MEASUREMENT AND CLASSIFICATION OF THE AUTISM PHENOTYPE

MEASUREMENT AND CLASSIFICATION OF THE HETEROGENEOUS AUTISM PHENOTYPE

By STELIOS GEORGIADES, B.A., M.A.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment of the

Requirements for the Degree Doctor of Philosophy

MCMASTER UNIVERSITY DOCTOR OF PHILOSOPHY (2013) HAMILTON, ONTARIO (HEALTH RESEARCH METHODOLOGY)

TITLE: Measurement and Classification of the Heterogeneous Autism Phenotype

AUTHOR: Stelios Georgiades, B.A., M.A.

SUPERVISOR: Professor M.H. Boyle

NUMBER OF PAGES: xiii, 120

ABSTRACT

Autism Spectrum Disorder (ASD) is a heterogeneous disorder with a high burden of suffering and economic cost to society. The current Thesis represents a systematic attempt to investigate ASD heterogeneity, as it relates to the measurement and classification of the clinical phenotype. The Thesis integrates information from multiple constructs (symptoms, traits, behaviours), methods (factor analysis, cluster analysis, and factor mixture modeling), populations (clinical and high-risk samples) and time points (at diagnosis and at age 6) for the investigation of the underlying structure of the ASD phenotype in young children. The Thesis consists of four interrelated empirical studies and one Editorial. Results can be organized into three overarching themes: 1) in preschool children with ASD core diagnostic symptoms (social communication deficits and repetitive behaviours) appear to overlap with other emotional/behavioural problems (attention, withdrawal, anxiety, aggression, emotional reactivity); 2) along the heterogeneous autism spectrum there appear to be distinct, relatively homogeneous subgroups of children; on average, children across these subgroups differ in their levels of symptom severity, adaptive skills, and emotional/behavioural problems; 3) the underlying structure of the ASD symptom phenotype changes as children grow and develop. Thesis findings lend support to a much-needed shift in our conceptual and methodological approach to the study of measurement and classification of autism pathology: that is, instead of a set of categorical symptoms that present early in childhood and remain static over the life span, ASD might be better understood as a complex and dynamic disorder, structured on both categorical and dimensional constructs that vary not only across individuals at any given point, but also within individuals across time.

ACKNOWLEDGEMENTS

The truth is I would have never been able to complete my PhD without the help and support of several people. First, I would like to thank the members of my PhD Supervisory Committee – Dr. Michael Boyle, Dr. Peter Szatmari, and Dr. Steven Hanna – for their guidance and support throughout this journey. I would also like to thank my colleagues and collaborators at the Offord Centre for Child Studies, McMaster University, and across Canada, as well as the sponsors of my work – Canadian Institutes of Health Research, Autism Speaks, Sinneave Family Foundation, and the Mayberry family. Next, I would like to thank the children and families who participate in our research – your positive approach to life is an inspiration! I would also like to thank my students – for teaching me that I know less than I think I know.

I am grateful to my uncle Kokos Hadjiantonis for providing "direction" along the way, my friend Costas Kalotaris for "pushing me" to do a PhD, as well as all my family and friends for their love and encouragement over the years.

I would like to especially thank my parents Pampos and Irene Georgiades for their unconditional love and support throughout my life. I know I wasn't an easy child to parent but the trust you have always shown in me gave me strength and confidence to become who I am today.

Last but not least, I would like to thank my wife Katholiki Georgiades for believing in me (more than I believe in myself) and for teaching me that hard work and patience is the best way of overcoming obstacles in life. Katholiki, you and our children Haris and Katerina are my inspiration in life!

This work is dedicated to my late grandmother Katina Hadjiantonis – for teaching me the value of FAMILY and for always encouraging me to follow my dreams.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS, CHARTS & DIAGRAMS

LIST OF TABLES

LIST OF ABBREVIATIONS AND SYMBOLS

ADI-R: Autism Diagnostic Interview - Revised

ADOS: Autism Diagnostic Observation Schedule

AGP: Autism Genome Project

AIC: Akaike Information Criterion

APA: American Psychiatric Asociation

ASD: Autism Spectrum Disorder

BIC: Bayesian Information Criterion

CBCL: Child Behavior Checklist

DSM: Diagnostic and Statistical Manual of Mental Disorders

EBRP: Emotional Behavioral Repetitive Problems

ES: Effect size

FA: Factor Analysis

FIRB: Fixated Interests and Repetitive Behaviour

FMM: Factor Mixture Modeling

HR: High-risk

ITSEA: Infant Toddler Social Emotional Assessment

LCA: Latent Class Analysis

LR: Low-risk

MP-R: Merrill Palmer-Revised Scales of Development

MSEL: Mullen Scales of Early Learning

PCA: Principal Component Analysis

PLS-4: Preschool Language Scale, 4th edition

PSI-SF: Parenting Stress Index – Short Form

RBS-R: Repetitive behaviors Scale – Revised

SCD: Social Communication Deficit

SD: Standard Deviation

SRS: Social Responsiveness Scale

VABS II: Vineland Adaptive Behavior Scales, Second Edition

DECLARATION OF ACADEMIC ACHIEVEMENT

This 'sandwich' Thesis consists of four empirical Studies and one Editorial. Three of the Studies and the Editorial have already been published in peer-review journals. One study is currently under review. The student (Stelios Georgiades) is the primary author on all of these studies and the main contributor to the following: conzeptualization and development of research ideas, questions and objectives, statistical analyses, interpretation of results, writing of manuscripts, revisions based on recommendations by the thesis committee members, study co-authors, and journal reviewers/editors. All of the work presented here is in line with the general framework outlined in the original thesis proposal and was completed during the period between September 2010 and May 2013. Finally, it is important to note that studies like the ones included in the current Thesis can only be achieved through multidisciplinary collaborations and data collection from across different sites; therefore the lisitng of multiple co-authors on each of the individual papers is well justified.

CHAPTER ONE
INTRODUCTION

Autism Spectrum Disorder

Autism was originally proposed as a distinct disorder by Dr. Leo Kanner back in 1943. In a classic paper Kanner introduced autism by noting: *"there have come to our attention a number of children whose condition differs so markedly and uniquely from anything reported so far, that each case merits – and I hope will eventually receive – a detailed consideration of its fascinating peculiarities"* (Kanner, p. 217; 1943). Around the same time, in an independent investigation, Dr. Hans Asperger (1944) published a paper describing a form of autism that later became known as Asperger's syndrome. Those two landmark papers represent to this date, the foundation for the clinical presentation of the complex disorder known as Autism Spectrum Disorder (ASD).

Using evidence from a twin study, Folstein and Rutter (1977) first proposed a genetic basis for autism which until that point was viewed as disorder associated with certain parenting styles (i.e., cold/distant mothers and uninvolved fathers). In 1991 Lord, Rutter and LeCouteur developed the Autism Diagnostic Interview (ADI) the first comprehensive semi-structured interview with the child's primary caregiver for the assessment of three core symptom domains – Social Impairment, Verbal and Nonverbal Communication, and Repetitive and Restricted Behaviours. In the absence of any biological markers of the disorder, those three phenotypic domains were used to operationalise the diagnostic criteria for autism in the 4th Edition of the Diagnostic and Statistical Manual (DSM-IV). In the DSM-IV, each of these three domains includes a set of symptoms rated on categorical, binary criteria (yes/no) to qualify for a diagnosis (APA, 1994; 2000). The DSM-IV classification distinguishes among three ASD subtypes, namely Autistic Disorder (AD), Asperger's Disorder (AS), and Pervasive Developmental Disorder Not Otherwise Specified (PDD-NOS).

According to a recent report by the American Psychiatric Association (APA, 2012) distinctions among ASD subtypes have been found to be inconsistent over time, variable across sites and often associated with severity, language level or intelligence rather than features of the disorder. Thus, the DSM-5 (APA, 2013) represents a significant shift in the diagnostic conceptualization of ASD. Specifically, because autism is defined by a common set of behaviors, it is represented as a single diagnostic category (i.e. ASD) rather than as multiple subtypes (Happe, 2011).

Epidemiology of ASD

Prevalence

1

Despite the fact that an experienced clinician can (in most cases) reliably assign an ASD diagnosis to a child as young as two years of age, the majority of children do not receive a final/formal diagnosis until much older (3 to 4 years of age; Lord et al., 2006). ASD is reported to occur in all racial, ethnic, and socioeconomic groups (CDC, 2012) and recent epidemiological studies suggest a dramatic increase in its prevalence (number of existing cases in a defined group of children at a specific point in time). The most recent ASD rates range from 1 in every 88 children in a US sample (CDC, 2012) to 1 in every 160 children worldwide (Elsabbagh et al., 2012), making ASD a more common neurodevelopmental disorder than previously thought. Some possible explanations for this prevalence increase include (but are not limited to) broadening of diagnostic concepts and criteria, growing awareness, improved detection of the disorder, and possibly a true increase in prevalence. Epidemiological studies have also agreed that ASD is more common in boys than girls by a ratio of roughly 4 to 1 (Fombonne et al., 2009).

Impact of the Disorder

ASD is known to have a pervasive impact on the person's social relationships, daily living activities and overall functioning. Although some children with ASD go on to achieve relative independence as adults most remain very dependent on their families and support services. In general, the majority of adults with ASD live with their parents or other relatives, do not appear to have close friends, and cannot maintain employment (Howlin et al., 2004). Moreover, while continuing to struggle with core autistic deficits, many individuals with ASD are faced with additional mental health problems starting in adolescence and continuing into adulthood (Rutter, 2012).

ASD creates an enormous burden of suffering for the families. Studies have shown that raising a child with ASD is associated with very high levels of family stress. Due to limited availability of publicly funded interventions families are often responsible for covering the cost of treatment and other related services for their child with autism (Cidav et al., 2012; Jarbrink et al., 2003).

Finally ASD has a huge economic cost for society. For example, recent studies report that annual ASD-related cost in the United States is estimated at $137 billion per year. According to these reports the cost includes indirect costs such as lost family income and productivity in addition to the direct costs of autism-associated care (Cidav et al., 2012).

Currently there is no cure for ASD. While some researchers are continuing their search for the causal mechanisms of the disorder with the hope of one day contributing to its cure, others are working on developing and evaluating the effectiveness of various treatments. A limited number of pharmacological studies have shown moderate improvements in associated features of autism (i.e., hyperactivity, irritability, among others) but there is currently no identified medication to treat the core symptoms of the

disorder (Rutter, 2012). Moreover, despite notable progress in developing effective behavioural and educational interventions, the response to those interventions is partial at best. Unfortunately, our understanding of the differential and heterogeneous response to treatment in ASD is sparse (Warren et al., 2011).

Etiology

ASD has an onset very early in life and is defined by impairment in the development of the brain (Posthuma & Polderman, 2013). Robust evidence supports the idea that ASD is a heritable disorder with a strong genetic basis (Rutter, 2012). As noted above, Folstein and Rutter (1977) were the first to report a difference in autism concordance rates between monozygous (MZ) and dyzgyous (DZ) twins. Subsequent studies (Bailey et al., 1995; Steffenburg et al., 1989) supported those findings and proposed heritability estimates greater than 90%. These estimates contributed to the widely accepted notion that most of the variance in the clinical presentation of autism is due to genetic factors.

Until recently, genetic studies searched for a single gene that would explain the causes of the disorder. However, rigorous investigations on the etiology of ASD using novel methods and technologies suggest that ASD is a complex disorder resulting from multiple genes and diverse causes (Szatmari, 2011). At the same time, an emerging body of literature is highlighting the increased importance of non-genetic factors in the causal mechanisms of the disorder. In a recent, large-scale twin study Hallmayer et al. (2011) reported that up to 55% of the variance in autism susceptibility can be explained by shared environmental factors. However, to date, our ability to identify specific genetic and/or environmental factors that can reliably explain substantial amounts of variance in the etiology of the disorder remains frustratingly limited (Rutter, 2012).

Complexity of the ASD Phenotype

Broader Autism Phenotype

Skuse et al. (2005) note that autistic traits are widely distributed in the general population and that the precise boundaries of the autism diagnosis are currently unclear. Moreover, numerous family studies have provided evidence for the existence of what is known as the "Broader Autism Phenotype" (BAP), a milder (sub-threshold) manifestation of autistic-like traits and characteristics in some relatives of individuals with autism but without serious degrees of impairment (Bailey et al. 1995; Bolton et al. 1994; Piven et al. 1997; Szatmari et al., 2008). Most studies of the BAP focus on parents of children with ASD, although there is a growing literature examining traits related to the BAP in siblings of probands with ASD (Constantino et al., 2006).

Overlap with Other Disorders

Evidence shows that there is a substantial overlap between ASD and other neurodevelopmental and mental disorders such as attention deficit hyperactivity disorder (ADHD), oppositional defiant disorder (ODD), anxiety and mood disorders (Bauminger et al., 2010; Brereton et al., 2006; Gadow et al., 2004). For example, it is believed that anywhere from 20 to 50% of children with ADHD also meet diagnostic criteria for ASD and 30-80% of children with ASD also meet criteria for ADHD (Rutter, 2012; Rommelse et al., 2010; Lichtenstein et al., 2010).

Heterogeneity

After decades of research it is now clear that ASD is a disorder more common and more complex than previously thought (Dawson, 2013). One of the main factors contributing to the limited success in studying the etiology, diagnosis, treatment, and prognosis of ASD, is the remarkable heterogeneity observed at the phenotypic level (Mandell, 2011). For example, while all children with ASD will share similar features that place them within the same "spectrum" they also exhibit notable differences: some are verbal while others are non-verbal; some have IQ in the typical range while others have very low IQ; some have high levels of repetitive behaviours while others only show difficulties in social communication; some are dealing with multiple commorbidities (medical and/or emotional/behavioural) while others are dealing primarily with the core autistic symptoms.

Based on the aforementioned, it becomes imperative that researchers provide systematic, comprehensive ways of "unpacking" this heterogeneity in ASD causes, symptom severity, symptom configuration, response to treatment, and developmental outcome. The development of useful measurement models that have the ability to reliably classify individuals with ASD into meaningful homogeneous subgroups can serve as a good starting point in our quest for understanding the remarkable heterogeneity seen in this disorder.

Measurement and Classification of the Heterogeneous ASD Phenotype

Clinical Approach

The current DSM-IV classification distinguishes among three ASD subtypes, namely Autistic Disorder (AD), Asperger's Disorder (AS), and Pervasive Developmental Disorder Not Otherwise Specified (PDD-NOS). Many investigators now believe that this conceptualization of ASD, which is based primarily on clinical expertise rather than empirical evidence, is limited in capturing the heterogeneity and complexity of clinical presentations of the disorder (Georgiades et al., 2007). Moreover, distinctions among ASD subtypes have been found to be inconsistent over time, variable across sites, and

often associated with severity of language deficits and intellectual impairment rather than a different manifestation of inherent ASD features (APA, 2012; Happe, 2011). Thus, in the DSM-5 (APA, 2013) rather than representing ASD as multiple subtypes, ASD is conceptualized as a single diagnostic category. So in a way, our inability to reliably classify children within autism spectrum has led us to abandon the use of subtypes in our clinical practice. This decision is viewed by some as "paradoxical" at a time when clinicians, researchers, and parents all seem to accept the notion that ASD is a heterogeneous disorder (Zwaigenbaum, 2012). Therefore, it is likely that categories and subgroups in children with autism will continue to be used, regardless of how the DSM changes its definition (Mandell, 2011; Rutter, 2012).

Empirical Approaches

In the absence of robust biological markers ASD continues to be defined and diagnosed using behavioural assessments. Thus, progress in all ASD-related research (i.e., diagnostics, genetics, neuroimaging, pharmacological trials, etc.) depends on our ability to accurately and reliably measure and classify the heterogeneous phenotype(s) associated with the disorder.

To date, researchers have used two general empirical approaches to investigate phenotypic heterogeneity in ASD. The first one is a *person-based* approach designed to explore specific symptom/behaviour profiles (i.e. subtypes) of individuals with ASD. Examples of this approach include Cluster Analysis (CA), Latent Class Analysis (LCA) and Taxometrics, among others. The second is a *variable-centered* approach designed to examine the patterns of association among ASD indicators (symptoms, traits, behaviours). Examples of this approach include Factor Analysis (FA) and Principal Component Analysis (PCA), among others.

It is important to note that although there is some confusion in the field about their use, the two methodological approaches are substantially different. The *person-centered* approach is based on a *categorical* conceptualization of disorder, as per the DSM-IV in which ASD is divided into three separate subgroups/subtypes – Autistic Disorder, Asperger's Disorder, or Pervasive Developmental Disorder Not Otherwise Specified (APA, 2000). In this approach, individuals are classified into qualitatively distinct and mutually exclusive subgroups; once assigned, all individuals within a given subgroup are treated as having similar severity and configuration of symptoms. On the other hand, the *variable-centered* approach is based on a *dimensional* conceptualization of the disorder, as per the DSM-5 in which autism is represented as a single diagnostic category (i.e. ASD) rather than as multiple subtypes (APA, 2012; Happe, 2011). In this approach, individuals are treated as having only quantitative differences on certain derived factors/dimensions; all individuals are treated as being part of the same group (i.e., spectrum) and are described using different levels of severity across symptom dimensions (Ruscio & Ruscio, 2004).

Novel statistical methods and recent data from well-designed, large-scale studies have changed the nature of the debate on "categorical versus dimensional" approaches by allowing researchers to explore the notion that "dimensions can be made into categories by defining thresholds" (Lord and Jones, 2012., p. 492). The key question that arises is how one can go about defining these thresholds along the autism spectrum in a way that meaningfully addresses the challenge of heterogeneity seen in ASD. A relatively new method called Factor Mixture Modeling (FMM) has the potential to (at least partially) answer that question empirically. FMM was designed for the study of complex phenotypes (in this case ASD); through the estimation of complex measurement models, it allows the simultaneous examination of continuous dimensions and latent categories/classes by using both FA and LCA. FMM is a general framework extending FA and LCA by combining the two as sub-models into a single general model (Lubke et al., 2007). Unlike LCA which specifies that observed variables have zero correlations with each other, FMM makes it possible to specify a dimensional factor model for each class, something that can better describe potential severity differences within class. Indeed, according to Hudziak (2007) a complimentary 'hybrid' approach that integrates both categorical and dimensional elements might be the best way forward towards a new diagnostic and classification system of mental disorders.

The Current Thesis

Rationale

Out of all child neurodevelopmental disorders, Autism Spectrum Disorder (ASD) is perhaps the most heterogeneous in terms of its etiology, clinical presentation and prognosis (Geschwind, 2009). To highlight this variability in the presentation of ASD, clinicians and parents will often say *"If you have seen one child with autism, you have seen one child with autism"*. The research presented in the current *Thesis* is driven from an aspiration to better understand this heterogeneity observed in children with ASD. It is my hope that this line of work will generate scientific evidence for the development of refined and comprehensive measurement and classification models for the complex clinical phenotype seen in children with ASD.

Objectives

The overarching objective of this *Thesis* was to create a comprehensive body of evidence that will contribute to our understanding of the phenotypic heterogeneity seen in ASD. The current *Thesis* focuses on issues related to the measurement and classification of ASD and represents a systematic investigation of the underlying structure of the clinical phenotype. The *Thesis* consists of four interrelated empirical studies and one Editorial. The empirical studies (Studies 1, 2, 3 & 4) demonstrate the application of three general methodological approaches – i.e., *variable-centered*, *person-centered*, and

combined *variable/person-centered* – to explore the underlying associations across conventionally used indicators (symptoms, traits, behaviours) related to ASD and to identify meaningful ways of grouping children who share similar profiles on these indicators. The Editorial builds on the four empirical studies to highlight the importance of studying heterogeneity in autism. The current *Thesis* represents the first ever attempt to integrate information from multiple *constructs* (symptoms, traits, behaviours), *methods* (factor analysis, cluster analysis, and factor mixture modeling), *populations* (i.e., clinical and high-risk samples) and *time points* (i.e., at diagnosis and then again at age 6) for the investigation of the underlying structure of the ASD phenotype.

Study 1 uses a *variable-centered* approach (principal component analysis) to examine the phenotypic overlap between core diagnostic features and emotional/behavioral problems in a sample of 335 newly diagnosed preschool children with ASD (Georgiades et al., 2011).

Study 2 uses a *person-centered* approach (cluster analysis) to prospectively investigate the emergence of autistic-like traits in unaffected (no ASD diagnosis) infant siblings of probands diagnosed with ASD. Two groups of children unaffected with ASD were assessed prospectively – 170 high-risk siblings of probands diagnosed with ASD and 90 low-risk controls with no family history of ASD (Georgiades et al., 2013).

Study 3 evaluates the ability of a combined (hybrid) *variable/person-centered* approach (factor mixture modeling) to describe the underlying structure of the core symptom phenotype of ASD. Data came from a sample of 391 newly diagnosed children (ages 2 to 5) participating in the Pathways in ASD study (Georgiades et al., 2013).

Study 4 examined if the factor mixture model found to best represent the latent class structure of core autism symptoms between ages 2 and 5 (see Study 3) is replicable at age 6. The sample consisted of 280 children with ASD participating in the Pathways study and had complete data at both diagnosis (ages 2 to 5) and at age 6 (Georgiades et al., under review).

The **Editorial** (Georgiades, Szatmari, & Boyle, 2013) identifies the major limitations of previous research and offers recommendations for studying heterogeneity as a general framework (rather than as a post-hoc outcome) that could guide the development, implementation, and interpretation of new study designs and measurements capable of "capturing" individual and subgroup differences within autism.

Importance

This *Thesis* will produce comprehensive empirical evidence for our understanding of the phenotypic heterogeneity seen in ASD. The underlying objective of this work is to develop systematic ways of identifying children with ASD that share similar, relatively

homogeneous clinical profiles and to see whether those profiles are stable over time. This evidence will be generated using an integrative approach including multiple *constructs* (symptoms, traits, behaviours), *methods* (factor analysis, cluster analysis, and factor mixture modeling), *populations* (clinical and high-risk samples), and *time points* (at diagnosis and at age 6) and has the potential to inform future research on the etiology, diagnosis, treatment and prognosis of ASD. The hope is that findings from future research will in turn inform clinical practice and policies that will reduce the burden of suffering and improve the quality of life for children with ASD and their families.

References

American Psychiatric Association (2013). Diagnostic and Statistical Manual of Mental Disorders (Fifth ed.). Arlington, VA: American Psychiatric Publishing.

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

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

American Psychiatric Association. (2012). *Report of the DSM-V Neurodevelopmental Disorders Work Group*. http://www.dsm5.org Accessed, June 6, 2012.

Asperger H (1944). "Die "Autistischen Psychopathen" im Kindesalter [Autistic psychopaths in childhood]" (in German). *Archiv für Psychiatrie und Nervenkrankheiten 117*: 76–136. doi:10.1007/BF01837709

Bailey, A., Le Couteur, A., Gottesman, I., Bolton, P., Simonoff, E., Yuzda, E., & Rutter, M. (1995). Autism as a strongly genetic disorder: Evidence from a british twin study. *Psychological Medicine, 25*(1), 63-77.

Bauminger, N., Solomon, M., & Rogers, S. J. (2010). Externalizing and internalizing behaviors in ASD. *Autism Research : Official Journal of the International Society for Autism Research, 3*(3), 101-112. doi: 10.1002/aur.131; 10.1002/aur.131

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

Brereton, A. V., Tonge, B. J., & Einfeld, S. L. (2006). Psychopathology in children and adolescents with autism compared to young people with intellectual disability. *Journal of Autism and Developmental Disorders, 36*(7), 863-870. doi: 10.1007/s10803-006-0125-y

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

Cidav, Z., Marcus, S. C., & Mandell, D. S. (2012). Implications of childhood autism for parental employment and earnings. *Pediatrics, 129*(4), 617-623. doi: 10.1542/peds.2011-2700; 10.1542/peds.2011-2700

Constantino J, Lajonchere C, Lutz M, Gray T, Abbacchi A, McKenne K, Singh D, Todd R. (2006). Autistic social impairment in the siblings of children with pervasive developmental disorders. *Am J Psychiatry.*163(2):294-296.

Dawson, G. (2013). Dramatic increase in autism prevalence parallels explosion of research into its biology and causes. *JAMA Psychiatry (Chicago, Ill.), 70*(1), 9-10. doi: 10.1001/jamapsychiatry.2013.488; 10.1001/jamapsychiatry.2013.488

Elsabbagh, M., Divan, G., Koh, Y. J., Kim, Y. S., Kauchali, S., Marcin, C., . . . Fombonne, E. (2012). Global prevalence of autism and other pervasive developmental disorders. *Autism Research : Official Journal of the International Society for Autism Research, 5*(3), 160-179. doi: 10.1002/aur.239; 10.1002/aur.239

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

Fombonne, E., Quirke, S., & Hagen, A. (2009). Prevalence and interpretation of recent trends in rates of pervasive developmental disorders. *McGill Journal of Medicine : MJM : An International Forum for the Advancement of Medical Sciences by Students, 12*(2), 73.

Gadow, K. D., DeVincent, C. J., Pomeroy, J., & Azizian, A. (2004). Psychiatric symptoms in preschool children with PDD and clinic and comparison samples. *Journal of Autism and Developmental Disorders, 34*(4), 379-393.

Georgiades, S., Szatmari, P., Boyle, M. (2013) EDITORIAL: The Importance of Studying Heterogeneity in Autism, *Neuropsychiatry: Future Medicine. Vol. 3, No. 2,* 123-125 , DOI 10.2217/npy.13.8

Georgiades, S., Szatmari, P., Boyle, M., Hanna, S., Duku, E., Zwaigenbaum, L., . . . Pathways in ASD Study Team. (2013). Investigating phenotypic heterogeneity in children with autism spectrum disorder: A factor mixture modeling approach. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 54*(2), 206-215. doi: 10.1111/j.1469-7610.2012.02588.x; 10.1111/j.1469-7610.2012.02588.x

Georgiades, S., Szatmari, P., Duku, E., Zwaigenbaum, L., Bryson, S., Roberts, W., . . . Pathways in ASD Study Team. (2011). Phenotypic overlap between core diagnostic features and emotional/behavioral problems in preschool children with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 41*(10), 1321-1329. doi: 10.1007/s10803-010-1158-9; 10.1007/s10803-010-1158-9

Georgiades, S., Szatmari, P., Zwaigenbaum, L., Bryson, S., Brian, J., Roberts, W., . . . Garon, N. (2013). A prospective study of autistic-like traits in unaffected siblings of

probands with autism spectrum disorder. *JAMA Psychiatry (Chicago, Ill.), 70*(1), 42-48. doi: 10.1001/2013.jamapsychiatry.1; 10.1001/2013.jamapsychiatry.1

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

Geschwind, D. H. (2009). Advances in autism. *Annual Review of Medicine, 60*, 367-380. doi: 10.1146/annurev.med.60.053107.121225; 10.1146/annurev.med.60.053107.121225

Hallmayer, J., Cleveland, S., Torres, A., Phillips, J., Cohen, B., Torigoe, T., . . . Risch, N. (2011). Genetic heritability and shared environmental factors among twin pairs with autism. *Archives of General Psychiatry, 68*(11), 1095-1102. doi: 10.1001/archgenpsychiatry.2011.76; 10.1001/archgenpsychiatry.2011.76

Happe, F. (2011). Criteria, categories, and continua: Autism and related disorders in DSM-5. *Journal of the American Academy of Child and Adolescent Psychiatry, 50*(6), 540-542. doi: 10.1016/j.jaac.2011.03.015; 10.1016/j.jaac.2011.03.015

Howlin, P., Goode, S., Hutton, J., & Rutter, M. (2004). Adult outcome for children with autism. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 45*(2), 212-229.

Hudziak, J. J., Achenbach, T. M., Althoff, R. R., & Pine, D. S. (2007). A dimensional approach to developmental psychopathology. *International Journal of Methods in Psychiatric Research, 16 Suppl 1*, S16-23. doi: 10.1002/mpr.217

Jarbrink, K., Fombonne, E., & Knapp, M. (2003). Measuring the parental, service and cost impacts of children with autistic spectrum disorder: A pilot study. *Journal of Autism and Developmental Disorders, 33*(4), 395-402.

Kanner L (1943). "Autistic disturbances of affective contact". Nerv Child 2: 217–50. Kanner, L (1968). "Reprint". *Acta Paedopsychiatr 35* (4): 100–36.

Lichtenstein, P., Carlstrom, E., Rastam, M., Gillberg, C., & Anckarsater, H. (2010). The genetics of autism spectrum disorders and related neuropsychiatric disorders in childhood. *The American Journal of Psychiatry, 167*(11), 1357-1363. doi: 10.1176/appi.ajp.2010.10020223; 10.1176/appi.ajp.2010.10020223

Lord, C., & Jones, R. M. (2012). Annual research review: Re-thinking the classification of autism spectrum disorders. *Journal of Child Psychology and Psychiatry, and*

Allied Disciplines, 53(5), 490-509. doi: 10.1111/j.1469-7610.2012.02547.x; 10.1111/j.1469-7610.2012.02547.x

Lord, C., Risi, S., DiLavore, P. S., Shulman, C., Thurm, A., & Pickles, A. (2006). Autism from 2 to 9 years of age. *Archives of General Psychiatry, 63*(6), 694-701. doi: 10.1001/archpsyc.63.6.694

Lubke, G. H., Muthen, B., Moilanen, I. K., McGough, J. J., Loo, S. K., Swanson, J. M., Smalley, S. L. (2007). Subtypes versus severity differences in attention-deficit/hyperactivity disorder in the northern finnish birth cohort. *Journal of the American Academy of Child and Adolescent Psychiatry, 46*(12), 1584-1593. doi: 10.1097/chi.0b013e31815750dd

Mandell, D. (2011). The heterogeneity in clinical presentation among individuals on the autism spectrum is a remarkably puzzling facet of this set of disorders. *Autism : The International Journal of Research and Practice, 15*(3), 259-261.

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

Posthuma, D., & Polderman, T. J. (2013). What have we learned from recent twin studies about the etiology of neurodevelopmental disorders? *Current Opinion in Neurology, 26*(2), 111-121. doi: 10.1097/WCO.0b013e32835f19c3; 10.1097/WCO.0b013e32835f19c3

Rommelse, N. N., Franke, B., Geurts, H. M., Hartman, C. A., & Buitelaar, J. K. (2010). Shared heritability of attention-deficit/hyperactivity disorder and autism spectrum disorder. *European Child & Adolescent Psychiatry, 19*(3), 281-295. doi: 10.1007/s00787-010-0092-x; 10.1007/s00787-010-0092-x

Ruscio, J., & Ruscio, A. M. (2004). Clarifying boundary issues in psychopathology: The role of taxometrics in a comprehensive program of structural research. *Journal of Abnormal Psychology, 113*(1), 24-38. doi: 10.1037/0021-843X.113.1.24

Rutter, M. (2012). Changing concepts and findings on autism. *Journal of Autism and Developmental Disorders,* doi: 10.1007/s10803-012-1713-7

Skuse, D. H., Mandy, W. P., & Scourfield, J. (2005). Measuring autistic traits: Heritability, reliability and validity of the social and communication disorders checklist. *The British Journal of Psychiatry : The Journal of Mental Science, 187*, 568-572. doi: 10.1192/bjp.187.6.568

Steffenburg, S., Gillberg, C., Hellgren, L., Andersson, L., Gillberg, I. C., Jakobsson, G., & Bohman, M. (1989). A twin study of autism in denmark, finland, iceland, norway and sweden. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 30*(3), 405-416.

Szatmari, P. (2011). Is autism, at least in part, a disorder of fetal programming? *Archives of General Psychiatry, 68*(11), 1091-1092. doi: 10.1001/archgenpsychiatry.2011.99; 10.1001/archgenpsychiatry.2011.99

Szatmari, P., Georgiades, S., Duku, E., Zwaigenbaum, L., Goldberg, J., & Bennett, T. (2008). Alexithymia in parents of children with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 38*(10), 1859-1865. doi: 10.1007/s10803-008-0576-4; 10.1007/s10803-008-0576-4

Warren, Z., McPheeters, M. L., Sathe, N., Foss-Feig, J. H., Glasser, A., & Veenstra-Vanderweele, J. (2011). A systematic review of early intensive intervention for autism spectrum disorders. *Pediatrics, 127*(5), e1303-11. doi: 10.1542/peds.2011-0426; 10.1542/peds.2011-0426

Zwaigenbaum, L. (2012). What's in a name: changing the terminology of autism diagnosis. Developmental Medicine and Child Neurology, 54, 871-872.

CHAPTER TWO
STUDY 1

TITLE: Phenotypic Overlap between Core Diagnostic Features and Emotional/Behavioral Problems in Preschool Children with Autism Spectrum Disorder

AUTHORS: Stelios Georgiades[1], Peter Szatmari[1], Eric Duku[1], Lonnie Zwaigenbaum[2], Susan Bryson[3], Wendy Roberts[4], Eric Fombonne[5], Pat Mirenda[6], Isabel Smith[3], Tracy Vaillancourt[7], Joanne Volden[2], Charlotte Waddell[8], Ann Thompson[1] & Pathways in ASD Study Team

[1]McMaster University, [2]University of Alberta, [3]IWK Health Centre, Dalhousie University, [4]The Hospital for Sick Children, University of Toronto, [5]McGill University, [6]The University of British Columbia, [7]University of Ottawa, [8]Simon Fraser University

CONTEXT AND IMPLICATIONS OF THIS STUDY: Evidence shows that there is a substantial overlap between ASD and other neurodevelopmental and mental disorders such as attention deficit hyperactivity disorder, oppositional defiant disorder, anxiety and mood disorders. This study uses a variable-centered approach (principal component analysis) to examine the phenotypic overlap between core diagnostic features and emotional/behavioral problems seen in other disorders in a sample of 335 newly diagnosed preschool children with ASD. Findings challenge the current definition of ASD which is based only on "core symptoms" and call for a systematic investigation that will evaluate the utility of an expanded set of indicators – emotional/behavioural problems and possibly IQ – as part of a more comprehensive measurement and classification framework for ASD.

CONFLICTS OF INTEREST: None

Abstract

This study examined the phenotypic overlap between core diagnostic features and emotional/behavioral problems in a sample of 335 preschool children with autism spectrum disorder (ASD). Results from principal component analysis (2 components; 49.70 % variance explained) suggested substantial phenotypic overlap between core diagnostic features and emotional/behavioral problems. Component I, Emotional Behavioral Repetitive Problems, was independent of the children's intellectual, adaptive functioning, and structural language abilities. Component II, Social Communication Deficits, was negatively related to the children's intellectual, adaptive functioning, and structural language abilities. Both components were positively related to parental stress. This exploratory study contributes to our understanding of the ASD phenotype and provides further support for including emotional/behavioral problems as part of the clinical characterization of children with ASD.

KEY WORDS: autism spectrum disorder, emotional/behavioral problems, phenotype, principal component analysis

Phenotypic Overlap between Core Diagnostic Features and Emotional/Behavioral Problems in Preschool Children with Autism Spectrum Disorder

Autism Spectrum Disorders (ASDs; also known as Pervasive Developmental Disorders), are defined in the 4th edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) as a group of neurodevelopmental disorders characterized by symptoms of social and communication impairment and by the presence of repetitive, restricted, stereotyped behaviours. In the DSM-IV, each of these three distinct categories includes a set of symptoms rated on binary criteria (yes/no) to qualify for a diagnosis (APA, 1994; 2000). The current DSM-IV classification distinguishes among three ASD subtypes, namely Autistic Disorder (AD), Asperger's Disorder (AS), and Pervasive Developmental Disorder Not Otherwise Specified (PDD-NOS). Many investigators now believe that this conceptualization of ASD, which is based primarily on clinical expertise rather than empirical evidence, is limited in capturing the variability and complexity of clinical presentations of the disorder (Frazier et al., 2008; Georgiades et al., 2007). As a result, the *DSM 5 Neurodevelopmental Disorders Work Group* is proposing the use of symptom *dimensions* in addition to symptom *categories* for the classification of ASD (APA, 2010).

Over the past decade, a better understanding of the basic structure of the core ASD clinical phenotype has emerged from a number of factor analytic studies (Bolte et al., 2001; Frazier et al., 2008; Georgiades et al., 2007; Boomsma et al., 2008; Gotham et al., 2008; Robertson et al., 1999; Snow et al., 2009; Szatmari et al., 2002; Tadevosyan-Leyfer et al., 2003; Tanguay et al., 1998; Van Lang et al., 2006; Kamp-Becker et al., 2009). Despite their methodological differences, most studies share at least two major findings: (a) the largest amount of variance in the core ASD clinical phenotype reflects symptoms and/or behaviors related to a single combined social-communication domain; and (b) a smaller amount of additional variance is accounted for by at least one other domain reflecting repetitive, restricted, stereotyped behaviors. Whereas most researchers agree that these two domains are essential to the core ASD phenotypic structure, the optimal number and most importantly the domain content areas needed to best describe the clinical presentation of the disorder warrant further investigation.

Although the aforementioned studies are informative in many ways, they are constrained by at least two methodological limitations. First, the use of a single ASD symptom-based instrument (e.g., the Autism Diagnostic Interview-Revised or Social Responsiveness Scale) may not fully capture the wide range of symptoms and/or behaviors that affect the daily life and development of children with ASD. Second, analyses that include participants across a wide age range (i.e., children, youth, and adults) do not take into account the potential variability in phenotypic structure across developmental stages.

To further our understanding of the associations among symptom and behavior domains in young children with ASD, we examined the underlying structure of a more comprehensive clinical phenotype that, in addition to the core diagnostic ASD features, includes information related to other emotional and behavioral (i.e., internalizing and externalizing) problems observed in a representative sample of newly diagnosed preschool children with ASD. The inclusion of emotional/behavioral problems in our investigation was based on several important reasons drawn from the ASD literature. First, the topic of "comorbidity" between other child disorders and ASD has attracted increasing attention in recent years. Specifically, a number of studies have documented the presence of symptoms/behaviours that are part of other disorders such as attention deficit hyperactivity disorder (ADHD), oppositional defiant disorder (ODD), anxiety and mood disorders in individuals with ASD (e.g., Bauminger et al., 2010; Brereton et al., 2006; Gadow et al., 2004; Witwer & Lecavalier, 2008). According to Pandolfi et al. (2009), it is often difficult to distinguish such co-occurring problem behaviors from the core diagnostic features of ASD. Second, these other emotional/behavioral problems are reported to increase the stress and reduce the quality of life experienced by families of children with ASD (Bauminger et al., 2010; Estes et al., 2009) so understanding the mechanism by which such "comorbidities" arise is important. Third, little is known about the early manifestations of these emotional/behavioral problems in preschool children with ASD, as opposed to their presentation later in childhood or adolescence (Bauminger et al., 2010; Pandolfi et al., 2009).

As noted by Brereton et al. (2006), children and youth with ASD experience high levels of emotional/behavioral problems beyond their core autism symptoms. If there is significant phenotypic overlap (i.e., shared variance) between ASD symptoms and these other emotional/behavioral problems, then use of the term "comorbidity" may be somewhat misleading, at least when describing very young children with ASD. Although the term "comorbidity" was introduced by Feinstein (1970) to describe the simultaneous existence of two or more *distinct* medical disorders in the same individual, it is often used to refer to a greater-than-chance co-occurrence of two or more psychiatric disorders or a specific constellation of symptoms. Thus, it is unclear whether the concomitant observation of these sets of symptoms reflects *distinct clinical entities* or *multiple manifestations of a single clinical entity* (Maj, 2005). This important question for the ASD field is deserving of careful empirical investigation.

In a recent study of the association between autistic-like and internalizing traits in a community sample of typically developing twins, Hallett et al. (2009) concluded that there is moderate phenotypic overlap between the two sets of traits. Based on these findings from the general population, Hallett et al. (2009) hypothesized that these traits might be more strongly associated in individuals with clinical levels of autistic symptoms and noted that this question should be examined in clinical samples. To the best of our knowledge, the current study is the first to explore the underlying structural association (and thus potential phenotypic overlap) between core diagnostic features and other

18

emotional/behavioral problems in such a clinical sample of preschool children diagnosed with ASD.

Methods

Participants and Procedure

The sample consisted of 335 newly-diagnosed preschool children with ASD participating in the *Pathways in ASD,* a Canadian longitudinal study examining the developmental trajectories of children with ASD (see Table 1 for sample demographic information). With a mean age of 39.8 months (SD = 9.0 months), children met criteria for a clinical diagnosis of ASD according to the DSM-IV arrived at by a multi-disciplinary team with expertise in the diagnosis. In addition all cases also met criteria for ASD according to both the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2002) and the Autism Diagnostic Interview – Revised (ADI-R; Rutter et al., 2003) using the Risi et al. (2006) criteria. To ensure the independence of observations, only one child per family was recruited to the study. The study was approved by the local Research Ethics Boards at each site. Families willing to participate went through an informed consent session prior to joining the study.

[place Table 1 about here]

Instruments

Autism Diagnostic Interview – Revised (ADI-R; Rutter et al., 2003*).* The ADI-R is a standardized semi-structured interview used in the differential diagnosis of ASD. It is designed to be used with a parent or caregiver who is familiar with the developmental history and current behavior of individuals over the age of 2 years. The ADI-R consists of three major domains: (a) language and communication, (b) reciprocal social interaction, and (c) restricted, repetitive, and stereotyped behaviors and interests. A cut-off point for each of the three domains provides a reliable diagnostic algorithm shown to be accurate in differentiating autism from other developmental disorders (Rutter et al., 2003).

Child Behavior Checklist (CBCL 1.5-5; Achenbach & Rescorla, 2000). The CBCL/1.5-5 is a well-validated measure of externalizing and internalizing behavior problems in typically-developing preschool children as well as children with ASD (Achenbach & Rescorla, 2000; Pandolfi et al., 2009). The CBCL obtains parent/caregiver ratings of 99 problem items that are empirically clustered into 6 domains: Emotionally Reactive, Anxious/Depressed, Somatic Complaints, Withdrawn, Attention Problems and Aggressive Behavior. The CBCL provides raw total scores as well as *T*-scores based on a normative sample of children aged 1.5-5 years. A *T*-score $\geq$ 70 indicates that the child is within "clinical range", that is above the 98th percentile (Achenbach & Rescorla, 2000).

Repetitive Behavior Scale-Revised (RBS-R; Bodfish et al., 1999; 2000). This is an empirically-derived clinical rating scale for measuring the presence and severity of a variety of forms of restricted, repetitive behaviors. The RBS-R is designed to provide a quantitative, continuous measure of the spectrum of repetitive behaviors. It comprises 43 items distributed across six conceptually-derived subscales: Stereotyped Behavior, Self-injurious Behavior, Compulsive Behavior, Routine Behavior, Sameness Behavior and Restricted Behavior. This scale is completed by parent/caregiver informants. Although the RBS-R was designed for somewhat older children, the 6-factor structure was recently validated in our sample with subscale internal consistency (alpha) ranging from .71 to .88 (see Mirenda et al., 2010).

Vineland Adaptive Behavior Scales, Second Edition (VABS II; Sparrow et al., 2005). The VABS II assesses child adaptive behavior in the communication, socialization, daily living skills and motor domains, and expresses overall functioning in the Adaptive Behavior Composite (ABC) score. The VABS II is administered to a parent or caregiver using a semi-structured interview format.

Merrill-Palmer-Revised Scales of Development (M-P-R; Roid & Sampers, 2004): This is an individually-administered measure of intellectual ability that is appropriate for children aged 2 to 78 months. The Developmental Index standard score comprises cognitive, receptive language and fine motor scales.

Preschool Language Scale, 4th edition (PLS-4; Zimmerman et al., 2002). This is an individually administered test of receptive and expressive structural language ability designed for children from birth through 6 years. This is a well-researched instrument with good reliability, validity, and utility (Zimmerman et al., 2002). The Total Language standard score was used in the analysis.

Parenting Stress Index – Short Form (PSI/SF; Abidin, 1995). This is a parent self-report questionnaire designed to identify potentially stressful parent-child dyadic relationship systems. The instrument yields a Total Stress Score (used in the current study), plus scale scores for components that reflect child and parent characteristics and their interactions that pinpoint sources of stress within the family.

[place Table 2 about here]

Data Analyses

Means and standard deviations of the total raw scores and *T*-scores were calculated to describe the levels/distributions of emotional/behavioral symptoms in the current sample. The percentages of children scoring within the "clinical range" (i.e., *T*-score $\geq$ 70; $\geq$ 98th percentile; Achenbach & Rescorla, 2000) on the six CBCL domains were also calculated. To illustrate the differences between the ASD sample and the

normative sample of similar age range, effect sizes were computed based on the ratio of the difference of *T*-scores from the two samples and the standard deviation of the normative sample (Achenbach & Rescorla, 2000).

To investigate the structure of the ASD phenotype, 15 domain scores – 3 domain total scores from the ADI-R, 6 subscale *T*-scores from the CBCL, and 6 total mean scores from the RBS-R – were used as indicators in principal component analysis (see Table 4 for a list of indicators). Data were available on these indicators for all children in our sample regardless of their verbal ability. Specifically, the communication sub-domain "B(V): Verbal Total" of the ADI-R that applied only to verbal children (i.e., is language dependent) was excluded from the analysis so that no imputation of the "verbal-only" items was required. Instead, the domain "B(NV): Nonverbal Total" was included as an indicator of nonverbal communication impairment for all children. Principal Component Analysis (PCA) with a Varimax Rotation and Kaiser Normalization was conducted using the 15 indicators. The PCA extraction method was selected over Exploratory Factor Analysis (EFA) because it is intended to reduce data while keeping as much information from the original set as possible (Norris & Lecavalier, 2010). Orthogonal (i.e., Varimax) rotation was selected over oblique (i.e., non-orthogonal) rotation because we were interested in deriving components that are distinct from each other. The selected extraction and rotation methods fit well with our long-term analytic plans to use empirically derived components to model the developmental trajectories of children on relatively distinct constructs. To ensure there were no differences in the metrics of the indicators from the four instruments, the analysis was conducted using both original scores and Z-scores (i.e., standardized scores). Components were constructed by selecting loadings over .40 (Tabachnick & Fidell, 2001).

To examine the potential associations between the derived components and other variables of interest, we calculated partial correlations between the component scores and children's intellectual abilities (indexed by MP-R), adaptive functioning (indexed by VABS II), and structural language abilities (indexed by PLS-4), as well as parental stress (indexed by PSI-SF). This analysis was conducted while controlling for the effect of children's ages.

Results

Table 3 shows the means and standard deviations of the total raw scores and *T*-scores of the six CBCL problem behavior subscales for the ASD and normative samples. This table also depicts the percentage of children from the ASD sample who scored within the "clinical range" on these domains. Specifically, 39.2% of children with ASD scored in the "clinical range" on the *Withdrawn* subscale, whereas 11.7% and 7.2% scored above the clinical cut-off point on the *Attention Problems* and *Emotionally Reactive* subscales, respectively. The effect sizes for the six subscales ranged from 0.1 to 2.7. Children with ASD scored especially higher than the normative sample on

Withdrawn (effect size 2.7), *Attention Problems* (effect size 1.1) and *Emotionally Reactive* (effect size 0.8).

[place Table 3 about here]

PCA results using original indicator scores were no different from those employing standardized scores. Therefore, original scores were used in all subsequent analyses. The following considerations were used to select the most appropriate component solution: (a) Eigenvalues; (b) scree plots; (c) percentage of variance explained; (d) minimum number of item cross-loadings; and (e) clinical interpretability of the derived components. Three Eigenvalues were larger than 1.0; thus a 3-component solution was first examined then rejected because of multiple cross-loadings on two of the components and lack of conceptual interpretability. Results indicated that the structure of the ASD phenotype in this preschool sample was optimally modeled by two distinct components (Table 4). The 2-component solution explained 49.70% of the variance and consisted of the following components: Component I, labeled Emotional, Behavioral, Repetitive Problems (EBRP), included the ADI-R repetitive behaviours indicator, all the RBS-R indicators, as well as the CBCL indicators with the exception of the Withdrawn subscale. Component II, labeled Social Communication Deficits (SCD), included the ADI-R indicators of social reciprocity and nonverbal communication impairment as well as the CBCL Withdrawn subscale. It should be noted that two indicators, the CBCL Withdrawn and Attention subscales, loaded relatively high on both components. Finally, as expected due to the orthogonal rotation method, the two components were independent of each other (p > .05).

[place Table 4 about here]

The partial correlations between the two derived components and other variables of interest, controlling for the effect of child age (i.e., 2- vs. 3- vs. 4-year-olds), are reported in Table 5. The EBRP component was not significantly correlated with children's intellectual, adaptive functioning or structural language abilities. In contrast, the SCD component was negatively related to intellectual, adaptive functioning and structural language abilities. Both components were positively related to parental stress.

[place Table 5 about here]

Discussion

The present study explored the phenotypic overlap (i.e., structural associations) between core diagnostic features and other emotional/behavioral problems in preschool children with ASD. At least three of the study findings warrant discussion and further investigation.

First, even at this young age, children with ASD present with increased levels of emotional/behavioral symptoms. Not surprisingly, children with ASD have elevated scores on the *Withdrawn, Attention Problems,* and *Emotionally Reactive* domains compared to the norms. This finding is in line with those from previous studies (Garon et al., 2009; Sikora et al., 2008). The presence of elevated scores on the CBCL at this young age highlights the need to identify and treat these challenging behaviours early, given evidence of the stability and intractability of these problems (e.g., Richter et al., 2010). Previous research suggests that one of the best ways to support families with children with ASD is to address these problems early by teaching parents how to manage them (e.g., Koegel & Koegel, 2006).

Second, the underlying structure of this expanded ASD clinical phenotype can be described using two independent components/domains, Emotional Behavioral Repetitive Problems (EBRP) and Social Communication Deficits (SCD). The SCD component identified in this study is similar in content to the social-communication component reported in numerous previous studies (Kamp-Becker et al., 2008; Snow et al., 2009), including a study by our group (Georgiades et al., 2007). This robust finding of a joint social-communication domain, which appears to hold even in this very young more age-homogeneous sample, has already been incorporated in the proposed revisions of the ASD section of the DSM 5 (APA, 2010). The EBRP component identified in the present study includes content from the fixated interests and repetitive behaviours (FIRB) component proposed in the revisions of the DSM 5 committee (APA, 2010). However, our EBRP component extends the traditional construct of repetitive behaviors by including other problem behaviours frequently seen in children with ASD such as anxiety, emotional reactivity, somatic complaints, attention problems and aggression. It is worth noting that the CBCL Withdrawn and Attention subscales loaded relatively high on both components. Interestingly, these are the same two indicators on which children with ASD have significantly elevated scores compared to the norms (see Table 3). Thus, one could hypothesize that these core deficits known to be elevated in children with ASD (Dawson et al., 2004) represent a common underlying theme between the two main phenotypic domains of ASD.

These findings suggest that, at least in very young children with ASD, emotional/behavioral problems should not be described as "comorbid" distinct clinical entities in relation to ASD symptoms. Rather, they can be conceptualized as multiple manifestations of a single clinical entity (Maj, 2005), in this case ASD. However, it is difficult to determine from cross-sectional data whether emotional/behavioral problems are truly "comorbid" or "part of" the ASD phenotype. This is especially true within this young age range, when these problems may represent relatively isolated symptoms (commonly seen in young children) rather than full-blown disorders. Accumulating longitudinal data from our study will allow us to examine whether these problems indeed co-vary over time with FIRB and/or SCD, that is, whether they influence one another or have independent trajectories. Substantial co-variation between these domains over time

would raise the possibility of a common set of risk factors. Evidence for such a possibility would challenge the widely held assumption that comorbid emotional/behavioral problems represent distinct content domains over and above the core features of the ASD clinical phenotype.

Third, there are differential associations between the two phenotypic components and other variables of interest. Specifically, the SCD component is inversely associated with children's abilities in the areas of intellect, structural language and adaptive functioning. This finding is in line with those reported in previous studies (Georgiades et al., 2007; Snow et al., 2009) and indicates that in children with ASD lower intellectual, language, and adaptive functioning abilities tend to accompany more social communication deficits. On the other hand, EBRPs appear to be independent of these variables. In general, these differential associations might be due to different etiological mechanisms responsible for each of the two ASD domains. If that is the case, this finding would provide support to the idea that no single unitary account can explain both social and non-social features of autism (Happe & Ronald, 2008). Finally, our finding that both components are associated with parental stress is noteworthy and speaks to the importance of early interventions targeting both core areas of children's symptoms, with the potential of reducing parental stress. As noted by Bauminger et al. (2010), the documented link between parental stress and children's emotional/behavioral problems is crucial and can inform intervention and support programs for parents of children with ASD.

Several limitations call for careful interpretation of the study findings. First, we acknowledge that the indicators used in the initial component analysis do not represent the full scope of the ASD-related phenotype. For example, it could be argued that adaptive functioning, language and intellectual ability are constructs that should be included as "parts of" rather than as "correlates" of the ASD phenotype. We believe that this is a possibility worth investigating further. However, for conceptual as well as practical reasons, we focused our current component analysis on symptoms and problem behaviors rather than on standardized assessments of abilities. Second, our study design (i.e., ASD sample) does not allow us to test whether the two domains (i.e., EBRP & SCD) distinguish ASD from other disorders (e.g., ADHD). Third, since our current data are limited to one time-point, we are unable to examine the temporal stability of the structural association between SCD and EBRP. Fourth, the cross-sectional nature of the data does not allow for "cause-and-effect" examinations regarding SCD and EBRP and the other variables (intellectual and structural language ability, adaptive functioning and parental stress). Fifth, data used in the component analysis were collected using parent report (i.e., ADI-R, CBCL & RBS-R), albeit with different methods (i.e., interview and rating scales), and thus are subject to biases associated with using a single informant. Although ADOS data were available, we chose not to include these in the principal component analysis due to the complex issue created by the administration of different, metrically incompatible modules in our sample. Furthermore, the ADOS calibrated

severity metric created by Gotham et al. (2009) would not be appropriate because it is a "total" score and does not allow for the investigation of the structural associations between different domains of ASD symptoms (i.e., SCD & FIRB), a main focus of this paper.

In conclusion, this exploratory study demonstrates that there is substantial phenotypic overlap between core diagnostic features and emotional/behavioral problems in young children with ASD. Study findings add to the existing ASD literature by emphasizing the importance of assessing general emotional/behavioral problems in conjunction with the core diagnostic symptoms in preschool children with ASD (Bauminger et al., 2010; Duarte et al., 2003; Sikora et al., 2008). Our findings are also consistent with the proposed DSM 5 revisions that recommend the use of two domains (SCD and FIRB) to characterize children with ASD (APA, 2010). Furthermore, these findings raise new questions about the potential mechanisms underlying the relationship between repetitive behaviors and emotional/behavioral problems in ASD. Therefore, it is important for future research to examine the developmental course and structural associations of these emotional/behavioral problems in relation to core ASD symptoms. Such an approach would be useful in illuminating the complex and, in our view, under-examined issue of "comorbidity" in individuals with ASD.

References

Abidin, R. (1995). *Parenting Stress Index, 3rd ed.* Odessa, Florida: Psychological Assessment Resources.

Achenbach, T. M., & Rescorla, L. A. (2000). *Manual for ASEBA Preschool Forms & Profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.

American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV).* Washington, DC: American Psychiatric Association.

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

American Psychiatric Association. Report of the DSM 5 Neurodevelopmental Disorders Work Group. http://www.dsm5.org. Accessed July 18, 2010.

Bauminger N, Solomon M, & Rogers S. J. (2010). Externalizing and internalizing behaviors in ASD. *Autism Research. 26*, 101-112.

Bodfish, J. W., Symons, F. J., Parker, D. E., & Lewis, M. H. (2000). Varieties in repetitive behavior in autism. *Journal of Autism and Developmental Disorders. 30*, 237–243.

Bodfish, J. W., Symons, F., & Lewis, M. (1999). *The Repetitive Behavior Scale: Test manual.* Morganton: Western Carolina Center Research Reports.

Bölte, S., Poustka, F. (2001). Factor structure of the Autism Diagnostic Interview-Revised (ADI-R): a study of dimensional versus categorical classification of autistic disorders. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie*. 29,221-9.

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

Brereton, A., Tonge, B., Einfeld, S. (2006). Psychopathology in children and adolescents with autism compared to young people with intellectual disability. *Journal of Autism Developmental Disorders.* 36, 863-870.

Dawson, G., Toth, K. , Abbott, R. , Osterling, J. , Munson, J. and Estes, A. (2004) Early social attention impairments in autism: Social orienting, joint attention, and attention to distress. *Developmental Psychology, 40*, 271-283

Duarte, C. S., Bordin, I. A. S., de Oliveria, A., & Bird, H. (2003). The CBCL and the identification of children with autism and related conditions in Brazil: Pilot findings. *Journal of Autism and Developmental Disorders, 33(6),* 703–707.

Estes, A., Munson, J., Dawson, G., Koehler, E., Zhou, X., & Abbott, R. (2009). Parenting stress and psychological functioning among mothers of preschool children with autism and developmental delay. *Autism.* 13, 375-387.

Feinstein, A.R. (1970). The pre-therapeutic classification of co-morbidity in chronic disease. *Journal of Chronic Disease.* 23, 455-468.

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.

Gadow, K., DeVincent, C., Pomeroy, J., & Azizian, A. (2004). Psychiatric symptoms in preschool children with PDD and clinic and comparison samples. *Journal of Autism and Developmental Disorders*. 34, 379–393.

Garon, N., Bryson, S., Zwaigenbaum, L., Smith, I. Brian, J. Roberts, W., Szatmari, P. (2009). Temperament and its relationship to autistic symptoms in a high-risk infant sib cohort. *Journal of Abnormal Child Psychology, 37*, 59-78.

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

Gotham, K., Pickles, A., & Lord, C., (2009). Standardizing ADOS scores for a measure of severity in autism spectrum disorders. *Journal of Autism and Developmental Disorders.* 39, 693-705.

Gotham, K., Risi, S., Dawson, G., Tager-Flusberg, H., Joseph, R., Carter, A., et al. (2008). A replication of the autism diagnostic observation schedule (ADOS) revised algorithms. *Journal of the American Academy of Child & Adolescent Psychiatry, 47*,642-651.

Hallett, V., Ronald, A., & Happé, F. (2009). Investigating the association between autistic-like and internalizing traits in a community-based twin sample. *Journal of the American Academy of Child and Adolescent Psychiatry,* 48, 618- 627.

Happé, F. & Ronald, A. (2008). 'Fractionable autism triad': A review of evidence from behavioural, genetic, cognitive and neural research. *Neuropsychology Review, 18*, 287-304.

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

Koegel, R. L. & Koegel, L. K. (2006). *Pivotal Response Treatments for Autism. Baltimore*, MD: Paul H. Brookes.

Lord, C., Rutter, M., DiLavore, P. C., & Risi, S. (2002). *ADOS: The Autism Diagnostic Observation Schedule*. Los Angeles, CA: Western Psychological Services.

Maj, M. (2005). Psychiatric comorbidity: an artifact of current diagnostic systems? *British Journal of Psychiatry*, *186*, 182- 184.

Mirenda, P., Smith, I., Vaillancourt, T., Georgiades, S., Duku, E., et al. (2010). Validating the Repetitive Behavior Scale-Revised in young children with autism spectrum disorder. *Journal of Autism and Developmental Disorders. (Epub ahead of print).*

Norris, M., & Lecavalier, L. (2009). Evaluating the use of exploratory factor analysis in developmental disability psychological research. *Journal of Autism and Developmental Disorders, 40*, 8-20.

Pandolfi, V., Magyar, C. I., & Dill, C. A. (2009). Confirmatory factor analysis of the Child Behavior Checklist 1.5–5 in a sample of children with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *39*, 986–995.

Richler, J., Huerta, M., Bishop, S.L., Lord, C. (2010) Developmental trajectories of restricted and repetitive behaviors and interests in children with autism spectrum disorders. *Developmental Psychopathology, 22*, 55-69.

Risi, S., Lord, C., Gotham, K., Corsello, C., Chrysler, C., Szatmari, P., Cook, E., Leventhal, B., & Pickles, A. (2006). Combining information from multiple sources in the diagnosis of autism spectrum disorders. *Journal of the American Academy of Child and Adolescent Psychiatry, 45*, 1094–1103.

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

Roid, G., & Sampers, J. (2004). *Merrill-Palmer-Revised Scales of Development*. Wood Dale, IL: Stoelting Co.

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

Sikora, D. M., Hall, T. A., Hartley, S. L., Gerrard-Morris, A. E., & Cagle, S. (2008). Does parent report of behavior differ across ADOS-G classifications: Analysis of scores from the CBCL and GARS. *Journal of Autism and Developmental Disorders, 38*, 440–448.

Snow, A.V., Lecavalier, L.A., & Houts, C. (2008). The structure of the Autism Diagnostic Interview-Revised: Diagnostic and phenotypic implications. *Journal of Child Psychology and Psychiatry, 50*, 734-742.

Sparrow, S. S., Cicchetti, D. V., & Balla, D. A. (2005). *Vineland Adaptive Behavior Scales: Second Edition (Vineland II), Survey Interview Form/Caregiver Rating Form*. Livonia, MN: Pearson Assessments.

Szatmari, P., Merette, C., Bryson, S., Thivierge, J., Roy, M-A., Cayer, M., et al. (2002). Quantifying dimensions in autism: A factor analytic study. *Journal of the American Academy of Child & Adolescent Psychiatry, 41*, 467-474.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics (4th ed.).* New York, NY: Harper Collins College Publishers.

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

Tanguay, P. E., Robertson. J., Derrick, A. (1998). A dimensional classification of autism spectrum disorder by social communication domains. *Journal of the American Academy of Child & Adolescent Psychiatry, 37*, 271-277.

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

Witwer, A. N., & Lecavalier, L. (2008). Examining the validity of autism spectrum disorder subtypes. *Journal of Autism and Developmental Disorders*, *38*, 1611–1624.

Zimmerman, I., Steiner, V., & Pond, R.E. (2002). *Preschool Language Scale, Fourth Edition (PLS-4)*. San Antonio, TX: The Psychological Corporation.

Table 1. Demographic information for sample (N=335)

Characteristic	N	%	Mean	SD
Sex				
Male	284	84.8		
Female	51	15.2		
Child's age at diagnosis (months)			39.8	9.0
Child's age group				
2-year-olds	123	36.7		
3-year-olds	143	42.7		
4-year-olds	69	20.6		
Ethnicity				
Caucasian	240	71.6		
Other	95	28.4		

Table 2. Descriptive statistics for sample (N=335)

Instrument scale/subscale scores	Mean	SD	Min.	Max.
ADI-R Social domain score	17.2	5.1	3	29
ADI-R Communication domain nonverbal score	11.9	3.3	0	21
ADI-R Behaviours domain score	5.0	2.2	0	12
RBS-R Stereotyped behavior mean score	5.0	3.3	0	15
RBS-R Self-injurious behavior mean score	2.1	2.9	0	19
RBS-R Compulsive behavior mean score	4.0	3.6	0	18
RBS-R Ritualistic behavior mean score	3.6	3.5	0	18
RBS-R Sameness behavior mean score	5.8	5.6	0	29
RBS-R Restricted behavior mean score	3.9	3.1	0	12
MP-R Developmental Index standard score	57.2	25.4	10	136
VABS II Adaptive Behavior Composite score	73.0	10.1	52	101
PLS-4 Total Language standard score	65.3	18.9	50	136

Notes. ADI-R: Autism Diagnostic Interview-Revised; RBS-R: Repetitive Behavior Scales-Revised; MP-R: Merrill Palmer-Revised Scales of Development Developmental Index standard score; VABS II: Vineland Adaptive Behavior Scales Second edition Adaptive Behavior Composite score; PLS-4: Preschool Language Scale, 4th edition Total Language standard score.

Table 3. Mean and standard deviations of total raw scores and *T*-scores, and percentage of children scoring in the clinical range in the current ASD sample and the CBCL 1.5-5 normative sample (Achenbach & Rescorla 2000)

	CBCL Normative Sample (N=700)				Pathways ASD Sample (N=335)					Effect size[2]
	Total raw scores		*T*-scores		Total raw scores		*T*-scores			
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	% within clinical range[1]	
Emotionally Reactive	2.4	2.2	54.0	5.7	4.0	3.3	58.3	9.0	**7.2**	**0.8**
Anxious/Depressed	2.9	2.3	54.2	5.7	3.1	2.6	55.0	6.7	1.5	0.1
Somatic Complaints	1.8	1.9	54.0	5.8	2.6	2.4	56.3	7.3	5.4	0.4
Withdrawn	1.5	1.7	54.1	5.8	6.0	2.9	69.6	9.4	**39.2**	**2.7**
Attention Problems	2.5	1.9	54.1	5.6	4.5	2.3	60.1	8.3	**11.7**	**1.1**
Aggressive Behavior	10.4	6.9	54.2	6.0	13.2	7.3	56.9	8.6	7.8	0.4

[1] *T*-score $\geq$ 70 (98th percentile; Achenbach & Rescorla, 2000); [2]Effect sizes were computed based on the ratio of the difference of *T*-scores from the two samples and the standard deviation of the normative sample.

Table 4. Two-component solution for the structure of the phenotype in a sample of 335 preschool children with Autism Spectrum Disorder

Indicators	Component	
	EBRP	SCD
ADI-R Social domain score	0.03	**0.84**
ADI-R Communication domain nonverbal score	-0.02	**0.69**
ADI-R Behaviours domain score	**0.42**	0.22
RBS-R Stereotyped behavior mean score	**0.63**	0.27
RBS-R Self-injurious behavior mean score	**0.56**	0.18
RBS-R Compulsive behavior mean score	**0.75**	0.05
RBS-R Ritualistic behavior mean score	**0.76**	0.02
RBS-R Sameness behavior mean score	**0.82**	0.08
RBS-R Restricted behavior mean score	**0.72**	0.14
CBCL Emotionally Reactive T-score	**0.75**	0.18
CBCL. Anxious/Depressed T-score	**0.63**	0.00
CBCL Somatic Complaints T-score	**0.69**	0.09
CBCL Withdrawn T-score	0.38	**0.60**
CBCL Attention Problems T-score	**0.45**	0.37
CBCL Aggressive Behavior T-score	**0.71**	0.15

Notes. Variance explained: 49.70%; Extraction Method: Principal Components; Rotation Method: Varimax with Kaiser Normalization. EBRP: Emotional Behavioral Repetitive Problems, SCD: Social Communication Deficits. Item loadings over .40 are in bold font.

Table 5. Partial correlations between EBRP and SCD component scores and other variables of interest, controlling for the effect of age (2- vs. 3- vs. 4-year-olds).

	Intellectual Ability (M-P-R)	Adaptive functioning (VABS II)	Structural language (PLS-4)	Parental stress (PSI-SF)
EBRP component score	-.01	-.11	.05	**.50****
SCD component score	**-.18****	**-.35****	**-.17***	**.38****

Notes. EBRP: Emotional Behavioral Repetitive Problems, SCD: Social Communication Deficits; MP-R: Merrill Palmer-Revised Scales of Development Developmental Index standard score; VABS II: Vineland Adaptive Behavior Scales Second edition Adaptive Behavior Composite score; PLS-4: Preschool Language Scale, 4th edition Total Language standard score; PSI-SF: Parenting Stress Index – Short Form Total Stress score; * p < .05; ** p < .01

CHAPTER THREE
STUDY 2

TITLE: A Prospective Study of Autistic-like Traits in Unaffected Siblings of Probands with Autism Spectrum Disorder

AUTHORS: Stelios Georgiades[1], M.A., Peter Szatmari[1], M.D., Lonnie Zwaigenbaum[2], M.D, Susan Bryson[3], Ph.D., Jessica Brian[4,5], Ph.D., Wendy Roberts[4,5], M.D., Isabel Smith[3], Ph.D., Tracy Vaillancourt[6], Ph.D., Caroline Roncadin[7], Ph.D., Nancy Garon[8], Ph.D.
[1]Offord Centre for Child Studies, McMaster University, [2]University of Alberta, [3]Dalhousie University/IWK Health Centre, [4]Bloorview Research Institute, [5]University of Toronto, [6]University of Ottawa, [7]Peel Children's Centre, [8]Mount Allison University.

CONTEXT AND IMPLICATIONS OF THIS STUDY: Evidence shows that autistic traits are widely distributed in the general population and that the precise boundaries of the autism diagnosis are currently unclear. Moreover, numerous family studies have provided evidence for the existence of what is known as the "Broader Autism Phenotype" (BAP), a milder (sub-threshold) manifestation of autistic-like traits and characteristics in some relatives of individuals with autism but without serious degrees of impairment. This study investigated the emergence of autistic-like traits in unaffected (no ASD diagnosis) infant siblings of probands diagnosed with ASD. Results showed that within the unaffected siblings there is a subgroup of children (19%) who exhibit autistic-like traits resembling a "broader autism phenotype" at the very young age of 12 months. These children go on to have elevated scores on social-communication impairment and internalizing problems by age 3. This study highlights the importance of close monitoring of later born infants in high-risk families with another child diagnosed with ASD.

CONFLICTS OF INTEREST: None

Abstract

Context. The presence of autistic-like traits in relatives of individuals with Autism Spectrum Disorder (ASD) is well-recognized, but the emergence of these traits early in development has not been studied. **Objective.** To prospectively investigate the emergence of autistic-like traits in unaffected (no ASD diagnosis) infant siblings of probands diagnosed with ASD. **Design.** Two groups of children unaffected with ASD were assessed prospectively – siblings of probands diagnosed with ASD (high risk; HR) and controls with no family history of ASD (low risk; LR). Scores on a measure of autistic-like traits at 12 months of age were used in a cluster analysis of the entire sample. **Setting.** A prospective study of infant siblings of ASD probands from three diagnostic centers in Canada. **Participants.** 170 HR and 90 LR children none of whom was diagnosed with ASD at age 3. **Main Measures.** The Autism Observation Scale for Infants was used to measure autistic-like traits and derive clusters at 12 months of age. Clusters were compared on ASD symptoms, cognitive abilities, and social-emotional difficulties at age 3. **Results.** Two clusters were identified. Cluster 1 (n=37; 14.2% of total sample) had significantly higher levels of autistic-like traits compared to Cluster 2. Within Cluster 1, 33 children came from the siblings (19.4% of HR group); only 4 came from the controls (4.5% of LR group). At age 3, children from Cluster 1 had more social-communication impairment (effect size ES>0.70; $p<0.001$), lower cognitive abilities (ES=-0.59; $p<0.005$), and more internalizing problems (ES=0.55; $p=0.01$). Compared to controls, HR siblings had a relative risk of 4.3 (95 % CI:1.6–11.9) for membership in Cluster 1. **Conclusions.** Study findings suggest the emergence of autistic-like traits resembling a "broader autism phenotype" by 12 months of age in approximately 19% of HR siblings who do not meet ASD diagnostic criteria at age 3.

A Prospective Study of Autistic-like Traits in Unaffected Siblings of Probands with Autism Spectrum Disorder

Introduction

There is now substantial evidence that autism and the related autism spectrum disorders (ASDs) aggregate within families and that the mechanism of that familial aggregation is likely influenced by both genetic and shared environmental factors.[1-5] Given this high recurrence risk of ASD within families, investigators have turned their attention to the examination of autistic-like traits and/or characteristics evident in relatives of probands with ASD who themselves do not meet criteria for an ASD diagnosis. This effort has been accelerated by recent twin studies demonstrating that the variation in autistic-like traits in the general population shares genetic variance with those at the extreme end of the distribution (i.e., those who receive a diagnosis of ASD).[6] Family studies have provided evidence that relatives of individuals diagnosed with ASD demonstrate mild symptoms of ASD but without clinically significant impairment.[7-11] These milder manifestations of ASD have been termed the "Broader Autism Phenotype" (BAP) and are believed to be associated with social, communication and cognitive deficits, restricted, rigid behavior patterns, as well as with certain personality characteristics and psychiatric difficulties.[7-18]

Most studies of the BAP focus on parents of children with ASD, although there is a growing literature examining traits related to the BAP in siblings of probands with ASD.[12, 19-22] For example, Constantino et al. reported that roughly 20% of siblings from families with more than one case of ASD showed elevated scores on the Social Responsiveness Scale, a measure of social impairment. This was not found in families with a single affected child. These traits were surprisingly common in female siblings, narrowing the sex ratio from that usually seen in ASD.[23] Although informative in many ways, previous studies may be limited by participation bias since they were conducted in the context of genetic research. In such studies, families with more autistic-like traits in parents or siblings might tend to enroll more commonly than families without such characteristics. In addition to ascertainment bias, previous studies may also be limited by measurement issues. In studies that rely on retrospective or cross-sectional data from a parent, the rater may be especially attuned to, and therefore report more, autistic-like traits in a sibling *after* receiving the diagnosis of ASD for the proband. Alternatively, a parent may deny traits in a second child to experience a sense of reassurance about that child's development. In either situation, this bias may not occur in controls, thus leading to differential misclassification[10] and a bias in the estimation of sibling risk.[5] Furthermore, retrospective reports do not easily address the possibility that a child may have exhibited aspects of ASD in infancy but then "lost" those symptoms/traits with development. For example, individuals who exhibit the BAP as adolescents or adults may in fact have had more severe symptoms in early childhood, meeting criteria for ASD earlier but

subsequently experienced a resolution or reduction of symptoms due either to development or to intervention.

The "Baby Sibs" research paradigm has generated new knowledge about the early signs and symptoms of ASD and may be able to address many of these methodological issues.[24] In this high-risk research design, families of children with a confirmed diagnosis of ASD (i.e., probands) are enrolled soon after the delivery of a new baby. That child (i.e., the "baby sib") is followed longitudinally with regular evaluations of social-emotional reciprocity, communication and play, and cognitive skills. Sibs who are subsequently diagnosed with ASD (usually by 36 months of age) can be compared to high-risk unaffected sibs and a low-risk comparison group (controls) to identify the nature and developmental course of the earliest signs of autism using measures independent of a parent.[25] The combination of prospective data collection and a blinded objective assessment provides an important justification for the use of this design to study autistic-like traits in unaffected siblings of children with ASD.

In general, baby sib studies have demonstrated that the onset of autism-specific symptoms is usually evident *after* 6 months of age, up to which point sibs who later develop ASD are either largely indistinguishable from controls, or exhibit non-specific delays (for example, in motor control).[26,27] Cardinal symptoms of ASD (e.g., reduced social communication) emerge by 12 to 18 months.[28] For example, our group reported that several behavioral markers at 12 months (as assessed using the Autism Observation Scale for Infants; AOSI[29]) were predictive of subsequent ASD classification in a cohort of 65 high-risk infants. These markers included atypical/reduced eye contact, social smiling, orienting to name, social interest, and reactivity, as well as the presence of atypical sensory-oriented behaviors such as intense visual inspection of play materials. Total scores were strongly predictive of ASD.[30] Bryson et al. reported the detailed case histories of the first 9 children diagnosed with ASD from within our high risk cohort.[26] Although there was substantial clinical heterogeneity among this group, the emergence of atypical social-communicative and play behaviors around the first birthday was a consistent finding.[26] Similar behavioral differences as identified on the AOSI have been independently observed at 12-14 months in high-risk infants subsequently diagnosed with ASD in other longitudinal cohorts.[22,28] Another important finding has been that the recurrence risk for ASD in these high-risk cohorts has been higher than anticipated, approximately 19% rather than the 5-9% reported in older studies that employed retrospective designs.[31] This may reflect in part a broader conception of the term "ASD" since most early studies only counted the more narrowly defined "autism" (as defined by earlier versions of the Diagnostic and Statistical Manual; DSM).

The 'Baby Sibs' design provides an ideal opportunity to study the emergence of autistic-like traits in unaffected siblings of probands with ASD. Here, using prospective data, the emergence of such traits in baby sibs who do not go on to develop full-blown ASD by age 3 can be studied. In a cross-sectional study, Stone et al.[21] reported that

younger siblings of children with ASD (mean age = 16 months) demonstrated lower scores across measures of social-communicative development, cognitive functioning, and higher scores on autistic symptoms relative to controls. According to Stone et al.[21], the "weaker performance" of the ASD sibling group may represent early-emerging features of the BAP, thus highlighting the importance of developmental surveillance for younger siblings.[21] However, no study, to our knowledge, has evaluated prospectively the emergence of autistic-like traits in infant siblings focusing specifically on those who do *not* develop ASD. It is possible that the group referred to by Stone et al.[21] as displaying a weaker performance is largely accounted for by those who develop the disorder.

The current study employs a "high-risk versus control" design to investigate prospectively the occurrence of autistic-like traits among high-risk (HR) unaffected baby sibs (i.e., those not diagnosed with ASD at age 3) and a low-risk (LR) control group of children with no family history of ASD. Our main hypothesis was that, as early as 12 months, elevated levels of autistic-like traits would be more prevalent among HR siblings than LR controls. Furthermore, we hypothesized that these elevated autistic-like traits (identified in a subgroup of HR siblings) at 12 months would be associated with ASD-related outcomes at age 3 (i.e., 24 months later).

Methods

Study participants

Participants were recruited from a prospective study of early development in ASD from three multidisciplinary autism diagnostic and treatment centers in Canada, including McMaster Children's Hospital in Hamilton, The Hospital for Sick Children in Toronto, and the IWK Health Centre in Halifax, and from clinicians in the surrounding regions. The study was approved by the Research Ethics Boards at the three participating centers and all parents gave written informed consent for their children to participate. A total of 222 high-risk (HR) infants (younger siblings of children with a confirmed diagnosis of ASD) and 91 low-risk (LR) infants (no family history of ASD) were followed to age 3 years. All participants were born at 36-42 weeks gestation, had a birth weight greater than 2500 g, and had no identifiable genetic or neurological disorders.

Clinical Best Estimate (CBE) diagnosis. At 3 years of age, an independent evaluation of each participant was conducted by an expert clinician blind to assessments from previous study visits. ASD diagnoses were assigned using DSM-IV-TR criteria, based on the best judgment of the clinician (all with at least 10 years of diagnostic experience), taking into account current information from the Autism Diagnostic Interview-Revised (ADI-R) and Autism Diagnostic Observation Schedule (ADOS) and concurrent assessment of cognitive, language and adaptive skills. The 3-year ADOS was completed blind to both prior visits and risk status for every participant. Parents were asked not to reveal prior diagnoses in the younger sibling or whether they had a child with

ASD. In most cases, the ADI-R was also completed by the clinician responsible for the blind 3-year assessment, although in some cases, the ADI-R was completed by an experienced research assistant who was not blind to the child's risk status. Even in those cases, the clinician reviewed the ADI-R blind to prior visits and risk status when establishing a best-estimate clinical diagnosis. Some children with a CBE diagnosis of ASD had sub-threshold algorithm scores on the ADI-R or ADOS, but met DSM-IV-TR criteria based on expert review of all available data. This approach is consistent with current best practice, informed by a solid evidence base indicating that both a structured diagnostic interview (such as the ADI-R) and interactive assessment (such as the ADOS) are essential to ASD diagnoses, but neither, on their own or in combination, are an adequate proxy for clinical judgment, particularly in this age group.[32,33]

For the purposes of this study, we excluded the 52 high-risk infants and 1 low-risk infant who were diagnosed with ASD at age 3 years (based on the ADI-R, ADOS and expert clinical assessment using DSM-IV-TR, blind to prior study assessments). Thus, the study sample consisted of 170 children from the HR sibs group (84 males; 49.4%) and 90 from the LR control group (45 males; 50%), for a total of 260 children, none of whom were diagnosed with ASD at 3 years based on CBE (see Supplementary eTable 1 for sample characteristics).

Assessment Instruments

ASD Symptoms/Traits
Autism Observation Scale for Infants– AOSI.[29] The AOSI is a semi-structured direct observational measure developed within the context of our longitudinal study to identify behavioural markers of autism in infancy. Standardized activities are used to allow the examiner to observe and systematically rate the occurrence/non-occurrence as well as the quality or severity of behaviors deemed to be informative of the earliest emergence of ASD. The assessment is designed to take 15-20 minutes, although administration times vary depending on the infant's ability to engage with the examiner, as well as his/her temperament, state, and developmental level. Behaviours are rated on a scale from 0 to 2 or 3, where 0 implies typical function and higher values indicate increasing deviation. Behavioural markers rated on the AOSI include: visual tracking and attentional disengagement, coordination of eye gaze and action, imitation, affective responses, early social-communicative behaviors, behavioral reactivity, and sensory-motor development. Ratings on the 16 items are summed to generate a total score. Further details on AOSI items and scoring are available in Bryson et al.[29]. The AOSI has excellent inter-rater reliability at 12 and 18 months (.94 for total score), good test-retest reliability at 12 months (.61 for total score),[29] and good predictive validity at 12 and 18 months for its original 16 items.[30] One of the co-authors of this paper is co-first author on a study involving a different sample investigating the relationship between AOSI scores and ASD outcomes. In the context of that study, a random sample of 54 AOSI assessments were further assessed through video coding by examiners blind to risk status

of participants to assess for rater bias. There were no differences between blind and unblind AOSI total scores, regardless of risk status (Dr. AM Estes, written communication, May 2012). In the current study the total AOSI score was used as an indicator of autistic-like traits in the cluster analysis.

Autism Diagnostic Observation Schedule – ADOS.[34] The ADOS uses standardized activities and 'presses' to elicit communication, social interaction, imaginative use of play materials, and repetitive behaviors, allowing the examiner to observe the occurrence/non-occurrence and severity of behaviors important to the diagnosis of ASD. Subscale scores for communication and social domains are based on a subset of items previously identified to best discriminate autism/ASD from other developmental disabilities, and summed to generate an overall diagnostic algorithm score. The ADOS consists of four modules, each of which is appropriate for individuals with differing language levels. A calibrated total severity metric that accounts for differences in age and module was used in the present study.[35]

Autism Diagnostic Interview-Revised – ADI-R.[36] The ADI-R is an investigator-directed interview used to elicit information about social development, verbal and non-verbal communication skills and repetitive, stereotyped interests and behaviors required to make a diagnosis of autism. The questions are designed to distinguish qualitative impairments from developmental delays, identify behaviors that would be considered deviant at any age, and examine current and most abnormal behaviors for those strongly influenced by maturational age. The ADI-R discriminates well between autism/ASD and other forms of developmental disability, and inter-rater reliability is excellent. For this study, we used domain scores for social impairment, communication skills and repetitive, stereotyped interests and behaviours.

Cognitive Skills and Abilities

The *Mullen Scales of Early Learning– MSEL.*[37] The MSEL consists of four scales: Visual Reception, Receptive Language, Expressive Language and Fine Motor (a fifth Gross Motor scale is only administered with children younger than 30 months). An Early Learning Composite (ELC) can be calculated based on scores from these four scales for children aged 0-69 months. Inter-rater and test-retest reliability are excellent.

Emotional and Behavioural Symptoms

Infant Toddler Social Emotional Assessment – ITSEA.[38] The ITSEA is a parent-report instrument, with subscales covering attachment status, task-mastery, emotion regulation, and coping behaviors. The acceptability, internal consistency, test-retest reliability, and validity of the ITSEA are excellent, as assessed in a diverse sample of 214 parents of typically developing children between the ages of 12 and 36 months. The ITSEA scales correlate well with laboratory measures of emotional regulation. Domain scores of externalizing behaviour, internalizing behaviour, dysregulation, and competence

were used in the present study to index variables previously found to be associated with ASD-related symptoms in very young children.[39]

Data Analysis

To explore the distribution of autistic-like traits, mean total scores from the Autism Observation Scale for Infants (AOSI) at age 12 months were compared on three sub-samples of interest: (1) high-risk siblings with a diagnosis of ASD at age 3 (HR-ASD; n=52); (2) high-risk siblings without a diagnosis of ASD at age 3 (HR-nonASD; n=170); and (3) low-risk controls without a diagnosis of ASD (LR-nonASD; n=90). Figure 1 depicts two findings: (a) the distribution of autistic-like traits appears to lie on a continuum/gradient of severity across these three sub-samples (HR-ASD > HR-nonASD > LR-nonASD; all differences are statistically significant; $p < 0.01$); and (b) there is notable heterogeneity of autistic-like traits within the three sub-samples, and the sub-sample distributions appear to overlap.

(insert Figure 1)

Next, in an attempt to better understand the observed heterogeneity of autistic-like traits in the nonASD population (the focus of the current study), mean total scores from the AOSI at age 12 months were used in cluster analysis (K-means) to test for the existence of a distinct sub-group of children with elevated scores. Our primary hypothesis was that this group would contain more HR siblings (of probands with ASD) than LR controls, and be associated with more ASD-like outcomes at 36 months. This methodology uses the technique of Euclidean distance to the mean of the cluster and an algorithm to minimize within-cluster variance and maximize variability between clusters in an analysis of variance (ANOVA)-like fashion. Specifically, ANOVA F-tests are conducted to examine differences between clusters on each variable used in the analysis; the magnitude of the F value is used to evaluate how well the variable discriminates between clusters. Cluster centers shift with each iteration. The process continues until cluster means do not shift more than a given cut-off value or the iteration limit is reached.[40]

A 2-cluster solution was specified to test our hypothesis that the sample consists of two distinct groups/clusters: those with high and those with low levels of autistic-like traits as measured by high and low total AOSI scores. As a second step, each participant was assigned to one of the two clusters based on probability scores. Cross-tabulation with chi-square test was used to describe the count of individuals from the two clusters within the HR sibling and LR control groups and to see if there was differential distribution of HR sibs and LR controls in the two clusters. The relative risk (RR) for cluster membership was calculated for the HR and LR groups.

Finally, t-tests were used to compare cluster mean scores on independently determined (a) ASD symptoms (indexed by the ADOS severity metric & ADI-R domain scores); (b) cognitive ability (indexed by the Mullen ELC); and (c) emotional and behavioral symptoms (indexed by the ITSEA domain scores) 24 months later at 3 years of age. Bonferroni correction for multiple comparisons was applied separately within the ASD symptoms domain (p value set at 0.0125) and the emotional and behavioral domain (*p* value set at 0.0125). Effect size (ES) was calculated for cluster mean differences on these variables using the formula[41]: *Effect size = [Mean (Cluster1) – Mean (Cluster 2)] / SD (Cluster 2).*

Results

Two distinct clusters with statistically significantly different scores on the AOSI at 12 months were derived. Cluster 1 (51.4% male) had more autistic-like traits with a total mean AOSI score of 10 (SD = 3.0) and consisted of 37 children (14.2% of total sample). Cluster 2 (49.3% male) had a total mean AOSI score of 2 (SD = 2.0) and consisted of 223 children (85.8 % of total sample). The relative proportion of HR and LR infants differed between the two clusters ($\chi^2_{(1)}$ = 10.8; p < .01). Within the HR group, 33 of 170 were assigned to Cluster 1 (19.4%) and 137 to Cluster 2; only 4 of 90 (4.5%) of the LR group were assigned to Cluster 1. Compared to LR infants, HR infants had a relative risk (RR) of 4.3 (95 % CI: 1.6–11.9) for membership in Cluster 1. There was no difference in the distribution of sex across Clusters (p > .05).

(insert Figure 2)

Mean comparisons (t-tests) showed that, at age 3 years, children in Cluster 1 (identified at 12 months) had significantly higher scores (i.e., more autistic-like behaviors) than children in Cluster 2 on measures of social (ES=0.86) and communication (ES=0.72) impairment, as indexed by the corresponding ADI-R domains (p < 0.001). Children in Cluster 1 also had lower cognitive levels, as indexed by the MSEL ELC (ES=-0.59; p < 0.005). Children in Cluster 1 scored higher on the internalizing problems domain of the ITSEA (ES=0.55; p=0.010). Finally, children in the two clusters did not differ significantly on the ADOS severity metric although the effect was in the expected direction, with Cluster 1 trending towards higher scores (ES=0.35; p=0.06).

(insert Table 1)

Discussion

Using prospective data, this study examined the presence of autistic-like traits at age 12 months among HR siblings (of probands with ASD) who do not go on to receive a diagnosis of ASD by age 3, in comparison to LR control infants. A cluster/subgroup of children had elevated levels of autistic-like traits at 12 months but did not meet CBE

criteria for an ASD diagnosis at 36 months. This cluster comprised approximately 19% of the HR unaffected siblings, or 15% of the entire HR group (i.e., including those receiving the ASD diagnosis by CBE at age 3). Cluster 1 was composed primarily of HR siblings of probands with ASD, (33 out of 37 or 89%) and included only 4 children from the LR controls. Compared to the rest of the children at age 3 (i.e., 24 months later), children within this cluster had, on average, more social and communication challenges, lower levels of cognitive functioning, and more internalizing problems.

Post-hoc analysis showed that cluster differences on levels of autistic-like traits were also evident at age 6 months (Cluster 1: Mean=8.48, SD=5.37 versus Cluster 2: Mean=5.24, SD=2.92; ES = 1.11, $P < 0.001$). However, it must me noted that available AOSI data at age 6 months was limited to only 23 of 37 children from Cluster 1 and 177 of 223 children from Cluster 2.

Study findings suggest that autistic-like traits emerge very early (by 12 months), although children in Cluster 1 ultimately experienced different outcomes than those HR sibs who were subsequently (i.e., at age 3) diagnosed with ASD. However, it is important to note that, with currently available data, we cannot yet confirm that these are the same children who present with the BAP at later ages (at least as it is currently understood). For example, rigidity, pragmatic language deficits, and circumscribed interests are a prominent part of the BAP but are not relevant traits to be measured at 12 months. Moreover, we cannot say that some of these sibs will not eventually (i.e., say at age 5) be diagnosed with ASD. However it would be reasonable to suggest that children within Cluster 1 present with features that could reflect the earliest manifestations of "a Broader Autism Phenotype"; whether or not this phenotype is the same (qualitatively and/or quantitatively) as the one defined in older populations remains an empirical question. Our continuing follow-up investigation of this sample will address these critical issues.

The question then that arises is whether HR unaffected siblings from Cluster 1 (n=33; used in current study) and HR affected/diagnosed siblings (n=52; not used in current study) start off (at 12 months of age) with similar or different levels of autistic-like traits. As depicted in Figure 1, post-hoc analysis showed that at 12 months of age the AOSI scores for the two groups are significantly different (unaffected sibs from Cluster 1: mean=9.7, SD=2.5 versus sibs with ASD diagnosis: mean=7.6, SD=5.1; p<0.01; ES=.46). We believe this finding (higher score in the unaffected sibs) is due to a statistical artifact (i.e., the cluster analysis applied only to unaffected sibs data derives a subgroup with extreme scores). Nevertheless, the finding that both HR infants diagnosed with ASD at 3 years *and* a subgroup not diagnosed with ASD have elevated symptom scores at 12 months relative to controls raises interesting questions about what factors might influence variation in subsequent trajectories and outcomes among the combined group of 12-month-old symptomatic HR siblings.

The current study has methodological strengths that enhance the validity of the results compared to previous studies that retrospectively assessed the BAP in unaffected siblings recruited largely through genetic studies. We had the advantage of following these siblings prospectively and of using a reliable and valid observer-based assessment of autistic-like traits designed for infants, the AOSI. The assignment to clusters was based on information gathered prior to outcome assessments at age 3 (which were blind to 12-month data), and we used a data-driven approach to assign children into clusters rather than using AOSI scores arbitrarily to divide the sample.

This study also has limitations that need to be taken into account. Most importantly, the assessors administering and scoring the AOSI at 12 months were on occasion not blind to HR sibling or LR control status. For example, some parents lacking child care brought the affected older child to the appointment. In addition, the sample sizes of the clusters (especially Cluster 1) are small. The fact that the ASD diagnosis at age 3 was based on the clinical best estimate (CBE) procedure means that some HR siblings (especially from in Cluster 1) might in fact have had "true" ASD at age 3, but may have been "missed" by the CBE. A descriptive analysis showed that of the 33 children in Cluster 1, 4 met the ADI-R cut-off for ASD and 10 met the ADOS cut-off for ASD at age 3; however none of these children met both ADI-R and ADOS criteria for ASD.

Although the cluster difference on the ADOS severity score (at 36 months) was not statistically significant (p=.06), the effect size (ES=.35) was in the expected direction. Several measurement issues may have influenced this result. One possible explanation is that since the ADOS severity metric used in the current analyses reflects a combined score from all ASD symptom domains, it cannot distinguish children from Cluster 1 who have elevated severity only on social-communication symptoms but not repetitive behaviours (see Table 1). An alternative explanation is that some children from Cluster 1 had more severe autism symptoms during infancy, but that those symptoms may have resolved by the time these children were 3 years of age. One final explanation for the lack of statistically significant cluster differences on the ADOS severity metric could be that the variability of the severity metric is (by definition) reduced in "unaffected' children (the metric was developed such that nonASD scores range from 1-3, whereas ASD scores range from 4-10).

Despite the limitations noted above, results from the current study, if replicated, could have important clinical and research implications. First, the results imply that genetic liability in these families is more normally distributed than restricted only to those with the disorder. Even children who do not go on to develop ASD could potentially benefit from surveillance and early intervention should there be impairment. Second, genetic, epigenetic and environmental studies of these children might help us to understand the mechanisms leading to variation in familial aggregation of ASD and related traits. Third, it will be important to follow these sibs to see whether they develop

other forms of psychopathology (such as internalizing disorders), and/or whether persistent social and communication difficulties, despite being sub-threshold for an ASD diagnosis, affect functional outcomes or quality of life.

The fact that the overall risk for ASD and these milder autistic-like traits now reaches almost 40% (based on Zwaigenbaum et al., 2011, and the current study) with an equal sex ratio has implications for our understanding of the genetic architecture of the disorder. It appears as if roughly 40% of infant sibs of autistic probands have autisticlike traits at 12 months. Traits persist in half of the group, who receive a diagnosis of ASD at age 3. In the other half, those traits attenuate to some extent so that at age 3, mild ASD-like traits, lower cognitive scores, and more internalizing symptoms are seen relative to LR controls, but these individuals fall below the diagnostic threshold for ASD. One unanswered question is "Why do autistic-like traits persist in one group and attenuate in the other?" This question will require further research and provides a focus on studying potential modifying mechanisms that might explain the variable expressivity between the Cluster 1 sibs and the ASD probands. These modifying factors might provide an opportunity for intervention in all children at risk for ASD and autistic-like traits. This study supports previous recommendations of continuous monitoring of all high-risk siblings of probands with ASD.

References

1. Bailey A, Le Couteur A, Gottesman I, Bolton P, Simonoff E, Yuzda E, Rutter M. Autism as a strongly genetic disorder: evidence from a British twin study. *Psychol Med.* 1995;25(1):63-77.
2. Constantino J, Todd R. Intergenerational transmission of subthreshold autistic traits in the general population. *Biol Psychiatry.* 2005;57(6):655-660.
3. Hallmayer J, Cleveland S, Torres A, Phillips J, Cohen B, Torigoe T, Miller J, Fedele A, Collins J, Smith K, Lotspeich L, Croen L, Ozonoff S, Lajonchere C, Grether J, Risch N. Genetic heritability and shared environmental factors among twin pairs with autism. *Arch Gen Psychiatry.* 2011;68(11):1095-1102.
4. Hoekstra R, Bartels M, Verweij H, Boomsa D. Heritability of autistic traits in the general population. *Arch Pediatr Adolesc Med.* 2007;161(4):372-377.
5. Szatmari P. Is autism, at least in part, a disorder of fetal programming? *Arch Gen Psychiatry.* 2011;68(11):1091-1092.
6. Lundström S, Chang Z, Råstam M, Gillberg C, Larsson H, Anckarsäter H, Lichtenstein P. Autism spectrum disorders and autistic-like traits: similar etiology in the extreme end and the normal variation. *Arch Gen Psychiatry.* 2012;69(1):46-52.
7. Bolton P, Macdonald H, Pickles A, Rios P, Goode S, Crowson M, Bailey A, Rutter M. A case-control family history study of autism. *J Child Psychol Psychiatry.* 1994;35(5):877-900.
8. Piven J, Palmer P, Jacobi D, Childress D, Arndt S. Broader autism phenotype: evidence from a family history study of multiple-incidence autism families. *Am J Psychiatry.* 1997;154(2):185-190.
9. Piven J. The broad autism phenotype: a complementary strategy for molecular genetic studies of autism. *Am J Med Genet.* 2001;105(1):34-35.
10. Szatmari P, MacLean J, Jones M, Bryson S, Zwaigenbaum L, Bartolucci G, Mahoney W, Tuff L. The familiar aggregation of the lesser variant in biological and nonbiological relatives of PDD probands: a family history study. *J Child Psychol Psychiatry.* 2000;41(5):579-586.
11. Szatmari P, Merette C, Emond C, Zwaigenbaum L, Jones M, Maziade M, Roy M, Palmour R. Decomposing the autism phenotype into familiar dimensions. *Am J Med Genet B Neuropsychiatr Genet.* 2008;147B(1):3-9.
12. Constantino J, Lajonchere C, Lutz M, Gray T, Abbacchi A, McKenne K, Singh D, Todd R. Autistic social impairment in the siblings of children with pervasive developmental disorders. *Am J Psychiatry.* 2006;163(2):294-296.
13. Landa R, Piven J, Wzorek M, Gayle J, Chase G, Folstein S. Social language use in parents of autistic individuals. *Psychol Med.* 1992;22(1):245-254.
14. Piven J, Wzorek M, Landa R, Lainhart J, Bolton P, Chase G, Folstein S. Personality characteristics of the parents of autistic individuals. *Psychol Med.* 1994;24(3):783-795.

15. Wolff S, Narayan S, Moyes B. Personality characteristics of parents of autistic children: a controlled study. *J Child Psychol Psychiatry.* 1988;28(2):143-153.
16. Murphy M, Bolton P, Pickles A, Fombonne E, Piven J, Rutter M. Personality traits of the relatives of autistic probands. *Psychol Med.* 2000;30(6):1411-1424.
17. Skuse D, Many P, Scourfield J. Measuring autistic traits: heritability, reliability and validity of the social and communication disorders checklist. *Br J Psychiatry.* 2005;187(6):568-572.
18. Yirmiya N, Shaked M. Psychiatric disorders in parents of children with autism: a meta-analysis. *J Child Psychol Psychiatry.* 2005;46(1):69-83.
19. Cassel T, Massinger D, Ibanez L, Haltigan J, Acosta S, Buchman A. Early social and emotional communication in the infant siblings of children with autism spectrum disorders: an examination of the broad phenotype. *J Autism Dev Disord.* 2007;37(1):122-132.
20. Goldberg M, Mostofky S, Cutting L, Mahone E, Astor B, Denckla M, Landa, R. Subtle executive impairment in children with autism and children with ADHD. *J Autism Dev Disord.* 2005;35(3):279-293.
21. Stone W, McMahon C, Yoder P, Walden T. Early social-communicative and cognitive development of younger siblings of children with autism spectrum disorders. *Arch Pediatr Adolesc Med.* 2007;161(4):384-390.
22. Landa R, Holman K, Garrett-Mayer E. Social and communication development in toddlers with early and later diagnosis of autism spectrum disorders. *Arch Gen Psychiatry.* 2007;64(7):853-864.
23. Constantino JN, Zhang Y, Frazier T, Abbacchi AM, Law P. Sibling recurrence and the genetic epidemiology of autism. *Am J Psychiatry.* 2012;167(11):1349-56.
24. Zwaigenbaum L, Bryson S, Lord C, Rogers S, Carter A, Carver L, Chawarska K, Constantino J, Dawson G, Dobkins K, Fein D, Iverson J, Klin A, Landa R, Messinger D, Ozonoff S, Sigman M, Stone W, Tager-Flusberg H, Yirmiya N. Clinical assessment and management of toddlers with suspected autism spectrum disorder: insights from studies of high-risk infants. *Pediatrics.* 2009;123(5):1383-1391.
25. Zwaigenbaum L, Thurm A, Stone W, Baranek G, Bryson S, Iverson J, Kau A, Klin A, Lord C, Landa R, Rogers S, Sigman M. Studying the emergence of autism spectrum disorders in high-risk infants: methodological and practical issues. *J Autism Dev Disord.* 2007;37(3):466-480.
26. Bryson S, Zwaigenbaum L, Brian J, Roberts W, Szatmari P, Rombough V, McDermott C. A prospective case series of high-risk infants who developed autism. *J Autism Dev Disord.* 2007;37(1):12-24.
27. Flanagan JE, Landa R, Bhat A, Bauman M. Head lag in infants at risk for autism: a preliminary study. *Am J Occup Ther.* In press.
28. Ozonoff S, Iosif A, Baguio F, Cook IC, Hill MM, Hutman T, Rogers SJ, Rozga A, Sangha S, Sigman M, Steinfeld MB, Young GS. A prospective study of the emergence of early behavioral signs of autism. *J Am Acad Child Adolesc Psychiatry.* 2010;49(3):256-266.

29. Bryson SE, Zwaigenbaum L, McDermott C, Rombough V, Brian J. The Autism Observation Scale for Infants: scale development and reliability data. *J Autism Dev Disord.* 2008;38(4):731-738.
30. Zwaigenbaum L, Bryson S, Rogers T, Roberts W, Brian J, Szatmari P. Behavioural manifestations of autism in the first year of life. *Int J Dev Neurosci.* 2005;23(2-3):143-152.
31. Ozonoff S, Young GS, Carter A, Messinger D, Yirmiya N, Zwaigenbaum L, Bryson S, Carver LJ, Constantino JN, Dobkins K, Hutman T, Iverson JM, Landa R, Rogers SJ, Sigman M, Stone WL. Recurrence risk for autism spectrum disorders: a baby siblings research consortium study. *Pediatrics.* 2011;128(3):e488-e495.
32. Baird G, Douglas HR, Murphy MS. Recognizing and diagnosing autism in children and young people: summary of NICE guidance. *BMJ.* 2011;343:d6360.
33. Kim SH, Lord C. Combining information from multiple sources for the diagnosis of autism spectrum disorders for toddlers and young preschoolers from 12 to 47 months of age. *J Child Psychol Psychiatry.* 2012;53(2):143-51.
34. Lord C, Risi S, Lambrecht L, Cook E, Leventhal B, DiLavore P, Pickles A, Rutter M. The Autism Diagnostic Observation Schedule-Generic: a standard measure of social and communication deficits associated with the spectrum of autism. *J Autism Dev Disord.* 2000;30(3):205-223.
35. Gotham K, Pickles A, Lord C. Standardizing ADOS scores for a measure of severity in autism spectrum disorders. *J Autism Dev Disord.* 2009;39(5):693-705.
36. Lord C, Rutter M, Le Couteur A. Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible persvaive developmental disorders. *J Autism Dev Disord.* 1994;24(5):659-685.
37. Mullen E. *Mullen Scales of Early Learning.* Circle Pines, MN: American Guidance Service; 1995.
38. Carter A, Briggs-Gowan M. *Manual of the Infant-Toddler Social-Emotional Assessment.* New Haven, CT: Yale University; 2000.
39. Georgiades S, Szatmari P, Duku E, Zwaigenbaum L, Bryson S, Roberts W, Fombonne E, Mirenda P, Smith I, Vaillancourt T, Volden J, Waddell C, Thompson A, Pathways in ASD Study Team. Phenotypic overlap between core diagnostic features and emotional/behavioral problems in preschool children with autism spectrum disorder. *J Autism Dev Disord.* 2011;41(10):1321-1329.
40. Tabachnick B, Fidell L. *Experimental designs using ANOVA.* Texas: Brooks/Cole Publishing; 2007.
41. Glass GV. Primary, secondary, and meta-analysis of research. *Educational Researcher.* 1976;5(10):3-8

Figure 1. Total mean Autism Observation Scale for Infants (AOSI) scores with standard deviation bars at 12 months of age for high-risk (HR) siblings with a diagnosis of autism spectrum disorder (ASD) at age 3 years, HR non-ASD siblings at age 3 years, and low-risk (LR) non-ASD groups. All differences between groups are statistically significant (P < .01 for all).
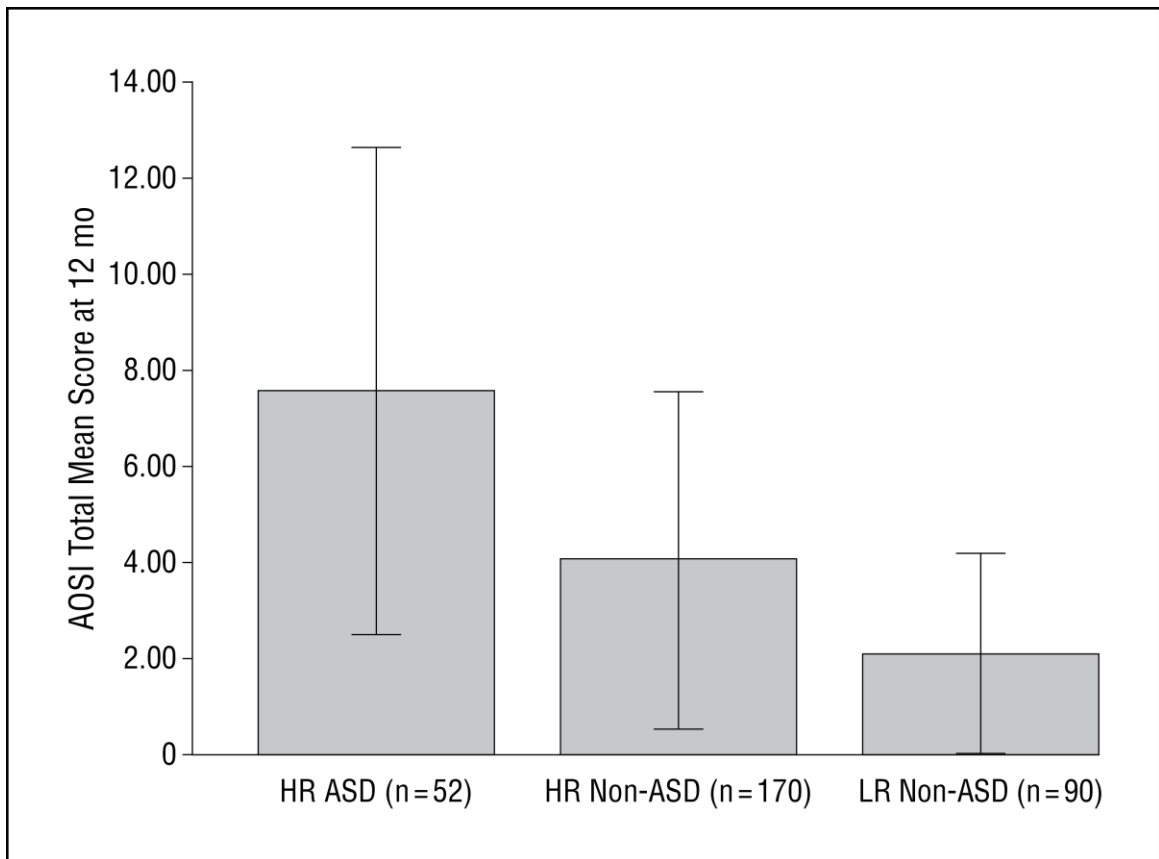
Figure 2. Total mean Autism Observation Scale for Infants (AOSI) scores with standard deviation bars at 12 months of age for high-risk (HR) autism spectrum disorder (ASD), cluster 1, and cluster 2 groups. The HR ASD group scores are shown for descriptive purposes so that the cluster analysis results can be placed within the distribution/context of autistic-like traits in all HR siblings. All differences between groups are statistically significant (P < .01 for all).
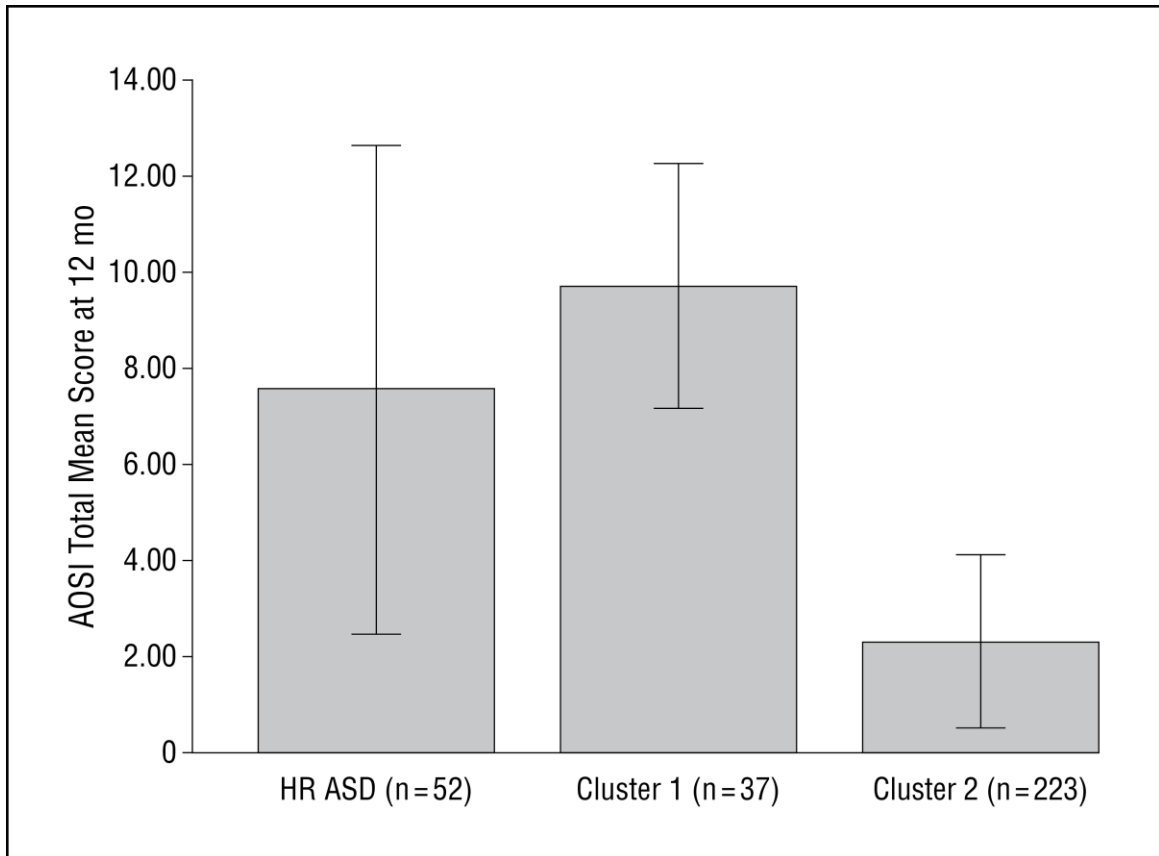
Table 1. Cluster Comparison on the ADOS, ADI-R, MSEL, and ITSEA Scores at Age 3

Outcome Measures	Cluster 1 (n=37)		Cluster 2 (n=223)			
	Mean	SD	Mean	SD	ES	p value*
ADOS severity metric	2.62	1.86	2.06	1.62	0.35	0.06
ADI-R social domain score	3.76	3.56	1.86	2.20	0.86	**< 0.001**
ADI-R communication domain score	3.22	3.33	1.66	2.17	0.72	**< 0.001**
ADI-R behaviours domain score	1.24	1.42	0.84	1.44	0.28	0.11
MSEL-ELC score	105.27	20.50	115.37	17.15	-0.59	**< 0.005**
ITSEA externalizing domain score	0.41	0.29	0.42	0.26	-0.04	0.83
ITSEA internalizing domain score	0.49	0.28	0.38	0.20	0.55	**0.01**
ITSEA dysregulation domain score	0.48	0.30	0.43	0.25	0.20	0.32
ITSEA competence domain score	1.50	0.34	1.59	0.27	-0.33	0.18

*p value cut-off set at 0.0125 based on a Bonferroni correction for multiple testing (significant values are in bold font). Cluster 1 (n = 37; 33 high risk HR sibs, 4 low risk LR controls); Cluster 2 (n = 223; 137 HR sibs, 86 LR controls); ADOS: Autism Diagnostic Observation Schedule; ADI-R: Autism Diagnostic Interview – Revised; MSEL: Mullen Scales of Early Learning – Early Learning Composite; ITSEA: Infant Toddler Social Emotional Assessment; ES: Effect size = [Mean (Cluster1) – Mean (Cluster 2)] / SD (Cluster 2)

Supplemental Table

eTable 1. Descriptive Characteristics for 260 Children Without a Diagnosis of ASD at Age 3 Y[a]

Descriptives	HR nonASD (n=170)	LR nonASD (n=90)
Sex, No. %		
Male	84 (49.4)	45 (50)
Female	86 (51.6)	45 (50)
Age at 12 mo (mean, SD mo)	12.49 (0.89)	12.44 (0.79)
Cognitive level (MSEL-ELC) at 12 months of age	110 (18.1)	120.5 (15.6)
Rate of language delays (≤1.5 SD below mean on MSEL subscales), %		
Expressive language	3.0	0.0
Receptive language	1.8	1.8
Expressive or receptive language	4.2	1.8

ASD: Autism Spectrum Disorder; ELC, Early Learning Composite; HR high-risk siblings; LR low-risk controls; MSEL ELC: Mullen Scales of Early Learning.
[a]Sample used in cluster analysis

CHAPTER FOUR
STUDY 3

TITLE: Investigating Phenotypic Heterogeneity in Children with Autism Spectrum Disorder: A Factor Mixture Modeling Approach

AUTHORS: Stelios Georgiades[1], Peter Szatmari[1], Michael Boyle[1], Steven Hanna[1], Eric Duku[1], Lonnie Zwaigenbaum[2], Susan Bryson[3], Eric Fombonne[4], Joanne Volden[2], Pat Mirenda[5], Isabel Smith[3], Wendy Roberts[6], Tracy Vaillancourt[7], Charlotte Waddell[8], Teresa Bennett[1], Ann Thompson[1], & Pathways in ASD Study Team
[1]McMaster University and Offord Centre for Child Studies, Hamilton, ON, Canada; [2]University of Alberta, Edmonton, AB, Canada; [3]Dalhousie University/IWK Health Center, Halifax, NS, Canada; [4]McGill University, Montreal, QC, Canada; [5]University of British Columbia, Vancouver, BC, Canada; [6]University of Toronto, Toronto, ON, Canada; [7]University of Ottawa, Ottawa, ON, Canada; [8]Simon Fraser University, Vancouver, BC, Canada.

CONTEXT AND IMPLICATIONS OF THIS STUDY: One of the main factors contributing to the limited success in studying the etiology, diagnosis, treatment, and prognosis of ASD is the remarkable heterogeneity observed at the phenotypic level. This study examined the underlying structure of the DSM 5 core symptom domains to answer the question of whether ASD is best represented by a categorical (i.e., classes), dimensional (i.e., dimensions), or hybrid (i.e., subgroups and dimensions) measurement model. Results showed that ASD can be best captured by a factor mixture model that describes the phenotype using three classes based on differential severity gradients on the core symptom dimensions. This study suggests that the symptom dimensions in the DSM 5 can be used to stratify children with ASD empirically into three relatively homogeneous classes/subgroups.

Tidmarsh, Larry Tuff, Kathryn Vaillancourt, Stephen Wellington, Isabelle Yun, and Li Hong Zhong.

CONFLICTS OF INTEREST: None

Abstract

Background: Autism Spectrum Disorder (ASD) is characterized by notable phenotypic heterogeneity, which is often viewed as an obstacle to the study of its etiology, diagnosis, treatment and prognosis. Based on empirical evidence, instead of three binary categories, the upcoming edition of the DSM 5 will use two dimensions — social communication deficits (SCD) and fixated interests and repetitive behaviours (FIRB) — for the ASD diagnostic criteria. Building on this proposed DSM 5 model, it would be useful to consider whether empirical data on the SCD and FIRB dimensions can be used within the novel methodological framework of Factor Mixture Modeling (FMM) to stratify children with ASD into more homogeneous subgroups. **Methods:** The study sample consisted of 391 newly-diagnosed children (mean age 38.3 months; 330 males) with ASD. To derive subgroups, data from the Autism Diagnostic Interview-Revised indexing SCD and FIRB were used in Factor Mixture Modeling; FMM allows the examination of continuous dimensions and latent classes (i.e., categories) by using both factor analysis (FA) and latent class analysis (LCA) as part of a single analytic framework. **Results:** Competing LCA, FA, and FMM models were fit to the data. Based on a set of goodness-of-fit criteria, a "2-factor/3-class" factor mixture model provided the overall best fit to the data. This model describes ASD using three subgroups/classes (Class 1: 34%, Class 2: 10%, Class 3: 56% of the sample) based on differential severity gradients on the SCD and FIRB symptom dimensions. In addition to having different symptom severity levels, children from these subgroups were diagnosed at different ages and were functioning at different adaptive, language, and cognitive levels. **Conclusions:** Study findings suggest that the two symptom dimensions of SCD and FIRB proposed for the DSM 5 can be used in Factor Mixture Modeling to stratify children with ASD empirically into three relatively homogeneous subgroups.

**Investigating Phenotypic Heterogeneity in Children with Autism Spectrum
Disorder: A Factor Mixture Modeling Approach**

Evidence shows that there is notable heterogeneity in the phenotypic presentation of Autism Spectrum Disorder (ASD), regarding both configuration and severity of behavioural symptoms (Geschwind, 2009; Wiggins et al., 2011). To date, researchers have used different methodological approaches to investigate this heterogeneity. A number of studies have used factor analysis (FA) methods to examine the underlying structure of the ASD phenotype (Boomsma et al., 2008; Frazier et al., 2008; Georgiades et al., 2007 & 2011; Kamp-Becker et al., 2009; Snow et al., 2009; Van Lang et al., 2006). As Snow et al. (2009) conclude, these studies have resulted in factor solutions that are not necessarily congruent with the three categorical domains of ASD as defined by the DSM-IV (i.e. social impairment, verbal/non-verbal communication impairment, & repetitive, restricted, stereotyped behaviours; APA, 2000); rather, several of these studies suggest that ASD is best conceptualized using two symptom dimensions, namely social-communication deficits (SCD) and fixated interests and repetitive behaviours (FIRB). This consistent finding has been incorporated in the proposed revisions of the ASD section of the upcoming DSM 5 (APA, 2011).

In parallel, several studies have attempted to identify homogeneous subgroups of individuals with ASD using empirical methods. To date, cluster analytic studies have proposed anywhere from one to four clusters (or subgroups) for ASD that differ largely on symptom severity and intellectual abilities (see Wiggins et al., 2011). Ingram et al. (2008) used taxometric methods (Ruscio & Ruscio, 2004) to determine which phenotypic domains would be most likely to divide a sample of ASD children into two discrete subgroups. Taxometric methods can test whether or not subjects in a given data set are best described in terms of two clusters or in terms of a single homogeneous population (Ruscio & Ruscio, 2004). Regarding ASD symptoms, results from the Ingram et al. study supported sub-grouping participants based only on variation in social communication (i.e., high versus low).

Munson et al. (2008) used latent class analysis and taxometric methods to classify children with ASD. In this study, evidence for multiple subgroups was found using both methods and these subgroups differed in level of intellectual functioning and patterns of verbal versus nonverbal ability. The Munson et al. (2008) study suggests that within the ASD group, there are distinct subtypes of autism which differ in severity of intellectual ability, patterns of cognitive strengths and weaknesses, and severity of autism symptoms.

More recently, Frazier et al. (2012) examined the structure of autism symptoms in a large sample of 14,744 children (8,911 ASD and 5, 863 non-ASD; ages 2 to 18), included in a national registry, the Interactive Autism Network. After comparing different categorical, dimensional, and hybrid (i.e., combined categorical and dimensional) models, the authors concluded that a hybrid model that included both a category (ASD versus

non-ASD) and two symptom dimensions (SCD and FIRB as proposed in the DSM 5) was more parsimonious than all other models and replicated across measures and subsamples. Although the Frazier et al. (2012) study is informative in many ways, it is limited by the reliance on questionnaire data (i.e., the Social Responsiveness Scale and the Social Communication Questionnaire), and the wide age range of the sample (2 to 18 years) as the structure of the ASD phenotype might be different across age groups (i.e., early childhood versus late adolescence). More importantly, the Frazier et al. study was based on a sample from an ASD registry of both ASD and non-ASD cases and therefore does not provide sufficient information on the phenotypic heterogeneity within the clinical ASD group alone. Interestingly, Frazier et al. (2012) noted that *"The two-factor/three-class FM model fit slightly better than all other models. However, the third class appears to overfit the symptom distribution by splitting ASD-affected youth according to extreme and less extreme groups across all SRS scales."* (page 32). Based on these results, further investigation of the distribution of symptom severity within the ASD group alone is warranted.

According to Rutter (2011), the complimentary use of categorical and dimensional classification has become the norm in most areas of medicine and the field of developmental psychopathology could also benefit from such an approach. A relatively new method called Factor Mixture Modeling (FMM) allows the examination of continuous dimensions and latent classes (i.e., categories) by using both factor analysis (FA) and latent class analysis (LCA; Muthen, 2004) in a single analysis. FMM is based on the idea that complex phenotypes require complex measurement models. One of the novel aspects of FMM in relation to taxometric methods is that FMM goes beyond class detection and allows the specification of hypothesis-based multidimensional factor models within each class. While taxometric methods have worked well to identify simple typologies (i.e., disorder is present versus absent), FMM has been developed to identify the underlying structure of more complex data where there may be a combination of multiple dimensions and more than two categories. Therefore, for the study of complex phenotypes, FMM may be superior to taxometric methods both in terms of class detection and class assignment (Lubke & Tueller, 2010). To date, FMM has been used successfully in the study of one other child psychiatric disorder; attention-deficit/hyperactivity disorder (ADHD; Lubke et al. 2007). As far as we are aware, FMM has never been applied to a sample of newly-diagnosed children with ASD.

Distinctions among ASD subtypes (i.e., autistic disorder, Asperger's disorder & pervasive developmental disorder not otherwise specified) have been found to be inconsistent over time, variable across sites, and often associated with severity of language deficits and intellectual impairment rather than a different manifestation of inherent ASD features such as SCD and FIRB symptoms (APA, 2011). Thus, the *Neurodevelopmental Disorders Work Group* for the upcoming version of the DSM 5 (anticipated release in 2013) is proposing a significant shift in the diagnostic conceptualization of ASD (APA, 2011). Rather than representing ASD as multiple

subtypes, ASD will be conceptualized as a single diagnostic category. Moreover, only two dimensions (instead of three categories) — social communication deficits (SCD) and fixated interests and repetitive behaviours (FIRB) — will be specified for the description of the ASD phenotype (APA, 2011). Each individual with ASD will be dimensionally described with these two domains (SCD & FIRB) using a severity gradient based on the level of "support required" by that individual (APA, 2011 & Happe, 2011). To the best of our knowledge, it is unclear how this gradient of ASD symptom severity will be defined empirically. Moreover, we are not aware of any (proposed) specific criteria on how to define informative subgroups/categories to compliment this dimensional approach.

Building on the DSM 5 model, it would be useful to consider whether empirical data on the SCD and FIRB symptom dimensions can be used within the novel methodological framework of FMM to stratify children with ASD into more homogeneous subgroups. Such a stratification could complement the dimensional approach as suggested by Rutter (2011) and provide the foundation for sub-grouping children for genetic, imaging, outcome and response to treatment studies.

Methods

Participants

The study sample consisted of 391 newly-diagnosed preschool children (mean age 38.3 months with SD of 8.7; 330 males) participating in a multisite longitudinal study (Pathways in ASD) examining the developmental trajectories of children with ASD (see Georgiades et al., 2011). All participants had a recent (i.e., within four months) clinical diagnosis of ASD, confirmed by the ADOS and the ADI-R, according to DSM-IV criteria (APA, 2000). The sampling procedure was based on consecutive referrals within specified geographic regions across five Canadian provinces. The study was approved by the local Research Ethics Boards at all sites and all parents gave written informed consent for their children to participate.

Assessment Measures

ASD Symptom Indicators

Autism Diagnostic Interview – Revised (ADI-R; Rutter et al., 2003). The ADI-R is a standardized semi-structured interview used in the diagnosis of ASD. It is designed to be employed with a parent or caregiver who is familiar with the developmental history and current behavior of individuals over the age of two years. The ADI-R is scored using an algorithm that is organized in three domain scales - social, communication, and repetitive behaviours. Currently, there are two versions of the ADI-R algorithm; one for children of ages two to four and one for children aged four and above. Since our sample comprised children aged two to five, both algorithms were used in our study. To ensure

comparability of scores across algorithm versions, 26 common algorithm items (scores of 3 recoded to 2) that apply to all children regardless of age or verbal abilities were selected for analyses. Algorithm items that were age-dependent or language-dependent were excluded. For a list of the 26 ADI-R algorithm items used in the current study see Table 1. There was no missing data on the ADI-R.

Class Correlates

Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2000). The ADOS uses standardized activities and 'presses' to elicit communication, social interaction, imaginative use of play materials and repetitive behaviors, allowing the examiner to observe the occurrence/non-occurrence and severity of behaviors important to the diagnosis of ASD. The ADOS consists of four modules, each of which is appropriate for individuals with differing language levels. A calibrated total severity metric that accounts for differences in age and module is used in the present study (Gotham et al., 2009).

Vineland Adaptive Behavior Scales, Second Edition (VABS II; Sparrow et al., 2005). The VABS II assesses child adaptive behavior in the communication, socialization, daily living skills and motor domains, and expresses overall functioning in the "Adaptive Behavior Composite" (ABC) score (used in current analyses). The VABS II is administered to a parent or caregiver using a semi-structured interview format.

Merrill-Palmer-Revised Scales of Development (MP-R; Roid & Sampers, 2004). This is an individually-administered measure of intellectual ability that is appropriate for children aged two to 78 months. The "Developmental Index standard score" (used in current analyses) comprises cognitive, receptive language and fine motor scales.

Preschool Language Scale – Fourth Edition (PLS-4; Zimmerman, et al, 2002). The PLS-4 is a language test used to identify children with language disorder between birth and 83 months or for older children (such as children with ASD) who function developmentally within this range. The "Total Language Score" is used in the current study.

Statistical Analyses

Factor Mixture Modeling

Factor Mixture Modeling (FMM) allows the simultaneous examination of continuous dimensions and latent classes (or categories, or subgroups) by using both factor analysis (FA) and latent class analysis (LCA; Muthen, 2004). FMM is a general framework extending FA and LCA by combining the two as sub-models into a single general model (Lubke et al., 2005). Unlike taxometric methods that can only test for dichotomous classes derived using data on only one dimension at a time, FMM permits

the specification (i.e., hypothesis) of a multidimensional factor model for each class (Lubke et al., 2007 & 2010). In FMM, individuals are stratified into discrete classes, but within each class, continuous latent factors account for potential differences in the severity of the disorder (Walton et al., 2011). Specific FMMs can be compared and evaluated using well-established indices of goodness-of-fit (Lubke & Muthen, 2005). In the current study, FMMs were applied to identify more homogeneous subgroups (or classes) of ASD using data indexing the SCD and FIRB severity dimensions of ASD within each class.

The 26 ADI-R algorithm items measuring ASD symptoms were subjected to a Principal Component Analysis (PCA) with Varimax (i.e., orthogonal) rotation to derive the most parsimonious model containing uncorrelated factors. Results indicated that compared to the 1, 3, and 4-factor solutions, the 2-factor solution (explaining 32% of the variance; see Table 1) was the most parsimonious solution in terms of both the *Scree Plot* criterion and a clear pattern of item loadings. Moreover, the specific 2-factor solution was selected for subsequent FMM analysis because of its conceptual interpretability and its consistency with established results of numerous factor-analytic studies in the literature as well as the current DSM 5 proposal for the structure of the ASD symptom phenotype (APA, 2011).

(insert Table 1)

Based on previous studies that have proposed the existence of one-to-four ASD subgroups, a total of four competing FMMs were tested using 26 ADI-R indicators. Models 2f1c (2 factors, 1 class), 2f2c, 2f3c, and 2f4c are FMMs with one, two, three and four classes, all with two factors. Specifically, the FMMs have two factors/dimensions (SCD & FIRB) with 20 of the ADI-R indicators forced to load only on the SCD factor and the remaining 6 indicators forced to load only on the FIRB factor (see Table 1 & Figure 1 below).

To confirm whether FMMs are a better overall fit to the data than structural models proposed in previous studies, five LCA models (with one-to-five classes) were evaluated in relation to the four FMMs described above. Finally, to confirm that the 2-factor model (shown in Table 1) used in the FMM analysis had a comparable fit to the data as other previously proposed solutions, five FA models (with one-to-five factors) were also tested.

The fit of all competing models to the data was tested simultaneously using established goodness-of-fit criteria such as the *Akaike Information Criterion (AIC)*, the *Bayesian Information Criterion (BIC)*, and the *Sample Size Adjusted BIC*. In general, lower values of AIC and BIC indicate a better model fit to the data (Lubke et al., 2007). Large simulation FMM studies have shown that the specific goodness-of-fit criteria can help researchers determine which model correctly depicts the data at hand (Lubke and

Spies, 2008). All models were run using *MPlus Version 5.0* statistical software (Muthen & Muthen, 2007).

Characterization of Classes

For the best fitting model, *factor scores* and *class assignment* were calculated for each individual child. Factor scores were calculated using the observed means of items that load on each factor; class assignment was implemented using modal assignment by placing subjects in the class with the highest posterior class probability (Lubke & Tueller, 2010). These scores were then used in post-hoc analyses to describe the derived classes using other child phenotypic indicators believed to be important for the characterization of ASD (APA, 2011; Volkmar et al., 2009).

Specifically, derived classes were described in relation to the child's age at diagnosis, adaptive functioning (indexed by the VABS II composite score), cognitive abilities (indexed by the M-P-R standard score) and language abilities (indexed by the PLS-4). To describe class profiles in terms of ASD symptoms, class mean scores were also compared on: (a) the two derived symptom factors, SCD and FIRB; (b) the original ADI-R algorithm domain scales (i.e., social, nonverbal communication, and repetitive behaviours); and (c) the ADOS severity metric. To better represent class variability on ASD symptom severity, a 2-dimensional (convex hull) plot of SCD by FIRB for the derived classes was created. For these analyses, effect sizes (ES) were estimated by computing class mean differences taking two classes at a time, divided by the overall standard deviation of all three classes combined. An effect size of 0.2 to 0.3 could be interpreted to be a "small" effect, around 0.5 a "medium" effect and 0.8 to infinity, a "large" effect (Cohen, 1992).

Cross-tabulation (chi-square analysis) was used to compare the proportion of children across classes by sex; for all other comparisons, one-way analysis of variance (ANOVA) was used. These analyses were conducted using the SPSS (version 19; 2011) statistical software.

Results

Factor Mixture Modeling

A direct statistical comparison of all competing models showed that the "2-factor/3-class" FMM provided the best fit to the data and was clearly superior across all models (FA, LCA, & FMM) based on all goodness-of-fit criteria - the AIC, BIC, and adjusted BIC (see Table 2). The specific FMM was estimated using a relatively small number of parameters, suggesting parsimony in the description of the underlying phenotypic structure of ASD. According to this FMM, ASD can be described in this sample using data on the two independent severity dimensions of SCD and FIRB to

stratify children into three relatively homogeneous classes (Class 1: 34%, Class 2: 10%, Class 3: 56% of the sample).

(insert Table 2 & Figure 1)

Characterization of Classes

Table 3 presents the ANOVA results and effect sizes (ES) for comparing classes on variables of interest (i.e. class correlates).

(insert Table 3)

In terms of ASD symptoms, on average, children assigned to *Class 1* (34% of the sample) score moderately high on social communication impairments (indexed by the SCD factor and the original ADI-R social and communication domains) and have the lowest scores of repetitive behaviours (indexed by the FIRB factor and the original ADI-R behaviours domain). Children assigned to *Class 2* (10% of the sample) have a reverse profile with the lowest scores on social communication impairments and moderately high scores of repetitive behaviours. Children assigned to *Class 3* (56% of the sample) have the highest scores on both social communication impairments and repetitive behaviours. The estimated effect sizes for class differences on ASD symptoms above were large (see Table 3). The between and within class variability on ASD symptoms is also shown in the 2-dimensional (convex hull) plot (Figure 3). It must be noted that class differences on the ADOS severity metric ranged from small to moderate.

(insert Figures 2 & 3)

Children assigned to *Class 2* were diagnosed at a later age on average (mean age=43.99; SD=9.18 months; $p < 0.01$) compared to children assigned to *Class 1* (mean age=38.42; SD=8.77 months) and *Class 3* (mean age=37.31; SD=8.1 months).

In terms of overall functioning (i.e., developmental level, language abilities and adaptive behaviour), on average, children in *Class 2* had the highest scores followed by children in *Class 1*; children assigned to *Class 3* had the lowest scores in relation to children from the other two classes ($p < 0.01$). From Table 3 we can see that the largest effect sizes were seen for differences between Classes 2 and 3.

Finally, there were no differences in distribution by sex (i.e., the proportion of males and females) across the three classes although the small proportion of females in the sample (16%) may limit our ability to detect statistically significant differences in this distribution across classes.

Discussion

This study used the novel method of FMM to stratify children with ASD into empirically derived subgroups based on their severity levels on the two diagnostic symptom domains of SCD and FIRB proposed for the DSM 5. Our findings confirm those from previous studies suggesting notable heterogeneity in the phenotypic presentation within the ASD spectrum even at this young age (see Munson et al., 2008). Our data suggest that there is evidence of three relatively homogeneous ASD subgroups or classes (Class 1: 34%, Class 2: 10%, Class 3: 56% of the sample) that lie on two spectra (i.e., SCD & FIRB) of ASD symptoms. Although the three subgroups/classes could be described using a total ASD severity gradient, this gradient does not follow the same pattern for both the SCD and FIRB symptom dimensions. Specifically, for the SCD dimension, Class 3 has the highest mean score followed by Class 1 and then Class 2; for the FIRB dimension Class 3 has the highest mean score followed by Class 2 and then Class 1 (see Figure 2). The fact that the class severity gradient pattern differs across dimensions speaks to the importance of treating SCD and FIRB as independent spectra that together make up the overall compound ASD phenotype. These data support the idea that the two ASD symptom domains of SCD and FIRB may potentially arise from largely independent (though possibly overlapping) underlying risk factors (Mundy & Skuse, 2008).

Statistically significant differences (see large effect sizes in Table 3) in ASD symptom severity as well as notable differences in profiles related to child functioning provide suggestive support for the potential utility of the three ASD subgroups proposed here. However, a closer inspection of the between and within class variability suggests that the three subgroups have overlapping distributions of both ASD symptoms (see Figure 3) and overall level of functioning (see Table 3). So even if the three subgroups are more homogeneous in relation to a single ASD spectrum, we still observe wide variability/heterogeneity within each subgroup. Therefore, until these subgroups are tested in genetic, imaging, outcome and treatment studies, it would be premature to claim that their statistically different profiles are clinically meaningful and/or useful.

Children in Class 2 were diagnosed (on average) at a later age than children from the other two classes. Although this finding is intriguing, it cannot be taken as evidence for later onset of ASD in this group because the age of diagnosis is directly connected to the age a child gets referred, as well as to the time a child spends on a wait list for a diagnostic assessment. This finding could be due to the more "subtle" presentation of ASD-related symptoms (i.e., low impairment on SCD) in children assigned to Class 2.

Data presented in this exploratory study do not offer definite answers to the complex issue of ASD heterogeneity; however, our empirical findings could be used to generate specific hypotheses related to the utility of the three derived ASD subgroups. For example, it is possible that children from the different ASD classes might follow different

developmental trajectories, which could be helpful in determining prognosis. Moreover, one could hypothesize that children from different classes would have a differential response to treatment (see Szatmari, 2011). Such research findings could offer clinicians flexible and practical solutions that allow for the utilization of dimensional symptom severity data that can be converted into categorical classes (e.g., mild, moderate, severe). This way clinicians will be able to reliably communicate the information to patients and colleagues and apply pre-specified inclusion/exclusion criteria for treatment purposes (Kamphuis and Noordhof, 2009; Kraemer et al., 2007). Furthermore, the empirical organization of children into more homogenous ASD classes could yield informative phenotypes for stratifying children in genetic studies and in studies in search of biological markers of ASD (Liu et al., 2011; Szatmari et al., 2007).

By exploring substantial data on the two symptom dimensions proposed for the DSM 5 (i.e., SCD and FIRB), we were able to derive more homogeneous classes of newly-diagnosed children with ASD based on severity levels. It is important to note that we chose not to refer to these ASD groups as "subtypes", something that would suggest (a priori) "qualitative" differences in etiology, diagnosis and prognosis (Witwer & Lecavalier, 2008). Rather, we chose to use the terms "classes" or "subgroups" that simply refer to empirical (i.e., data-driven), potentially informative groupings of children, in this case with similar scores on ASD symptom severity dimensions and other related phenotypes.

Limitations

The present study is of an exploratory nature and has several limitations which call for a cautious interpretation of findings. First, FMM analyses were based on parent-report data from the ADI-R and are subject to potential reporting biases. Although data on the direct observation ADOS were available, they were not used in FMM analysis because it was not possible to overcome the complex measurement issues arising from the administration of different ADOS modules to different children. In addition, the ADOS severity metric that can account for differences in modules was not used in the models tested in the current study. This metric is comprised of a total score and does not provide separate scores on the SCD and FIRB symptom dimensions – the core indicators in our study (the lack of notable class differences on the ADOS severity metric supports our decision at the beginning of the study not to use it in the FMM). Second, this study uses cross-sectional data from a sample of newly-diagnosed children in a limited age range; thus, the specific findings cannot be generalized to older children, and must be interpreted within the context of the diagnostic process. Third, to ensure comparability of scores across ADI-R algorithm versions, items that were age-dependent or language-dependent were excluded. As a result, it is possible that important phenotypic information related to language and/or age was missed in the current analysis. For example, the exclusion of language-dependent items from the analyses prohibits the exploration of an additional language-related dimension, a construct known to be important in the clinical

characterization of ASD. Fourth, the fact that the FMM method artificially imposes a common factor structure in each class, does not allow for the examination of potentially different factor structures within different classes/subgroups. For example, one could hypothesize that the phenotypic structure of ASD might have a different appearance among higher functioning children from Class 2 compared to lower functioning children from Class 3. Fifth, the FMMs with more than 3 classes did not converge, preventing any test for the potential superiority of more complex models with 4 or more classes. In fact, in the LCA tests, the 4-class model was a better fitting model than the 3-class model (the same was true for the 3- versus 2-factor solutions; see Table 2). This could be perceived as a limitation of the data; for example, variability in symptom presentation may be restricted by the narrow age range in our sample (i.e., ages 2 to 5 years), something that may in turn hinder the ability of the FMM procedure to identify structural models that are more complex than the "2f/3c" solution. Sixth, the use of ordinal (0 to 2) algorithm items from the ADI-R might have had an effect on the estimation of model parameters in FMM; preferably, normally distributed indicators derived from continuous measures of ASD symptoms should be used in future research. Perhaps the biggest limitation of this study is the absence of a construct validity criterion against which the utility of the proposed ASD model can be tested. However, as longitudinal and genetic data on this sample become available, we plan to evaluate the ability of this model to predict specific developmental trajectories, response to treatment, and genetic markers of ASD.

Conclusion

Heterogeneity within the autism spectrum is perhaps the biggest obstacle to research and translation of research into clinical practice (Newschaffer et al., 2002). Abandoning the "single entity" approach to autism is a necessary step to overcome that obstacle (Happe, Ronald & Plomin, 2006). Moreover, as Rutter (2011) notes, although empirical findings indicate that most mental disorders operate in a dimensional manner, it is still useful to continue using categories (in a complementary way) since they can be quite informative for clinical practice as well as for stratification purposes in clinical, intervention, biological and genetic research.

We propose here a factor mixture model that uses dimensional severity scores on the SCD and FIRB symptom spectra to stratify children with ASD into three relatively homogeneous subgroups. Children from these subgroups have different severity levels of ASD symptoms, are diagnosed at different ages and function at different adaptive, language, and cognitive levels. However, as noted by Szatmari (2011), rather than focusing on assigning labels to these three ASD subgroups, we should focus instead on identifying markers that capture diversity — in developmental trajectories, in responses to treatment, and in the genetic heterogeneity inherent in ASD.

We believe that the proposed "2-factor/3-class" ASD model could inform the ongoing work of the DSM 5 revisions (APA, 2011). For example, as noted by Lord and

Jones (2012), even if the two proposed symptom dimensions (SCD & FIRB) can be informative for the characterization of children with ASD, quantitative measures are needed to accurately map these dimensions. The observed (in the current study) between and within class variability on ASD symptoms as well as the differential but overlapping class distributions of child functioning indicators must be taken into consideration when defining "severity levels" and "clinical specifiers" for the revised ASD criteria in the DSM 5 (APA, 2011). Ideally, carefully designed DSM 5 field trials will incorporate these observations and test the relevant hypotheses empirically. In the mean time, findings from the current study serve as a renewal of our quest for understanding the complex issue of ASD phenotypic heterogeneity, and thus contribute to the study of the etiology, diagnosis, treatment and prognosis of ASD.

References

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

American Psychiatric Association. (2011). *Report of the DSM-V Neurodevelopmental Disorders Work Group.* http://www.dsm5.org Accessed, June 6, 2011.

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

Cohen, J. (1992). "A power primer". *Psychological Bulletin,* 112: 155–159

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.

Frazier, T.W., Youngstrom, E.A., Speer, L., Embacher, R., Law P., Constantino J., Findling, .L., Eng, C. (2012). Validation of proposed DSM-5 criteria for autism spectrum disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, 51, 28-40.

Georgiades, S., Szatmari, P., Duku, E., Zwaigenbaum, L., Bryson, S., Roberts, W., Fombonne, E., Mirenda, P., Smith, I., Vaillancourt, T., Volden, J., Waddell, C., Thompson, A., & Pathways in ASD Study Team. (2011) Phenotypic Overlap between Core Diagnostic Features and Emotional/Behavioral Problems in Preschool Children with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders.* 41, 1321-1329.

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

Geschwind, D. (2009). Advances in autism. *Annual Review of Medicine*, 60, 367-380.

Gotham, K., Pickles, A., Lord, C. (2009). Standardizing ADOS scores for a measure of severity in autism spectrum disorders. *Journal of Autism and Developmental Disorders.* 39, 693-705.

Happe, F. (2011). Criteria, categories, and continua: Autism and related disorders in DSM-5. *Journal of the American Academy of Child & Adolescent Psychiatry*. 50, 540-542.

Happé, F., Ronald, A. & Plomin, R. (2006). Time to give up on a single explanation for autism. *Nature Neuroscience*, 9, 1218-1220.

Hudziak J.J., & Achenbach T.M., (2007). A dimensional approach to developmental psychopathology. *International Journal of Methods in Psychiatric Research*, 16 suppl 1:S16–S23

Ingram, G. D., Takahashi, T. N., & Miles, J. H. (2008). Defining autism subgroups: a taxometric solution. *Journal of Autism and Developmental Disorders*, 38, 950-960.

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

Kamphuis, J.H. & Noordhof, A. (2009). On categorical diagnoses in DSM-V: cutting dimensions at useful points? *Psychological Assessment*, 21, 294-301

Kraemer, H. C. (2007). DSM categories and dimensions in clinical and research contexts. *International Journal of Methods in Psychiatric Research*, 16 Suppl 1, S8-S15.

Liu, X., Georgiades, S., Duku, E., Thompson, A., Devlin, B., Cook, E.H., Wijsman, E.M, Paterson, A.D., & Szatmari, P. (2011). Identification of genetic loci underlying the phenotypic constructs of autism spectrum disorders. *Journal of the American Academy of Child & Adolescent Psychiatry*. 50, 687-696.

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

Lord, C., Rutter, M., DiLavore, P.C., & Risi, S. (2002). *Autism Diagnostic Observation Schedule.* Los Angeles, CA: Western Psychological Services.

Lord, C. & Jones, R. M. (2012). Annual research review: re-thinking the classification of autism spectrum disorders. *Journal of Child Psychology and Psychiatry, 53(5),* 490-509. doi:10.1111/j.1469-7610.2012.02547.x.

Lubke, G., & Tueller, S. (2010). Latent class detection and class assignment: A comparison of the MAXEIG taxometric procedure and factor mixture modeling approaches. *Structural Equation Modeling*, 17, 605-628.

Lubke, G.H. & Spies, J. (2008). Choosing a 'correct' factor mixture model: Power, limitations, and graphical data exploration. In G. R. Hancock & K. M. Samuelsen (Eds.), *Advances in latent variable mixture models.* Charlotte, NC: Information Age Publishing, 343-362.

Lubke, G.H., & Muthén, B.O. (2005). Investigating population heterogeneity with factor mixture models. *Psychological Methods*, 10, 21-39.

Lubke, G.H., Muthén, B.O, Moilanen, I.K, McGough, J.J., Loo, S.K., Swanson, J. M., Yang, M. H., Taanila, A., Hurtig, T., Jarvelin, M.R., & Smalley, S.L. (2007). Subtypes vs. severity differences in Attention Deficit Hyperactivity Disorder in the Northern Finnish Birth Cohort (NFBC). *Journal of the American Academy of Child & Adolescent Psychiatry*, 46, 1584-1593.

Mandy, W.P.L., & Skuse, D.H. (2008). Research review: What is the association between the social-communication element of autism and repetitive interests, behaviours and activities. *Journal of Child Psychology and Psychiatry*, 49, 795–708.

Munson, J., Dawson, G., Sterling, L., Beauchaine, T., Zhou, A., Koehler, E., 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 on Mental Retardation, 113,* 427-438.

Muthén B. (2004). Latent variable analysis: growth mixture modeling and related techniques for longitudinal data; In: *Handbook of Quantitative Methodology for the Social Sciences.* Kaplan D, editor. Thousand Oaks, Sage; pp. 345–368.

Muthen, L.K., & Muthen, B.O. (2007). *Mplus Version 5.0 User's Guide*. Los Angeles: Muthen & Muthen.

Newsschaffer, C. J., Fallin, D. & Nora, L. L. (2002). Heritable and nonheritable risk factors for autism spectrum disorders. *Epidemiologic Reviews* 24(2), 137-153.

Roid, G., & Sampers, J. (2004). *Merrill-Palmer-Revised Scales of Development*. Wood Dale, IL: Stoelting Co.

Ruscio, J., & Ruscio, A. M. (2004). Clarifying boundary issues in psychopathology: The role of taxometrics in a comprehensive program of structural research. *Journal of Abnormal Psychology*, 113(1), 24–38.

Rutter, M. (2011), Research Review: Child psychiatric diagnosis and classification: concepts, findings, challenges and potential. *Journal of Child Psychology and Psychiatry*, 52: 647–660.

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

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

Sparrow, S.S., Cicchetti, D.V., & Balla, D.A., (2005). *Vineland Adaptive Behavior Scales: Second Edition (Vineland II), Survey Interview Form/Caregiver Rating Form*, Livonia, MN: Pearson Assessments.

SPSS for Windows, Rel. 19 (2011). Chicago: SPSS Inc.

Szatmari, P. (2011). New recommendations on autism spectrum disorder. *BMJ*; 342:d2456

Szatmari, P., Maziade, M., Zwaigenbaum, L., Merette, C., Roy, M-A., Joober, R., & Palmour, R. (2007). Informative phenotypes for genetic studies of psychiatric disorders. American *Journal of Medical Genetics Part B*, 144B, 581–588.

Van Lang, N., Boomsma, A., Sytema, S., Bildt, A., Kraijer, D., Ketelaars, C., et al. (2006). Structural equation analysis of a hypothesised symptom model in the autism spectrum. *The Journal of Child Psychology and Psychiatry*. 47, 37-44.

Volkmar, F.R., State, M., & Klin, A. (2009) Autism and autism spectrum disorders: Diagnostic issues for the coming decade. *Journal of Child Psychiatry and Psychology*. 50(1-2), 108-115.

Walton, K. E., Ormel, J., & Krueger, R. F. (2011). The dimensional nature of externalizing behaviors in adolescence: evidence from a direct comparison of categorical, dimensional, and hybrid models. *Journal of Abnormal Child Psychology*, 39(4), 553-561.

Wiggins, L. D., Robins, D. L., Adamson, L. B., Bakeman, R., & Henrich, C (2011). Support for a dimensional view of autism spectrum disorders in toddlers. *Journal of Autism and Developmental Disorders* (Epub ahead of print).

Witwer, A. N., & Lecavalier, L. (2008). Examining the validity of autism spectrum disorder subtypes. *Journal of Autism and Developmental Disorders*. 38, 1611–1624.

Table 1. The 2-factor structure of the 26 ADI-R algorithm items indexing ASD symptoms (N=391).

	#	ADI-R diagnostic algorithm items (item #)	SCD factor	FIRB factor
SCD	1	Direct gaze (item 50)	.405	.238
	2	Social smiling (item 51)	.509	.147
	3	Range of social expressions used to communicate (item 57)	.535	.131
	4	Interest in children (item 62)	.650	.020
	5	Response to approaches of other children (item 63)	.617	.031
	6	Showing and directing attention (item 52)	.656	.147
	7	Offering to share (item 53)	.546	.001
	8	Seeking to share enjoyment with others (item 54)	.608	.140
	9	Use of other's body to communicate (item 31)	.283	.145
	10	Offering comfort (item 55)	.607	.133
	11	Quality of social overtures (item 56)	.636	.100
	12	Inappropriate facial expressions (item 58)[a]	.063	.536
	13	Appropriateness of social responses (item 59)	.616	.025
	14	Pointing to express interest (item 42)	.626	.009
	15	Nodding (item 43)	.507	.097
	16	Head shaking (item 44)	.547	-.047
	17	Conventional/instrumental gestures (item 45)	.647	.045
	18	Spontaneous imitation of actions (item 47)	.583	.179
	19	Imaginative play (item 48)	.521	.199
	20	Imitative social play (item 61)	.507	-.006
FIRB	21	Unusual preoccupations (item 67)	.047	.414
	22	Compulsions/rituals (item 70)	.018	.348
	23	Hand and finger mannerisms (item 77)	-.007	.547
	24	Other complex mannerisms or stereotyped body movements (item 78)	.088	.595
	25	Repetitive use of objects or interest in parts of objects (item 69)	.203	.529
	26	Unusual sensory interests (item 71)	.100	.619

Notes: ADI-R: Autism Diagnostic Interview-Revised; In the factor mixture modeling (FMM) analysis, the first 20 items were "forced" to load on the Social Communication Deficits (SCD) factor; the remaining 6 items loaded on the Fixated Interests and Repetitive Behaviours (FIRB) factor. [a] Item 58 (inappropriate facial expressions) was the only item that did not load as expected; however, for the FMM analysis it was "forced" to load on the SCD factor for practical reasons.

Table 2. ASD structural symptom model comparisons, fit indices, and class proportions (N=391)

	Number of Classes (c) or Factors (f)	Log Likelihood	Number of Free Parameters	AIC	BIC	Adjusted BIC	Class percentages
LCA models	1c	-11124.957	52	22353.915	22560.288	22395.294	100%
	2c	-10326.232	79	20810.464	21123.992	20873.329	32%, 68%
	3c	-10135.720	106	20483.440	20904.123	20567.790	31%, 22%, 47%
	4c	-10049.148	133	20364.297	20892.135	20470.132	20%, 14%, 41%, 25%
	5c	-10011.462	160	20342.925	20977.918	20470.246	11%, 19%, 23%, 22%, 25%
FA models	1f	-10212.711	78	20581.423	20890.982	20643.492	
	2f [a]	-10138.238	103	20482.476	20891.253	20564.439	
	3f	-10057.198	127	20368.395	20872.421	20469.456	
	4f	-10015.669	150	20331.337	20926.644	20450.701	
	5f	No convergence					
FMMs	2f1c	-10187.397	78	20530.795	20840.354	20592.864	100%
	2f2c	-10143.391	81	20448.782	20770.247	20513.238	70%, 30%
	2f3c	**-10073.055**	**84**	**20314.109**	**20647.480**	**20380.953**	**34%, 10%, 56%**
	2f4c	No convergence					

Notes: *ASD: Autism Spectrum Disorder; LCA: latent class analysis; FA: factor analysis; FMM: factor mixture model; AIC: Akaike Information Criterion; BIC: Bayesian Information Criterion. The best fitting model from a direct comparison of all models (FA, LCA, & FMM) and across all goodness-of-fit criteria is presented in bold font. [a] The specific 2-factor model is based on the principal component analysis depicted in Table 1.*

Table 3. Means, standard deviations and effect sizes for the 3 classes of children with ASD on variables of interest

	Mean (SD)			Effect Size		
	Class 1	**Class 2**	**Class 3**	**C1 vs. C2**	**C1 vs. C3**	**C2 vs. C3**
SCD dimension	1.18 (0.36)[a]	0.72 (0.19)[a]	1.51 (0.26)[a]	1.18	-0.86	-2.04
FIRB dimension	0.67 (0.37)	1.08 (0.34)[b]	1.18 (0.38)[b]	-0.93	-1.17	-0.24
ADI-R social domain scale	15.27 (5.05)[a]	10.59 (3.18)[a]	19.56 (3.94)[a]	0.90	-0.83	-1.73
ADI-R communication (nonverbal) domain scale	11.30 (3.17)[a]	9.05 (3.60)[a]	12.81 (3.07)[a]	0.67	-0.45	-1.12
ADI-R repetitive behaviours domain scale	3.74 (2.04)	5.78 (1.97)[b]	5.88 (2.12)[b]	-0.88	-0.93	-0.04
ADOS severity metric	7.23 (1.80)[c]	7.65 (1.69)	7.79 (1.63)	-0.24	-0.33	-0.08
VABS II adaptive behaviour	75.57 (9.60)[d]	79.49 (10.14)[d]	70.30 (9.424)	-0.39	0.52	0.91
M-P-R developmental index score	60.94 (26.75)[d]	64.97 (25.38)[d]	52.03 (24.34)	-0.16	0.35	0.50
PLS-4 total language score	68.70 (20.44)[d]	72.57 (21.69)[d]	61.76 (17.11)	-0.20	0.36	0.57

Notes: ASD: Autism Spectrum Disorder; ADI-R: Autism Diagnostic Interview-Revised; SCD: Social Communication Deficit; FIRB: Fixated Interests and Repetitive Behaviour;VABS II: Vineland Adaptive Behavior Scales, Second Edition; M-P-R: Merrill-Palmer-Revised Scales of Development; PLS-4: Preschool Language Scale – Fourth Edition.

Class 1: 34% of sample; Class 2: 10% of sample; Class 3: 56% of sample;

[a] All three classes are significantly different from each other ($p<0.05$); [b] Each of these classes is significantly different from Class 1($p<0.05$). [c] This class is significantly different than the other two Classes 2 and 3 ($p<0.05$); [d] Each of these classes is significantly different from Class 3($p<0.05$).

Effect size C1 vs. C2 = [mean(C1) – mean(C2)]/overall SD, Effect size C1 vs. C3 = [mean(C1) – mean(C3)]/overall SD, Effect size C2 vs. C3 = [mean(C2) – mean(C3)]/overall SD

Figure 1. Factor mixture model of the ASD symptom phenotype with "2 factors/3 classes" (N=391)
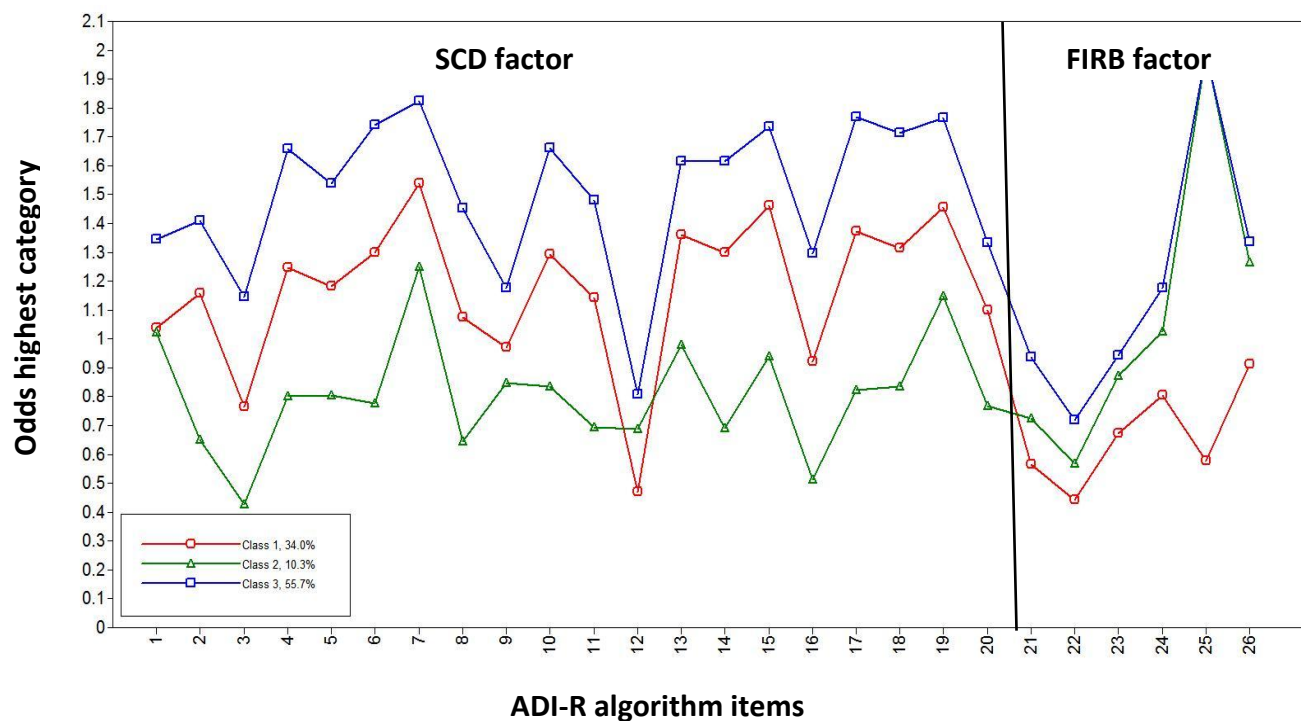
Notes: The horizontal axis represents the 26 algorithm items from the ADI-R. Items 1-20 load on the SCD factor and items 21-26 load on the FIRB factor. The vertical axis represents the probability of scoring in the highest response category/class for each item in proportion to scoring in any of the other categories/classes for 391 children with ASD in the "two factor/three class" factor mixture model. ADI-R: Autism Diagnostic Interview-Revised; SCD: Social Communication Deficits; FIRB: Fixated Interested and Repetitive Behaviours.

Figure 2. Class profiles using mean scores of SCD and FIRB symptom dimensions (N=391)
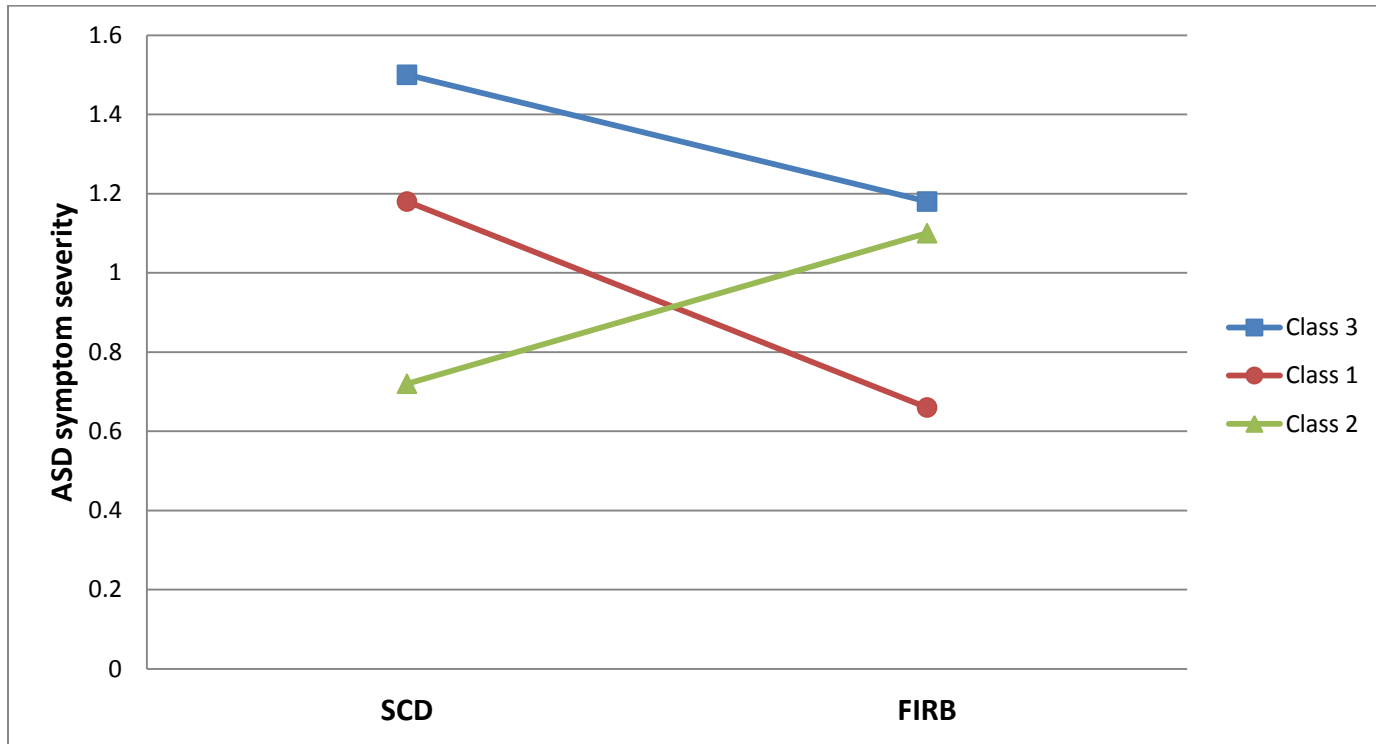
Notes: *SCD: Social Communication Deficits; FIRB: Fixated Interested and Repetitive Behaviours; Class 1: 34% of sample; Class 2: 10% of sample; Class 3: 56% of sample.*

Figure 3. SCD by FIRB 2-dimensional (convex hull) plot for the 3 derived ASD classes (N=391)
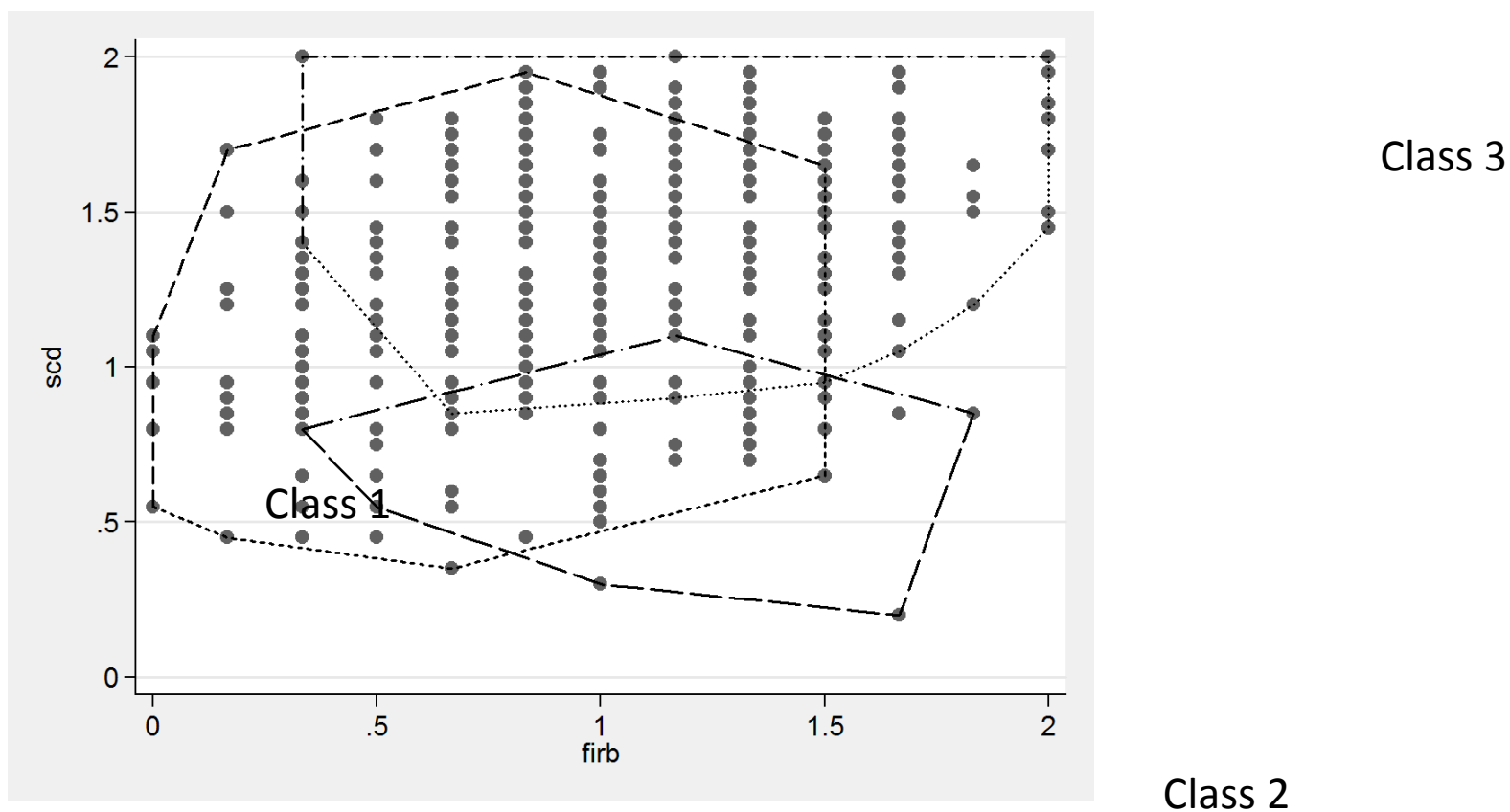
Notes: The horizontal axis represents scores on the Fixated Interested and Repetitive Behaviours (FIRB) symptom dimension; The vertical axis represents scores on the Social Communication Deficits (SCD) symptom dimension; Class 1 (34% of sample); Class 2 (10% of sample); Class 3 (56% of sample).

KEY POINTS:

- Autism Spectrum Disorder (ASD) is characterized by notable phenotypic heterogeneity, which is often viewed as an obstacle to the study of its etiology, diagnosis, treatment and prognosis.
- This study used the novel method of Factor Mixture Modeling (FMM) that allows for the integration of both categories and dimensions to stratify children with ASD into relatively more homogeneous subgroups.
- Results showed that children with ASD can be classified into three subgroups based on their severity on the symptom dimensions of social communication deficits (SCD) and fixated interests and repetitive behaviours (FIRB). Children within these subgroups were diagnosed at different ages and were functioning at different adaptive, language, and cognitive levels.
- Clinically, it is possible that children from these subgroups might follow different developmental trajectories and/or or have a differential response to treatment.
- Study findings can inform the ongoing work on the DSM 5 revisions for ASD.

CHAPTER FIVE
STUDY 4

TITLE: Modeling the Phenotypic Architecture of Autism Symptoms from Time of Diagnosis to Age 6

AUTHORS: Stelios Georgiades[1], Michael Boyle[1], Peter Szatmari[1], Steven Hanna[1], Eric Duku[1], Lonnie Zwaigenbaum[2], Susan Bryson[3], Eric Fombonne[4], Joanne Volden[2], Pat Mirenda[5], Isabel Smith[3], Wendy Roberts[6], Tracy Vaillancourt[7], Charlotte Waddell[8], Teresa Bennett[1], Ann Thompson[1], & Pathways in ASD Study Team

[1]McMaster University and Offord Centre for Child Studies, Hamilton, ON, Canada; [2]University of Alberta, Edmonton, AB, Canada; [3]Dalhousie University/IWK Health Center, Halifax, NS, Canada; [4]McGill University, Montreal, QC, Canada; [5]University of British Columbia, Vancouver, BC, Canada; [6]University of Toronto, Toronto, ON, Canada; [7]University of Ottawa, Ottawa, ON, Canada; [8]Simon Fraser University, Vancouver, BC, Canada

CONTEXT AND IMPLICATIONS OF THIS STUDY: This study examined if the measurement model found to best represent the latent class structure of core autism symptoms at diagnosis (ages 2 to 5; see Study 3) is replicable at age 6. Results showed that at age 6 a more parsimonious model with only 2 (instead of 3) classes and 2 dimensions provided the best fit to the data. Findings suggest that there is a substantive change in the phenotypic architecture of autism symptoms during the first few years after diagnosis and speak to the importance of repeated classification assessments of symptoms, functional skills, and behaviours as children develop.

CONFLICTS OF INTEREST: None

SUBMITTED TO: Journal of Autism and Develomental Disorders

Abstract

This study examined the underlying latent class structure of autism symptoms – social communication deficits, and fixated interests and repetitive behaviours – from the time of diagnosis to age 6 in a sample of 280 children with Autism Spectrum Disorder (ASD). Factor mixture modeling (FMM) was performed on 26 items from the Autism Diagnostic Interview – Revised algorithm. The FMM was repeated twice, at time of diagnosis (Time 1) and at age 6 (Time 2). At Time 1, a "2-factor/3-class" model provided the best fit to the data (Class 1=35%; Class 2=11%; Class 3=54% of the sample). At Time 2 a more parsimonious "2-factor/2-class" model provided the best fit to the data (Class A=32%, Class B=68% of the sample). According to this model, 6-year old children with ASD can be classified in two classes characterized by significantly different levels of symptom severity, adaptive functioning, and emotional/behavioural problems. These findings demonstrate the dynamic nature of the ASD phenotype and speak to the importance of repeated classification assessments of symptoms, functional skills, and behaviours as children develop.

Modeling the Phenotypic Architecture of Autism Symptoms from Time of Diagnosis to Age 6

Autism Spectrum Disorder (ASD) is a complex and heterogeneous disorder (Geschwind 2011; Rutter 2012; Wiggins et al. 2011). The DSM IV and DSM IV-TR (American Psychiatric Association [APA] 1994 & 2000) categorical criteria used to assign a diagnosis do not appear to sufficiently and reliably capture the clinical heterogeneity and symptom variability observed in children with ASD (Happe 2011). As a result, autism researchers have long debated whether the diagnostic criteria for the disorder should be based on a "categorical" or a "dimensional" approach. However, novel statistical methods that can simultaneously compare, combine, and evaluate both approaches have shifted the research focus on how to best transform dimensions into categories by defining thresholds (Lord and Jones 2012). According to Ozonoff (2012), this is the idea behind the recent revision of the ASD diagnostic criteria. Specifically, in the DSM 5, the three diagnostic subtypes are merged into a single "umbrella category" of ASD, and the symptom phenotype is described dimensionally (instead of categorically) using two core symptom domains – social communication deficits (SCD) and fixated interests and repetitive behaviours (FIRB). Each individual is assigned a severity indicator for each dimension based on the level of support required, i.e., "requiring very substantial support," "requiring substantial support," or "requiring support" (APA 2013).

Since clinicians, researchers, and parents all seem to accept the notion that ASD is a heterogeneous disorder, it is likely that categories and subgroups in children with autism will continue to be used, regardless of how the DSM changes its definition (Mandell 2011; Rutter 2011). We and others argue that while the DSM 5 criteria might reflect an improvement in the diagnostic process (i.e., distinguishing ASD from non-ASD cases), empirical data on the DSM 5 symptom dimensions must be examined in rigorous analyses to further inform our understanding of the underlying/latent structure and phenotypic heterogeneity of the disorder. A better understanding of the underlying structure of psychopathology can inform the development of refined and comprehensive assessment methods, targeted interventions, and future research in search of the aetiology of complex disorders such as ASD (Frazier et al. 2010).

Recent studies have shown that data on the two symptom dimensions (SCD and FIRB) of the DSM 5 can be used to distinguish ASD from non-ASD cases with good sensitivity and specificity (Frazier et al. 2012; Huerta et al. 2012). As far as we are aware, only one study has used data on the SCD and FIRB dimensions to systematically model ASD heterogeneity *within* the autism spectrum by classifying 391 newly-diagnosed children into three classes that differ in symptom severity and configuration as well as in levels of adaptive functioning and age at diagnosis (Georgiades et al. 2013). In another relevant study, using data from a large online national registry (11,507 children; 6,901 with ASD and 4,606 non-ASD; ages 2 to 18), Frazier et al. (2010) proposed a hybrid model that included two categories (ASD versus non-ASD) as well as two symptom

dimensions (SCD and FIRB) for describing the structure of ASD. Frazier et al. (2010) noted that although ASD can be adequately represented as a single category that is qualitatively distinct from non-ASD cases, it is possible that the autism spectrum (i.e., ASD cases only) can be further subdivided into two groups with "extreme" and "less extreme" severity scores on autism-related symptoms. So the question that remains is whether symptom severity *within* the autism spectrum is best reflected as a single continuum or as distinct classes.

It is important to note that all the studies on this topic have used cross-sectional data and thus were not able to draw conclusions about the structure of the autism symptom phenotype at different developmental stages. Therefore, longitudinal research is needed to better understand how symptom heterogeneity unfolds as children with ASD develop. The current study aimed to (partially) address this gap in the literature. The primary objective of this study was to use longitudinal data on the SCD and FIRB dimensions collected as part of the *Pathways in ASD* study to model the underlying latent class structure of core autism symptoms from the time of diagnosis to age 6. Specifically, this study examined whether the "2-factor/3-class" factor mixture model found to best represent the latent class structure of autism symptoms between ages 2 and 5 (Georgiades et al. 2013) is replicable at age 6. Another objective was to replicate and validate the derived latent class structure of autism symptoms at age 6 using data from an independent sample of children with ASD participating in the Autism Genome Project (AGP).

Methods

Participants

Descriptive statistics for study participants are shown in Table 1.

Sample 1: Pathways in ASD study

Data for the primary analyses came from a longitudinal sample of 280 children participating in the *Pathways in ASD* study (www.asdpathways.ca) from five regional clinical centres across Canada. All participants had a clinical diagnosis of ASD, confirmed by the ADOS and the ADI-R, according to DSM-IV criteria (APA 2000). Children were first assessed within a period of 4 months after receiving an ASD diagnosis (Time 1). The follow-up assessment took place when the children reached 6 years of age (Time 2). Because the Time 1 assessment was linked to diagnosis (i.e. a study inclusion criterion), the range of children's age at that point varied substantially from 2 to 5 years; this was not the case for Time 2 where all children were 6 years of age (i.e., between 6.0 and 6.11 years old). The 280 children included in the current study were selected because they had complete longitudinal data at both time points (Time 1 & Time 2); these 280 children are a subsample of the original sample of 391 children reported in the Georgiades et al. (2013) cross-sectional study. It is important to note that the subsample of 280 children included in the longitudinal analysis did not differ from the original

sample of 391 children on baseline characteristics (SCD and FIRB symptom levels, age at diagnosis, sex distribution; p > .05 for all). However, the children in the subsample used in the longitudinal analysis (n=280) had significantly lower scores on the SCD (mean: 1.29) and FIRB (mean: 0.95) and were diagnosed at an older age (mean: 38.9 months) compared to those for whom no data were available at 6 years of age (n=111; SCD mean: 1.39; FIRB mean: 1.11; age at diagnosis: 36.9; p < .05). The study was approved by the local Research Ethics Boards at all sites and all participants gave written informed consent to participate.

Sample 2: Autism Genome Project

Data for the replication analysis came from the Autism Genome Project (AGP), a collaborative research consortium of scientists from Europe and North America studying genetic mechanisms underlying autism susceptibility. Although the assessment process across participating AGP sites varied to some extent, all children had a clinical diagnosis of ASD based on DSM IV-TR criteria (American Psychiatric Association, 2000) that was confirmed by at least the ADI-R (and in most cases the ADOS). For purposes of replicating and validating the results derived using data from the *Pathways in ASD study* (Sample 1), a subsample (N=517; Sample 2) of children within their 6th year of age were selected from the larger AGP sample (see Liu et al., 2011, for details). Informed parental consent was obtained for all participants in the study, and Research Ethics Boards gave ethical approval for the research procedures.

(insert Table 1)

Assessment Measures

Autism Symptoms - Class Indicators

Data on the autism symptom class indicators were available for both the Pathways sample (Sample 1) and the AGP sample (Sample 2).

Autism Diagnostic Interview – Revised (ADI-R; Rutter et al. 2003). The ADI-R is a standardized semi-structured interview with a parent or caregiver used in the diagnosis of ASD. The ADI-R is scored using an algorithm that is organized in three domain scales - social, communication, and repetitive behaviours. Currently, there are two versions of the ADI-R algorithm; one for children ages 2-4 and one for children ages 4 and above. Despite the fact that children in the present study were 6 years of age at the time of the second assessment point, due to the longitudinal nature of the analysis, we chose to only include data on the 26 algorithm items (scores of 3 recoded to 2) that are common across both versions of the instrument and apply to all children regardless of age or verbal abilities. In other words, algorithm items that were age-or language-dependent were excluded. There were no missing data on the ADI-R.

Class Correlates

Data on class correlates were available only for the Pathways sample (Sample 1).

Social Responsiveness Scale (SRS; Constantino and Gruber 2005). The SRS is a 65-item parent/caregiver rating scale providing a picture of a child's atypical social behaviour, including social awareness, social information processing, reciprocal communication, social anxiety or avoidance, and autistic preoccupations and traits. Higher total scores indicate greater severity of impairment. The SRS has good test-retest reliability (range: .77 to .85) and internal consistency (range: .77 to .90) in clinical samples of children with ASD. The SRS total t-score was used in the current analyses.

Repetitive Behavior Scale-Revised (RBS-R; Bodfish et al. 1999; 2000). The RBS-R is a parent/caregiver rating scale designed to provide a quantitative measure of the presence and severity of various forms of restricted, repetitive behaviours. It comprises 43 items distributed across six conceptually-derived subscales: Stereotyped Behavior, Self-injurious Behavior, Compulsive Behavior, Routine Behavior, Sameness Behavior and Restricted Behavior. The RBS-R total mean score was used in current analyses.

Vineland Adaptive Behavior Scales, Second Edition (VABS II; Sparrow et al. 2005). The VABS II is a semi-structured interview with a parent or caregiver that assesses child adaptive behavior in the communication, socialization, daily living skills and motor domains. The total "Adaptive Behavior Composite" (ABC) score was used in current analyses.

Child Behavior Checklist (CBCL 1.5-5; Achenbach and Rescorla 2000). The CBCL/1.5-5 obtains parent/caregiver ratings of 99 items of externalizing and internalizing behavior problems observed in typically-developing preschoolers, as well as children with ASD (Achenbach and Rescorla 2000; Pandolfi et al. 2009). The CBCL total t-score based on a normative sample of children aged 1.5-5 years was used in current analyses.

Statistical Analyses

Primary Analysis – Sample 1

Time 1 (at diagnosis)

A total of 26 algorithm items from the Autism Diagnostic Interview – Revised (ADI-R) were used to index the SCD and FIRB dimensions in factor mixture modeling (FMM) to evaluate the latent structure of autism symptoms at the time of diagnosis (Time 1; mean age 40.3 months). This analysis was completed as a preliminary step to confirm

the structural model derived in the larger Pathways cohort (N=391; Georgiades et al. 2013) in the subsample (N=280) used in the current study.

FMM is a general framework that allows the simultaneous examination of both factor analysis (FA) and latent class analysis (LCA; Muthen 2004). In FMM, individuals are stratified into discrete classes; but within each class, continuous latent factors account for potential differences in the severity of the disorder (Walton et al. 2011). Specific FMMs can be compared and evaluated using well-established indices of goodness-of-fit (Lubke and Muthen 2005). FMM is flexible with respect both to the number of latent classes and to modeling the observed variables within class. In a Monte Carlo simulation study, Nylund et al. (2007) reported that non-convergence in factor mixture analysis occurred in less than 5% of cases and thus can be considered an indication of model misfit. In case of non-convergence this criterion will be used as evidence that the model with one fewer class is statistically superior. In the current study, FMMs were applied to identify more homogeneous subgroups (or classes) of ASD using data indexing the SCD and FIRB symptom severity dimensions of ASD within each class. Table 2 shows the item composition for the SCD and FIRB dimensions, as used in the FMM analysis.

(insert Table 2)

A highest posterior probability score based on the similarity of the child's individual response vector to the estimated overall class mean was used to assign to each child membership in one of the ASD classes derived from the FMM analysis. SCD and FIRB mean scores were calculated for each ASD class. Individual standardized scores on the SCD and FIRB dimensions were presented in a 2-dimensional scatterplot using markers to indicate the different ASD classes.

Time 2 (at age 6)

The FMM analysis (see above) using the same 26 ADI-R algorithm items was repeated in follow-up data collected when children reached age 6 (Time 2; mean age 79.4 months).

Each child was assigned membership in one of the ASD classes. SCD and FIRB mean scores were calculated for each ASD class. Individual standardized scores on the SCD and FIRB dimensions were presented in a 2-dimensional scatterplot using markers to indicate the different ASD classes. The ASD classes were also compared on other variables of interest (i.e., class correlates) such as adaptive functioning (VABS II), emotional/behavioural problems (CBCL), repetitive behaviours (RBS-R), child's age at diagnosis and child's sex.

Longitudinal (Time 1 to Time 2)

Cross-tabulation between class membership at T1 and class membership at T2 was used to create a new variable representing all possible *ASD class pathways* from T1 to T2. Then, to describe the longitudinal stability and/or change of ASD symptoms, SCD and FIRB mean scores at T1 and at T2 were calculated and plotted for all *ASD class pathways.*

Replication Analysis – Sample 2

To replicate and validate the results from Sample 1, the same FMM analysis (see above) was performed using data on 6-year old participants from the AGP sample. Because the AGP is a cross-sectional study, earlier data on the 6-year-olds were not available. Therefore, the replication test was done only for Time 2 data (age 6) from the Pathways study.

RESULTS

Primary Analysis – Sample 1

Time 1 (at diagnosis)

At Time 1, a "2-factor/3-class" factor mixture model provided the best statistical fit to the data (Class 1=35%; Class 2=11%; Class 3=54% of the sample). This model suggests that, during the period soon after diagnosis, the latent structure of autism symptoms can be described using three distinct classes of children with different levels of severity on the SCD and FIRB dimensions. This model confirms the results previously reported by Georgiades et al. (2013) in their cross-sectional study.

(insert Table 3 & Figure 1)

Visual inspection of a 2-dimensional scatterplot (Figure 2) showed that the three latent classes were composed of children with different ranges of scores on SCD and FIRB symptom dimensions. However, since class membership was calculated using probability scores rather than actual scores, some overlap between the three classes was observed, indicating that even if these classes of children with ASD are relatively homogeneous, they are not entirely distinct.

(insert Figure 2)

Time 2 (at age 6)

The "2-factor/3-class" structural model derived using Time 1 data failed to converge at Time 2. Instead, a more parsimonious "2-factor/2-class" model provided the best statistical fit to the data at Time 2 (Class A=32%; Class B=68%). This model (see

Table 3) suggests that, by the time children reach 6 years of age, the underlying latent structure of autism symptoms can be adequately described using two (instead of three) distinct classes of children with significantly different levels of severity on the SCD and FIRB dimensions.

(insert Figure 3)

Visual inspection of a 2-dimensional scatterplot (Figure 4) showed that the two latent classes were composed of children with different ranges of scores on the SCD and FIRB symptom dimensions. As in the Time 1 FMM model described previously, some overlap between the two classes was observed; however, in this case the overlap seemed to be reduced, indicating that, as children develop they tend to "cluster" in fewer classes that appear to be "further apart" in terms of score ranges (i.e., more distinct).

(insert Figure 4)

As seen in Table 4, compared to children from Class B, children in Class A had significantly less severe autistic symptoms, higher adaptive functioning skills, and presented with fewer emotional/behavioural problems (p < .01 for all). Children across the two classes did not differ in terms of the age at which they were diagnosed (p > .05).

(insert Table 4)

Furthermore, there was a statistically significant difference (p < .01) in the way boys and girls were distributed across the two ASD classes; girls tended to be assigned to the less severe, higher functioning class (Class A: 61.5%; Class B: 38.5%), while the reverse was true for boys (Class A: 27.4%; Class B: 72.6%).

Longitudinal (Time 1 to Time 2)

Figure 5 is a schematic depicting the six possible *ASD class pathways* from T1 to T2. Results show that, for children in Class 1 at T1, 53.5% (n=53) were assigned to Class A at T2 and the rest (n=46) to Class B. For Children in Class 2 at T1, 34.6% (n=9) were assigned to Class A at T2 and the rest (n=17) to Class B. Finally, for children in Class 3 at T1, 18.1% (28) were assigned to Class A and the rest (n=127) to Class B.

(insert Figure 5)

Figures 6a and 6b show the longitudinal changes in mean SCD and FIRB symptom severity for the six identified *ASD class pathways,* and provide a descriptive summary of the different (aggregated) patterns of change from Time 1 to Time 2. These results suggest differential (across children at any given point) but also longitudinal (within children over time) change in SCD and FIRB symptom severity.

(insert Figures 6a & 6b)

Replication Analysis – Sample 2

FMM analysis using an independent sample of 6-year-olds from the Autism Genome Project (Sample 2) replicated the latent class structure of autism symptom derived in the *Pathways in ASD study* (Sample 1) at Time 2. Once again, the "2-factor/3-class" model failed to converge and the "2-factor/2-class" model provided the best statistical fit to the data (see Table 3; Class A=31%; Class B=69%). As noted above, this model describes the underlining structure of autism symptoms using two distinct classes of children with different levels of severity on the SCD and FIRB dimensions.

Discussion

Numerous studies have examined the latent class structure of the heterogeneous autism phenotype using cross-sectional data (e.g., Frazier et al. 2010; Georgiades et al. 2013). However, since ASD is a developmental disorder, longitudinal research is needed to better understand how symptom heterogeneity unfolds as children develop. The current study builds on previous cross-sectional work in the *Pathways in ASD* cohort (see Georgiades et al. 2013) and represents the first longitudinal investigation (from time of diagnosis to age 6) of the underlying latent class structure (phenotypic architecture) of autism symptoms.

Study findings suggest that there is a change, statistically, in the underlying latent class structure of autism symptoms during the first few years after diagnosis. Specifically, it appears that in our cohort assessed shortly after diagnosis (ages 2 to 5), symptom structure can be described using three latent classes, but that by age 6 this structure can be modeled with only two distinct, relatively homogeneous latent classes of children with ASD. This "2-factor/2-class" structure of the autism symptom phenotype at age 6 is robust and replicable across samples (i.e., Pathways and AGP studies). The two latent classes of children identified at age 6 have different levels of symptom severity; they also differ in terms of levels of adaptive functioning skills and emotional/behavioural problems.

The multi-faceted nature of variability of autism symptoms could not have been captured without the use of the FMM approach. For example, conventional analytic techniques can only describe the structure of ASD symptoms as either categorical (i.e., latent class analysis) or dimensional (i.e., factor analysis). By using the FMM we were able to classify children into empirically derived classes (i.e., categories) with different mean levels of ASD symptoms but at the same time allow for the description of severity within each class (i.e., dimensions). By repeating the FMM analysis at two time points we were able to demonstrate (at a statistical level) changes in class structure (see Figure 5) as well as changes within individual children over time (see Figures 6a and 6b).

Taken together, these findings suggest that ADI-R data indexing the symptom dimensions in the DSM 5 (SCD and FIRB) can be used to provide empirical support for the following: (a) symptom severity of children within the autism spectrum can be best modeled, statistically, using (at least) two latent classes rather than a single continuum; (b) symptom heterogeneity in children with ASD decreases over time such that by the time children turn 6 years of age, they can be classified, statistically, into two distinct, relatively homogeneous classes – a "higher severity" class (68% of sample), and a "lower severity" class (32% of sample); and (c) at age 6 there is a tendency for girls to be assigned to the class that is composed of children with less severe symptoms. This interesting finding provides some support to the notion that girls with ASD are somehow protected and that their symptom severity reduces with development, at least during the preschool years (Robinson et al. 2013).

The main question that arises from these results is the extent to which the observed change in the underlying class structure of the autism symptom phenotype — from 3 classes at diagnosis to 2 classes at age 6 — is a substantive finding with meaningful clinical implications, rather than a methodological artifact. To explore this question, we conducted post hoc analysis by examining the concordance of class membership at Time 1 (ages 2 to 5) with class membership at Time 2 (age 6). For this analysis a "2-factor/2-class" model was forced on Time 1 data. Results showed that there was low concordance (i.e., classification agreement) across the two time points ($\varphi = .22$). Specifically, 25% of children in Class 2 (more severe) at Time 1 were classified in Class A (less severe) at Time 2; conversely, 51% of children in Class 1 (less severe) at Time 1 were classified in Class B (more severe) at Time 2. This low classification agreement could be interpreted as evidence for a substantive change in the latent class structure of the ASD phenotype over time; it can also be taken as one possible explanation for the non-convergence of the "2-factor/3-class" model at Time 2. At the same time, one could argue that low classification agreement between Time 1 and Time 2 class membership at the individual child level, does not necessarily translate into evidence for differences in the underlying structure of the ASD phenotype at the latent class level.

Another important finding that warrants further investigation is the effect of the *child's age at diagnosis* on the underlying structure of autism symptoms at any given point across development. In the Georgiades et al. (2013) study, *age at diagnosis* was found to be associated with class membership (i.e., children diagnosed at an older age had a higher probability of being assigned to the lower severity, higher functioning class). Because that study used data collected shortly after diagnosis, the specific finding could possibly be attributed to the documented negative association between symptom severity and the age at which children are diagnosed (i.e., more severely affected children tend to be diagnosed earlier). Therefore, it is important to take the *child's age at diagnosis* into account when interpreting structural models of the ASD phenotype. However, in the current study where all children were within their sixth year of age (Time 2), the two derived classes did not differ statistically by the *child's age at diagnosis.* Therefore,

differences between classes possibly reflect symptom severity differences that are relatively independent of age confounding effects resulting from referral and/or diagnostic procedures (as might have been the case at Time 1).

Results from our exploratory longitudinal analysis showed that, in addition to symptom heterogeneity *across* children at any given time point (i.e., ASD classes), there appears to be heterogeneity *within* children over time. For example, Figures 5 and 6 depict the differential *ASD class pathways* indexed by "substantive movement" of children from a "higher severity class" (i.e., Class 3) at Time 1 to a "lower severity class" (i.e., Class A) at Time 2. This interesting finding provides support for the potential utility of a detailed investigation of the individual and contextual factors associated with within-child heterogeneity of symptom severity and functioning over time, the main objective of the *Pathways in ASD* study.

Implications and Limitations

This study has important implications for future research that could in turn produce evidence to inform clinical practice. In general, study findings support the use of the SCD and FIRB dimensions of the DSM 5 (as indexed here by the ADI-R factors) to classify children with ASD based on their level of symptom severity. However, at this point it is not clear what criteria will be used in the DSM 5 to assign children to one of the three severity levels, i.e., "requiring very substantial support," "requiring substantial support," or "requiring support" (APA, 2013). From a practical point of view, we suspect that finding a reliable and valid way of distinguishing that third (middle) subgroup characterised by "moderate" symptom severity (i.e., see Class 2 at Time 1; 11% of sample) might prove to be a challenge for researchers and clinicians. At the same time, the data-driven statistical classification of our ASD sample into a "higher severity class" (68% of sample) and a "lower severity class" (32% of sample) could serve as an example and/or methodological template of how such distinctions could be made.

Another implication is related to the demonstrated use of novel methodological approaches (i.e., FMM) as tools in our efforts to disentangle the heterogeneity of symptom presentation in complex disorders such as ASD. Moreover, study findings support the use of a dimensional approach to the description of the ASD phenotype but suggest that severity indices of core autism symptoms (i.e., SCD & FIRB) provide a better fit (statistically) if organized into two classes rather than as a single continuum. Finally, study findings support the realization that heterogeneity in symptom presentation changes both at the individual child level but also at the latent class level over time. The question that remains and needs to be addressed by future studies is how much of that change is due to true developmental manifestations of ASD (e.g., maturation effects or treatment received) and how much is related to measurement error in our assessment instruments and/or our analytic approaches (for a discussion on this see Editorial by Georgiades, Szatmari, and Boyle 2013).

Although the proposed empirical classification of children with ASD into a "higher severity class" and a "lower severity class" has the potential to inform clinical practice, statistical results from observational (i.e., non-experimental) studies cannot be directly interpreted as clinically meaningful. So despite the fact that this study contributes to our understanding of the measurement and classification of ASD, several limitations must be noted and used as guidelines for future research. The first limitation is the reliance on data from ADI-R, a semi-structured parent interview that was not designed specifically to quantify ASD symptom severity, or to capture the complexity of the developmental nature of these symptoms. Future research must examine the structure of the ASD phenotype using additional, observational measures such as the ADOS. The second limitation is the exclusion of ADI-R algorithm items that differ by age and/or verbal ability. This was done to ensure comparability of scores across ADI-R algorithm versions. It is likely that important phenotypic information linked to developmental level (as it relates to language) was not captured in the current analysis. Future studies could explore the possibility of "matching" items across ADI-R versions that are equivalent but differentially age-appropriate. The third limitation is related to the sampling of children diagnosed between ages 2 and 5. It is possible that children who receive a diagnosis after the age of 5 might have a very different phenotypic profile that does not fit any of the two classes proposed here. Fourth, as noted earlier, the children in the subsample used in the current longitudinal analysis had significantly lower scores on the SCD, FIRB, and were diagnosed at an older age compared to those for whom no data were available at 6 years of age. Although this raises concerns about the effect of missing data on the estimation of the factor mixture models, post-hoc analysis showed that the distribution of children in the three ASD classes at Time 1 was similar across the two subsamples (i.e. those included in longitudinal analysis versus those who had only Time 1 data). Fifth, although the study documents change patterns in the underlying latent class structure of the ASD phenotype, it does not examine the factors that might be associated with those patterns. Future studies must carefully explore the associations (and if possible causal mechanisms) between changes in the underlying structure of symptoms and individual child and contextual factors (i.e., family, services received, etc.).

Conclusions

Study findings suggest that there is a change, statistically, in the underlying latent class structure of autism symptoms during the first few years after diagnosis. Specifically, it appears that in our cohort assessed shortly after diagnosis (ages 2 to 5), symptom structure can be described using three latent classes, but that by age 6 this structure can be modeled with only two distinct, relatively homogeneous classes (subgroups) of children with ASD. These findings underscore the developmental nature of ASD and provide support to the idea of repeated classification assessments as children develop, especially at key points of transition (e.g., transition into the school system around age 6). The results also lend support to a much-needed shift in our conceptual and methodological approach to the study of measurement and classification of autism pathology: that is,

instead of a set of categorical symptoms that present early in childhood and remain static over the life span, ASD might be better understood as a complex and dynamic disorder, structured on both categorical and dimensional constructs that vary not only across individuals at any given point, but also within individuals across time.

References

Achenbach, T. M., & Rescorla, L. A. (2000). *Manual for ASEBA preschool forms & profiles.* Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.

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

American Psychiatric Association (2013). Diagnostic and Statistical Manual of Mental Disorders (Fifth ed.). Arlington, VA: American Psychiatric Publishing.

Bodfish, J. W., Symons, F., & Lewis, M. (1999). *The Repetitive Behavior Scale: Test manual.* Morganton: Western Carolina Center Research Reports.

Bodfish, J. W., Symons, F. J., Parker, D. E., & Lewis, M. H. (2000). Varieties in repetitive behavior in autism. *Journal of Autism and Developmental Disorders, 30,* 237–243.

Constantino, J. N., & Gruber, C. P. (2005). *Social Responsiveness Scale*. Los Angeles: Western Psychological Services.

Frazier, T.W., Youngstrom, E.A., Sinclair, L., Kubu, C.S., Law, P., Rezai, A., ... & Eng, C. (2010). Autism spectrum disorders as a qualitatively distinct category from typical behavior in a large, clinically ascertained sample. *Assessment*, *17,* 308-320.

Frazier, T.W., Youngstrom, E.A., Speer, L., Embacher, R., Law, P., Constantino, J., … & Eng, C. (2012). Validation of proposed DSM-5 criteria for autism spectrum disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 51,* 28–40.

Georgiades, S., Szatmari, P., Boyle, M. (2013) EDITORIAL: The Importance of Studying Heterogeneity in Autism, *Neuropsychiatry: Future Medicine. Vol. 3, No. 2,* 123-125 , DOI 10.2217/npy.13.8

Georgiades, S., Szatmari, P., Boyle, M., Hanna, S., Duku, E., Zwaigenbaum, L., ... & Pathways in ASD Study Team. (2013). Investigating phenotypic heterogeneity in children with autism spectrum disorder: a factor mixture modeling approach. *Journal of Child Psychology and Psychiatry, 54,* 206-215.

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

Happe, F. (2011). Criteria, categories, and continua: Autism and related disorders in DSM-5. *Journal of the American Academy of Child & Adolescent Psychiatry, 50,* 540–542.

Huerta, M., Bishop, S.L., Duncan, A., Hus, V. & Lord, C. (2012). Application of DSM-5 criteria for autism spectrum disorder to three samples of children with DSM-IV diagnoses of pervasive developmental disorders. *American Journal of Psychiatry, 169,* 1056-1064.

Lord, C., & Jones, R. M. (2012). Annual research review: Rethinking the classification of autism spectrum disorders. *Journal of Child Psychology and Psychiatry, 53,* 490–509.

Lubke, G.H., & Muthen, B.O. (2005). Investigating population heterogeneity with factor mixture models. *Psychological Methods, 10,* 21–39.

Mandell, D. (2011). The heterogeneity in clinical presentation among individuals on the autism spectrum is a remarkably puzzling facet of this set of disorders. *Autism : The International Journal of Research and Practice, 15*(3), 259-261.

Muthen, B. (2004). Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data; In D Kaplan (Ed.), *Handbook of quantitative methodology for the social sciences* (pp. 345–368). Sage: Thousand Oaks.

Nylund, K.L., Asparouhov, T., & Muthen, B. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling. A Monte Carlo simulation study. *Structural Equation Modeling, 14,* 535-569.

Ozonoff, S. (2012). Editorial perspective: autism spectrum disorders in DSM-5 – an historical perspective and the need for change. *Journal of Child Psychology and Psychiatry, 53,* 1092-1094.

Pandolfi, V., Magyar, C. I., & Dill, C. A. (2009). Confirmatory factor analysis of the child behavior checklist 1.5-5 in a sample of children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 39,* 986-995.

Robinson, E. B., Lichtenstein, P., Anckarsater, H., Happe, F. & Ronald, A. (2013). Examining and interpreting the female protective effect against autistic behavior. *Proceedings of the National Academy of Sciences of the United States of America: Psychological and Cognitive Sciences*, doi: 10.1073/pnas.1211070110.

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

Rutter, M. (2011). Research review: child psychiatric diagnosis and classification: concepts, findings, challenges and potential. *Journal of Child Psychology and Psychiatry, 52,* 647–660.

Rutter, M. (2012). Changing concepts and findings on autism. *Journal of Autism and Developmental Disorders,* doi: 10.1007/s10803-012-1713-7

Sparrow, S.S., Cicchetti, D.V., & Balla, D.A. (2005). *Vineland Adaptive Behavior Scales: Second Edition (Vineland II), Survey Interview Form/Caregiver Rating Form.* Livonia, MN: Pearson Assessments.

Walton, K. E., Ormel, J., & Krueger, R. F. (2011). The dimensional nature of externalizing behaviors in adolescence: Evidence from a direct comparison of categorical, dimensional, and hybrid models. *Journal of Abnormal Child Psychology, 39*, 553–561.

Wiggins, L. D., Robins, D. L., Adamson, L. B., Bakeman, R., & Henrich, C. (2011). Support for a dimensional view of autism spectrum disorders in toddlers. *Journal of Autism and Developmental Disorders, 42,* 191–200.

Table 1. Descriptive statistics for Sample 1 (Pathways sample) and Sample 2 (AGP sample)

	Time 1 (at diagnosis)	Time 2 (at age 6)	
	Sample 1 **(N=280)**	**Sample 1** **(N=280)**	**Sample 2** **(N=517)**
Child's sex			
Male	86%	86%	83.4%
Female	14%	14%	16.6%
Child's mean age in months (SD)	40.3 (9.1)	79.4 (4.0)	76.8 (3.4)
SCD mean score (SD)	1.29 (.38)	1.31 (.45)	1.41 (.38)
FIRB mean score (SD)	.95 (.42)	1.01 (.46)	1.06 (.44)
SCD internal consistency – alpha (20 items)	.87	.91	.87
FIRB internal consistency – alpha (6 items)	.49	.52	.46

Notes: SCD: Social Communication Deficit; FIRB: Fixated Interests and Repetitive Behaviour; SD: Standard Deviation

Table 2. ADI-R algorithm item composition of the SCD and FIRB symptom dimensions

SCD dimension	**FIRB dimension**
1 Direct gaze (item 50)	**21** Unusual preoccupations (item 67)
2 Social smiling (item 51)	**22** Compulsions/rituals (item 70)
3 Range of social expressions used to communicate (item 57)	**23** Hand and finger mannerisms (item 77)
4 Interest in children (item 62)	**24** Other complex mannerisms or stereotyped body movements (item 78)
5 Response to approaches of other children (item 63)	**25** Repetitive use of objects or interest in parts of objects (item 69)
6 Showing and directing attention (item 52)	**26** Unusual sensory interests (item 71)
7 Offering to share (item 53)	
8 Seeking to share enjoyment with others (item 54)	
9 Use of other's body to communicate (item 31)	
10 Offering comfort (item 55)	
11 Quality of social overtures (item 56)	
12 Inappropriate facial expressions (item 58)	
13 Appropriateness of social responses (item 59)	
14 Pointing to express interest (item 42)	
15 Nodding (item 43)	
16 Head shaking (item 44)	
17 Conventional/instrumental gestures (item 45)	
18 Spontaneous imitation of actions (item 47)	
19 Imaginative play (item 48)	
20 Imitative social play (item 61)	

Notes: ADI-R: Autism Diagnostic Interview-Revised; In the factor mixture modeling (FMM) analysis, the first 20 items were "forced" to load on the Social Communication Deficits (SCD) dimension; the remaining 6 items loaded on the Fixated Interests and Repetitive Behaviours (FIRB) dimension.

Table 3. ASD symptom factor mixture models, fit indices, and class proportions from time of diagnosis (Time 1) to age 6 (Time 2).

|	Sample	Number of Factors (f) & Classes (c)	Log Likelihood	Number of Free Parameters	AIC	BIC	Adjusted BIC	Class percentages
DIAGNOSIS (Time 1)	**Sample 1** (mean age 40.3 months)	2f1c	-7353.960	78	14863.920	15147.434	14900.101	100%
		2f2c	-7321.904	81	14805.808	15100.226	14843.380	30%, 70%
		2f3c	**-7276.529**	**84**	**14721.058**	**15026.381**	**14764.022**	**35%, 11%, 54%**
		2f4c	no convergence					
AGE 6 (Time 2)	**Sample 1** (mean age 79.4 months)	2f1c	-7458.442	78	15072.883	15356.397	15109.064	100%
		2f2c	**-7359.711**	**81**	**14881.421**	**15175.839**	**14918.994**	**32%, 68%**
		2f3c	no convergence					
	Sample 2 (mean age76.8 months)	2f1c	-13864.461	78	27884.922	28216.270	27968.683	100%
		2f2c	**-13785.544**	**81**	**27733.088**	**28077.179**	**27820.070**	**31%, 69%**
		2f3c	no convergence					

Note: LCA: FMM: factor mixture model; AIC: Akaike Information Criterion; BIC: Bayesian Information Criterion. The best fitting models are presented in bold font. Sample 1: Pathways in ASD (N=280); Sample 2: Autism Genome Project (N=517).

Table 4. Means, standard deviations and significance levels for the two ASD classes on variables of interest (Time 2, age 6, Sample 1; N=280)

	Class A (32% sample)	Class B (68% sample)	
			p value
Child' sex			
Male	27.4%	72.6%	<0.01
Female	61.5%	38.5%	
Child's mean age at diagnosis (SD)	39.4 (8.7)	38.6 (8.7)	.48
SCD mean score (SD)	1.0900 (.45)	1.42 (.41)	<0.01
FIRB mean score (SD)	.65 (.36)	1.19 (.40)	<0.01
SRS total t-score (SD)	66.24 (14.43)	72.25 (14.62)	<0.01
RBS-R total mean score (SD)	.32 (.35)	.52 (.41)	<0.01
VABS adaptive behaviour composite score (SD)	80.57 (12.10)	75.03 (14.53)	<0.01
CBCL total problems t-score (SD)	48.75 (12.71)	54.87 (12.15)	<0.01

Notes: ASD: Autism Spectrum Disorder; ADI-R: SCD: Social Communication Deficit; FIRB: Fixated Interests and Repetitive Behaviour; SRS: Social Responsiveness Scale; RBS-R: Repetitive behaviors Scale – Revised; VABS II: Vineland Adaptive Behavior Scales, Second Edition; CBCL: Child Behavior Checklist; SD: Standard Deviation

Fig. 1 Factor mixture model of the ASD symptom phenotype with '2-factors/3-classes' at Time 1 (mean age 40.3 months; Sample 1, N = 280). The horizontal axis represents the 26 algorithm items from the ADI-R. Items 1–20 load on the SCD factor and items 21–26 load on the FIRB factor. The vertical axis represents the mean class score for each item.
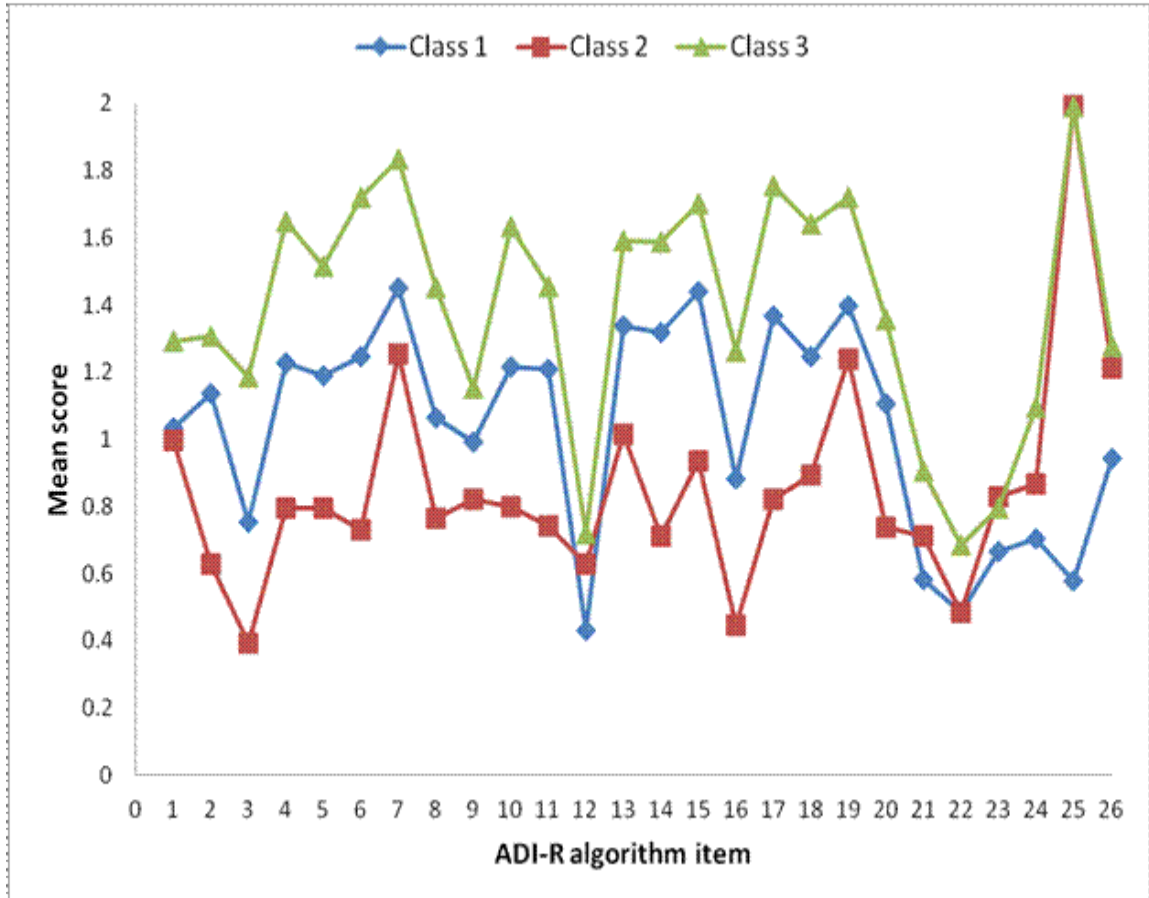
Fig. 2. Scatterplot of SCD by FIRB scores for the three derived ASD classes at diagnosis (Time 1)
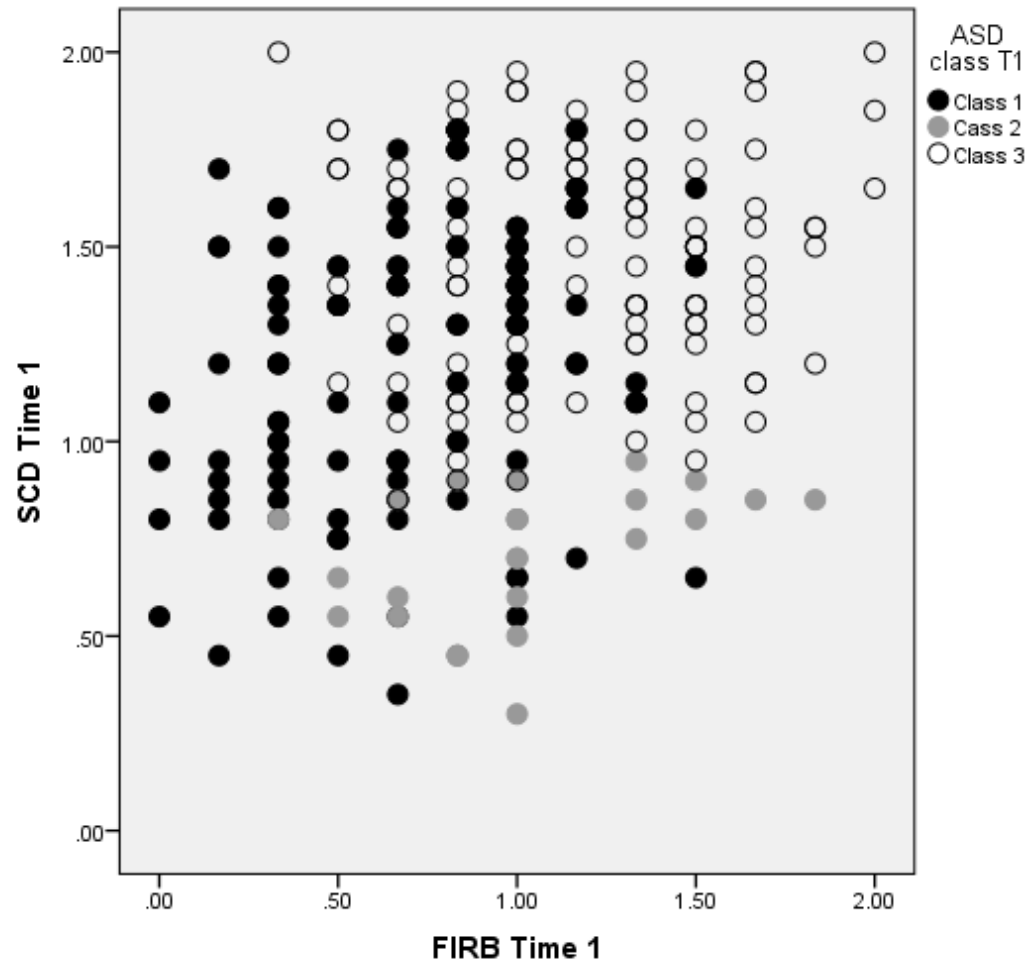
Fig. 3. Factor mixture model of the ASD symptom phenotype with '2-factors/2-classes' at Time 2 (mean age 79.4 months; Sample 1, N = 280). The horizontal axis represents the 26 algorithm items from the ADI-R. Items 1–20 load on the SCD factor and items 21–26 load on the FIRB factor. The vertical axis represents the mean class score for each item.

Fig. 4. Scatterplot of SCD by FIRB scores for the two derived ASD classes at age 6 (Time 2, Sample 1)
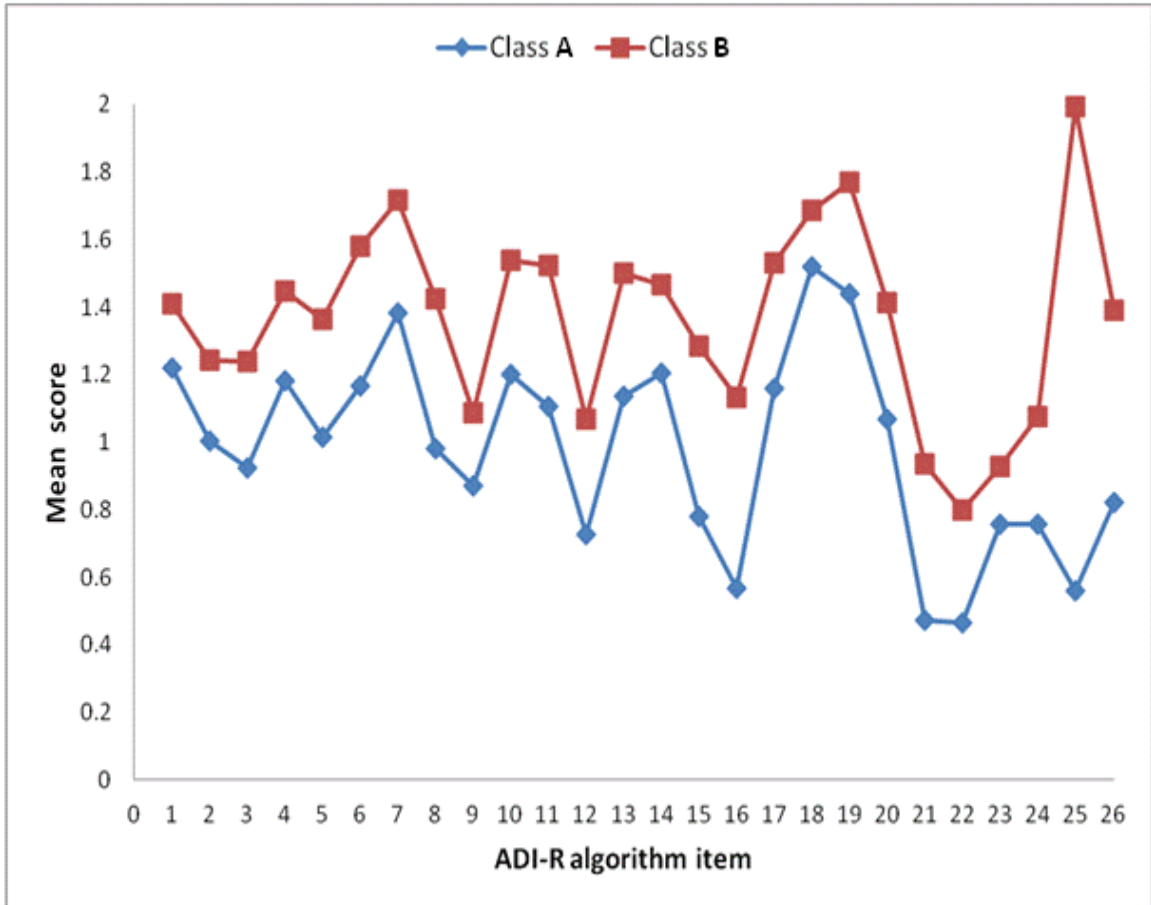
Fig. 5. Schematic of *ASD latent class membership* and *ASD class pathways* from Time 1 (at diagnosis; mean age 40.3 months) to Time 2 (mean age 79.4 months; N=280)
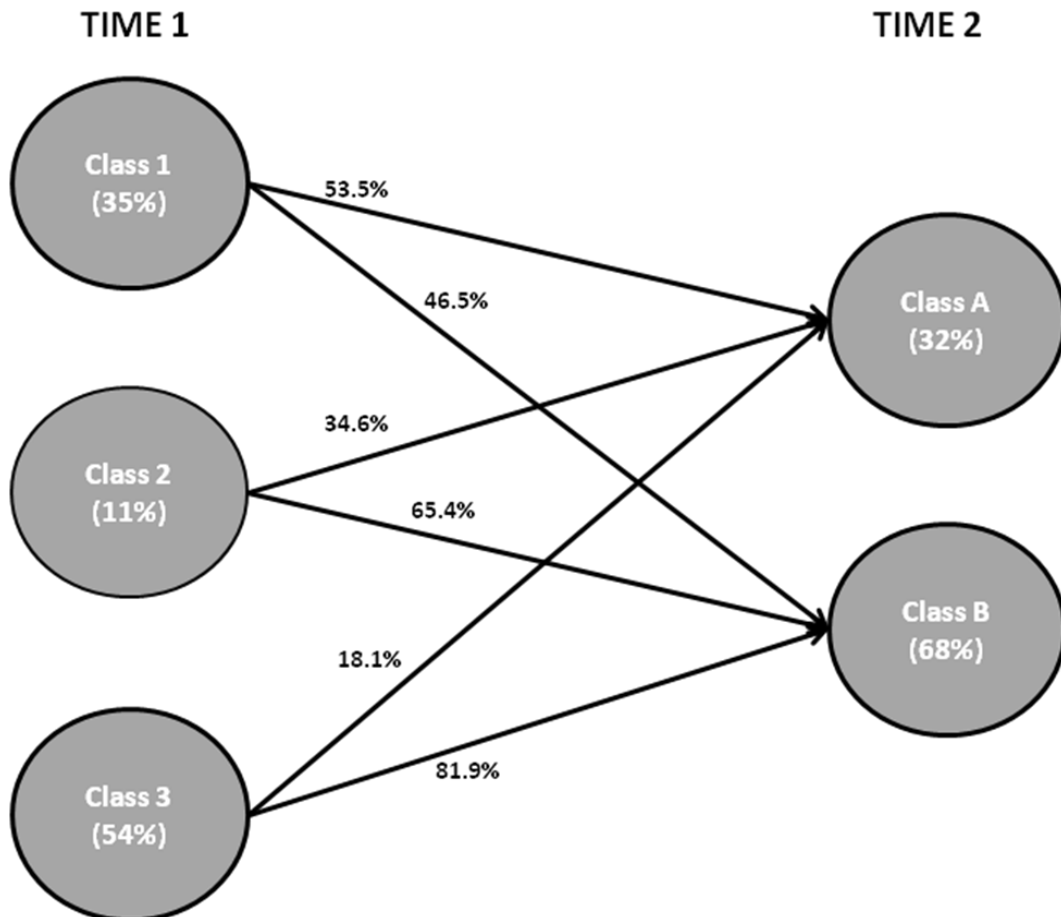
Fig. 6. a & b Longitudinal change (T1 to T2) in SCD and FIRB symptom severity for the six *ASD class pathways* (N=280)
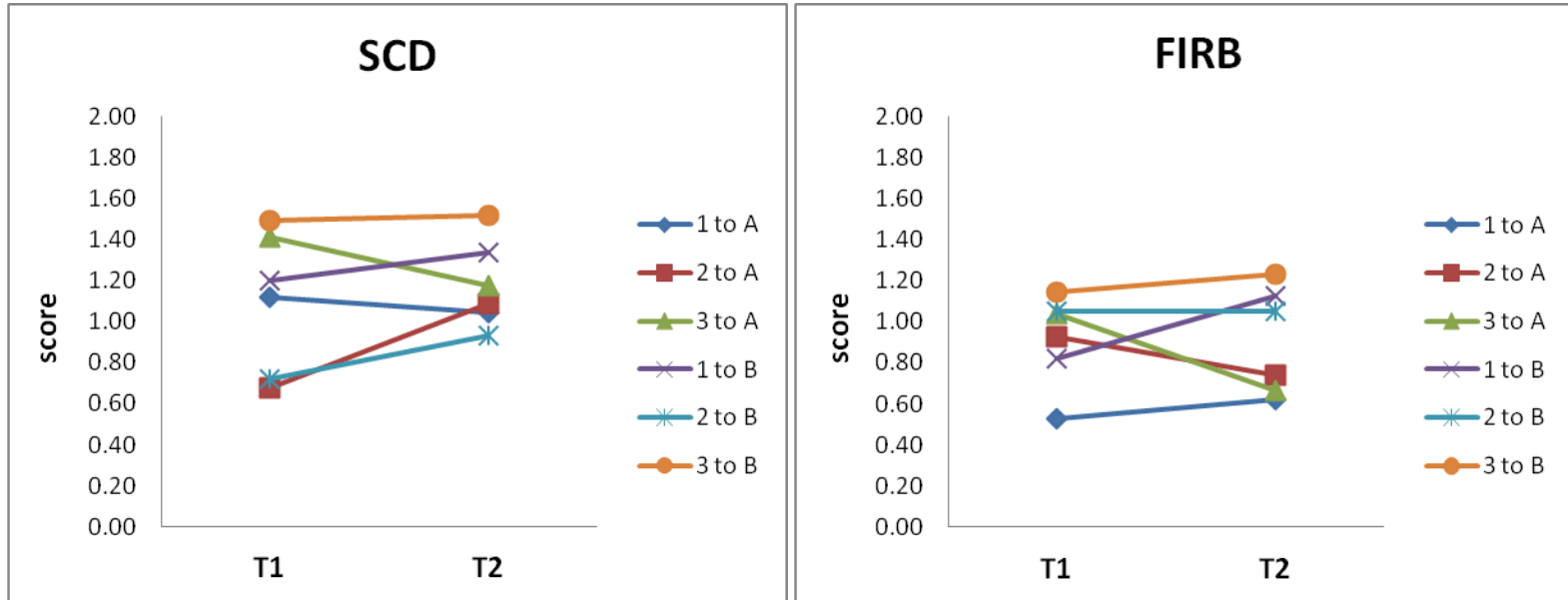
Class Pathways (% of sample)	SCD severity change	FIRB severity change
1 to A (18.9 %)	Stable	Stable
2 to A (3.2%)	Major Increase	Major Decrease
3 to A (10%)	Minor Decrease	Major Decrease
1 to B (16.4%)	Minor Increase	Major Increase
2 to B (6.1%)	Increase	Decrease
3 to B (45.4 %)	Minor Decrease	Minor Increase

CHAPTER SIX
EDITORIAL

TITLE: The Importance of Studying Heterogeneity in Autism

AUTHORS: Stelios Georgiades[1], Peter Szatmari[1], Michael Boyle[1]

[1]Offord Centre for Child Studies, McMaster University

CONTEXT AND IMPLICATIONS: This Editorial identifies the major limitations of previous research (including the four empirical papers in the current Thesis) and offers recommendations for studying heterogeneity in autism. The main message is that the time is right for a much-needed shift in our conceptual and methodological approach to the study of measurement and classification of autism pathology: that is, instead of a set of categorical symptoms that present early in childhood and remain static over the life span, ASD might be better understood as a complex and dynamic disorder, structured on both categorical and dimensional constructs that vary not only across individuals at any given point, but also within individuals across time.

CONFLICTS OF INTEREST: None

EDITORIAL: The Importance of Studying Heterogeneity in Autism

Among the child-onset psychiatric disorders, autism is perhaps the most serious, intractable and challenging to address. One of the factors contributing to this challenge is its heterogeneity observed along a spectrum of pathology [1,2]. The current diagnostic and classification system (DSM-IV) classifies children with autism in one of three subtypes − Autistic Disorder, Asperger's Disorder, or Pervasive Developmental Disorder Not Otherwise Specified. However, the inability to establish the reliability and validity of these subcategories empirically is moving the DSM towards replacing the existing subtypes with a severity gradient under the diagnostic umbrella of one Autism Spectrum Disorder [201]. The proposed changes to the way we classify autism may represent a scientific advance in how we might understand this condition but it is also a sobering reminder that, despite progress, our knowledge about autism is both fragile and sparse. Categorizing the clinical heterogeneity in children with autism is still of critical importance, regardless of how the DSM changes its definition [3]. Unfortunately, the indicators that we use to represent autism as a heterogeneous condition come from a mixture of fallible inferences and observations vulnerable to error − both systematic and random. Our failure to identify valid and reliable biological markers or other indicators less prone to error represents a serious impediment that needs to be addressed in future research. Therefore, we argue here that a better understanding of heterogeneity in autism itself could generate useful information for the study of aetiology, diagnosis, treatment, and prognosis of the disorder [4].

But what do we mean by heterogeneity? Heterogeneity denotes diversity or variability; it describes dissimilar parts that are somehow connected. We think of autism as a disorder that causes deficits in patterns of cognitive, emotional, behavioural and social functioning that are manifested differently across subgroups of children. This generic, lay definition of heterogeneity provides a descriptive foundation for building a scientific framework to systematically study heterogeneity in autism. First, autism can be associated with a diversity of functional qualities − i.e., some children with autism are verbal, others are nonverbal; some have high IQ, others have low IQ. Second, autism can be conceptualized as symptom configurations from different domains, exhibiting different severity levels − i.e., some children with autism have severe social communication deficits and mild fixated interests and repetitive behaviours; other children exhibit the reverse profile. Some children present with "comorbid" or "associated" symptoms (anxiety, attention deficits, etc.) while other children primarily exhibit only core autistic symptoms [5]. Third, contrary to previous theories, recent findings show that autism is a disorder resulting from diverse causes − i.e., updated genetic findings identify multiple genetic variants both at the same and different loci as being associated with autism [6] and recent twin studies suggest that, in addition to genes, environmental factors play an important role in the causal mechanisms of the disorder [7]. Fourth, autism is perhaps a classic example of a heterogeneous disorder in which dissimilar parts are somehow connected − i.e., despite the differences described above (functional qualities, symptom

type and severity, causal factors, etc.) autism is still viewed as one entity, with all affected individuals placed within a spectrum of pathology – the autism spectrum disorder [201]. So the widely-accepted (but understudied) picture of autism as a heterogeneous disorder appears to be a valid one. Interestingly, at a time when scientists and policy makers are discussing the idea of personalized medicine for other disorders [8], clinicians and therapists of children with autism are still having a very hard time answering pressing questions from parents related to individualized treatment and specific outcomes. Although we know there is variability in prognosis, in a comprehensive review of early intervention studies, Warren et al. [9] concluded that our ability to predict response to treatment and outcome is currently very limited and warrants further investigation.

We believe the time is right for a more scientifically rigorous approach that will lead to a better understanding of autism heterogeneity. Such an approach would not be based on the arbitrary classification of static diagnostic subtypes but rather on the systematic evaluation of the clinical and research utility of phenotypic and genotypic markers that vary across subgroups of children. This will be of particular importance as we move into a new generation of autism research studies. After decades of research using single-method/design case studies we now find ourselves entering an era of autism research with large, costly studies involving multiple methods and technologies (phenotypic, cognitive/experimental, genetics, epigenetics, genomics, neuroimaging, pharmacogenetics, randomized control trials, etc.); these studies are aiming to not only "describe the clinical picture" but also understand the "underlying mechanisms" associated with causation, manifestation, development, and response to treatment in individuals with autism [10].

Although notable progress has been achieved, the integration and interpretation of data from multi-method, multi-design studies of autism has proved to be a major challenge. As a general rule (that could of course be confirmed by a limited number of exceptions) these ambitious research studies have "failed" to find strong and/or replicable effects. Some of the usual explanations for this phenomenon are related to methodological issues – small sample size, assessment and measurement, etc. – that become even more complex by the heterogeneous nature of the disorder. We think a new research paradigm is needed as we move forward: rather than conducting studies that compare "autism cases" with "typically developing individuals" we have to focus on understanding the meaning of individual and subgroup differences within the autism spectrum. For this new research paradigm to be successful, future studies must focus on the development and evaluation of appropriate measures that could be used to (a) operationalize autism as a heterogeneous entity; and (b) collect data to evaluate the reliability, validity, and utility of this new conceptualization of autism. Such measures need to be equivalent across subgroups of interest (i.e., children and youth, males and females, verbal and non-verbal, severe and mild cases of autism); these measures also need to be "sensitive to change" and have the ability to "capture" possible treatment effects.

In closing, we highlight the importance of studying heterogeneity in autism itself and propose a conceptual and methodological shift to future research: instead of viewing heterogeneity as a post-hoc, observable outcome of our generic measurements (most of which were originally designed to distinguish autism from non-autism cases), we believe heterogeneity could provide a general framework that will guide the development, implementation, and interpretation of new study designs and measurements and that will have the ability to "capture" individual and subgroup differences within autism; ultimately, these differences should be robust enough to provide informative "links" between the different levels of autism – i.e., phenotype and genotype – and account for a substantial amount of the variability observed in studies of autism causes, diagnosis, treatment, and prognosis.

References

1   Georgiades S, Szatmari P, Boyle M *et al.* Investigating phenotypic heterogeneity in children with autism spectrum disorder: a factor mixture modeling approach. *J. Child Psychol. Psyc.* 54(2), 206-215 (2013).

2   Geschwind D. Advances in autism. *Annu. Rev. Med.* 60, 367-380 (2009).

3   Mandell D. Editorial. *Autism* 15(3), 259-261(2011).

4   Szatmari P. New recommendations on autism spectrum disorder. *Brit. Med. J.* 342:d2456 (2011).

5   Georgiades S, Szatmari P, Duku E *et al.* Phenotypic overlap between core diagnostic features and emotional/behavioral problems in preschool children with autism spectrum disorder. *J. Autism Dev. Disord.* 41(10), 1321-1329 (2011).

6   Sullivan PF, Daly MJ, O'Donovan M. Genetic architectures of psychiatric disorders: the emerging picture and its implications. *Nat. Rev. Genet.* 13(8), 537-551 (2012).

7   Hallmayer J, Cleveland S, Torres A *et al.* Genetic heritability and shared environmental factors among twin pairs with autism. *Arch. Gen. Psychiat.* 68(11), 1095-1102 (2011).

8   Ruberg SJ, Chen L, Wang Y. The mean does not mean as much anymore: finding sub-groups for tailored therapeutics. *Clin. Trials* 7(5), 574-583 (2010).

9   Warren Z, McPheeters ML, Sathe N, Foss-Feig JH, Glasser A, Veenstra-VanderWeele J. A systematic review of early intensive intervention for autism spectrum disorders. *Pediatrics* 127(5), 1303–1311 (2011).

10   Dawson G. Dramatic increase in autism prevalence parallels explosion of research into its biology and causes. *JAMA Psychiatry* 70(1), 9-10 (2013).

Websites

201   American Psychiatric Association. Report of the DSM-V neurodevelopmental disorders work group. http://www.dsm5.org/progressreports/pages/0904reportofthedsm-vneurodevelopmentaldisordersworkgroup.aspx

CHAPTER SEVEN
DISCUSSION

Autism Spectrum Disorder (ASD) is a common disorder with a high burden of suffering and economic cost to society. According to Dawson (2013) the dramatic increase in autism prevalence has led to an explosion of research into its biology and causes. However, despite the significant progress achieved through research to date, our understanding of what causes the disorder and how the disorder unfolds over time is frustratingly limited (Rutter, 2012). One of the factors contributing to this challenge is the heterogeneity – in symptoms (i.e., social communication deficit and fixated interest and repetitive behaviours) and behaviours (i.e., withdrawl, emotional reactivity, etc.) – observed along the spectrum of autism pathology (Mandell, 2011).

The current *Thesis* represents a much needed systematic attempt to investigate ASD heterogeneity, as it relates to the measurement and classification of the clinical phenotype. Specifically, while all children with ASD share similar features that place them within the same "spectrum" they also exhibit notable differences in symptom severity and configuration, behavioural commorbidities, IQ, and level of functioning (i.e., language and daily living skills).

Currently, our ability to distinguish children with ASD from those without the disorder is quite good. However, the classification of children within the autism spectrum into subtypes has proven to be a challenge (Happe, 2011). This inability to "unpack" the clinical heterogeneity in meaningful and reliable ways might reflect the current state of knowledge related to measurement and classification of ASD. Unfortunately, in the absence of robust biological markers for the disorder, the conventional indicators we use to measure the ASD phenotype are derived from behavioural data (semi-structured interviews with parents, direct observation of the child, questionnaires, etc.) vulnerable to error. Moreover, our understanding of how these phenotypic indicators come together to form the variable manifestations of ASD is limited.

Through four interrelated empirical studies applying three different but complimentary methodological approaches – i.e., *variable-centered*, *person-centered*, and combined *variable/person-centered* – the current *Thesis* explored the underlying associations across phenotypic indicators (symptoms, traits, behaviours) conventionally used to measure ASD.

Study 1 (Georgiades et al., 2011) examined the phenotypic overlap between core diagnostic features and emotional/behavioral problems in a sample of 335 newly diagnosed preschool children with ASD. Results suggested substantial phenotypic overlap between conventional diagnostic symptoms and problem behaviours in children with ASD, suggesting that ASD might in fact "share" phenotypes with other child-onset neurodevelopmental disorders such as ADHD and OCD (POND Network, 2013). This

exploratory study contributes to our understanding of the complexity of the ASD phenotype and supports the inclusion of additional "clinical specifiers" (e.g., emotional/behavioral problems) as part of the diagnostic characterization of children with ASD (APA, 2013). In a way, these findings challenge the current definition of ASD which is based only on "core symptoms" (i.e., SCD & FIRB) and call for a systematic investigation that will evaluate the utility of an expanded set of indicators – emotional/behavioural problems and possibly IQ – as part of a more comprehensive measurement and classification framework for ASD. Future research on this topic will hopefully provide evidence on whether these additional indicators often seen in children with ASD are "fundamental features" of the disorder or whether they should be kept as "external", "associated features" of ASD.

Study 2 (Georgiades et al., 2013) investigated the emergence of autistic-like traits in unaffected (no ASD diagnosis) infant siblings of probands diagnosed with ASD. Two groups of children unaffected with ASD were assessed prospectively – 170 high-risk siblings of probands diagnosed with ASD and 90 low-risk controls with no family history of ASD. Results showed that within the unaffected siblings there is a subgroup of children (19%) who exhibit autistic-like traits resembling a "broader autism phenotype" at the very young age of 12 months. These children go on to have elevated scores on social-communication impairment and internalizing problems by age 3. This study highlights the importance of close monitoring of later born infants in high-risk families with another child diagnosed with ASD. At the same time, it is important to note that although early detection of sub-threshold autistic-like traits/behaviours in high-risk populations can be useful from a prevention point of view, it comes with certain risks and limitations. For example, in the absence of biological markers of the disorder, the use of behavioural assessments (such as the AOSI) for the identification of elevated levels of symptoms and/or behaviours in 12-month old high-risk children is likely subject to increased measurement bias and error. Moreover, one must be careful with "labeling" these children as having a (or any) form of autism (i.e., broader autism phenotype) because of the stigma associated with this terminology. This issue could be exacerbated in the absence of effective interventions for these milder deficits believed to be associated with autism during the first years of life.

Study 3 (Georgiades et al., 2013) examined the underlying structure of the DSM 5 core symptom domains – Social Communication Deficits (SCD) and Fixated Interests and Repetitive Behaviours (FIRB). By analyzing SCD and FIRB data on 391 newly diagnosed children this study attempted to answer the question of whether ASD is best represented by a categorical (i.e., classes), dimensional (i.e., dimensions), or hybrid (i.e., subgroups and dimensions) measurement model. Results showed that ASD can be best captured by a factor mixture model that describes the ASD phenotype using three classes (Class 1: 34%, Class 2: 10%, Class 3: 56% of the sample) based on differential severity gradients on the SCD and FIRB symptom dimensions. Children from these classes were diagnosed at different ages and were functioning at different adaptive, language, and

cognitive levels. This study suggests that the two symptom dimensions of SCD and FIRB in the DSM 5 can be used to stratify children with ASD empirically into three relatively homogeneous classes/subgroups.

Study 4 (Georgiades et al., under review) examined if the measurement model found to best represent the latent class structure of core autism symptoms at diagnosis (ages 2 to 5; see Study 3) is replicable at age 6. Results showed that at age 6 a more parsimonious model with only 2 (instead of 3) classes and 2 dimensions provided the best fit to the data (Class A=32%, Class B=68% of the sample). According to this factor mixture model, 6-year old children with ASD can be classified in two distinct classes characterized by significantly different levels of severity on the SCD and FIRB symptom dimensions, different adaptive functioning skills and emotional/behavioural problems. Furthermore, there is a difference in the way boys and girls are distributed across the two ASD classes. Finally, children across the two classes did not differ in terms of the age at which they were diagnosed. Findings from this study suggest that there is a substantive change in the phenotypic architecture of autism symptoms during the first few years after diagnosis. Specifically, it appears that there is a reduction in symptom heterogeneity in children with ASD from the time of diagnosis to age 6. These findings demonstrate the dynamic nature of the ASD phenotype and speak to the importance of repeated classification assessments of symptoms, functional skills, and behaviours as children develop.

The **Editorial** (Georgiades, Szatmari, & Boyle, 2013) identifies the major limitations of previous research and offers recommendations for studying heterogeneity in autism. Points made in this Editorial serve as the foundation for the section "Implications and Future Research" below.

General Limitations

Despite the contributions of this *Thesis* to our understanding of the measurement and classification of the heterogeneous ASD phenotype, findings from the four empirical studies presented here must be interpreted with caution and within the context of the following overarching limitations (for more details on specific limitations please see the Discussion section in each of the four empirical thesis papers).

First, all studies used assessment measures and instruments that were originally designed for diagnostic purposes; that is to distinguish children with ASD from those without the disorder. Therefore, findings related to heterogeneity and variability *within* the ASD group are limited by the scope of these diagnostic measures. Second, although statistical techniques (factor analysis, cluster analysis, factor mixture modeling) can be useful in deriving empirical subgroups of children with (or at-risk for) ASD, we cannot conclude that these subgroups are valid and/or meaningful until we can prove their utility – in predicting aetiology, prognosis and/or response to treatment. Third, in the statistical

techniques used in this *Thesis*, despite the reliance on widely accepted indices/criteria, the selection of the "best fitting model" is subject to error and bias related to factors such as initial item/indicator pool, sample characteristics (size, clinical versus high-risk, versus population), data collection and report bias (i.e., parent report, clinician bias, etc.). Fourth, the measurement of symptoms, traits, and behaviours associated with autism was based on the assumption that the assessment instruments were equivalent across subgroups within the study samples – i.e., in males and females, in verbal and non-verbal children, in younger and older children, among others.

Implications and Future Research

Taken together, this *Thesis* has provided empirical evidence for moving the field of measurement and classification of ASD forward. Specifically, findings presented in the current *Thesis* shed light on the underlying structure of the ASD phenotype and the complex ways autism-related variables (symptoms, traits, behaviours) come together to make up the heterogeneous clinical profiles seen in preschool children with (or at-risk for) the disorder. By better understanding the phenotypic architecture of the ASD phenotype we can begin to refine our measurement classification models in order to identify meaningful subgroups of children who share homogeneous clinical profiles that might in turn be associated with different etiological factors, response to treatment, and/or developmental outcomes.

The work presented in this *Thesis* has implications both for research and clinical practice. In general, the use of statistical methods to identify relatively homogeneous subgroups of children with similar profiles of symptoms and/or behaviours contributes to our understanding of the heterogeneous autism phenotype. However, it is important to note that the potential utility of these subgroups was not investigated here and thus needs to be tested in future research. For example, future studies could examine if children from different ASD subgroups respond differently to different treatments; they could also test the utility of these subgroups in aetiological studies in search of different genetic factors associated with ASD. It will be important to see whether children from different subgroups follow different developmental trajectories as they get older (Szatmari, 2011). Finally, since ASD is a dynamic disorder with different manifestations at different developmental points, it will be interesting to test the notion that, as they develop, children might move from on subgroup to another. Once such movement is identified we can look at the factors potentially associated with classification change patterns.

The overarching theme that arises from the current *Thesis* is that the time is right for a more scientifically rigorous approach that will lead to a better understanding of autism heterogeneity. Such an approach would not be based on the arbitrary classification of static diagnostic subtypes but rather on the systematic evaluation of the clinical and research utility of phenotypic markers that vary across subgroups of children. The current *Thesis* also lends support to a much-needed shift in our conceptual and methodological

approach to the study of measurement and classification of autism pathology: that is, instead of a set of categorical symptoms that present early in childhood and remain static over the life span, ASD might be better understood as a complex and dynamic disorder, structured on both categorical and dimensional constructs that vary not only across individuals at any given point, but also within individuals across time.

A new research paradigm is needed as we move forward: rather than conducting studies that compare "autism cases" with "typically developing individuals" we have to focus on understanding the meaning of individual and subgroup differences within the autism spectrum. For this new research paradigm to be successful, future studies must focus on the development and evaluation of appropriate measures that could be used to (a) operationalize autism as a heterogeneous entity; and (b) collect data to evaluate the reliability, validity, and utility of this new conceptualization of autism. Such measures need to be equivalent across subgroups of interest (i.e., children and youth, males and females, verbal and non-verbal, severe and mild cases of autism); these measures also need to be "sensitive to change" and have the ability to "capture" possible treatment effects.

In closing, it is proposed that instead of viewing ASD heterogeneity as a post-hoc, observable outcome of our generic measurements (most of which were originally designed to distinguish autism from non-autism cases), heterogeneity could be used as a general framework that will guide the development, implementation, and interpretation of new study designs. Such studies will use appropriate measurements that will have the ability to "capture" individual and subgroup differences within autism; ultimately, these differences should be robust enough to provide informative "links" between the different levels of autism – i.e., phenotype and genotype – and account for a substantial amount of the variability observed in studies of autism causes, diagnosis, treatment, and prognosis.

References

Dawson, G. (2013). Dramatic increase in autism prevalence parallels explosion of research into its biology and causes. *JAMA Psychiatry (Chicago, Ill.), 70*(1), 9-10. doi: 10.1001/jamapsychiatry.2013.488; 10.1001/jamapsychiatry.2013.488

Georgiades, S., Szatmari, P., Boyle, M. (2013) EDITORIAL: The Importance of Studying Heterogeneity in Autism, *Neuropsychiatry: Future Medicine. Vol. 3, No. 2,* 123-125 , DOI 10.2217/npy.13.8

Georgiades, S., Szatmari, P., Boyle, M., Hanna, S., Duku, E., Zwaigenbaum, L., . . . Pathways in ASD Study Team. (2013). Investigating phenotypic heterogeneity in children with autism spectrum disorder: A factor mixture modeling approach. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 54*(2), 206-215. doi: 10.1111/j.1469-7610.2012.02588.x;

Georgiades, S., Szatmari, P., Duku, E., Zwaigenbaum, L., Bryson, S., Roberts, W., . . . Pathways in ASD Study Team. (2011). Phenotypic overlap between core diagnostic features and emotional/behavioral problems in preschool children with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 41*(10), 1321-1329. doi: 10.1007/s10803-010-1158-9;

Georgiades, S., Szatmari, P., Zwaigenbaum, L., Bryson, S., Brian, J., Roberts, W., . . . Garon, N. (2013). A prospective study of autistic-like traits in unaffected siblings of probands with autism spectrum disorder. *JAMA Psychiatry (Chicago, Ill.), 70*(1), 42-48. doi: 10.1001/2013.jamapsychiatry.1; 10.1001/2013.jamapsychiatry.1

Georgiades, S., Boyle, M., Szatmari, P., Hanna, S., Duku, E., Zwaigenbaum, L., … Pathways in ASD Study Team. Modeling the Phenotypic Architecture of Autism Symptoms From Time of Diagnosis to Age 6 (under review).

Happe, F. (2011). Criteria, categories, and continua: Autism and related disorders in DSM-5. *Journal of the American Academy of Child and Adolescent Psychiatry, 50*(6), 540-542. doi: 10.1016/j.jaac.2011.03.015;

Mandell, D. (2011). The heterogeneity in clinical presentation among individuals on the autism spectrum is a remarkably puzzling facet of this set of disorders. *Autism : The International Journal of Research and Practice, 15*(3), 259-261.

Rutter, M. (2012). Changing concepts and findings on autism. *Journal of Autism and Developmental Disorders,* doi: 10.1007/s10803-012-1713-7

Szatmari P. (2011) New recommendations on autism spectrum disorder. *Brit. Med. J.* 342:d2456.