	12.34	75.15	12.42	-3.79 (2755)	<.001	-.21
Socialization Standard Score	69.26	12.71	71.19	12.55	-2.76 (2755)	.006	-.153
Communication Standard Score	74.63	14.11	77.45	14.61	-3.49 (2755)	<.001	.194
Daily Living Skills Standard Score	73.72	13.52	76.81	13.88	-4.03 (2755)	<.001	.224

	Girls		Boys		*t (2756)*	*p*	Cohen's *d*
	M	*SD*	*M*	*SD*			
Autism Symptoms							
ADI-R							
Reciprocal Social Interaction Domain Total Score	20.55	6.00	20.30	5.67	.786 (2754)	.432	.044
Verbal Communication Domain Total Score	16.69	4.42	16.47	4.24	.878 (2421)	.380	.052
Non-Verbal Communication Domain Total Score	9.31	3.69	9.26	3.40	.265 (2754)	.791	.015
Age first noticed not right (years)	1.75	1.18	1.83	1.15	-1.21 (2710)	.226	-.068
ADOS Core Descriptive							
Calibrated Severity Score (CSS)	7.46	1.72	7.44	1.67	.251 (2681)	.802	.014
Communication and Social Total Score	13.77	4.45	13.26	4.10	2.182 (2755)	.029	.121
Social Affect Total Score	11.60	4.20	11.10	3.97	2.214 (2681)	.027	.125
ADOS Module 1							
CSS for Mod 1	7.58	1.60	7.46	1.41	.660 (512)	.510	.079
Communication and Social Total for Mod 1	15.24	3.59	15.17	3.42	.169 (513)	.866	.020
Social Affect Total for Mod 1	14.43	3.31	14.37	3.05	.165 (513)	.869	.020

	Girls		Boys		*t (2756)*	*p*	Cohen's *d*
	M	*SD*	*M*	*SD*			
ADOS Module 2							
CSS for Mod 2	7.83	1.59	7.51	1.48	1.940 (598)	.053	.213
Communication and Social Total for Mod 2	16.71	4.60	15.65	4.29	2.208 (599)	.028	.243
Social Affect Total for Mod 2	12.89	4.03	11.99	3.82	2.123 (599)	.034	.233
ADOS Module 3							
CSS for Mod 3	7.21	1.80	7.40	1.80	-1.328 (1563)	.184	-.104
Communication and Social Total for Mod 3	11.67	3.55	11.87	3.59	-.689 (1563)	.491	-.054
Social Affect Total for Mod 3	9.63	3.61	9.74	3.56	-.416 (1563)	.677	-.033
ADOS Module 4							
Communication and Social Int Total for Mod 4	10.67	2.40	11.55	2.99	-.851 (72)	.397	-.303
Communication Total for Mod 4	2.89	.93	3.52	1.32	-1.386 (72)	.170	-.493
Social Interaction Total for Mod 4	7.78	2.28	8.03	2.14	-.331 (72)	.742	-.118

	Girls		Boys		*t (2756)*	*p*	Cohen's *d*
	M	*SD*	*M*	*SD*			
Restricted and Repetitive Behaviour and Interests							
ADI-R							
Total Score for the RRBI domain	6.13	2.46	6.58	2.50	-3.23 (2754)	.001	.070
ADOS							
RRB Totals from Modules 1-3	3.99	2.17	3.96	2.04	.232 (2755)	.817	.013
Restricted and Repetitive Behavior Total for Mod 1	5.24	1.92	5.34	1.82	-.441 (513)	.660	-.053
Restricted and Repetitive Behavior Total for Mod 2	5.02	2.05	4.89	1.81	.656 (599)	.512	.072
Restricted and Repetitive Behavior Total for Mod 3	2.96	1.74	3.30	1.81	-2.424 (1563)	.015	-.190
Restricted and Repetitive Behavior Total for Mod 4	2.11	1.83	1.60	1.21	1.111 (72)	.270	.395
RBS-R							
Total Score on the RBS-R	27.03	17.30	27.15	17.41	-.125 (2753)	.900	-.007
Stereotyped Behavior subscale	4.21	3.35	4.49	3.36	-1.52 (2753)	.130	-.084
Self-Injurious Behavior subscale	2.38	3.00	2.04	2.85	2.13 (2753)	.033	.118
Compulsive Behavior subscale	4.24	3.84	4.09	3.94	.695 (2754)	.487	.039

	Girls		Boys		*t (2756)*	*p*	Cohen's *d*
	M	*SD*	*M*	*SD*			
Ritualistic Behavior subscale	5.13	4.09	5.03	3.94	.470 (2753)	.638	.026
Sameness Behavior subscale	7.98	5.99	7.72	5.98	.768 (2752)	.442	.043
Restricted Behavior subscale	3.09	2.68	3.76	2.84	-4.283 (2752)	<.001	.238
Avg Ritualistic/Sameness Behavior subscale	.71	.53	.69	.54	.609 (2738)	.542	.034
Avg Stereotypic Behavior subscale	.30	.37	.26	.36	2.05 (2741)	.041	.114
Avg Self-Injurious Behavior subscale	.63	.51	.71	.53	-2.64 (2731)	.008	.147
Avg Compulsive Behavior subscale	.52	.49	.49	.50	0.95 (2747)	.344	.053
Avg Restricted Interests subscale	1.17	.81	1.25	.82	-1.74 (2751)	.083	.097
SRS (parent)							
Mannerism subscale score	18.73	7.15	18.67	6.80	.146 (2744)	.884	.008

	Girls		Boys		*t (2756)*	*p*	Cohen's *d*
	M	*SD*	*M*	*SD*			
Socio-emotional and behavioural functioning							
CBCL (parent)							
Internalizing Total score for all ages	59.80	10.04	60.40	9.48	-1.13 (2748)	.260	.063
Externalizing Total score for all ages	57.52	10.02	56.43	10.69	1.85 (2748)	.064	.103
TRF (teacher)							
Internalizing Total score for all ages	60.40	9.17	60.63	8.37	-.34 (1239)	.736	.027
Externalizing Total score for all ages	61.22	8.54	58.38	8.15	4.28 (1238)	<.001	.346
Social responsiveness difficulties							
SRS (parent)							
Parent Total raw score	99.95	27.30	97.72	26.95	1.474 (2744)	.140	.082
Awareness subscale total	12.75	3.77	12.54	3.65	1.056 (2744)	.291	.059
Cognition subscale total	18.98	5.42	18.48	5.65	1.610 (2741)	.108	.090
Communication subscale total	34.36	9.84	33.31	9.93	1.886 (2741)	.059	.106
Motivation subscale total	15.12	5.87	14.73	5.73	1.209 (2744)	.227	.068

Table 12

Means, Standard Deviations, and Multivariate One-Way Analysis of Variance between boys and girls, comparing adaptive behaviour for SFARI dataset covarying for IQ (n=2757; girls n=375).

Construct	Girls		Boys		$F(1, 2755)$	η_p^2
	M	*SD*	*M*	*SD*		
MANCOVA - Using IQ as a covariate						
Vineland Communication Domain SS	69.22	12.71	71.21	12.55	.249	<.001
Vineland Daily Living Skills SS	74.58	14.09	77.48	14.59	.294	<.001
Vineland Socialization SS	73.70	13.53	76.84	13.87	2.643	.001

Note. Interaction term between Sex and IQ was not significant. Interaction term was removed from the model.

Table 13

Means, Standard Deviations, and Multivariate One-Way Analysis of Variance between boys and girls, comparing autism symptomatology for SFARI dataset covarying for IQ (n=2682; girls n=366).

Construct	Girls		Boys		$F(1, 2680)$	η_p^2
	M	*SD*	*M*	*SD*		
MANCOVA - Using IQ as a covariate						
ADI-R Reciprocal Social Interaction	20.64	5.95	20.30	5.69	.590	<.001
ADOS CSS	7.46	1.72	7.44	1.67	.356	<.001
ADOS Communication + Social Total	13.84	4.47	13.31	4.12	.108	<.001
ADOS Social Affect Total	11.60	4.20	11.10	3.97	.002	<.001

Note. Interaction term between Sex and IQ was not significant. Interaction term was removed from the model.

Table 14

Means, Standard Deviations, and Multivariate One-Way Analysis of Variance between boys and girls, comparing restrictive and repetitive behaviour and interests for SFARI dataset covarying for IQ (n=2754; girls n=375).

Construct	Girls		Boys		$F(1, 2752)$	η_p^2
	M	*SD*	*M*	*SD*		
MANCOVA - Using IQ as a covariate						
ADI-R RRBI Total score	6.13	2.46	6.58	2.50	11.30*	.004
ADOS RRBI Total score	3.98	2.16	3.96	2.04	4.08	.001
RBS-R Total score	27.03	17.33	27.16	17.41	.672	<.001

*p =.001

Significant finding for the ADI-R RRBI Total score. The interaction term was significant and removed from the model. The analysis was run again using a categorized FSIQ variable. Results were very similar to the original analysis, so the original model was retained and is presented here.

Table 15

Means, Standard Deviations, and Multivariate One-Way Analysis of Variance between boys and girls, comparing restrictive and repetitive behaviour and interests for subscales of the RBS-R based on study by Lam and Aman (2007) in the SFARI dataset covarying for IQ (n=2704; girls n=369).

Construct	Girls		Boys		$F(1, 2700)$	η_p^2
	M	*SD*	*M*	*SD*		
MANCOVA - Using IQ as a covariate						
Avg Ritualistic/Sameness Behavior subscale	.71	.53	.69	.54	.501	<.001
Avg Stereotypic Behavior subscale	.30	.37	.26	.36	1.801	.001
Avg Self-Injurious Behavior subscale	.63	.52	.71	.53	18.049*	.007
Avg Compulsive Behavior subscale	.52	.50	.49	.50	.315	<.001
Avg Restricted Interests subscale	1.17	.80	1.25	.82	5.025	.002

*p <.001

Note. Interaction term between Sex and IQ was not significant. Interaction term was removed from the model.

Table 16

Means, Standard Deviations, and Multivariate One-Way Analysis of Variance between boys and girls, comparing restrictive and repetitive behaviour and interests for the mannerism subscale of the SRS in the SFARI dataset covarying for IQ (n=1362; girls n=194).

Construct	Girls		Boys		$F(1, 1359)$	η_p^2
	M	*SD*	*M*	*SD*		
MANCOVA - Using IQ as a covariate						
Mannerisms Parent	18.75	7.14	18.87	6.79	.516	<.001

*p <.001

Note. Interaction term between Sex and IQ was not significant. Interaction term was removed from the model.

Table 17

Means, Standard Deviations, and Multivariate One-Way Analysis of Variance between boys and girls, comparing subscales of the SRS for parents in the SFARI dataset covarying for IQ (n=2740; girls n=368).

Construct	Girls		Boys		$F(1, 2736)$	η_p^2
	M	*SD*	*M*	*SD*		
MANCOVA - Using IQ as a covariate						
Awareness Raw score (Parent)	12.71	3.75	12.53	3.65	.010	<.001
Cognition Raw score (Parent)	18.98	5.43	18.47	5.65	.186	<.001
Communication Raw score (Parent)	34.31	9.84	33.31	9.93	.777	<.001
Motivation Raw score (Parent)	15.10	5.88	14.72	5.73	.318	<.001

Note. Interaction term between Sex and IQ was not significant. Interaction term was removed from the model.

Table 18

Means, Standard Deviations, and Multivariate One-Way Analysis of Variance between boys and girls, comparing internalizing and externalizing t scores for parents on the CBCL in the SFARI dataset covarying for IQ (n=2750; girls n=373).

Construct	Girls		Boys		$F(1, 2747)$	η_p^2
	M	*SD*	*M*	*SD*		
MANCOVA - Using IQ as a covariate						
Internalizing T score (Parent)	59.80	10.05	60.40	9.48	.254	<.001
Externalizing T score (Parent)	57.53	10.04	56.43	10.69	3.272	.001

*p <.001

Note. Interaction term between Sex and IQ was not significant. Interaction term was removed from the model.

Table 19

Means, Standard Deviations, and Multivariate One-Way Analysis of Variance between boys and girls, comparing internalizing and externalizing t scores for teachers on the CBCL in the SFARI dataset covarying for IQ (n=1240; girls n=179).

Construct	Girls		Boys		$F(1, 1237)$	η_p^2
	M	*SD*	*M*	*SD*		
MANCOVA - Using IQ as a covariate						
Internalizing T score (Teacher)	60.40	9.17	60.64	8.37	.138	<.001
Externalizing T score (Teacher)	61.22	8.54	58.38	8.15	16.251*	.013

*p <.001

Note. Interaction term between Sex and IQ was not significant. Interaction term was removed from the model.
~Interaction for the Externalizing T score (TRF – Teacher; *Full_scale_iq) p=.049

Table 20

LPA Fit Statistics for Girls based on Model 1 in the SP dataset (Equal Variances and Covariances Fixed to Zero)

Classes	AIC	BIC	Entropy
2	3813.21	3872.53	0.81
3	3612.77	3694.33	0.83
4	3523.76	3627.56	0.87
5	3490.82	3616.86	0.86

Table 21

LPA Fit Statistics for Boys based on Model 1 in the SP dataset (Equal Variances and Covariances Fixed to Zero)

Classes	AIC	BIC	Entropy
2	18275.20	18359.37	0.79
3	17479.26	17594.99	0.82
4	17194.80	17342.09	0.80
5	17020.11	17198.97	0.81

Table 22

Class Membership for Girls in the SP dataset (n=301)

Classes	*n* for class 1	*n* for class 2	*n* for class 3	*n* for class 4	*n* for class 5
2	165	136	--		
3	149	66	86	--	
4	131	9	70	91	--
5	118	26	69	72	16

Table 23

Class Membership for Boys in the SP dataset (n=1424)

Classes	*n* for class 1	*n* for class 2	*n* for class 3	*n* for class 4	*n* for class 5
2	688	735	--		
3	363	309	751	--	
4	256	503	549	115	--
5	214	204	516	352	137

Table 24

Three-Profile Model Results in SP Dataset for Girls (n=301)

Variable	Class 1 (*n*=149)	Class 2 (*n*=66)	Class 3 (*n*=86)
Child Age	-0.32 (.10)	-0.39 (.16)	0.83 (.15)
VABS - Socialization SS	0.01 (.12)	1.27 (.24)	-1.01 (.10)
VABS - Communication SS	0.03 (.14)	1.23 (.20)	-1.00 (.08)
VABS - Daily Living Skills SS	0.15 (.10)	1.11 (.21)	-1.12 (.15)
CARS - Total	0.13 (.13)	-1.11 (.19)	0.65 (.11)

Table 25

Three-Profile Model Results in SP Dataset for Boys (n=1423)

Variable	Class 1 (*n*=363)	Class 2 (*n*=309)	Class 3 (*n*=751)
Child Age	0.74 (.06)	-0.24 (.07)	-.26 (.04)
VABS - Socialization SS	-1.10 (.05)	1.16 (.08)	.05 (.04)
VABS - Communication SS	-1.04 (.04)	1.21 (.06)	-.01 (.05)
VABS - Daily Living Skills SS	-1.14 (.06)	1.17 (.07)	.06 (.04)
CARS - Total	0.79 (.06)	-.95 (.06)	0.02 (.04)

Table 26

Three-Profile Model in SP Dataset Unstandardized Means for Girls

Variable	Overall Sample Average for Girls	Class 1	Class 2	Class 3
Child Age	4 yrs 1 mo	3 years 6 months (-)	3 years 6 months (-)	4 years 11 months (+)
VABS - Socialization SS	63	62 (mild low)	80 (bord low)	46 (mod low)
VABS - Communication SS	61	66 (mild low)	79 (bord low)	48 (mod low)
VABS - Daily Living Skills SS	63	63 (mild low	75 (bord low)	53 (mod low)
CARS - Total	30	31 (mild to mod)	25 (min to no)	33 (mild to mod)

Table 27

LPA Fit Statistics for Girls based on Model 1 in the SFARI dataset (Equal Variances and Covariances Fixed to Zero)

Classes	AIC	BIC	Entropy
2	19818.97	20058.51	0.92
3	19425.92	19747.93	0.89
4	19058.94	19463.42	0.90
5	18801.26	19288.19	0.91
6	18627.03	19196.43	0.92
7	18525.73	19177.60	0.91

Table 28

LPA Fit Statistics for Boys based on Model 1 in the SFARI dataset (Equal Variances and Covariances Fixed to Zero)

Classes	AIC	BIC	Entropy
2	127071.65	127423.97	0.91
3	123458.12	123931.72	0.89
4	121038.20	121633.09	0.89
5	119617.64	120333.82	0.88
6	118506.19	119343.67	0.88
7	117565.86	118524.63	0.89

Table 29

Class Membership for Girls in the SFARI dataset (n=375)

Classes	*n* for class 1	*n* for class 2	*n* for class 3	*n* for class 4	*n* for class 5	*n* for class 6	*n* for class 7
2	218	157	--				
3	122	105	148	--			
4	94	76	95	110	--		
5	61	49	86	71	108	--	
6	55	46	84	42	61	87	--
7	55	44	81	37	64	39	55

Table 30

Class Membership for Boys in the SFARI dataset (n=2382)

Classes	*n* for class 1	*n* for class 2	*n* for class 3	*n* for class 4	*n* for class 5	*n* for class 6	*n* for class 7
2	766	1616	--				
3	625	1115	642	--			
4	437	497	619	829	--		
5	594	411	297	506	574	--	
6	232	464	498	378	301	509	--
7	284	407	339	239	438	528	147

Table 31

Five-Class Model Results SFARI Sample for Girls (n=375)

Variable	Class 1 (*n*=61)	Class 2 (*n*=49)	Class 3 (*n*=86)	Class 4 (*n*=71)	Class 5 (*n*=108)
Age	-0.49 (.14)	0.39 (.24)	-0.05 (.22)	0.34 (.23)	-0.07 (.18)
Non-Verbal IQ	0.46 (.18)	-1.18 (.14)	-0.85 (.14)	0.36 (.22)	0.71 (.10)
Verbal IQ	0.35 (.16)	-1.21 (.14)	-0.92 (.18)	0.49 (.18)	0.76 (.10)
VABS - Socialization SS	0.44 (.18)	-1.17 (.17)	-0.61 (.20)	-0.08 (.19)	0.81 (.14)
VABS - Communication SS	0.58 (.19)	-1.29 (.16)	-0.68 (.17)	0.04 (.19)	0.76 (.16)
VABS - Daily Living Skills SS	0.50 (.18)	-1.19 (.17)	-0.61 (.15)	-0.02 (.21)	0.74 (.14)
ADI-R - Socialization	-0.13 (.19)	0.98 (.12)	0.47 (.17)	-0.10 (.12)	-0.68 (.13)
ADOS - CSS	0.95 (.15)	0.72 (.26)	0.26 (.31)	-0.64 (.24)	-0.68 (.17)
ADOS - Comm Social	0.86 (.18)	0.89 (.24)	0.48 (.35)	-0.73 (.20)	-0.83 (.14)
ADOS - Social Affect	0.79 (.19)	1.02 (.22)	0.50 (.33)	-0.78 (.20)	-0.83 (.13)
ADI-R - RRB Total	-0.01 (.25)	0.58 (.14)	0.09 (.14)	0.29 (.19)	-0.51 (.13)
ADOS - RRB Total	0.22 (.16)	0.81 (.21)	0.56 (.25)	-0.55 (.14)	-0.60 (.13)
RBS - Ritual Same Avg	-0.04 (.35)	0.84 (.29)	-0.44 (.14)	0.88 (.32)	-0.57 (.10)
RBS - SIB Avg	-0.21 (.16)	1.04 (.31)	-0.28 (.12)	0.43 (.24)	-0.40 (.08)
RBS - Stereotypy Avg	-0.22 (.22)	1.36 (.26)	-0.02 (.13)	0.25 (.22)	-0.64 (.11)
RBS - Compulsive Avg	0.17 (.32)	0.74 (.27)	-0.33 (.13)	0.62 (.32)	-0.57 (.07)
RBS - Restricted Avg	-0.18 (.28)	0.85 (.21)	-0.13 (.13)	0.71 (.29)	-0.62 (.12)
SRS - parent total	-0.26 (.23)	1.09 (.13)	0.08 (.15)	0.59 (.16)	-0.78 (.16)
CBCL - Internalizing t score	-0.09 (.28)	0.15 (.17)	-0.35 (.13)	0.66 (.16)	-0.16 (.19)
CBCL - Externalizing t score	0.10 (.27)	0.56 (.19)	-0.29 (.14)	0.43 (.14)	-0.36 (.17)

Table 32

Five-Class Model Results SFARI Sample for Boys (n=2382)

Variable	Class 1 (*n*=594)	Class 2 (*n*=411)	Class 3 (*n*=297)	Class 4 (*n*=506)	Class 5 (*n*=574)
Age	-0.17 (.06)	0.15 (.05)	-0.20 (.06)	-0.33 (.09)	0.47 (.10)
Non-Verbal IQ	-0.03 (.07)	-1.54 (.05)	0.29 (.10)	0.70 (.05)	0.36 (.08)
Verbal IQ	-0.14 (.08)	-1.58 (.04)	0.35 (.11)	0.72 (.04)	0.46 (.08)
VABS - Socialization SS	0.04 (.07)	-1.24 (.05)	-0.10 (.08)	1.08 (.06)	-0.06 (.08)
VABS - Communication SS	0.04 (.08)	-1.38 (.06)	0.14 (.10)	1.04 (.08)	-0.06 (.10)
VABS - Daily Living Skills SS	-0.01 (.08)	-1.25 (.05)	0.13 (.11)	0.96 (.07)	-0.02 (.10)
ADI-R - Socialization	0.12 (.06)	0.86 (.04)	0.15 (.09)	-0.80 (.06)	-0.10 (.06)
ADOS - CSS	0.86 (.06)	0.31 (.06)	-0.14 (.14)	-0.39 (.11)	-0.69 (.06)
ADOS - Comm Social	0.88 (.06)	0.80 (.06)	-0.30 (.14)	-0.56 (.09)	-0.82 (.05)
ADOS - Social Affect	0.79 (.06)	0.98 (.05)	-0.35 (.14)	-0.62 (.09)	-0.78 (.05)
ADI-R - RRB Total	0.01 (.06)	0.04 (.05)	0.58 (.06)	-0.31 (.07)	-0.07 (.08)
ADOS - RRB Total	0.42 (.06)	0.83 (.05)	-0.13 (.10)	-0.37 (.06)	-0.63 (.05)
RBS - Ritual Same Avg	-0.24 (.06)	-0.04 (.07)	1.49 (.13)	-0.51 (.06)	-0.06 (.12)
RBS - SIB Avg	-0.27 (.04)	0.45 (.06)	0.93 (.16)	-0.42 (.03)	-0.16 (.06)
RBS - Stereotypy Avg	-0.18 (.06)	0.75 (.07)	1.08 (.15)	-0.61 (.05)	-0.38 (.09)
RBS - Compulsive Avg	-0.19 (.06)	0.25 (.07)	1.29 (.18)	-0.46 (.04)	-0.24 (.09)
RBS - Restricted Avg	-0.12 (.06)	0.22 (.08)	1.15 (.09)	-0.58 (.06)	-0.12 (.10)
SRS - parent total	-0.13 (.06)	0.66 (.06)	0.91 (.08)	-0.96 (.07)	0.05 (.09)
CBCL - Internalizing t score	-0.22 (.06)	-0.10 (.05)	0.83 (.07)	-0.45 (.09)	0.26 (.11)
CBCL - Externalizing t score	-0.22 (.06)	0.07 (.05)	0.82 (.07)	-0.47 (.08)	0.17 (.10)

Table 33

Five-Class Model Unstandardized Means SFARI Sample for Girls - Simplified

Variable	Overall Sample Average for Girls	Class 1	Class 2	Class 3	Class 4	Class 5
Age	9 years 2 months	7 years 5 months	10 years 8 months	8 years 11 months	10 years 4 months	8 years 11 months
Non-Verbal IQ	78	90 (avg)	47 (mod ID)	56 (mild ID)	88 (avg)	97 (avg)
Verbal IQ	74	85 (avg)	35 (severe ID)	44 (mod ID)	90 (avg)	98 (avg)
VABS - Socialization SS	69	75 (bord low)	54 (mod low)	62 (mild low)	68 (mild low)	80 (bord low)
VABS - Communication SS	75	83 (bord low)	56 (mild low)	65 (mild low)	75 (bord low)	85 (avg)
VABS - Daily Living Skills SS	74	81 (bord low)	57 (mild low)	65 (mild low)	74 (bord low)	84 (bord low)
ADI-R - Socialization	21	20 (-)	26 (+)	23 (+)	20 (=)	16 (-)
ADOS - CSS	7	9 (high)	9 (high)	<8 (mod)	6 (mod)	6 (mod)
ADOS - Comm Social	14	18 (+)	18 (+)	16 (+)	11 (-)	10 (-)
ADOS - Social Affect	12	15 (+)	16 (+)	14 (+)	8 (-)	8 (-)
ADI-R - RRB Total	6	6 (=)	8 (+)	6 (=)	7 (+)	5 (-)
ADOS - RRB Total	4	4 (=)	6 (+)	5 (+)	3 (-)	3 (-)
RBS - Ritual Same Avg	.7	0.7 (=)	1 (+)	0.5 (-)	1.2 (+)	0.4 (-)
RBS - SIB Avg	.3	0.2 (-)	0.7 (+)	0.2 (-)	0.5 (+)	0.2 (-)
RBS - Stereotypy Avg	.6	0.5 (-)	1.4 (+)	0.6 (=)	0.8 (+)	0.3 (-)
RBS - Compulsive Avg	.5	0.6 (+)	0.9 (+)	0.4 (-)	0.8 (+)	0.2 (-)
RBS - Restricted Avg	1.2	1.0 (-)	1.9 (+)	1.1 (-)	1.7 (+)	0.7 (-)
SRS - parent total	100	92 (mod)	131 (severe)	103 (severe)	115 (severe)	78 (mod)
CBCL - Internalizing t score	60	59 (normal)	61 (Bord)	57 (normal)	66 (clinical)	58 (normal)
CBCL - Externalizing t score	58	58 (normal)	63 (Bord)	55 (normal)	62 (Bord)	53 (normal)

Table 34

CARS Polychoric Correlations for Girls (n = 301)

Item	1	2	3	4	5	6	7	8	9	10	11	12	13	15
1	-													
2	-.11	-												
3	-.23	.00	-											
4	-.02	-.12	-.12	-										
5	-.14	-.04	-.10	-.08	-									
6	-.12	-.12	-.01	-.06	-.13	-								
7	-.06	.02	-.04	-.01	.00	-.04	-							
8	-.08	.00	-.09	-.05	-.07	-.07	-.10	-						
9	-.11	-.05	-.07	-.02	-.05	-.03	-.10	-.07	-					
10	-.08	-.05	-.07	.01	-.05	-.03	-.06	-.10	.01	-				
11	-.20	-.18	-.09	-.17	-.13	-.10	-.21	-.13	-.19	-.24	-			
12	-.04	.02	-.09	-.02	-.04	-.09	-.12	-.07	-.10	-.23	.07	-		
13	-.10	-.03	-.08	-.03	-.06	-.02	-.11	-.06	-.04	-.14	-.04	-.14	-	
15	-.07	.04	-.12	-.02	.02	-.09	-.17	.04	-.06	-.26	.03	-.05	-.01	-

Table 35

CARS Polychoric Correlations for Boys (n = 1423)

Item	1	2	3	4	5	6	7	8	9	10	11	12	13	15
1	-													
2	-.10	-												
3	-.03	.00	-											
4	-.14	-.06	-.12	-										
5	-.12	-.08	-.13	-.09	-									
6	-.11	-.04	-.03	-.06	-.09	-								
7	-.10	.00	-.12	-.03	-.03	-.10	-							
8	-.08	.03	-.09	-.02	-.06	-.07	-.04	-						
9	-.08	-.06	-.05	-.03	-.06	-.02	-.08	-.04	-					
10	-.08	-.06	-.08	-.07	-.07	-.05	-.06	-.07	-.05	-				
11	-.17	-.22	-.15	-.11	-.14	-.11	-.16	-.19	-.17	-.12	-			
12	-.07	.04	-.11	-.02	.00	-.10	-.04	-.05	-.07	-.15	.01	-		
13	-.07	-.04	-.08	-.05	-.08	-.08	-.09	-.07	-.05	-.13	-.02	-.16	-	
15	-.06	.00	-.13	-.04	.00	-.08	-.05	-.05	-.11	-.14	-.01	-.11	-.03	-

Table 36

CARS CFA Model Fit for Girls and Boys

Model	df	X^2	CFI	TLI	RMSEA	SRMR
1. Three-factor model – Girls	71	**186.9**	**.959**	**.948**	**.074**	**.054**
2. Three-factor model - Boys	71	**833.7**	**.946**	**.930**	.087	**.048**

Note. Acceptable fit indices are bolded.

Table 37

CARS Standardized Factor Loadings for 3-Factor model for Girls and Boys

	Girls	Boys
Social Interaction Factor		
1. Relate to people	**.82**	**.84**
3. Emotional response	.10	.20
7. Visual response	**.79**	**.74**
15. General impressions	**.83**	**.73**
Communication Factor		
2. Imitation	**.77**	**.77**
5. Object use	**.63**	**.63**
11. Verbal communication	**.59**	**.63**
12. Non-verbal communication	**.82**	**.83**
Emotion Regulation		
3. Emotional response	**.73**	**.83**
4. Body use	**.67**	**.67**
5. Object use	.23	.15
6. Adapt to change	**.50**	**.39**
8. Listening response	**.64**	**.71**
9. Taste, smell, touch	**.57**	**.62**
10. Fear or nervous	**.46**	**.54**
13. Activity level	**.54**	**.56**
15. General impressions	.05	.16

Note. significant loadings ($p < .001$) are bold

Table 38

ADI-R Polychoric Correlations for Girls (n = 367)

Item	31	33	34	35	36	37	38	39	42	43	44	45	47	48	49	50	51	52	53	54
31	-																			
33	.29	-																		
34	.36	.26	-																	
35	.33	.28	.44	-																
36	-.27	.00	-.08	-.36	-															
37	.35	.39	.36	.18	-.15	-														
38	.13	.24	-.09	-.18	.25	.05	-													
39	.13	.40	.07	-.02	.09	.22	.32	-												
42	.25	.19	.30	.30	.04	.11	.01	.11	-											
43	.38	.30	.28	.37	-.38	.29	-.06	.09	.51	-										
44	.33	.31	.22	.32	-.27	.28	-.03	.17	.46	.95	-									
45	.29	.19	.26	.30	-.07	.10	.01	.07	.61	.66	.59	-								
47	.17	.21	.28	.23	-.02	.03	.03	.13	.45	.41	.44	.50	-							
48	.19	.17	.20	.26	.06	.08	-.01	.05	.40	.43	.40	.45	.59	-						
49	.26	.28	.36	.38	-.06	.21	-.03	.07	.39	.48	.39	.47	.47	.65	-					
50	.12	.12	.13	.04	.11	.10	.10	.12	.30	.19	.22	.30	.32	.27	.22	-				
51	.28	.22	.11	.16	-.03	.08	.09	.19	.35	.35	.34	.49	.38	.32	.37	.34	-			
52	.21	.25	.20	.33	-.03	.26	.00	.14	.51	.47	.48	.49	.51	.46	.57	.36	.48	-		
53	.25	.25	.33	.23	.02	.15	.01	.21	.42	.43	.35	.49	.41	.39	.42	.21	.35	.49	-	
54	.12	.27	.24	.30	-.06	.04	-.03	.18	.52	.45	.44	.57	.51	.41	.46	.36	.47	.62	.48	-
55	.29	.24	.35	.30	-.10	.07	.02	.15	.46	.61	.55	.54	.58	.45	.43	.29	.43	.55	.49	.52
56	.24	.23	.20	.22	.02	.11	.14	.25	.38	.36	.38	.41	.36	.37	.32	.47	.40	.46	.34	.48
57	.07	.23	.15	.16	.01	.11	.16	.25	.27	.39	.41	.45	.35	.36	.35	.30	.49	.42	.35	.49
58	.16	.21	.22	.15	.21	.20	.02	.23	.27	.18	.22	.19	.26	.26	.21	.18	.26	.29	.29	.30
59	.31	.27	.14	.17	-.01	.04	.12	.26	.40	.33	.35	.48	.34	.30	.35	.38	.57	.45	.40	.51
61	.09	.14	.12	.14	.08	.00	.11	.08	.31	.35	.34	.40	.42	.40	.43	.25	.33	.47	.35	.48
62	.19	.13	.31	.17	.00	.13	.03	.11	.40	.36	.33	.42	.36	.33	.43	.20	.47	.41	.35	.42
63	.47	.25	.25	.33	-.14	.22	.06	.21	.39	.44	.41	.41	.35	.34	.48	.20	.38	.42	.37	.35
64/65	.24	.27	.24	.28	-.06	.17	.04	.08	.35	.34	.34	.38	.41	.45	.57	.29	.38	.42	.37	.40
67	.30	.15	.07	-.08	-.01	.26	.35	.17	.03	.11	.14	.11	-.04	.11	.12	.18	.11	.11	.23	.06
68	-.05	.06	-.13	-.03	.24	-.07	.08	.06	-.06	-.11	-.05	.01	.01	.01	.00	-.02	.04	.04	.04	-.02
69/71	.30	.22	.25	.15	.11	.21	.17	.15	.22	.28	.27	.30	.23	.29	.27	.30	.24	.25	.24	.12
70	.14	.13	-.01	.01	.04	.06	.20	.46	.02	.07	.16	.06	.08	.06	.15	.04	.10	.02	.08	.12
77/78	.36	.31	.30	.09	-.01	.27	.05	.00	.13	.17	.07	.22	.16	.20	.25	.20	.08	.12	.18	.09

Item	55	56	57	58	59	61	62	63	64/65	67	68	69/71	70	77/78
55	-													
56	.48	-												
57	.40	.45	-											
58	.28	.30	.24	-										
59	.44	.44	.46	.30	-									
61	.41	.36	.43	.17	.26	-								
62	.39	.32	.25	.19	.40	.40	-							
63	.39	.32	.21	.23	.45	.34	.57	-						
64/65	.38	.32	.33	.23	.35	.43	.48	.48	-					
67	.11	.04	.16	.18	.25	.10	.00	.13	.18	-				
68	.06	.04	.09	.13	.10	.05	.04	-.05	.02	-.01	-			
69/71	.24	.26	.22	.23	.25	.21	.17	.30	.26	.13	.04	-		
70	.15	.15	.21	.14	.18	.18	.15	.18	.19	.19	.06	.22	-	
77/78	.17	.12	.12	.15	.14	.10	.07	.23	.22	.17	-.07	.19	.06	-

Table 39

ADI-R Polychoric Correlations for Boys (n = 2356)

Item	31	33	34	35	36	37	38	39	42	43	44	45	47	48	49	50	51	52	53	54
31	-																			
33	.24	-																		
34	.21	.30	-																	
35	.28	.32	.50	-																
36	-.07	.02	-.09	-.12	-															
37	.34	.34	.20	.27	-.02	-														
38	.04	.14	-.02	.00	.21	.11	-													
39	.17	.32	.10	.11	.17	.25	.19	-												
42	.21	.13	.29	.29	.03	.10	.00	.15	-											
43	.31	.19	.22	.23	-.12	.21	.00	.11	.48	-										
44	.24	.13	.18	.21	-.08	.15	.01	.10	.50	.89	-									
45	.18	.12	.28	.28	-.02	.09	.01	.09	.57	.57	.53	-								
47	.15	.15	.23	.27	.02	.04	-.01	.05	.42	.34	.31	.46	-							
48	.18	.17	.28	.28	.03	.10	-.05	.15	.38	.37	.34	.42	.48	-						
49	.23	.16	.31	.34	.06	.15	-.01	.10	.39	.36	.33	.43	.44	.64	-					
50	.10	.08	.17	.16	.12	.04	.06	.02	.33	.19	.17	.32	.27	.23	.29	-				
51	.16	.11	.22	.19	.10	.05	.07	.09	.37	.36	.33	.47	.35	.30	.35	.43	-			
52	.24	.15	.33	.36	.03	.08	-.01	.10	.51	.44	.40	.49	.46	.48	.51	.32	.44	-		
53	.19	.14	.30	.24	.09	.08	-.03	.10	.37	.32	.31	.43	.37	.37	.45	.30	.43	.49	-	
54	.15	.14	.32	.27	.07	.10	.01	.10	.46	.37	.36	.48	.40	.40	.45	.36	.47	.63	.49	-
55	.18	.10	.27	.23	.07	.06	.06	.13	.40	.39	.36	.47	.41	.37	.39	.28	.42	.50	.43	.48
56	.17	.12	.23	.26	.15	.07	.04	.10	.44	.34	.32	.44	.36	.32	.36	.50	.43	.47	.36	.47
57	.13	.09	.23	.23	.10	.02	.03	.09	.37	.33	.29	.42	.36	.33	.37	.40	.49	.46	.38	.49
58	.23	.17	.11	.18	.22	.14	.07	.18	.19	.19	.15	.17	.19	.19	.22	.15	.22	.25	.22	.17
59	.22	.15	.24	.34	.08	.08	.04	.14	.34	.34	.27	.40	.35	.34	.39	.38	.52	.44	.41	.45
61	.12	.07	.18	.18	.16	.01	.06	.09	.36	.27	.29	.41	.37	.37	.41	.29	.38	.43	.38	.40
62	.18	.11	.27	.21	.02	.05	.00	.10	.35	.29	.24	.37	.30	.35	.43	.30	.39	.44	.39	.41
63	.27	.21	.28	.28	.06	.14	.07	.10	.30	.33	.27	.34	.27	.35	.47	.28	.38	.42	.40	.41
64/65	.25	.19	.27	.28	.04	.12	.04	.12	.35	.34	.30	.40	.34	.38	.52	.29	.34	.45	.41	.43
67	.11	.18	.07	.07	.11	.14	.10	.26	.05	.09	.10	.05	-.01	.10	.08	.02	.03	.05	.03	.06
68	-.07	.05	-.03	-.10	.28	-.03	.14	.14	-.07	-.14	-.12	-.08	.00	-.08	-.03	.07	.07	-.06	-.03	.00
69/71	.30	.31	.24	.23	.04	.19	.08	.21	.21	.27	.24	.21	.20	.27	.27	.13	.15	.26	.21	.24
70	.10	.14	.07	.07	.16	.08	.14	.37	.13	.05	.06	.09	.06	.09	.11	.08	.11	.07	.09	.09
77/78	.17	.11	.12	.14	.07	.09	.03	.19	.15	.16	.15	.10	.13	.13	.17	.09	.11	.16	.15	.14

Item	55	56	57	58	59	61	62	63	64/65	67	68	69/71	70	77/78
55	-													
56	.42	-												
57	.45	.44	-											
58	.30	.21	.23	-										
59	.39	.44	.48	.22	-									
61	.38	.38	.37	.19	.35	-								
62	.35	.37	.37	.18	.41	.43	-							
63	.32	.38	.33	.20	.44	.37	.62	-						
64/65	.37	.35	.40	.24	.43	.48	.56	.59	-					
67	.12	.07	.09	.16	.06	.05	.03	.08	.12	-				
68	.03	.03	.06	.09	.01	.02	.00	.01	.00	.02	-			
69/71	.21	.18	.16	.23	.22	.19	.21	.25	.27	.25	.06	-		
70	.14	.12	.14	.19	.17	.10	.09	.12	.15	.17	.16	.20	-	
77/78	.14	.19	.13	.20	.14	.11	.11	.18	.17	.15	.07	.24	.15	-

Table 40

ADI-R CFA Model Fit for Girls

Model	df	X^2	CFI	TLI	RMSEA	SRMR
1. Three-factor model	524	**1284.7**	**.919**	**.913**	**.063**	.090
2. Two-factor DSM model	526	**1310.6**	**.916**	**.911**	**.064**	.091
3. Three-factor Play/Peer model	524	**1222.7**	**.925**	**.920**	**.060**	.088
4. Four-factor model	521	**1048.4**	**.944**	**.939**	**.053**	**.081**

Note: Model 4 had significantly better fit than Model 1, X^2 diff (3) = 107.94, $p < 0.001$; Model 2, X^2 diff (5) = 119.81, $p < 0.001$; and Model 3, X^2 diff (3) = 97.081, $p < 0.001$.

Table 41

ADI-R CFA Model Fit for Boys

Model	df	X^2	CFI	TLI	RMSEA	SRMR
1. Three-factor model	524	**7008.2**	.868	.858	**.072**	**.075**
2. Two-factor DSM model	526	**7248.0**	.863	.854	**.074**	**.074**
3. Three-factor Play/Peer model	524	**6261.7**	.877	.868	**.070**	**.072**
4. Four-factor model	521	**4987.2**	**.909**	**.902**	**.060**	**.066**

Note: Model 4 had significantly better fit than Model 1, X^2 diff (3) = 690.23, $p < 0.001$; Model 2, X^2 diff (5) = 788.07, $p < 0.001$; and Model 3, X^2 diff (3) = 722.8, $p < 0.001$.

Table 42

ADI-R Standardized Factor Loadings for Model 4 for Girls and Boys (4-Factor Model)

	Girls	Boys
Social Factor		
31. Use of other's body	**.44**	**.35**
50. Direct gaze	**.47**	**.51**
51. Social smiling	**.66**	**.65**
52. Showing/directing attention	**.76**	**.77**
53. Offering to share	**.64**	**.65**
54. Seek to share enjoyment	**.75**	**.73**
55. Offering comfort	**.75**	**.66**
56. Quality of social overtures	**.63**	**.66**
57. Range of facial expressions	**.62**	**.65**
58. Inappropriate facial expressions	**.40**	**.36**
59. Appropriateness of social responses	**.68**	**.66**
Communication Factor		
34. Social chat (verbal item)	**.33**	**.35**
35. Reciprocal conversation (verbal item)	**.49**	**.51**
42. Pointing to express interests	**.72**	**.74**
43. Nodding	**.97**	**.90**
44. Head shaking	**.94**	**.86**
45. Conventional instrumental gestures	**.84**	**.81**
Play/Peer Factor		
47. Spontaneous imitation of actions	**.72**	**.62**
48. Imaginative play	**.70**	**.67**
49. Imaginative peer play	**.77**	**.74**
61. Imitative social play	**.62**	**.63**
62. Interest in children	**.64**	**.68**
63. Response to children's approaches	**.67**	**.69**
64/65. Group play / Friendships	**.69**	**.72**
RRBI Factor		
33. Stereotyped utterances (verbal item)	**.65**	**.53**
36. Inappropriate questions (verbal item)	-.09	.17
37. Pronominal reversal (verbal item)	**.43**	**.38**
38. Neologisms (verbal item)	.20	.15
39. Verbal rituals (verbal item)	**.47**	**.46**
67. Unusual preoccupations	**.36**	**.31**
68. Circumscribed interests	.04	.05
69/71. Repetitive use of objects/Sensory interests	**.61**	**.70**
70. Compulsions/Rituals	**.35**	**.38**
77/78. Hand Finger movements/ Complex Mannerisms	**.42**	**.43**

Note. significant loadings ($p < .001$) are bold

Table 43

RBS-R Polychoric Correlations for Total Sample (n = 2757)

Item	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	-																			
2	.53	-																		
3	.47	.40	-																	
4	.40	.40	.44	-																
5	.35	.33	.39	.49	-															
6	.32	.36	.38	.46	.44	-														
7	.26	.29	.28	.32	.31	.32	-													
8	.27	.30	.20	.38	.37	.36	.68	-												
9	.26	.26	.23	.35	.39	.32	.74	.73	-											
10	.24	.23	.16	.24	.27	.28	.43	.36	.40	-										
11	.28	.25	.18	.23	.22	.25	.44	.41	.43	.40	-									
12	.25	.23	.18	.22	.23	.28	.34	.35	.39	.41	.50	-								
13	.23	.28	.19	.32	.32	.36	.33	.35	.35	.32	.41	.36	-							
14	.08	.11	.07	.10	.15	.17	.18	.18	.23	.27	.44	.53	.30	-						
15	.17	.20	.17	.22	.25	.31	.20	.24	.23	.12	.17	.20	.23	.14	-					
16	.21	.28	.22	.29	.33	.35	.20	.23	.21	.17	.19	.21	.25	.18	.63	-				
17	.21	.20	.19	.21	.24	.27	.20	.22	.25	.17	.24	.20	.26	.27	.40	.46	-			
18	.27	.23	.23	.18	.24	.26	.18	.20	.18	.14	.20	.17	.24	.11	.40	.57	.51	-		
19	.19	.23	.13	.27	.26	.27	.13	.24	.23	.06	.10	.18	.20	.05	.51	.41	.31	.36	-	
20	.06	.15	.03	.10	.18	.19	.15	.16	.23	.10	.22	.20	.17	.21	.40	.34	.29	.26	.29	-
21	.28	.25	.28	.28	.32	.30	.23	.25	.22	.23	.20	.22	.28	.13	.43	.48	.38	.52	.44	.27
22	.29	.32	.30	.36	.42	.40	.28	.34	.30	.22	.23	.28	.29	.15	.32	.41	.35	.34	.31	.27
23	.17	.16	.17	.21	.25	.30	.16	.24	.20	.14	.12	.14	.13	.13	.31	.33	.27	.25	.32	.22
24	.14	.20	.14	.22	.23	.32	.21	.23	.27	.11	.15	.20	.20	.14	.41	.45	.34	.39	.35	.37
25	.18	.21	.12	.17	.24	.23	.19	.24	.24	.13	.15	.24	.17	.16	.42	.43	.38	.41	.36	.37
26	.19	.21	.17	.27	.29	.31	.22	.25	.27	.20	.22	.22	.23	.18	.37	.41	.29	.35	.35	.35
27	.15	.16	.12	.21	.24	.25	.15	.25	.26	.13	.13	.17	.19	.16	.38	.37	.28	.33	.37	.35
28	.11	.14	.09	.13	.14	.18	.12	.17	.19	.10	.14	.15	.08	.14	.24	.25	.24	.28	.31	.29
29	.20	.19	.18	.18	.20	.26	.20	.23	.25	.14	.13	.19	.18	.16	.60	.55	.37	.47	.38	.43
30	.21	.16	.17	.19	.19	.26	.21	.22	.27	.17	.20	.18	.15	.13	.24	.27	.20	.27	.18	.25
31	.16	.22	.11	.23	.24	.28	.25	.32	.29	.15	.23	.20	.20	.18	.31	.31	.23	.21	.30	.35
32	.33	.34	.28	.37	.34	.31	.32	.33	.31	.22	.26	.26	.21	.13	.42	.43	.38	.39	.43	.31
33	.20	.21	.08	.19	.21	.21	.23	.26	.26	.18	.21	.20	.19	.18	.34	.35	.27	.31	.33	.33

34	.21	.21	.14	.19	.23	.29	.19	.24	.31	.14	.17	.23	.22	.17	.35	.38	.33	.38	.32	.34

Item	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
21	-																			
22	.43	-																		
23	.34	.24	-																	
24	.43	.33	.48	-																
25	.52	.32	.42	.58	-															
26	.46	.33	.40	.51	.58	-														
27	.45	.29	.41	.48	.50	.56	-													
28	.38	.22	.30	.37	.37	.40	.54	-												
29	.50	.33	.40	.52	.55	.52	.48	.36	-											
30	.29	.20	.32	.31	.35	.41	.38	.29	.45	-										
31	.32	.27	.31	.39	.39	.42	.46	.36	.46	.50	-									
32	.50	.38	.33	.36	.47	.51	.42	.30	.47	.37	.45	-								
33	.39	.25	.34	.42	.49	.59	.45	.35	.52	.41	.49	.58	-							
34	.38	.31	.33	.42	.46	.48	.45	.38	.56	.48	.49	.47	.57	-						

Item	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
35	.24	.28	.20	.30	.29	.31	.24	.29	.30	.22	.15	.20	.28	.12	.39	.42	.32	.40	.43	.25
36	.23	.22	.25	.27	.24	.34	.25	.25	.20	.18	.17	.20	.23	.14	.34	.38	.27	.26	.25	.22
37	.19	.19	.12	.22	.23	.27	.25	.31	.29	.22	.25	.18	.19	.17	.32	.31	.19	.21	.23	.31
38	.27	.23	.17	.24	.26	.31	.28	.31	.33	.19	.20	.23	.21	.18	.39	.39	.33	.34	.34	.32
39	.23	.21	.13	.17	.19	.27	.22	.26	.30	.17	.18	.21	.12	.19	.35	.35	.32	.36	.34	.32
40	.17	.20	.15	.17	.22	.29	.14	.18	.23	.12	.16	.17	.12	.19	.32	.28	.23	.25	.26	.35
41	.22	.22	.19	.24	.32	.32	.24	.24	.34	.14	.16	.19	.23	.16	.37	.36	.29	.29	.28	.40
42	.26	.29	.29	.34	.50	.43	.24	.28	.29	.20	.25	.21	.28	.14	.37	.40	.33	.32	.35	.27
43	.28	.29	.33	.37	.49	.41	.23	.27	.27	.20	.17	.15	.29	.08	.32	.37	.28	.35	.39	.11

Item	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
35	.51	.32	.37	.42	.46	.56	.46	.35	.51	.38	.38	.67	.63	.55	-					
36	.42	.28	.36	.37	.34	.39	.36	.34	.39	.34	.33	.34	.31	.38	.41	-				
37	.33	.21	.31	.39	.38	.44	.46	.37	.45	.51	.67	.38	.50	.52	.43	.39	-			
38	.46	.30	.43	.50	.55	.56	.50	.37	.56	.51	.55	.48	.59	.59	.54	.47	.68	-		
39	.45	.26	.37	.48	.54	.54	.48	.41	.54	.43	.50	.49	.55	.54	.54	.41	.59	.81	-	
40	.30	.23	.30	.35	.30	.34	.43	.48	.34	.30	.42	.31	.30	.33	.31	.40	.40	.41	.39	-
41	.39	.28	.30	.42	.37	.42	.43	.35	.43	.34	.39	.38	.34	.35	.37	.39	.38	.44	.41	.63
42	.40	.36	.24	.31	.31	.33	.34	.24	.34	.24	.31	.38	.27	.36	.37	.34	.28	.33	.30	.40
43	.39	.35	.23	.29	.26	.33	.28	.21	.29	.20	.25	.36	.27	.29	.44	.33	.23	.30	.26	.30

Item	41	42	43
41	-		
42	.46	-	
43	.36	.70	-

Table 44

RBS-R CFA Model Fit for Girls and Boys based on 5-Factor model by Lam and Aman (2007)

Model	df	X^2	CFI	TLI	RMSEA	SRMR
1. Five-factor model – Girls	655	**1001.1**	**.947**	**.943**	**.038**	**.081**
2. Five-factor model - Boys	655	**4209.2**	**.920**	**.914**	**.048**	**.059**

Note. Acceptable fit indices are bolded.

Table 45

RBS-R Standardized Factor Loadings for 5-Factor model (Lam & Aman, 2007) for Girls, and Boys.

	Girls	Boys
Ritualistic Sameness Factor		
26. Travel	**.71**	**.73**
27. Play	**.66**	**.69**
28. Communication	**.54**	**.56**
30. Objects	**.62**	**.60**
31. Becomes upset	**.72**	**.71**
32. Insists walking	**.68**	**.74**
33. Insists sitting	**.66**	**.72**
34. Dislikes changes	**.71**	**.71**
35. Insists door	**.65**	**.76**
37. Resists change	**.75**	**.74**
38. Insists routine	**.86**	**.86**
39. Insists place	**.80**	**.81**
Self-Injurious Behaviour Factor		
7. Hits self body	**.68**	**.76**
8. Hits self object	**.93**	**.79**
9. Hits self object	**.78**	**.86**
10. Bites self	**.49**	**.56**
11. Pulls	**.63**	**.63**
12. Rubs	**.52**	**.64**
13. Inserts finger	**.71**	**.59**
14. Skin picking	**.54**	**.44**
Stereotypy Factor		
1. Whole body	**.64**	**.53**
2. Head	**.55**	**.57**
3. Hand finger	**.53**	**.49**
4. Locomotion	**.56**	**.62**
5. Object usage	**.69**	**.66**
6. Sensory	**.65**	**.68**
22. Touch tap	**.61**	**.65**
42. Preoccupation	**.79**	**.76**
43. Fascination movement	**.80**	**.71**
Compulsive Behaviour Factor		
15. Arranging	**.70**	**.73**
16. Complete	**.75**	**.77**
17. Washing	**.66**	**.60**
18. Checking	**.67**	**.65**
19. Counting	**.57**	**.64**
20. Hoarding	**.54**	**.57**
Restricted Interests Factor		
36. Likes piece music	**.70**	**.68**
40. Fascination subject	**.63**	**.72**
41. Strongly attached	**.65**	**.77**

Note. significant loadings ($p < .001$) are bold