December 2017

Concordance and Discordance: Cognitive and Neuropsychological Performance of Twins with Autism Spectrum Disorder

Elizabeth P. McKernan
Syracuse University

Abstract

This study sought to examine the cognitive and neuropsychological performance of monozygotic twin pairs who vary in concordance for ASD. The study used secondary data collected from a larger pilot study, resulting in a final sample of 28 children, all of whom were monozygotic twins (IQ range: 36 – 109). First, the study sought to determine whether pairwise concordance for ASD among monozygotic twins has changed with the use of new diagnostic criteria. McNemar tests found that pairwise concordance rates were not significantly different across *DSM-IV-TR* and *DSM-5* diagnostic criteria, but were significantly greater with the use of either parent report or direct observation as compared to both measures for *DSM-5* criteria. Secondly, the study examined the relationship between IQ and ASD symptomatology using ordinal logistic regression, finding that a decrease in IQ did not predict greater severity of ASD symptoms as measured by the Calibrated Severity Score (Gotham, Pickles, & Lord, 2009), but did predict greater symptom severity as measured by the total algorithm score of the Autism Diagnostic Interview-Revised (Lord, Rutter, & Le Couteur, 1994). Third, the study examined IQ performance in relation to concordance, finding that the relationship between concordance and IQ was not significant for any area score. Finally, the study examined relative strengths and weaknesses of neuropsychological performance for children with ASD, with no significant differences in performance found for any of the NEPSY subtests examined (Korkman, Kirk, and Kemp, 1998). Results from this study support the idea that the phenotypic characteristics of autism spectrum disorder are heterogeneous even among children with the same genotype.

Keywords: autism spectrum disorder, monozygotic twins, concordance, intellectual functioning

CONCORDANCE AND DISCORDANCE: COGNITIVE AND NEUROPSYCHOLOGICAL PERFORMANCE OF TWINS WITH AUTISM SPECTRUM DISORDER

by

Elizabeth P. McKernan

B.A., Villanova University, 2014

THESIS

Submitted in partial fulfillment of the requirements for the degree of Master of Science in *Psychology.*

Syracuse University

December 2017

TABLE OF CONTENTS

PAGE

INTRODUCTION…….. 1

Diagnosis of ASD……... 4

Family Studies of ASD……... 6

Twin Studies of ASD…….. 7

Cognitive Functioning in ASD……... 9

IQ Profiles of ASD……. 11

Neuropsychological Profiles of ASD……. 12

Aims of the Current Study……. 18

METHODS…… 20

Participants…… 20

Measures…… 21

Autism Diagnostic Interview-Revised…….. 22

Autism Diagnostic Observation Schedule-Generic…….. 23

NEPSY…….. 24

Stanford-Binet Intelligence Scale, Fourth Edition…… 25

Procedure…….. 26

RESULTS……. 29

Aim 1…… 29

Relaxed standard…….. 30

Strict standard……... 31

Relaxed vs. strict standard……. 32

Aim 2……. 33

ADOS……... 34

ADI-R……... 35

Aim 3…….. 36

Aim 4……... 39

DISCUSSION…….. 43

Aim 1…….. 43

Aim 2……... 47

Aim 3…….. 48

Aim 4…….. 50

Limitations…… 51

Implications…… 52

TABLES…… 57

FIGURE……... 67

REFERENCES…… 68

VITA…….. 84

Concordance and Discordance: Cognitive and Neuropsychological Performance of Twins with Autism Spectrum Disorder

Autism spectrum disorder (ASD) is a lifelong neurodevelopmental disorder which is evident in early childhood and is characterized by impairment in social communication and social interaction, as well as the presence of restricted, repetitive patterns of behavior, interests, or activities (American Psychiatric Association, 2013). Impairment in social communication and social interaction may include deficits in a variety of areas, such as social-emotional reciprocity, nonverbal communication behaviors, or difficulty in understanding or establishing social relationships (American Psychiatric Association, 2013). Notably, however, symptoms differ across individuals and across time (Lord, Corsello, & Grzadzinski, 2014). A diagnosis of ASD requires that symptoms are present within the early developmental period, although they may not become fully evident until social demands exceed the individual's limited capacities or may be hidden later in life by learned strategies (American Psychiatric Association, 2013). Consistent with the conceptualization of autism as a spectrum, behavioral characteristics of ASD vary in presentation, severity, and resulting functional limitation.

Although previous research has suggested general areas of strength and limitation in the cognitive and neuropsychological functioning of individuals with an autism spectrum disorder, there is a paucity of literature on the neuropsychological features of related individuals with ASD. Understanding the cognitive and neuropsychological profiles of individuals who share similar genotypes may provide further information about the genetic contributions inherent to these abilities. A detailed, in-depth investigation of the cognitive and neuropsychological profile of monozygotic twins who vary in concordance for ASD may assist in confirming previous

findings regarding the strengths and weaknesses thought to be common to ASD, and may aid in clarifying the influence of ASD on these abilities.

The purpose of the current study was to examine the cognitive and neuropsychological performance of monozygotic twin pairs who vary in the degree of concordance for ASD. Originally, diagnoses of the twin pairs in this study were determined using diagnostic criteria from the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision* (*DSM-IV-TR*; American Psychiatric Association, 2000). Because the diagnostic construct of ASD has changed with the use of the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* (*DSM-5*; American Psychiatric Association, 2013), it was possible that the concordance of these twin pairs had changed. Thus, the study first sought to determine whether the degree of concordance for ASD among monozygotic twins had changed with the use of new diagnostic criteria (*Aim 1*). Although pairwise concordance rates for ASD among monozygotic twins have been estimated to be as high as 95 percent (Rosenberg et al., 2009), to date, no twin studies have been conducted specifically using *DSM-5* diagnostic criteria for ASD. Examining concordance rates across different diagnostic criteria in a sample of monozygotic twins may provide particular insight into the way in which autism spectrum disorder is currently defined.

Secondly, the present study examined the relationship between intelligence quotient (IQ) performance and ASD symptomatology (*Aim 2*). Previous research by Mitchell and colleagues (2009) using the monozygotic twins of this cohort has suggested that IQ accounted for nearly twice the variance in severity scores in co-twins than it did in twins with autism, as measured by the Autism Diagnostic Interview-Revised (ADI-R; Lord, Rutter, & Le Couteur, 1994), but severity was not significantly associated with IQ scores for twins with autism when measured with the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2000). In this previous

study, severity was measured via the total score on these instruments. Using the Calibrated Severity Score (CSS; Gotham, Pickles, & Lord, 2009) as an estimate of the severity of autism symptomatology as indexed by the ADOS, coupled with domain algorithm scores of the ADI-R, might help to specify whether IQ performance actually predicts ASD symptom severity. Although previous research has suggested that cognitive skill is not predictive of ASD symptom severity (Wilson et al., 2014), further exploration of this relationship in a sample of monozygotic twins varying in concordance for ASD may provide further evidence for this finding.

Third, the current study examined cognitive strengths and weaknesses in individuals with ASD (*Aim 3*). Consistent with prior research (Bishop, 1989; Brook & Bowler, 1992; de Bruin, Verheij, & Ferdinand, 2006; Lincoln, Allen, & Kilman, 1995; Tager-Flusberg, 1981), it was hypothesized that on a group level, individuals with ASD would display particular strengths in the areas of quantitative and abstract visual reasoning, with weaknesses in verbal reasoning. In addition, exploratory analyses of the IQ performance of concordant and discordant twin pairs were undertaken in order to clarify the relationship between ASD diagnosis and cognitive performance.

Finally, the present study examined neuropsychological performance in individuals with ASD (*Aim 4*). In fitting with prior research (Damarla et al., 2010; Jolliffe & Baron-Cohen, 1997; Shah & Frith, 1983), it was hypothesized that on a group level, individuals with ASD would display particular strengths on subtests that assess visuospatial abilities, and weaknesses on subtests that assess language abilities. In addition, exploratory analyses of the neuropsychological performance of concordant and discordant twin pairs as measured by performance on four individual subtests of the NEPSY (Korkman, Kirk, & Kemp, 1998) were

undertaken in order to clarify the relationship between ASD diagnosis and neuropsychological performance.

Diagnosis of ASD

Autism is a relatively new diagnosis, as it was first described in 1943 by Dr. Leo Kanner, who labeled the disorder as "autistic disturbances of affective contact," characterized by what he termed extreme aloneness and insistence on sameness (Kanner, 1943). Although Kanner's basic descriptions remain highly relevant, diagnostic criteria for autism have become more specifically defined and have changed substantially over time. The *Diagnostic and Statistical Manual of Mental Disorders, Third Edition* (*DSM-III*; American Psychiatric Association, 1980) was the first classification system to explicitly distinguish autism from childhood schizophrenia, which were previously considered the same. The revised version of this system, *DSM-III-R* (American Psychiatric Association, 1987) created a more precise definition of autistic disorder, and also included a new category, "Pervasive Developmental Disorder-Not Otherwise Specified" (PDD-NOS), for individuals who did not meet the full criteria for autistic disorder. The *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* (*DSM-IV*; American Psychiatric Association, 1994), expanded the category of pervasive developmental disorders to create the diagnostic subcategories of Asperger's disorder and Rett's syndrome, and further delineated the criteria for autistic disorder. Like autistic disorder, Asperger's disorder was characterized by social deficits and restricted patterns of behavior and/or interests, but was differentiated by a lack of language impairment and cognitive delay (American Psychiatric Association, 1994). The diagnosis was contentious, however, as it was unknown to what extent Asperger's disorder truly differed from autistic disorder with normal intelligence (Ghaziuddin, 2010). As a result, many

suggested that Asperger's syndrome be excluded from future diagnostic classification systems (Ritvo, Ritvo, Guthrie, & Ritvo, 2008; Tryon, Mayes, Rhodes, & Waldo, 2006).

The current diagnostic classification system, *DSM-5* (American Psychiatric Association, 2013) moved away from this broader definition of autism by collapsing the pervasive developmental disorder subcategories into a single category: autism spectrum disorder. The transition from the *DSM-IV-TR* to the *DSM-5* diagnostic criteria generated a great deal of controversy among individuals in both the scientific and the lay community. Many expressed concern that individuals who formerly were diagnosed with Asperger's disorder or PDD-NOS would no longer meet criteria due to changes in the diagnostic construct of ASD (Kulage, Smaldone, & Cohn, 2014; Mattila et al., 2011; McPartland, Reichow, & Volkmar, 2012). McPartland, Reichow, and Volkmar (2012) applied proposed *DSM-5* diagnostic criteria for ASD to a sample of 933 participants evaluated during the *DSM-IV* field trial, 657 of whom carried a clinical diagnosis of an ASD, and 276 of whom were diagnosed with a non-autistic disorder. McPartland et al. (2012) found that overall specificity was high, with 94.9 percent of individuals accurately excluded from the spectrum; however, sensitivity varied by diagnostic subgroup and cognitive ability. Asperger's disorder and PDD-NOS had the lowest sensitivity values at 0.25 and 0.28, respectively, and sensitivity was 0.46 for individuals with an IQ greater than or equal to 70, suggesting that individuals who lack cognitive impairment and an autistic disorder diagnosis would be excluded from the newly conceptualized autism spectrum. Similarly, Mattila et al. (2011) applied *DSM-5* draft criteria to an epidemiologic sample of 5,484 eight-year-old children in Finland, finding that sensitivity was lower for those with Asperger's syndrome and high-functioning autism.

In contrast, other research has suggested that most individuals with a previous diagnosis of a pervasive developmental disorder would retain their diagnosis under the new conceptualization of ASD within *DSM-5* (Huerta, Bishop, Duncan, Hus, & Lord, 2012; Kim et al., 2014). Kim et al. (2014) used total-population prevalence data from a cohort of 55,266 Korean children born from 1993 to 1999, and found that 98 percent of individuals with a *DSM-IV* diagnosis of autistic disorder and 92 percent of individuals with a *DSM-IV* diagnosis of Asperger's disorder met *DSM-5* criteria for ASD. Similarly, Huerta et al. (2012) examined 4,453 children in three data sets with *DSM-IV* pervasive developmental disorder diagnoses and found that 65.49, 77.91, and 91.16 percent of children met full *DSM-5* criteria for ASD. In addition, the researchers found that *DSM-5* criteria had greater specificity compared to *DSM-IV* criteria for Asperger's disorder and PDD-NOS.

Examining concordance rates across *DSM-IV-TR* and *DSM-5* criteria in a sample of monozygotic twins enables comparison across diagnostic criteria, and thus allows confirmation that the individuals included in the sample indeed still have an ASD diagnosis. In addition, comparing previous and current diagnostic classification systems allows exploration of the degree to which previous studies of cognitive and neuropsychological functioning in ASD populations are still relevant in relation to the way in which autism spectrum disorder is currently defined.

Family Studies of ASD

Family studies provide evidence for a significant genetic component in the etiology of autism. For instance, previous research has demonstrated that the risk of autism in siblings of autistic probands is approximately 45 times greater than that in the general population (Lord, Leventhal, & Cook, 2001). Ozonoff and colleagues (2011) conducted a prospective longitudinal

study of 664 infants at risk for ASD, defined as those with an older biological sibling with ASD. At 36 months of age, 132 infants met criteria for ASD, yielding an estimated recurrence rate of 18.7 percent. In addition, male gender and multiplex family status were significant, independent predictors of an ASD diagnosis at 36 months, with a 2.8-fold increase in the risk of ASD for male infants, and an additional twofold increase in risk for infants who had multiple older affected siblings relative to those who had only one affected sibling (Ozonoff et al., 2011).

Family studies have also provided evidence that relatives may show milder forms of individual components of autism symptomatology, termed the broader autism phenotype (Piven et al., 1997; Piven & Palmer, 1999). These family members may display social or language deficits that are qualitatively similar to those of the proband with autism, but that are milder in severity. The traits of the broader autism phenotype are assumed to be continuously distributed throughout the general population; however, studies have found that rates of the broader autism phenotype among family members of individuals with ASD are much higher (Pisula & Ziegart-Sadowska, 2015). In developing the Broader Phenotype Autism Symptoms Scale (BPASS; Dawson et al., 2007), Dawson and colleagues utilized a sample of 690 individuals from 201 families having two or more children with an autism spectrum disorder to examine the four symptom domains of social motivation, social expressiveness, conversational skills, and flexibility/range of interests. Using a family history interview, the researchers found that 50 percent of parents were identified as expressing at least one broader autism phenotype feature (Dawson et al., 2007). Taken together, findings of these studies suggest that family members of individuals with ASD may display some of the same behaviors that are known to characterize ASD, thus implying genetic factors in the disorder's phenotypic presentation.

Twin Studies of ASD

In general, research remains inconclusive regarding the relative contribution of genetic and environmental influences in the etiology of ASD (Gaugler et al., 2014; Hallmayer et al., 2011; Klei et al., 2012). Twin studies are particularly helpful in terms of providing a way in which to clarify the respective contributions of genetics and environment, as they minimize variance from genetic and family factors (Kates et al., 2004). Previous research includes mixed findings, with anywhere from zero to 24 percent of dizygotic twin pairs classified as concordant, and 36 to 91 percent of monozygotic twin pairs classified as concordant (Bailey et al., 1995; Deng et al., 2015; Folstein & Rutter, 1977; Rosenberg et al., 2009; Steffenburg et al., 1989). Ronald and Hoekstra (2011) point out that the observed lesser degree of concordance among dizygotic than among monozygotic twins is to be expected due to nonadditive genetic effects, de novo mutations, and chromosomal abnormalities, all of which will inflate the similarity of monozygotic twin pairs. However, the lack of perfect concordance observed within monozygotic twin pairs suggests that the nonshared environment also has an influence.

In a review of twin studies of ASD, heritability estimates were high, with median values for proband-wise concordance equaling 76 percent for monozygotic twins and 0 percent for dizygotic twins (Ronald & Hoekstra, 2011). Other research, however, has placed greater emphasis on environmental factors in the etiology of ASD (Frazier et al., 2014; Hallmayer et al., 2011). In an attempt to reconcile these discrepant findings, Tick, Bolton, Happé, Rutter, and Rijsdijk (2016) performed a meta-analysis of all published twin studies of ASD, which yielded seven studies that met their minimal inclusion criteria of systematic ascertainment of subjects. Group heritability estimates were substantial, ranging from 64 to 91 percent, depending on the threshold on the liability to or population prevalence of ASD that was used. This suggests that a large amount of the variance in ASD is attributable to genetic factors.

Although much research has focused on the broader autism phenotype in family members of individuals with ASD, few studies have specifically examined the broader autism phenotype in monozygotic twins. Le Couteur and colleagues (1996) found that characteristics of the broader autism phenotype were noted in 77 percent of monozygotic co-twins who were discordant for autism, compared with only 5 percent of dizygotic twin pairs, providing further support for the idea that the autism phenotype derives from a similar genetic liability.

A case report by Kates et al. (1998) of a pair of monozygotic twin boys discordant for autism described the unaffected twin as displaying the broader autism phenotype, noting difficulties in social interaction and play, although he did not meet diagnostic criteria for autism. Kates and colleagues described the cognitive and neuropsychological functioning of the discordant twins, and noted that their neuropsychological profiles were relatively consistent with one another. When standard scores were converted to *z* scores, for instance, the twins scored within one standard deviation of each other on most tasks. Although the performance of the twin without an autism diagnosis was superior to his brother, both twins performed below the mean on most tasks of neuropsychological functioning (Kates et al., 1998). One potential interpretation of these results is that deficits in neuropsychological functioning may be common to both strictly defined autism and the broader autism phenotype, such that it may be a general characteristic of an ASD diagnosis. Kates et al. (1998) note that because this study involved a single twin pair, further research involving additional discordant twin pairs is warranted in order to more fully explain the neuropsychological performance of the narrow and broad phenotypes for autism, and to specify whether there are particular neuropsychological strengths and weaknesses common to ASD.

Cognitive Functioning in ASD

Overall, people with ASD form a heterogeneous group in which cognitive abilities and adaptive functioning vary widely (Kanne et al., 2011; Reinvall et al., 2013). In both ASD and non-ASD populations, IQ is a strong predictor of outcomes in terms of school, work, and social functioning (Firkowska-Mankiewicz, 2011; Holwerda, van der Klink, Groothoff, & Brouwer, 2012). Studies of the estimated comorbidity of ASD and intellectual disability have reported discrepant numbers, with estimates of 50 to 70 percent of ASD cases also qualifying for an intellectual disability (Fombonne, 2003). In a study of prevalence in an Italian population sample, LaMalfa, Lassi, Bertelli, Salvini, and Placidi (2004) reported that 70 percent of persons with ASD also had an intellectual disability. In contrast, current research from the Centers for Disease Control and Prevention (CDC) suggests that the co-occurrence of intellectual disability and ASD is much lower. Intellectual disability, defined as an IQ score less than or equal to 70, was found to be present in 31.6 percent of children with an ASD; however, 43.9 percent of children were classified in the average or above average range, defined as an IQ greater than 85 (Centers for Disease Control and Prevention [CDC], 2016).

Although intellectual disability and ASD may not display as high of a degree of comorbidity as previously thought, former studies have suggested that IQ influences the disorder's phenotypic presentation, particularly with regard to restricted, repetitive behaviors as well as challenging behaviors (Bishop, Richler, & Lord, 2006; Matson & Shoemaker, 2009). Previous research has suggested that intelligence may act as a moderator in the cognitive presentation of ASD, with different affected cognitive processes in individuals with above-average versus below-average IQs. Consequently, further research is warranted regarding the variety of cognitive processes that may underlie ASD at different ranges of intelligence (Rommelse et al., 2015). In addition, it remains unknown whether intelligence directly predicts

the severity of ASD symptoms. Thus, examining IQ and ASD severity in a sample of monozygotic twins varying in concordance for ASD with a range of IQ scores may aid in clarifying the nature of this relationship.

IQ Profiles of ASD

On Wechsler intelligence tests, a characteristic profile of strengths and weaknesses has been found on various subtests, with strong performance on tasks such as Block Design, and weaker performance on Comprehension tests for individuals with ASD (de Bruin, Verheij, & Ferdinand, 2006; Lincoln, Allen, & Kilman, 1995). Previous research has indicated that verbal IQ often lags behind performance IQ for many individuals with an autism spectrum disorder, possibly due to these individuals' difficulties with language (Charman, Pickles, Simonoff, Chandler, & Loucas, 2011; Mayes & Calhoun, 2003; Tager-Flusberg & Joseph, 2003).

Findings have been mixed with regard to the supposed verbal IQ-performance IQ discrepancy previously demonstrated. Siegel, Minshew, and Goldstein (1996) examined profile characteristics of the Wechsler Intelligence Scale for Children-Revised (WISC-R; Wechsler, 1974) and the Wechsler Adult Intelligence Scale-Revised (WAIS-R; Wechsler, 1981) in 81 high-functioning children and adults with autism, but did not find the previously reported pattern of higher performance IQ relative to verbal IQ. Similarly, Ehlers and colleagues (1997) used the WISC-R in a sample of 120 children with Asperger syndrome, autistic disorder, and attention disorders, but suggested that other factors such as Verbal Comprehension, Perceptual Organization, and Freedom from Distractibility accounted for the observed variance on the measure, rather than a true split between verbal IQ and performance IQ scores in these children.

There have also been discrepant findings in terms of the profile of performance on certain subtests. For instance, Nader, Jelenic, and Soulières (2015) examined the Wechsler Intelligence

Scale for Children-Fourth Edition (WISC-IV; Wechsler, 2003) in a sample of 51 children with autistic disorder, 15 children with Asperger's syndrome, and 42 children with typical development. Consistent with previous research, children with autism exhibited significant strengths on the Block Design, Matrix Reasoning, and Picture Concepts subtests when compared to the average performance on all subtests. Significant group weaknesses were observed on the Comprehension, Digit Span, Letter-Number Sequencing, and Coding subtests (Nader, Jelenic, & Soulières, 2015). However, this pattern has not been consistently replicated across studies. In a study of 156 children aged 10 to 14 years, Charman et al. (2011) found a mixed pattern, with children with ASD exhibiting poorer performance on the Vocabulary and Comprehension subtests, but without peak performance on the Block Design or Object Assembly subtests.

The lack of consistent findings with regard to IQ profiles of ASD warrants further investigation. Most of the previous research has examined the Wechsler scales of intelligence, with little information about the existence of a prototypical IQ profile in ASD as measured by other IQ tests. Thus, examining the strengths and weaknesses of individuals with ASD using a different IQ measure might provide convergent evidence for the existence of a prototypical IQ profile in ASD. In addition, using a sample of monozygotic twins who vary in concordance for ASD may help to clarify the relationship between IQ and an ASD diagnosis. For instance, it can be inferred that IQ and an ASD diagnosis are orthogonal if twins are discordant for ASD but display similar IQ performance, or if twins are concordant for ASD but display dissimilar IQ performance. In this manner, these analyses may provide evidence regarding whether or not certain patterns of cognitive performance are specific to ASD, as well as whether IQ is related to having a diagnosis of ASD.

Neuropsychological Profiles of ASD

Previous research has attempted to characterize the neuropsychological profile of individuals with autism spectrum disorders, but findings have been mixed. For instance, Minshew, Goldstein, and Siegel (1997) examined the neuropsychological functioning of 33 individuals with autism of average intelligence who were matched to neurotypical controls on the basis of chronological age and full-scale IQ. Subjects were assessed using a neuropsychological battery developed for the purposes of the study that included subtests from several standardized measures, such as the Luria-Nebraska Tactile Scale (Golden, 1980), Wechsler Adult Intelligence Scale-Revised (WAIS-R; Wechsler, 1981), Kaufman Test of Educational Achievement (K-TEA; Kaufman & Kaufman, 1985), Trail Making Tests A and B (Reitan & Wolfson, 1993), and the Wisconsin Card Sorting Test (Grant & Berg, 1948). Altogether, the battery was made up of 9 different domains: attention, sensory perception, motor, reasoning, visual-spatial, simple and complex language, and simple and complex memory. Participants with autism exhibited poorer performance as compared to control subjects in the motor function, complex language, complex memory, and reasoning domains, but significantly superior performance in the simple language domain. Performance in the attention, sensory perception, visual-spatial, and simple memory domains was not significantly different between groups. As a result, Minshew, Goldstein, and Siegel (1997) suggest that both impaired and intact neuropsychological abilities are relevant determinants of the phenotypic presentation of autism. Deficits tended to appear in the most complex abilities of certain domains, but not in the simpler abilities of those domains. The researchers suggest that based on these results, autism is primarily a disorder of late information processing, in which information acquisition is spared, but later neural events in information processing are affected. In this view, the specific

neuropsychological processes that will be implicated in ASD are those that require reasoning that is more complex.

Narzisi, Muratori, Calderoni, Fabbro, and Urgesi (2013) used an Italian version of the NEPSY-II (Korkman, Kirk, & Kemp, 2007) to describe the neurocognitive functioning of 22 boys with high-functioning ASD. The NESPY-II is a developmental neuropsychological battery that assesses the areas of executive function, memory, language comprehension, sensorimotor functions, visuospatial processing, and social perception abilities. In terms of attention and executive function, the high-functioning ASD group had significantly lower performance for all subtests except for Visual Attention and Design Fluency. As might be expected due to the characteristic difficulties with language observed in ASD, the high-functioning ASD group had lower performance on language subtests as compared to controls, but only significantly lower performance on the Oromotor Sequences task, which assesses an individual's articulatory coordination in repeating a verbally presented sequence. ASD participants displayed lower performance on all memory and learning subtests except for Memory for Designs and Memory for Names. For sensorimotor functions, children with ASD had significantly lower performance in imitation of hand postures and manual motor sequences, but not in finger tapping. Children with ASD also had significantly lower performance in terms of social perception, which was due to their lower scores in the verbal items of the theory of mind test. Finally, children with ASD displayed significantly lower performance on Design Copying and Arrows, two visuospatial processing subtests that measure the ability to copy two-dimensional geometric figures and to judge line orientation and directionality, respectively.

Narzisi et al. (2013) point out that the intact performance of children and adolescents with high-functioning ASD on certain subtests may be indicative of alterations in multiple

cognitive subsystems, rather than a general cognitive deficit. Thus, the researchers' findings suggest that specific neuropsychological processes may be affected in individuals with ASD, while others may be spared or even superior to individuals with typical development. This fits with the idea of autism as a disorder of late information processing as posited by Minshew et al. (1997), in which information acquisition is spared but neural events involved in information processing are affected.

Reinvall, Voutilainen, Kujala, and Korkman (2013) examined the cognitive and neuropsychological profiles of 30 adolescents with Asperger syndrome who were matched to 30 typically developing adolescents on the basis of age, gender, and maternal education using the Finnish versions of the Wechsler Intelligence Scale for Children, Third Edition (WISC-III; Wechsler, 1991) and NEPSY-II. Adolescents with Asperger syndrome had significantly lower scores than typically developing adolescents on the Auditory Attention and Response Set, Memory for Faces, Visuomotor Precision, and Design Copying subtests on the NEPSY-II. No other significant group differences were found on any subtests. No significant correlations were found between ASD symptom severity and impaired cognitive functioning in the sample of adolescents with higher functioning ASD, which suggests an orthogonal relationship between IQ and the severity of characteristic behaviors of ASD. There was however, a relationship between cognitive and neuropsychological functioning, as lower WISC-III scores were related to impaired neuropsychological performance as measured by NEPSY-II scores (Reinvall et al., 2013). These findings suggest that although IQ and neuropsychological skills are significantly related, IQ and ASD symptom severity are not.

Although these findings share some overlap, there are a notable number of discrepancies. For instance, while Narzisi et al. (2013) found that boys with high-functioning ASD

demonstrated poorer performance compared to typically developing controls on the imitation of hand postures and manual motor sequences on the NEPSY-II, Reinvall et al. (2013) found no differences on these subtests between adolescents with Asperger syndrome and adolescents with typical development. Similarly, although Narzisi et al. (2013) and Reinvall et al. (2013) both found that participants with high-functioning ASD demonstrated significantly poorer performance on the Design Copying subtest of the NEPSY-II, Narzisi et al. (2013) found that boys with high-functioning ASD demonstrated poorer performance on Arrows, another visuospatial subtest, whereas Reinvall et al. (2013) found no difference between groups. Results of these studies are summarized in Table 1. These disparate findings may be due in part to the differences in the samples, as well as differences in the versions of the test that were used. For instance, Narzisi et al. (2013) used an Italian version of the NEPSY-II, and Reinvall et al. (2013) used a Finnish version of the test. Though the observed differences in findings may be a result of different norms, it is reasonable to assume that neuropsychological abilities that are characteristic of ASD would be similar across cultural populations. Thus, discrepant findings regarding the neuropsychological performance of individuals with ASD preclude definitive conclusions regarding characteristic strengths and weaknesses of performance as measured by the NEPSY.

Wilson et al. (2014) compared the cognitive profiles of adults with Asperger syndrome and high-functioning autism to evaluate the utility of cognitive measures in predicting diagnostic subtype. A group of 178 male adults with an ASD and average intelligence completed a variety of tasks that the authors selected for the study based on domains that the existing literature suggests are abnormal in ASD. The Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999) was used to assess IQ; the Karolinska Directed Emotional Faces task (Lundqvist, Flykt, & Ohman, 1998) and the "Reading the Mind in the Eyes" task (Baron-Cohen

et al., 2001) were used to assess emotion processing; the Frith-Happé Animations Test (Abell, Happé, & Frith, 2000) was used to assess theory of mind; the Go-No Go task of Rubia et al. (2001) was adapted to assess executive functions; the Embedded Figures Test (Witkin, Oltman, Raskin, & Karp, 1971) was used to assess central coherence; and the Nonword Repetition Task of Gathercole, Willis, Baddeley, and Emslie (1994) was adapted to assess phonological memory.

No significant differences were observed with regard to any of the cognitive measures between adults with high-functioning autism and Asperger syndrome, suggesting that cognitive profiles are not different between the two groups. The researchers also investigated whether a cognitive profile, as assessed by the combination of tests, could predict diagnostic subtype, and found that individuals could be accurately classified as ASD or control subjects with 78 percent sensitivity and 85 percent specificity; however, cognitive profiles did not classify individuals with high-functioning autism vs. Asperger's syndrome any better than chance. This finding may provide support for the idea that Asperger's syndrome does not differ from autistic disorder without cognitive impairment. Similar to Reinvall et al. (2013), Wilson et al. (2014) found that ASD symptom severity was not correlated with the cognitive factors that were examined. This suggests that the severity of ASD symptoms is not related to cognitive skill, nor is cognitive skill predictive of ASD symptom severity.

Taken together, the extant literature regarding neuropsychological functioning in ASD suggests that both strengths and weaknesses are present within this population. Rather than a general deficit, individuals with ASD demonstrate impaired performance compared to individuals with typical development within certain domains, but similar or even enhanced performance within other domains; however, discrepancies within these findings make it difficult

to specify which domains are consistent strengths and which are consistent weaknesses in individuals with ASD.

It is important to note that much of the research regarding neuropsychological profiles in ASD has been conducted using samples of individuals who had a *DSM-IV* diagnosis of Asperger syndrome or who have high-functioning ASD. For instance, Wilson et al. (2014) specifically compared the cognitive profiles of Asperger syndrome and high-functioning autism; Narzisi et al. (2013) compared 22 children with high-functioning ASD to a group of 44 healthy control children matched two to one for age, gender, race, and education; and Reinvall et al. (2013) compared the neurocognitive profiles of adolescents with Asperger syndrome to those of typically developing adolescents. Although these studies matched participants with ASD to control subjects based on age and gender, none specifically matched on IQ, nor did any of these studies examine individuals with ASD with a broader range of symptoms, cognitive abilities, or levels of functioning. Thus, examining monozygotic twins with a range of cognitive abilities who vary in concordance for ASD may be informative in specifying this neuropsychological profile in a wider range of clinical presentations.

Aims of the Current Study

The purpose of the current study was to examine the cognitive and neuropsychological performance of monozygotic twin pairs who vary in the degree of concordance for ASD. Because the diagnostic construct of autism has changed from the *DSM-IV-TR,* which was originally used to characterize the twins included in this study, to the current diagnostic system in use, the *DSM-5*, it was necessary to examine whether the twins included in the study indeed met diagnostic criteria for ASD. Thus, the study first sought to determine whether the degree of concordance for ASD among monozygotic twins had changed with the use of new diagnostic

criteria (*Aim 1*). It was hypothesized that pairwise concordance rates would not be significantly different with the use of *DSM-5* criteria.

Secondly, the present study examined the relationship between IQ performance and ASD symptomatology (*Aim 2*). It was hypothesized that IQ would not be a significant predictor of the severity of ASD symptoms, as indexed by the ADOS's Calibrated Severity Score (Gotham, Pickles, & Lord, 2009) and the diagnostic algorithm domain scores of the ADI-R (Lord, Rutter, & Le Couteur, 1994).

Third, the current study sought to examine whether a particular IQ profile was specific to individuals with ASD (*Aim 3a*). It was hypothesized that individuals with ASD would demonstrate relative strengths with regard to the abstract/visual reasoning and quantitative reasoning domains, and relative weaknesses in terms of the verbal reasoning and short-term memory domains of the Stanford-Binet Intelligence Scale, Fourth Edition (SB-IV; Thorndike, Hagen, & Sattler, 1986). In addition, the study examined the relationship between IQ and ASD diagnosis within concordant and discordant twin pairs (*Aim 3b*). It was hypothesized that an ASD diagnosis and IQ performance would be orthogonal, such that concordant twins would not consistently demonstrate similar patterns of IQ performance, or conversely, that discordant twins would not consistently demonstrate dissimilar patterns of IQ performance.

Finally, the present study also undertook an exploratory analysis of several individual subtests of the NEPSY (Korkman, Kirk, & Kemp, 1998). In particular, the study examined whether certain strengths and weaknesses in neuropsychological performance were specific to individuals with ASD (*Aim 4a).* It was hypothesized that individuals with ASD would demonstrate relative strengths in terms of the Design Copying and Memory for Faces subtests, but relative weaknesses with respect to the Imitating Hand Positions and Phonological

Processing subtests of the NEPSY. In addition, the current study examined the relationship between ASD diagnosis and NEPSY performance within concordant and discordant twin pairs (*Aim 4b*). It was hypothesized that an ASD diagnosis and NEPSY performance would be significantly related, such that concordant twins would demonstrate more similar patterns of NEPSY performance than would discordant twins.

Methods

Participants

The original study group consisted of 54 children, including 18 pairs of monozygotic twins and 18 singleton age- and gender-matched typically developing peers. Due to the research questions pertinent to this study, the group of singleton age- and gender-matched typically developing peers were excluded from analyses. In addition, four monozygotic twin pairs were excluded from analyses due to a lack of complete data that are required to examine the current research questions. The final study group consisted of 28 children, all of whom were monozygotic twins.

Participants included 28 children between the ages of 5 and 12 years (14 pairs of monozygotic twins). The majority of participants identified as White (86%), with a smaller percentage identified as Black or African American (7%) and Asian (7%). The mean age of the monozygotic twin pairs was 8.0 years (SD=1.96 years; range=5.6-12.2 years). Twelve twin pairs were boys. Participant characteristics (age, IQ, and diagnostic measure scores) are provided in Table 2. Socioeconomic status was measured using the Two Factor Index of Social Standing (Hollingshead & Redlich, 1958), which takes into account the amount of formal education a respondent has completed, as well as the respondent's occupation, keyed according to the occupational titles used by the United States Census in 1970. The mean socioeconomic status of

the sample using this metric was 48.78 (SD=12.41; range=22.0-63.5). This would correspond to at least one parent having completed one to three years of college, and at least one parent having an occupation such as a meeting planner or computer programmer. This corresponds to a mid-level socioeconomic status, as scores on this metric are divided into five social classes, ranging from 20 to 134.

Families of children with ASD were recruited through the Autism Society of America, the National Alliance for Autism Research, the Kennedy Krieger Center for Autism and Related Disorders, and clinical neurologists at the Kennedy Krieger Institute and Johns Hopkins Hospital. Subjects with a genetic disorder or a history of status epilepticus were excluded. Recruitment methods are described in detail elsewhere (Kates et al., 2004). Data for the present study were collected in a multiyear period from 1998-2001. In order to be included in this cohort, a child had to meet criteria for autism on the Autism Diagnostic Interview-Revised (ADI-R; Lord, Rutter, & Le Couteur, 1994) and score within one point of the criteria for autism on the Autism Diagnostic Observation Schedule-Generic (ADOS-G; Lord et al., 2000). Each monozygotic twin pair in the study included at least one child with a diagnosis of autism based on these criteria. Parents of all subjects signed informed consent forms approved by the institutional review board of Johns Hopkins Hospital.

DNA fingerprinting probes were used to confirm zygosity in all twin pairs. Eight independent loci were tested for each twin pair. DNA profiles for all twin pairs were statistically identical at every locus, indicating that the probability of monozygosity for each twin pair was approximately 99.99%.

Measures

Autism Diagnostic Interview-Revised. The Autism Diagnostic Interview-Revised (ADI-R; Lord, Rutter, & Le Couteur, 1994) is a standardized, semi-structured diagnostic interview for caregivers. The ADI-R is a 93-item measure administered by an experienced clinical interviewer. Questions about the developmental history and current behavior of the individual suspected of having ASD fall under three domains: language; reciprocal social interaction; and restricted, repetitive, and stereotyped behaviors or interests. Items are scored for their most abnormal manifestation between the ages of 4 and 5 years, as well as the child's current behavior. The ADI-R is a reliable and valid instrument for making diagnoses of ASD for children of preschool age (Gray, Tonge, & Sweeney, 2008; Lord, Rutter, & Le Couteur, 1994), children with pervasive developmental disabilities (Lecavalier et al., 2006), and children and adolescents with mental retardation (de Bildt et al., 2004). In addition, the ADI-R has demonstrated effectiveness in differential diagnosis of autism from other developmental disorders (Lord, Rutter, & Le Couteur, 1994; Tadevosyan-Leyfer et al., 2003).

In a group of twenty preschool-age children (10 with autistic disorder and 10 with an intellectual disability or language impairment), multirater weighted kappa levels exceeded .70 for 12 out of 15 social algorithm items, .69 for all communication items, and .63 for all 7 restricted and repetitive behavior items. For algorithm items, intraclass correlations for domain scores ranged from .93 to .97. In terms of internal consistency, Cronbach's alpha was .95 for the 15 items in the social area and .69 for the restricted and repetitive behaviors area. Validity was examined in a group of fifty preschoolers (25 with autistic disorder and 25 with an intellectual disability or language impairment). All social and nonverbal communication algorithm items yielded significant differences for diagnosis (Lord, Rutter, & Le Couteur, 1994).

Autism Diagnostic Observation Schedule-Generic. The Autism Diagnostic Observation Schedule-Generic (ADOS-G; Lord et al., 2000) is a play-based, semi-structured, standardized assessment of social interaction, communication, and play for individuals suspected of having an autism spectrum disorder. The measure consists of four modules, which are designed to be appropriate for individuals of different developmental and language levels. Subsets of items in each module are used to create diagnostic algorithms, which provide classification based on exceeding thresholds on the two domains of social behavior and communication. The ADOS-G has demonstrated effectiveness in discriminating autism from nonspectrum disorders, with sensitivity values ranging from .93-1.0 and specificity values also ranging from .93-1.0 (Lord et al., 2000).

Reliability of the measure across the four modules was examined in a sample of 98 children and adolescents with 12 examiners. Overall, interrater item reliability was high, with mean multirater weighted kappa levels of .78, .70, .65, and .66 for Modules 1, 2, 3, and 4, respectively. Test-retest reliability was assessed for a sample of 27 participants who were administered the same ADOS-G module twice by two different examiners within an average of 9 months. Intraclass correlations indicated excellent stability for the communication (.73), social (.78) and total (.82) domains, and good stability for stereotyped behaviors and restricted interests (.59) domains. Internal consistency was assessed using Cronbach's alpha, and was highest for the social domain (.86-.91 for each module), and ranged from .74-.84 for the communication domain. For the social-communication totals, Cronbach's alphas ranged from .91-.94 for all modules. Analyses of variance and specific comparisons comparing the distributions for social domains and social-communication totals across diagnostic groups were significantly different for all modules (Lord et al., 2000).

NEPSY. The NEPSY (Korkman, Kirk, & Kemp, 1998) is a developmental neuropsychological battery designed specifically for children ages 3 to 12 years. It consists of 27 subtests, which assess five domains: attention and executive functions; language; memory and learning; sensorimotor; and visuospatial processing. Each domain is composed of core subtests graded using a scaled score (mean=10; standard deviation=3). Together, these subtests yield a Core Domain Score (mean=100; standard deviation=15) for each of the five domains. Subtests composing the Attention and Executive Functions domain include Tower, Auditory Attention and Response Set, Visual Attention, Statue, Design Fluency, and Knock and Tap. The Language domain is made up of seven subtests: Body Part Naming, Phonological Processing, Speeded Naming, Comprehension of Instructions, Repetition of Nonsense Words, Verbal Fluency, and Oromotor Sequences. The Sensorimotor Functions domain includes Fingertip Tapping, Imitating Hand Positions, Visuomotor Precision, Manual Motor Sequences, and Finger Discrimination subtests. The Visual-Spatial domain includes four subtests: Design Copying, Arrows, Block Construction, and Route Finding. Finally, the Memory and Learning domain is made up of the Memory for Faces, Memory for Names, Narrative Memory, Sentence Repetition, and List Learning subtests. The NEPSY is based upon Luria's (1976) theoretical model of neuropsychological functioning, which posits that human mental processes involve groups of specific brain areas that work in concert as a functional system.

Internal consistency reliability coefficients indicate that the majority of the NEPSY subtests have moderate to high internal consistency or stability. Subtests with the highest reliability coefficients include Phonological Processing (.91), Memory for Names (.89), and List Learning (.91). Subtests with the lowest reliability coefficients are those on which test-retest correlation was used, such as Design Fluency (.59), Verbal Fluency (.74), and Fingertip Tapping

(.71). The NEPSY Core Domain Scores exhibit moderately high reliability scores, ranging from .88 to .91 for children three to four years of age and from .79 to .87 for children five to twelve years of age. Stability of performance on the NEPSY subtests and Core Domain Scores was examined in a sample of 168 children who were given the full NEPSY on two occasions ranging from 2 to 10 weeks apart. Pearson correlation coefficients were used to obtain stability coefficients between scores obtained at the first test session and scores obtained at the second test session, and then were corrected for sampling error on the first testing. Overall, corrected stability coefficients ranged from .68 for the Attention/Executive Functions domain, .78 for the Language domain, .77 for the Sensorimotor domain, .72 for the Visuospatial domain, and .90 for the Memory and Learning domain (Korkman, Kirk, & Kemp, 1998).

In order to assess the clinical utility and discriminant validity of the NEPSY, data were collected on groups with various neurological and developmental disabilities. In a sample of 23 children with a *DSM-IV* clinical diagnosis of autistic disorder matched to controls for age, sex, parent education, and race/ethnicity, children with autism demonstrated a deficit in Attention/Executive Functions relative to controls. Specific impairments were demonstrated on the Memory for Faces, Narrative Memory, and List Learning subtests (Korkman, Kirk, & Kemp, 1998).

Stanford-Binet Intelligence Scale, Fourth Edition. The Stanford-Binet Intelligence Scale, Fourth Edition (SB-IV; Thorndike, Hagen, & Sattler, 1986) is a standardized assessment of intelligence and cognitive abilities. There are 15 subtests that make up the SB-IV: Vocabulary, Comprehension, Absurdities, Verbal Relations, Pattern Analysis, Copying, Matrices, Paper Folding and Cutting, Quantitative, Number Series, Equation Building, Bead Memory, Memory for Sentences, Memory for Digits, and Memory for Objects. Raw scores are converted into

scaled scores called standard age scores (SASs) for all subtests (mean=50; standard deviation=8). Area scores and the Composite score have a mean of 100 and a standard deviation of 16.

Overall, the SB-IV possesses good reliability. Internal consistency coefficients range from .86 to .97 for Verbal Reasoning, .85 to .97 for Abstract/Visual Reasoning, .80 to .97 for Quantitative Reasoning, and .86 to .95 for Short-Term Memory. Reliability for individual subtests ranges from .73 for Memory for Objects to .94 for Paper Folding and Cutting. Test-retest coefficients for children of preschool age are excellent for the Composite (r=.91) and adequate for the area scores (r=.71-.78). Test-retest coefficients for children of elementary school age are also adequate, with r=.90 for the Composite and r=.87, .67, .81, and .51 for the respective areas of Verbal Reasoning, Abstract/Visual Reasoning, Short-Term Memory, and Quantitative Reasoning (Youngstrom, Glutting, & Watkins, 2003). The SB-IV demonstrates a high degree of concurrent validity with other IQ measures, including the Differential Ability Scales (DAS; Elliott, 1990), Kaufman Assessment Battery for Children (K-ABC; Kaufman & Kaufman, 1983), Stanford-Binet, Form L-M (SB-LM; Thorndike, 1973), Wechsler Adult Intelligence Scale-Third Edition (WAIS-III; Psychological Corporation, 1997), Wechsler Intelligence Scale for Children-Third Edition (WISC-III; Wechsler, 1991), and Wechsler Preschool and Primary Scale of Intelligence (WPPSI; Wechsler, 1967). The SB-IV yields a single, global estimate of cognitive ability in the form of the Composite score. The Composite score of the SB-IV has strong support from factor analyses of the SB-IV subtests. In addition, the standard age scores permit interpretation of subtest profile performance (Youngstrom, Glutting, & Watkins, 2003).

Procedure

The current study was part of a larger project in which magnetic resonance imaging was also conducted. A medical history form and the Autism Behavior Checklist (Krug, Arick, & Almond, 1980) were mailed to interested families to screen twin pairs for potential participation in the study. If at least one twin scored above 57 on the Autism Behavior Checklist, the family was contacted by telephone. Participants with a genetic disorder or a history of status epilepticus were excluded. The ADI-R was administered separately for each co-twin in a series of telephone interviews with a parent of the twins by a project staff member. If the diagnosis of autism was confirmed by the ADI-R for at least one co-twin in each twin pair, the family was invited for additional assessment, at which time the ADOS-G was administered separately to each co-twin. Because diagnoses were originally assigned to the twins in this cohort using *DSM-IV* diagnostic criteria for autism, changes are needed in order for the data to correspond to the current conceptualization of autism spectrum disorder as specified by the *DSM-5*. First, scores on the ADOS-G (Lord et al., 2000) were transferred to the Autism Diagnostic Observation Schedule, Second Edition (ADOS-2; Lord et al., 2012) in order to reflect the most up-to-date version of the instrument. Similarly, scores on the ADI-R (Lord et al., 1994) were transferred to the most recent algorithm (Rutter, Le Couteur & Lord, 2003) to correspond to the current use of this instrument.

Unlike the ADOS-G, the ADOS-2 includes a measure of symptom severity in the form of the Calibrated Severity Score (CSS; Gotham, Pickles, & Lord, 2009). The CSS is an estimate of the severity of autism symptomatology based on which ADOS module was administered, the individual's score on that module, the individual's chronological age, and the individual's language level. The ADOS CSS was originally calculated by calibrating ADOS-2 algorithm totals using age- and language-driven cells chosen on the basis of theoretically-driven expectations for specific age ranges with similar developmental impairments (Gotham, Pickles,

& Lord, 2009). Within each of these cells, severity scores were based on the raw total percentiles that corresponded to each of three possible ADOS classifications for diagnosis. As a result, the CSS ranges from 1 to 10 with scores less than 4 indicating a non-spectrum level of functioning (Ankenman, Elgin, Sullivan, Vincent, & Bernier, 2014). The CSS was calculated for each child in the present sample by summing their scores on the Social Affect and Restricted, Repetitive Behavior domains of the ADOS-2 and comparing this value to the table of CSS scores designated for each module and age level. For all modules, a CSS of 0, 1, or 2 indicates minimal to no evidence of autism spectrum-related symptoms; a CSS of 3 or 4 indicates a low level of autism spectrum-related symptoms; a CSS of 5, 6, or 7 indicates a moderate level of autism spectrum-related symptoms; and a CSS of 8, 9, or 10 indicates a high level of autism spectrum-related symptoms. The CSS has been shown to be a valid indicator of autism symptom severity that is stable over 12 to 24 months (Shumway et al., 2012). The CSS was used as a metric of the severity of ASD symptomatology for the purposes of this study.

Previous work by Huerta et al. (2012) matched individual items from the ADOS-2 and the ADI-R to *DSM-IV-TR* criteria for pervasive developmental disorders and *DSM-5* criteria for autism spectrum disorder. These item mappings were used in the current study in order to operationalize diagnostic criteria. Following the conventions of Huerta and colleagues (2012), symptom counts were used to determine how many children met the criteria for a *DSM-IV-TR* diagnosis of pervasive developmental disorder or *DSM-5* diagnosis of autism spectrum disorder. For each item, a score of 1, 2, or 3 indicates the presence of a symptom, whereas a score of zero indicates the absence of a symptom. *DSM-IV-TR* and *DSM-5* guidelines were then followed to determine whether each participant met or did not meet the criteria for a pervasive developmental disorder or autism spectrum disorder, respectively.

Results

Due to the small sample size of 28 children and the number of questions posed and subjected to statistical significance testing across the four aims, this study is limited in terms of its statistical power. Given the nature of the data, there was a theoretical basis to the exploration of the following aims; however, these findings are very tentative. Because of the limitations posed by the lack of statistical power, all results should be considered highly preliminary and in need of replication.

Aim 1

In order to determine whether the degree of concordance for ASD among this sample of monozygotic twins had changed with the use of new diagnostic criteria, diagnostic classifications were assigned in one of two ways (relaxed standard or strict standard), following the ADOS-2 and ADI-R item mappings of *DSM-IV-TR* and *DSM-5* criteria created by Huerta et al. (2012). Using a relaxed standard, a score greater than zero on an item from *either* the ADOS-2 or the ADI-R was taken as evidence of the presence of a behavior, whereby the presence of a single behavior was considered to be sufficient to meet a criterion. For example, a current score of 2 on item 51 of the ADI-R, Social Smiling, was considered sufficient to meet Section A, criterion 1 of *DSM-5* diagnostic criteria: deficits in social-emotional reciprocity. Under the strict standard, classifications were assigned by requiring evidence from *both* the ADOS-2 and the ADI-R. Using the strict standard, a score greater than zero on an item from the ADOS-2 and a score greater than zero on an item from the ADI-R were taken as evidence of the presence of a behavior. For instance, a current score of 2 on the ADI-R for Social Smiling, as well as a score of 1 on the ADOS-2 for Shared Enjoyment in Interaction, were necessary to meet Section A, criterion 1 (deficits in social-emotional reciprocity) of *DSM-5* diagnostic criteria. These

procedures were followed for each child using the ADOS-2 and ADI-R item mappings for *DSM-IV-TR* and *DSM-5* criteria as established by Huerta and colleagues (2012).

Once diagnoses were determined following these procedures, McNemar tests were performed to determine whether there was a significant difference in sample pairwise concordance rate between *DSM-IV-TR* and *DSM-5* diagnostic criteria. Effect sizes were calculated using the phi coefficient. As McNemar tests assess the significance of the difference between two correlated proportions and are appropriate for use on paired nominal data, and because concordance rates are based on matched-pair samples, McNemar tests were an appropriate statistic to use. Because the phi coefficient is a measure of association between two binary variables, it is an appropriate measure of effect size for the goodness of fit in a 2x2 contingency table, such as that in which twin pairs are classified as concordant or discordant according to two diagnostic classification systems (*DSM-IV-TR* and *DSM-5*). A phi coefficient of .1 is considered a small effect, .3 is considered a medium effect, and .5 is considered a large effect (Cohen, 1988).

Relaxed standard. Under *DSM-IV-TR* diagnostic criteria, concordance was defined as both twins having a diagnosis of any pervasive developmental disorder (autistic disorder, Asperger's disorder, or PDD-NOS). Using evidence from either the ADI-R or the ADOS-2, twelve of the 14 twin pairs were classified as concordant. Pairwise concordance rate was calculated as the proportion of twin pairs with two affected children out of all twin pairs in the sample. Using this procedure with *DSM-IV-TR* diagnostic criteria yielded a pairwise concordance rate of 85.71 percent. For the two twin pairs who were classified as discordant, one twin in each pair had a diagnosis of autistic disorder, while the other lacked a diagnosis of any pervasive developmental disorder.

Under *DSM-5* diagnostic criteria, concordance was defined as both twins having a diagnosis of autism spectrum disorder. Using evidence from either the ADI-R or the ADOS-2, eleven of the 14 twin pairs were classified as concordant, yielding a pairwise concordance rate of 78.57 percent. For the three twin pairs who were classified as discordant, one twin in each pair had a diagnosis of ASD, while the other lacked an ASD diagnosis. Two of the three discordant twin pairs were also classified as discordant according to *DSM-IV-TR* criteria; the third discordant twin pair were previously characterized as concordant.

In this previously concordant pair, one twin met all *DSM-5* criteria for Section A, but none of the criteria for Section B, suggesting impairment in social communication and social interaction but a lack of restricted, repetitive patterns of behavior, interests, or activities. This individual may meet criteria for social communication disorder, a diagnosis new to the *DSM-5* that is defined in terms of impairment of pragmatics and the social use of nonverbal and verbal communication. If the definition of concordance is widened to include social communication disorder, there was no change in pairwise concordance rates from *DSM-IV-TR* to *DSM-5* diagnostic criteria when using evidence from either the ADI-R or the ADOS-2.

A McNemar test demonstrated that there was no significant difference in sample pairwise concordance rate between *DSM-IV-TR* and *DSM-5* diagnostic criteria (85.71% and 78.57%, respectively) when evidence was required from either the ADI-R or the ADOS-2, χ^2 =8.556, df=1, p=1.000, φ=.782.

Strict standard. Concordance decreased when evidence was required from both the ADI-R and the ADOS-2. When evidence was required from both the ADOS-2 and the ADI-R for all *DSM-IV-TR* diagnostic subdomains which had relevant items from both measures, eight of the 14 twin pairs were classified as concordant, resulting in a pairwise concordance rate of 57.14

percent. When evidence was required from both the ADOS-2 and the ADI-R for all *DSM-5* diagnostic subdomains in which there were relevant items from both measures, five of the 14 twin pairs were classified as concordant, resulting in a pairwise concordance rate of 35.71 percent. A McNemar test demonstrated that there was no significant difference in the *DSM-IV-TR* sample pairwise concordance rate of 57.14 percent and the *DSM-5* sample pairwise concordance rate of 35.71 percent, χ^2 =5.833, df=1, *p*=.250, φ=.645.

When expanding the definition of concordance to include social communication disorder for the one individual who met all criteria for section A and none of the criteria for section B, seven twin pairs were classified as concordant, yielding a pairwise concordance rate of 50 percent for *DSM-5* criteria, as compared to 57.14 percent for *DSM-IV-TR* criteria, when requiring evidence from both the ADOS-2 and the ADI-R. A McNemar test revealed no significant difference in the *DSM-IV-TR* sample pairwise concordance rate of 57.14 percent and the *DSM-5* sample pairwise concordance rate of 50 percent, χ^2 =0, df=1, *p*=1.0, φ=-.07.

Relaxed vs. strict standard. Differences in methodology were also investigated in order to determine if the use of information from either the ADI-R or the ADOS-2 (relaxed standard) as opposed to information from both measures (strict standard) had an effect on concordance rates. There was no significant difference in *DSM-IV-TR* pairwise concordance rates with the use of evidence from either the ADOS-2 or ADI-R (85.71%) as compared to both the ADOS-2 and the ADI-R (57.14%), x^2 = 3.111, df = 1, *p* =.125, φ = .471. However, results of a second McNemar test demonstrated that *DSM-5* pairwise concordance rates were significantly different when evidence of symptoms was required from either measure (78.57%) versus both measures (35.71%), χ^2 =2.121, df=1, *p*=.031, φ=.389 (see Figure 1).

For the relaxed criteria analyses, evidence from either parent report or clinical observation was sufficient for assigning diagnoses. With the use of *DSM-IV-TR* diagnostic criteria, out of the 26 children who were classified as having a pervasive developmental disorder diagnosis, the ADI-R was sufficient for 10 of these diagnoses, and the remaining 16 of these diagnoses were due to evidence from both the ADOS and the ADI-R. With the use of DSM-5 diagnostic criteria, out of the 25 children who were classified as having an autism spectrum disorder diagnosis, the ADI-R was sufficient for 15 of these diagnoses, and the remaining 10 of these diagnoses were due to evidence from both the ADOS and the ADI-R.

Aim 2

In order to explore the relationship between IQ and ASD symptomatology, ordinal logistic regression was used in order to determine whether IQ, as assessed by the Stanford-Binet Intelligence Scale, Fourth Edition, was a significant predictor of the severity of ASD symptoms, as indexed by the Autism Diagnostic Observation Schedule (Lord et al., 2000) and the Autism Diagnostic Interview-Revised (Lord, Rutter, & Le Couteur, 1994). Because three individuals were missing data for the SB-IV, the remaining 25 of the 28 children were included in this analysis. In addition, because one key assumption of an ordinal logistic regression is that there is no multicollinearity in the independent variables, only the Test Composite score was examined as a predictor of ASD symptom severity, due to the high degree of collinearity among area scores of the SB-IV (see Table 3). A specific kind of ordinal logistic regression, called a cumulative odds ordinal logistic regression with proportional odds, models the relationship between a predictor variable and the propensity to be in an ordered category. A proportional odds model is a regression model specifically for ordinal dependent variables. This was an appropriate model to use, as the severity of ASD symptoms is measured on an ordinal scale, because severity

can be ordered (i.e., least severe to most severe) but intervals between the values quantifying severity may not be equally spaced.

The Calibrated Severity Score (CSS; Gotham, Pickles, & Lord, 2009) was used to assess severity of ASD symptoms as indexed by the Autism Diagnostic Observation Schedule. The CSS ranges from 1 to 10 with scores less than 4 indicating a non-spectrum level of functioning (Ankenman, Elgin, Sullivan, Vincent, & Bernier, 2014). For the ADI-R, ASD symptom severity was assessed using total algorithm scores summed across all domains, with higher algorithm scores indicating a greater number of symptoms and by extension, greater severity. Thus, the outcome variable was severity, and IQ score was the predictor variable. A cumulative odds ordinal logistic regression with proportional odds was conducted separately for the ADOS and the ADI-R, with the Stanford-Binet Intelligence Scale Test Composite score entered as a predictor variable. All analyses were conducted using IBM SPSS Statistics (Version 24).

ADOS. A cumulative odds ordinal logistic regression with proportional odds was conducted to determine the effect of IQ on symptom severity as indexed by the Calibrated Severity Score (CSS) of the ADOS-2. Four groups were created on the basis of CSS scores. A CSS of 1 or 2 was classified as minimal to no evidence of ASD; a CSS of 3 or 4 was classified as low; a CSS of 5, 6, or 7 was classified as moderate; and a CSS of 8, 9, or 10 was classified as high. These four categories of the ordinal dependent variable resulted in the creation of three dichotomous variables, reflecting three cumulative splits of the categories of the ordinal dependent variable. A cumulative odds ordinal logistic regression with proportional odds requires the assumption of proportional odds, meaning that each independent variable has an identical effect at each cumulative split of the ordinal dependent variable. The assumption of proportional odds was met, as assessed by a full likelihood ratio test comparing the fitted model

to a model with varying location parameters, $\chi^2(2)=3.84$, $p=.147$. The deviance goodness-of-fit test indicated that the model was a good fit to the observed data, $\chi^2(59)=41.74$, $p=.957$, but most cells were sparse with zero frequencies in 70.2% of cells. The final model did not significantly predict the dependent variable over and above the intercept-only model, $\chi^2(1)=3.36$, $p=.067$. A decrease in Test Composite score was not significantly associated with an increase in the odds of greater ASD symptom severity as measured by the CSS, with an odds ratio of .966, 95% CI [.929, 1.004], Wald $\chi^2(1)=3.01$, $p=.083$.

ADI-R. A second cumulative odds ordinal logistic regression with proportional odds was conducted to determine the effect of IQ score on symptom severity as indexed by the total algorithm score for the ADI-R. Three groups were created on the basis of total ADI-R algorithm scores. Because the cutoff score for ASD for all sections of the ADI-R is 29, all scores below 29 were classified as "minimal to no evidence of ASD." A median split was used to classify the remainder of the scores. Scores falling below the median of 54 but above 29 were classified as "low," and scores falling above the median of 54 were classified as "high." These three categories of the ordinal dependent variable resulted in the creation of two dichotomous variables, reflecting two cumulative splits of the categories of the ordinal dependent variable. A cumulative odds ordinal logistic regression with proportional odds requires the assumption of proportional odds, meaning that each independent variable has an identical effect at each cumulative split of the ordinal dependent variable. The assumption of proportional odds was met, as assessed by a full likelihood ratio test comparing the fitted model to a model with varying location parameters, $\chi^2(1)=1.99$, $p=.158$. The deviance goodness-of-fit test indicated that the model was a good fit to the observed data, $\chi^2(39)=37.57$, $p=.535$, but most cells were sparse with zero frequencies in 61.9% of cells. However, the final model significantly predicted the

dependent variable over and above the intercept-only model, $\chi^2(1)=6.65$, $p=.010$. A decrease in IQ Test Composite score was significantly associated with an increase in the odds of greater ASD symptom severity as measured by the ADI-R total algorithm score, with an odds ratio of .950, 95% CI [.910, .991], Wald $\chi^2(1)=5.55$, $p=.018$.

Aim 3

In order to examine strengths and weaknesses of IQ performance among individuals with ASD, two series of *t*-tests were conducted, both of which included corrections for multiple comparisons. First, to determine whether particular area scores represent strengths or weaknesses relative to the population means, a series of one-sample *t*-tests were conducted on the SB-IV for individuals with ASD. Specifically, analyses were conducted for the 22 individuals who met *DSM-5* diagnostic criteria for ASD using the relaxed standard. Descriptive statistics for these 22 individuals are presented in Table 4. Analyses were conducted using R (R Core Team, 2016). The mean scores of these 22 individuals for each SB-IV area (Verbal Reasoning, Abstract/Visual Reasoning, Quantitative Reasoning, Short-Term Memory, and the Test Composite) were compared to the population mean score of 100. Bonferroni corrections were applied to all analyses to adjust for multiple comparisons. The mean Verbal Reasoning ($t(21)=-4.54$, $p <.0005$), Abstract/Visual Reasoning ($t(21)=-7.0$, $p < .0005$), Quantitative Reasoning ($t(21)=-3.55$, $p < .0005$), Short-Term Memory ($t(21)=-6.68$, $p < .0005$), and Test Composite ($t(21)=-6.04$, $p < .0005$) scores for the 22 children with ASD were significantly lower than the population mean score of 100.

Next, to determine whether particular area scores represent relative strengths or weaknesses in relation to other areas for the participants with ASD, paired *t*-tests were conducted on each possible pairing of area scores. Analyses were conducted using R (R Core Team, 2016).

Bonferroni corrections were applied to all analyses in order to correct for multiple comparisons. It was hypothesized that children with ASD would demonstrate relative strengths with regard to the Abstract/Visual Reasoning and Quantitative Reasoning areas, and relative weaknesses in terms of the Verbal Reasoning and Short-Term Memory areas. One-tailed *t*-tests were conducted according to these hypotheses. The average Verbal Reasoning score was not significantly lower than the average Abstract/Visual Reasoning, Quantitative Reasoning, Short-Term Memory or Test Composite scores for the 22 participants with ASD ($t(21)=-.24$, $p=1.0$; $t(21)=-2.40$, $p=.13$; $t(21)=2.26$, $p=.35$; and $t(21)=1.42$, $p=1.0$, respectively). The average Abstract/Visual Reasoning score was not significantly different than the average Quantitative Reasoning, Short-Term Memory, or Test Composite score ($t(21)=-3.05$, $p=.06$; $t(21)=2.12$, $p=.23$; $t(21)=1.54$, $p=1.0$). The average Quantitative Reasoning score was significantly greater than the average Short-Term Memory score, $t(21)=4.47$, $p=.001$. The average Quantitative Reasoning score was significantly different than the average Test Composite score, $t(21)=5.83$, $p < .0005$. The average Short-Term Memory score was not significantly different than the average Test Composite score, $t(21)=-1.84$, $p=.80$. In summary, only the hypothesis that children with ASD would demonstrate strengths with regard to the Quantitative Reasoning area was supported, as the average Quantitative Reasoning score was significantly greater than both the average Short-Term Memory score and the average Test Composite score.

In order to determine whether IQ performance was significantly related to concordance, point-biserial correlations were conducted between each area score and concordance. A point-biserial correlation measures the strength and direction of the association between a continuous variable and a dichotomous variable. Because IQ is measured as a continuous variable and concordance is measured as a dichotomous variable, a point-biserial correlation was the most

appropriate statistic to use to assess the relationship between the two. Prior to calculating the correlations, several assumptions were checked. Visual examination of boxplots for each SB-IV area score revealed two outliers in the form of values below the first quartile in the distributions of both the Abstract/Visual Reasoning and the Quantitative Reasoning scores. The Shapiro-Wilk test of normality was used in order to assess whether area scores were normally distributed. The Quantitative Reasoning data did not appear to follow a normal distribution, W=.90, p=.02. Levene's test for homogeneity of variance was used in order to determine whether area scores had equal variances. Results of Levene's test for homogeneity of variance were nonsignificant for all area scores. Because the point-biserial correlation is robust to mild non-normality (Fowler, 1987), analyses were completed with and without including the outliers. Thus, seven point-biserial correlations were conducted, with all Verbal Reasoning, Abstract/Visual Reasoning, Quantitative Reasoning, Short-Term Memory, and Test Composite scores examined in relation to concordance, as well as two point-biserial correlations which excluded the outlier Abstract/Visual Reasoning and Quantitative Reasoning scores. Analyses were conducted using IBM SPSS Statistics (Version 24).

The point-biserial correlations between Verbal Reasoning, Abstract/Visual Reasoning, Quantitative Reasoning, Short-Term Memory, Test Composite scores and concordance were not significant (r_{pb}=-.05, n=25, p=.807; r_{pb}=-.27, n=25, p=.186; r_{pb}=-.15, n=25, p=.478; r_{pb}=-.10, n=25, p=.618; and r_{pb}=-.14, n=25, p=.490, respectively). When excluding the two outlier values for the Abstract/Visual Reasoning and Quantitative Reasoning tests, the point-biserial correlations remained nonsignificant (r_{pb}=-.22, n=23, p=.306 and r_{pb}=-.05, n=23, p=.812, respectively). Thus, there was not a significant relationship between IQ performance and concordance for any area score.

In order to clarify the relationship between ASD diagnosis and cognitive performance, exploratory analyses of the IQ performance of each individual twin pair were undertaken for all twin pairs with complete SB-IV data. Specifically, paired Bayesian *t*-tests were used for the twelve twin pairs with complete IQ data in order to determine whether the twins in each pair exhibited significantly different IQ strengths and weaknesses. All analyses were conducted using JASP (Version 0.8.0.0). The data were examined by estimating a Bayes factor using Bayesian Information Criteria (Wagenmakers, 2007), comparing the fit of the data under the null hypothesis and the alternative hypothesis. Lee and Wagenmakers (2013) created their own descriptive, approximate classification scheme for the interpretation of Bayes factors (BF_{10}), ranging from decisive evidence for the null hypothesis or alternative hypothesis, to no evidence. These criteria were used in the current study in order to interpret the degree of evidence for each hypothesis. Results are presented in Table 5.

Taken together, these results tentatively suggest that Verbal Reasoning, Quantitative Reasoning, and Test Composite scores were significantly different among the monozygotic twin pairs included in this sample.

Finally, paired Bayesian *t*-tests were used to determine whether concordant twins and discordant twins exhibited significantly different IQ strengths and weaknesses. Of the twelve twin pairs included in this analysis, 9 were concordant and 3 were discordant. All analyses were conducted using JASP (Version 0.8.0.0). Results are presented in Table 6.

Taken together, these results tentatively suggest that Verbal Reasoning, Quantitative Reasoning, and Test Composite scores were significantly different among this study's sample of concordant twins, but not its sample of discordant twins.

Aim 4

In order to examine strengths and weaknesses of neuropsychological performance among individuals with ASD, two series of *t*-tests were conducted, both including Bonferroni corrections for multiple comparisons. First, to determine whether particular subtest scores represented strengths or weaknesses relative to the population means, a series of one-sample *t*-tests were conducted on the NEPSY for individuals with ASD. Specifically, analyses were conducted for subjects who met *DSM-5* diagnostic criteria for ASD using the relaxed standard and who had data for the Design Copying, Phonological Processing, Memory for Faces, and Imitating Hand Positions subtests of the NEPSY. Descriptive statistics for these individuals are presented in Table 7. Analyses were conducted using R (R Core Team, 2016). The mean scores of these individuals for each of these four NEPSY subtests (Design Copying, Phonological Processing, Memory for Faces, and Imitating Hand Positions) were compared to the population mean scaled scores (mean=10) for the NEPSY standardization sample. The mean Design Copying, Phonological Processing, Memory for Faces, and Imitating Hand Positions scaled scores for children with ASD were significantly lower than the mean scores for the population ($t(22)=-7.11$, $p<.0005$; $t(22)=-3.94$, $p=.002$; $t(20)=-6.95$, $p<.0005$; and $t(20)=-5.56$, $p<.0005$, respectively).

Next, to determine whether particular NEPSY subtests represented relative strengths or weaknesses in relation to other subtests for the participants with ASD, paired *t*-tests were conducted on each possible pairing of NEPSY subtests. Analyses were conducted using R (R Core Team, 2016). Bonferroni corrections were applied to all analyses in order to adjust for multiple comparisons. It was hypothesized that children with ASD would demonstrate relative strengths in terms of the Design Copying and Memory for Faces subtests, and relative weaknesses with respect to the Imitating Hand Positions and Phonological Processing subtests.

As such, one-tailed *t*-tests were conducted for all tests which followed these directional hypotheses. The average Design Copying scaled score was not significantly greater than the average Phonological Processing, Memory for Faces, or Imitating Hand Positions scaled score for participants with ASD ($t(22)=-2.40$, $p=1.0$; $t(20)=-1.04$, $p=1.0$; and $t(20)=-1.74$, $p=1.0$, respectively). The average Phonological Processing scaled score was not significantly different than the averaged Memory for Faces or Imitating Hand Positions scaled score ($t(20)=1.99$, $p=.36$ and $t(20)=.80$, $p=1.0$, respectively). The average Memory for Faces scaled score was not significantly different than the average Imitating Hand Positions scaled score, $t(18)=-1.45$, $p=.99$. Thus, no significant differences in performance were found with regard to any NEPSY subtest for children with ASD.

In order to determine whether neuropsychological performance was significantly related to concordance, point-biserial correlations were conducted between each NEPSY subtest score and concordance. Because NEPSY scores are measured as continuous variables and concordance is measured as a dichotomous variable, a point-biserial correlation was the most appropriate statistic to use to assess the relationship between the two. Prior to calculating the correlations, several assumptions were checked. Visual examination of boxplots for the distributions of each NEPSY subtest score revealed two outliers, in the form of values above the third quartile, in the Memory for Faces scores. The Shapiro-Wilk test of normality was used in order to assess whether subtest scores were normally distributed. The Design Copying data did not appear to follow a normal distribution, $W=.92$, $p=.04$. The Memory for Faces data also did not appear to follow a normal distribution, $W=.90$, $p=.02$. Levene's test for homogeneity of variance was used in order to determine whether subtest scores had equal variances. Results of Levene's test for homogeneity of variance were nonsignificant for all NEPSY subtest scores. Because the point-

biserial correlation is robust to mild non-normality (Fowler, 1987), analyses were completed with and without including the outliers. Thus, five point-biserial correlations were conducted, with all Design Copying, Phonological Processing, Imitating Hand Positions, and Memory for Faces scores examined in relation to concordance, as well as a point-biserial correlation which excluded the outlier Memory for Faces scores. Analyses were conducted using IBM SPSS Statistics (Version 24).

The point-biserial correlation between Design Copying and concordance was r_{pb}=-.42, n=26, p=.031. The point-biserial correlations between Phonological Processing, Imitating Hand Positions, Memory for Faces scores and concordance were nonsignificant (r_{pb}=-.15, n=26, p=.454; r_{pb}=-.19, n=24, p=.365; r_{pb}=-.19, n=24, p=.366, respectively). When excluding the two outlier values, the point-biserial correlation between Memory for Faces and concordance remained nonsignificant (r_{pb}=-.07, n=22, p=.747). Thus, there was only a significant relationship between concordance and performance on the Design Copying subtest of the NEPSY.

Finally, in order to clarify the relationship between ASD diagnosis and neuropsychological performance, exploratory analyses of the NEPSY performance of each individual twin pair were undertaken. Table 8 presents NEPSY subtest scores, as well as IQ scores, by concordance. Specifically, paired Bayesian *t*-tests were used for the nine twin pairs with complete NEPSY data in order to determine whether the twins in each pair exhibited significantly different neuropsychological strengths and weaknesses. All analyses were conducted using JASP (Version 0.8.0.0). The data were examined by estimating a Bayes factor using Bayesian Information Criteria (Wagenmakers, 2007), comparing the fit of the data under the null hypothesis and the alternative hypothesis. Lee and Wagenmakers (2013) criteria for the

interpretation of Bayes factors (BF_{10}) were used in the current study to interpret the degree of evidence for each hypothesis. Results are presented in Table 9.

Taken together, the results of this analysis suggested that NEPSY subtest scores were not significantly different among the monozygotic twin pairs included in this sample.

Finally, paired Bayesian *t*-tests were used to determine whether concordant twins and discordant twins exhibited significantly different performance on these four NEPSY subtests. All analyses were conducted using JASP (Version 0.8.0.0). Results are presented in Table 10.

Taken together, these results suggested that NEPSY subtest performance was not significantly different for concordant and discordant twins, with the exception of the Memory for Faces subtest, in which discordant twins seemed to display significant differences in performance.

Discussion

Aim 1

Probands that were ascertained on the basis of *DSM-IV-TR* criteria as having a diagnosis of autistic disorder also met the diagnostic criteria for autism spectrum disorder using the *DSM-5.* Pairwise concordance was examined by requiring evidence of symptomatology from both the ADOS-2 and the ADI-R, in addition to the relaxed standard of using evidence from either one instrument or the other. Huerta et al. (2012) found that specificity of *DSM-5* diagnostic criteria increased, although sensitivity decreased, when evidence was required from both parent report and clinical observation. This is not surprising, given that additional sources of information will likely provide further evidence and thus may increase variability. As noted by Huerta et al. (2012), however, it is not practical to establish classifications solely using information from the ADOS, as there are no relevant items for certain diagnostic subdomains. For instance, using the item assignments created by Huerta et al. (2012), there are no ADOS items that are applicable for

sections B2 and B3 of DSM-5 criteria. The fewer ADOS items in comparison to ADI-R items likely limited the number of classifications that were made on the basis of this measure in the present study.

Because the changes in diagnostic criteria from *DSM-IV-TR* to *DSM-5* altered the diagnostic construct of autism spectrum disorder itself, there were widespread concerns that individuals who formerly had a diagnosis of an autism spectrum disorder would not retain it (Kulage et al., 2014; McPartland et al., 2012; Wing et al., 2011). In the present study, however, this does not appear to be the case. In this small sample of monozygotic twins, it appears that the *DSM-5* classification system is capturing most of the same individuals as did the *DSM-IV-TR* system.

It is important to note, however, that these results would likely have been different if both twins had a *DSM-IV-TR* diagnosis of any pervasive developmental disorder. Because the first twin ascertained was required to have autistic disorder rather than any pervasive developmental disorder, as follows from the requirement of meeting criteria for autism on the ADI-R and scoring within one point of the criteria for autism on the ADOS-G, the present results are limited in their ability to generalize to the entire autism spectrum.

As might be expected, differences in methodology changed concordance rates among the monozygotic twin pairs in this sample. In general, the use of both parent report and clinical observation is considered to be the gold standard in diagnostic practice for ASD, as neither source is as informative alone as it is in combination. In the present study, pairwise concordance rates were higher if diagnoses were assigned on the basis of information from either the ADOS-2 or the ADI-R, rather than requiring evidence from both measures to meet diagnostic criteria. This "relaxed" standard is much more relevant to clinical practice, as the two measures can provide

differing snapshots of behavior. For instance, the ADI-R, but not the ADOS, includes specific items related to certain *DSM-5* criteria, such as peer interactions. In addition, restricted and repetitive behaviors may exist outside of the testing context, but may not be observed during the administration of the ADOS. The significant difference in the pairwise concordance rate for *DSM-5* criteria when requiring evidence from both the ADOS and ADI-R versus either the ADOS or the ADI-R suggests that the ASD diagnostic measures that are used can have a considerable influence on the resulting clinical diagnosis.

In the present study, social communication disorder was not included as being concordant with ASD, as it is a separate diagnostic entity in the *DSM-5*. However, some evidence suggests that social communication disorder is highly related to ASD (Swineford, Thurm, Baird, Wetherby, & Swedo, 2014). If the definition of concordance was broadened to include the presence of social communication disorder in the co-twin of an individual with ASD, there was little change in pairwise concordance rates among the monozygotic twins included in this sample. Although previous research has provided evidence in support of the independence of the domains of social communication and restricted, repetitive patterns of behavior (Ronald et al., 2006), further research is needed regarding the differential diagnosis of social communication disorder and ASD in order to clarify the degree of overlap between these disorders.

Both the *DSM-IV-TR* and the *DSM-5* diagnostic criteria remain consistent in supporting a genetic influence inherent to ASD, with respective sample pairwise concordance rates of 57.14 percent and 35.71 percent using evidence from both the ADOS-2 and the ADI-R. However, it should be noted that there are several shortcomings involved in using pairwise concordance rates. As illustrated by McGue (1992), the pairwise concordance rates of the twins in this sample are not directly comparable to the overall population prevalence of ASD. In addition, the sample

pairwise rate cannot be generalized to estimate the population pairwise concordance of ASD. Future research involving the genetic and epigenetic mechanisms contributing to monozygotic twin ASD concordance will be particularly useful to clarify genetic factors that confer susceptibility to this neurodevelopmental disorder.

An important distinction should be drawn between clinical diagnosis and the way in which diagnoses were assigned in this study. Usually, diagnosis is reliant on a combination of direct observation of the child's behavior, caregiver description of the child's developmental history, and clinical judgment. In the present study, diagnoses were made following *DSM-IV-TR* and *DSM-5* diagnostic criteria on the basis of symptom counts. This would not be appropriate in clinical practice, but was adopted in the present study in order to remain consistent in methodology.

One limitation inherent in this aim is that direct translation between former and current measures was not always possible. Although most items from the ADOS-G (Lord et al., 2000) were easily transferred to the ADOS-2 (Lord et al., 2012), one item did not have an exact match. Amount of social overtures/maintenance of attention—item B8 on ADOS-2 Module 3—mapped onto *DSM-IV-TR* Section A1 criteria, but there was no similar item on ADOS-G Module 3. As a result, this item was not included in assigning diagnoses for participants who were administered this module. Importantly, this did not impact whether or not participants met diagnostic criteria or influence the degree of concordance for any twin pair. In addition, one twin pair was administered an ADOS-G Module 3, but their behaviors were scored using an ADOS-G Module 2 algorithm. Three items did not correspond between ADOS-G Modules 2 and 3 (B4: response to name; B5: showing; B7: response to joint attention) and thus were scored as a zero for this twin pair. This did not affect diagnoses or concordance, as there were several other relevant ADOS-G

items within Section A, criterion 1 (deficits in social-emotional reciprocity) and criterion 2 (nonverbal communicative behaviors used for social interaction) of *DSM-5* diagnostic criteria. Thus, although this twin pair was classified as discordant, it was not due to the missing items between Modules 2 and 3.

Aim 2

When the Calibrated Severity Score (CSS; Gotham, Pickles, & Lord, 2009) was used to assess severity of ASD symptoms as indexed by the Autism Diagnostic Observation Schedule, a decrease in IQ as measured by the Test Composite of the Stanford-Binet Intelligence Scale, Fourth Edition (SB-IV; Thorndike, Hagen, & Sattler, 1986) did not predict an increase in the odds of greater severity of ASD symptoms. This finding is in fitting with previous work using the monozygotic twins of this cohort by Mitchell et al. (2009), who found no association between IQ scores and symptom severity when measured using the total score on the ADOS. This provides convergent evidence for the conclusions of Wilson et al. (2014), who suggested that cognitive ability is not predictive of ASD symptom severity.

Counter to this finding, however, when the total algorithm score of the Autism Diagnostic Interview-Revised was used to assess severity of ASD symptoms, a decrease in IQ was significantly associated with an increase in the odds of greater severity of ASD symptoms. Because the ADI-R is able to capture a wider range of symptoms due its use of a developmental history as reported by a parent or caregiver, rather than focusing solely on current behavior as rated by an expert observer in the ADOS, it is likely that these two measures provide differing snapshots of symptoms, which may affect the degree to which IQ predicts the outcome of ASD symptom severity. In addition, because ADI-R scores are the result of parent/caregiver ratings, it is possible that IQ differentially affects ADI-R scores to a greater extent than it does ADOS scores, in that the parent or caregiver may provide a wider perspective of their child's behavior

which incorporates both symptoms and impairment, rather than solely a measure of ASD symptoms.

It is important to consider that scores on the Autism Diagnostic Interview-Revised were not normalized for the purpose of estimating the severity of core autism features. Gotham, Pickles, and Lord (2009) note that although higher scores often indicate that an individual has a greater number of items which may represent greater impairment, nonverbal children are not scored on roughly one-quarter of the total ADI-R items, which results in summary scores, particularly for the communication domain, being restricted by non-random missing data. This was not the case for any of the twin pairs in the current sample; however, the use of total ADI-R algorithm scores still represents a limitation in the interpretation of this result. In addition, a median split likely does not take into account the degree of ASD symptom severity as measured by the ADI-R, but was adopted in the present study as there is no existing metric to qualify severity on the basis of ADI-R algorithm scores.

Aim 3

In the present sample, children with ASD demonstrated significantly lower IQ scores across all domains of the Stanford-Binet Intelligence Scale, Fourth Edition as compared to the normative sample. This result is not surprising, as a growing body of research has found that cognitive measures, such as the Wechsler scales, which rely on language and the comprehension of verbal instructions often underestimate the abilities of individuals with ASD (Dawson, Soulieres, Gernsbacher, & Mottron, 2007). Nonverbal measures of intelligence, such as the Raven's Progressive Matrices (Raven, 1998) have been shown to be a better assessment of cognitive abilities in many individuals with ASD (Bolte et al., 2009; Courchesne et al., 2015; Dawson et al., 2007). Because the SB-IV relies on verbal instructions, it is possible that this measure does not appropriately assess these individuals' abilities.

When area scores were examined relative to one another for participants with ASD, the average Quantitative Reasoning score was significantly different than both the average Short-Term Memory score and the average Test Composite score, in fitting with the hypothesis that children with ASD would demonstrate relative strengths with regard to the Quantitative Reasoning domain. Again, it is possible that individuals with ASD demonstrated a significant strength in this area due to the lesser emphasis on verbal abilities within this domain.

In addition, the relationship between concordance and IQ was not significant for any area score, suggesting that cognitive performance was not related to the presence or absence of an ASD diagnosis in this sample of monozygotic twins.

Finally, when comparing the IQ performance within twin pairs, an exploratory Bayesian analysis suggested significant differences for Verbal Reasoning, Quantitative Reasoning, and Test Composite scores among the monozygotic twin pairs included in the sample. When examining concordant and discordant pairs separately, the Bayesian analysis revealed that concordant twins, but not discordant twins, demonstrated significant differences for Verbal Reasoning, Quantitative Reasoning, and Test Composite scores. Originally, it was hypothesized that an ASD diagnosis and IQ performance would be orthogonal, such that concordant twins would not consistently demonstrate similar patterns of IQ performance and conversely, that discordant twins would not consistently demonstrate dissimilar patterns of IQ performance. These results support this hypothesis, providing preliminary evidence that a diagnosis of ASD and IQ are not necessarily related. It is important to note, however, that there is a larger proportion of concordant twin pairs than discordant twin pairs in the present sample, which likely drove the difference among the overall sample. Thus, these results should not be considered conclusive.

Aim 4

In the present sample, children with ASD demonstrated significantly lower NEPSY scores across the Design Copying, Phonological Processing, Imitating Hand Positions, and Memory for Faces subtests as compared to the normative sample. Reinvall et al. (2013) previously found in a sample of adolescents with Asperger syndrome that lower WISC-III scores were related to impaired neuropsychological performance as measured by the NEPSY-II. This may explain the lower performance on these NEPSY subtests, as the mean Test Composite score for the present sample was 73.41.

The hypothesis that children with ASD would demonstrate relative strengths on the Design Copying and Memory for Faces subtests and relative weaknesses on the Imitating Hand Positions and Phonological Processing subtests of the NEPSY was not supported. When examining relative performance on these four NEPSY subtests for children with ASD, no significant differences in performance were found with regard to any NEPSY subtest. In general, the performance of the children included in the sample was consistently low across these four subtests, which may explain the lack of relative strengths and weaknesses.

The relationship between concordance and neuropsychological performance as measured by the NEPSY was significant only for the Design Copying subtest. The Design Copying subtest, part of the visual-spatial domain of the NEPSY, assesses the ability to copy two-dimensional geometric figures. It is important to note that there are motor demands inherent to this subtest. Consequently, it is possible that visual-spatial abilities are uniquely related to concordance, or it may be that related but distinct task demands (such as motor abilities) also play a role. With the exception of the Design Copying subtest, however, these findings suggest that neuropsychological performance is largely independent of the presence or absence of an ASD diagnosis among the twins included in this sample.

Finally, when comparing NEPSY performance within twin pairs, no significant differences were found for any subtest. When examining concordant and discordant twin pairs separately, scores on the Memory for Faces subtest were significantly different for discordant twins; however, the Bayes factor was small and thus provides only anecdotal evidence for this finding. No other differences were significant for either concordant or discordant twins for the Design Copying, Phonological Processing, or Imitating Hand Positions subtests of the NEPSY. Originally, it was hypothesized that an ASD diagnosis and NEPSY performance would be significantly related, such that concordant twins would demonstrate more similar patterns of NEPSY performance than would discordant twins. These results counter this hypothesis, providing no evidence for such a relationship.

Limitations

There are a number of limitations inherent to the present study. Because the children were selected for this sample on the basis of their scores on the ADI-R and the ADOS, and the present study assessed concordance by using scores on these same measures, ascertainment is not ideal. It would be preferable to use *DSM-IV* clinical diagnoses to select participants for the sample, and then to use the instrument scores as a proxy measure of concordance.

A central limitation of this study is with regard to statistical power. Due to the small number of participants and the large number of research questions subjected to statistical significance testing, the present study is underpowered to detect an effect. Thus, in the current study, the probability of making a Type I error, or rejecting the null hypothesis when it is in fact true, is low, but the probability of making a Type II error is high. Because a power analysis was not conducted, the sample size required to detect an effect of a given size with a given level of confidence is not known. Although the present study found no difference in ASD concordance rates from *DSM-IV-TR* to *DSM-5*; found no significant relationship between concordance and

IQ; found no significant differences in performance for any NEPSY subtest for individuals with ASD; and no significant differences in NEPSY performance between concordant and discordant twin pairs, it is possible that there were indeed such differences or relationships that the present study did not detect.

Implications

Although this study is necessarily limited in its ability to draw firm conclusions due to the lack of statistical power, there are a number of clinical and research implications applicable to these results. Because the current sample included participants with lower cognitive functioning, the present sample is likely more representative of the overall population of ASD. Many research studies include only high-functioning individuals with ASD, including those formerly diagnosed with Asperger syndrome, and thus are limited in their ability to generalize to the entire autism spectrum. Because approximately 1.2 million children in the United States have an autism spectrum disorder, and intellectual disability is found in approximately 31.6 percent of children with an ASD (CDC, 2016), the results of the present study provide further information about cognitive and neuropsychological functioning that is more relevant to the approximately 350,400 children in the United States with ASD and a comorbid intellectual disability.

Perhaps most notably, results of this study suggest that pairwise concordance rates for ASD have not substantially changed from *DSM-IV-TR* to *DSM-5* diagnostic criteria. Given the potential for the composition of the autism spectrum to change with the reconceptualization of autism into a single diagnostic entity, these results provide some support that the *DSM-5* system is capturing most of the same individuals as did the previous diagnostic classification system; however, the significant difference in pairwise concordance rate for *DSM-5* criteria when requiring evidence from both the ADOS and ADI-R versus either the ADOS or the ADI-R suggests that concordance rates are higher when only requiring information from one diagnostic

instrument. Clinicians and researchers alike should be judicious in their selection of diagnostic instruments, and in keeping with best practices (Huerta & Lord, 2012), should include an assessment of multiple domains of functioning and behavior, incorporate both child testing and parent interviews, and consider developmental factors when making a diagnosis.

The finding of a significant difference in concordance rates when using either as opposed to both diagnostic instruments has important implications for public policy. In general, the use of both the ADOS and the ADI-R is considered to be the gold standard in ASD diagnosis; however, in clinical practice, both measures may not be used when making a diagnosis. Clinicians who are limited in terms of time and resources may elect to administer one instrument rather than both, which makes the "relaxed standard" reported in the present study much more relevant to clinical diagnosis. For instance, in clinical practice, it may be more common for an evaluation to include only the ADOS rather than both the ADOS and ADI-R, and for the resulting diagnosis to be based upon this instrument. Because the ADI-R and the ADOS incorporate different sources of information (e.g., parent or caregiver versus clinician), use of one measure rather than both provides a diagnostic decision that is based on a more limited sample of information. Thus, it is possible that prevalence rates of ASD may be over- or under-estimated due to differences in the method used for diagnosis.

The finding of a significant difference in concordance rates based on the source of information has particular implications for American public schools in terms of educational classifications for autism. Under the Individuals with Disabilities Education Act (IDEA; 2004), autism is defined as "a developmental disability significantly affecting verbal and nonverbal communication and social interaction, generally evident before age three, that adversely affects a child's educational performance." This definition does not require the use of gold standard

diagnostic measures for diagnosis; however, given the results of the present data, one can hypothesize that educational classification rates of ASD will be higher with the use of only one source of information as opposed to using multiple measures which provide an assessment of current symptoms as well as a developmental history.

Relatedly, the finding that IQ is not predictive of ASD symptom severity as measured by the ADOS-2 Calibrated Severity Score, but is predictive of ASD symptom severity as measured by the ADI-R, suggests that the type of instrument, as well as how it is used, can have an important impact on the conclusions that clinicians and researchers are able to draw. For instance, the ADI-R was not normed as a measure of ASD symptom severity, and as such, it is inappropriate to base estimates of symptom severity on this measure. As part of a clinical assessment, it is often necessary to include a statement of severity, including the resulting functional impairment that the individual experiences. Knowing that other measures that provide autism severity ratings tend to yield scores that are strongly correlated with IQ (Spiker et al., 2002; Szatmari et al., 2003), it is important for clinicians and researchers alike to use measures that do not conflate the two. Because the ADOS-2 Calibrated Severity Score provides a measure of ASD severity with independence from characteristics such as age and verbal IQ (Gotham, Pickles, & Lord, 2009), one can be more confident in the present study's conclusion that cognitive ability is not predictive of ASD symptom severity.

Though the present study is limited in its ability to draw firm conclusions about genetic influences due to the lack of a dizygotic twin sample, one potential implication of these findings concerns the relationship between cognitive performance and concordance. The finding that monozygotic twins who were concordant for ASD demonstrated significant differences in IQ performance, but that those who were discordant did not, may suggest that cognitive ability and

ASD are orthogonal, even in individuals with identical genotypes. Although previous research has shown relatives of individuals with ASD to have similar behaviors, such as social or language deficits, that are qualitatively similar to those of the proband with ASD (Dawson et al., 2007; Pizula & Ziegart-Sadowska, 2015), the present results imply that this relationship does not apply to cognitive ability. Clinicians who evaluate individuals with ASD should be mindful that not only is the disorder's phenotypic presentation heterogeneous among the general population, but that its presentation, severity, and associated characteristics are equally as varied among individuals with the same genotype.

The importance of treating each person with ASD as an individual cannot be understated; implementation of individualized interventions and treatment plans is essential. In educational settings, individualization becomes all the more important as students with ASD who receive special education services from public schools are provided with an individualized education program (IEP), as specified by IDEA (2004). Because an IEP functions as a blueprint for services, it is essential that such services for children with ASD are truly individualized, instead of applied as in a general fashion. Likely due to the heterogeneity of needs and abilities within the population of individuals with ASD, there is no single approach or program that will be effective for all students with ASD (Heflin & Simpson, 1998; Iovannone et al., 2003). For instance, some students with ASD may need minor accommodations in order to function effectively in the school setting, while others may need a substantial amount of support. Regardless, school personnel should be cognizant not only of the areas in which a student may have difficulty, but also areas in which the student will likely excel. The incorporation of relative strengths as well as areas for growth within an IEP may provide more accurate targets for intervention. Having an understanding of both the cognitive and neuropsychological strengths

and weaknesses of children with ASD can serve to improve their educational experiences during their school-age years which consequently can have a positive impact on their future outcomes.

Table 1

Summary of Previous Studies of NEPSY Performance in ASD Samples

Study	**Measures Used**	**Sample Composition**	**Significant Weaknesses**
Narzisi et al. (2013)	NEPSY-II (Italian version)	22 boys with high-functioning ASD compared to typically developing control children	7 out of 9 Attention and Executive Functions subtests; Oromotor Sequences; 5 of 7 Memory and Learning subtests; Imitating Hand Postures; Manual Motor Sequences; Theory of Mind; Design Copying; Arrows
Reinvall et al. (2013	NEPSY-II (Finnish version)	30 adolescents with Asperger syndrome compared to 30 typically developing adolescents	Auditory Attention and Response Set; Memory for Faces; Visuomotor Precision; Design Copying

Table 2

Participant Characteristics (N=28)

Score	Mean	Standard Deviation	Range
Age (years)	8.0	1.96	5.58 – 12.2
IQ*	77	21.29	36 – 109
ADOS-2 Total	12.11	6.34	1 – 24
ADOS-2 Comparison Score	5.21	2.38	1 – 10
ADI-R Social Interaction	16.0	8.78	0 – 29
ADI-R Verbal Communication	13.0	6.06	0 – 22

*3 participants were missing IQ data.

Table 3

Correlations Between SB-IV Area Scores

Area	Verbal Reasoning SAS	Abstract/ Visual Reasoning SAS	Quantitative Reasoning SAS	Short-Term Memory SAS	Test Composite SAS
Verbal Reasoning SAS	1	.			
Abstract/Visual Reasoning SAS	.746**	1			
Quantitative Reasoning SAS	.787**	.877**	1		
Short-Term Memory SAS	.872**	.726**	.776**	1	
Test Composite	.934**	.884**	.922**	.925**	1

**Correlation is significant at the 0.01 level (2-tailed).

Table 4

Stanford-Binet IV Scores for Children with ASD

Area	Mean	Standard Deviation
Verbal Reasoning	76.05	24.75
Abstract/Visual Reasoning	76.91	15.48
Quantitative Reasoning	83.86	21.34
Short-Term Memory	70.45	20.76
Test Composite	73.41	20.64

Table 5

Paired Bayesian t-tests of IQ Performance of Twin Pairs

Area	BF_{10}	Interpretation
Verbal Reasoning	9.43	Moderate evidence for H_α
Abstract/Visual Reasoning	.45	Anecdotal evidence for H_0
Quantitative Reasoning	3.05	Anecdotal evidence for H_α
Short-Term Memory	1.16	Anecdotal evidence for H_0
Test Composite	2.17	Anecdotal evidence for H_α

Table 6

Paired Bayesian t-tests of IQ Performance of Concordant and Discordant Twin Pairs

Area	Diagnosis	BF_{10}	Interpretation
Verbal Reasoning	Concordant	4.29	Moderate evidence for H_α
	Discordant	1.09	Anecdotal evidence for H_0
Abstract/Visual Reasoning	Concordant	.41	Anecdotal evidence for H_0
	Discordant	.55	Anecdotal evidence for H_0
Quantitative Reasoning	Concordant	2.04	Anecdotal evidence for H_α
	Discordant	.80	Anecdotal evidence for H_0
Short-Term Memory	Concordant	.54	Anecdotal evidence for H_0
	Discordant	1.03	Anecdotal evidence for H_0
Test Composite	Concordant	1.53	Anecdotal evidence for H_α
	Discordant	.82	Anecdotal evidence for H_0

Table 7

NEPSY Subtest Scores for Children with ASD

Subtest	Mean	Standard Deviation
Design Copying	5.09	3.32
Phonological Processing	6.57	4.18
Memory for Faces	5.62	2.89
Imitating Hand Positions	6.19	3.14

Table 8

IQ Standard Scores and NEPSY Subtest Scaled Scores by Concordance

Twin Pair	DSM-5 Diagnosis (relaxed standard)	ADOS-2 Comparison Score	SB-IV Test Composite	NEPSY Design Copying	NEPSY Phonological Processing	NEPSY Memory for Faces	NEPSY Imitating Hand Positions
1A	ASD	7	-	1	1	2	-
1B	ASD	3	75	4	3	5	9
2A	ASD	10	72	4	15	6	5
2B	ASD	7	105	10	11	8	8
3A	ASD	6	59	2	2	3	3
3B	ASD	6	67	6	2	5	8
4A	ASD	7	53	1	1	3	2
4B	None	1	109	14	11	10	13
5A	ASD	6	36	1	1	-	1
5B	ASD	4	36	1	1	1	1
6A	ASD	7	58	2	7	4	6
6B	ASD	6	73	7	9	5	6
7A	ASD	2	104	12	11	13	13
7B	ASD	7	96	4	6	6	8
8A	ASD	6	83	5	6	5	4
8B	ASD	6	77	3	7	4	3
9A	ASD	6	50	7	6	-	8
9B	None	1	105	14	14	13	12
10A	ASD	8	80	5	6	6	6
10B	ASD	6	104	11	15	12	11
11A	ASD	6	-	-	-	-	-
11B	ASD	8	-	-	-	-	-
12A	ASD	6	49	2	7	6	-
12B	ASD	1	79	9	9	4	6
13A	ASD	5	87	6	8	5	9
13B	ASD	2	90	7	9	8	7
14A	None	2	96	12	10	3	6
14B	ASD	4	82	7	8	7	6

- : denotes missing data.

Table 9

Paired Bayesian t-tests of NEPSY Performance of Twin Pairs

Subtest	**BF_{10}**	**Interpretation**
Design Copying	1.02	Anecdotal evidence for H_0
Phonological Processing	.64	Anecdotal evidence for H_0
Memory for Faces	.64	Anecdotal evidence for H_0
Imitating Hand Positions	.64	Anecdotal evidence for H_0

Table 10

Paired Bayesian t-tests of NEPSY Performance of Concordant and Discordant Twin Pairs

Subtest	Diagnosis	BF_{10}	Interpretation
Design Copying	Concordant	.74	Anecdotal evidence for H_0
	Discordant	.64	Anecdotal evidence for H_0
Phonological Processing	Concordant	.37	Anecdotal evidence for H_0
	Discordant	.84	Anecdotal evidence for H_0
Memory for Faces	Concordant	.38	Anecdotal evidence for H_0
	Discordant	1.44	Anecdotal evidence for H_α
Imitating Hand Positions	Concordant	.37	Anecdotal evidence for H_0
	Discordant	.90	Anecdotal evidence for H_0

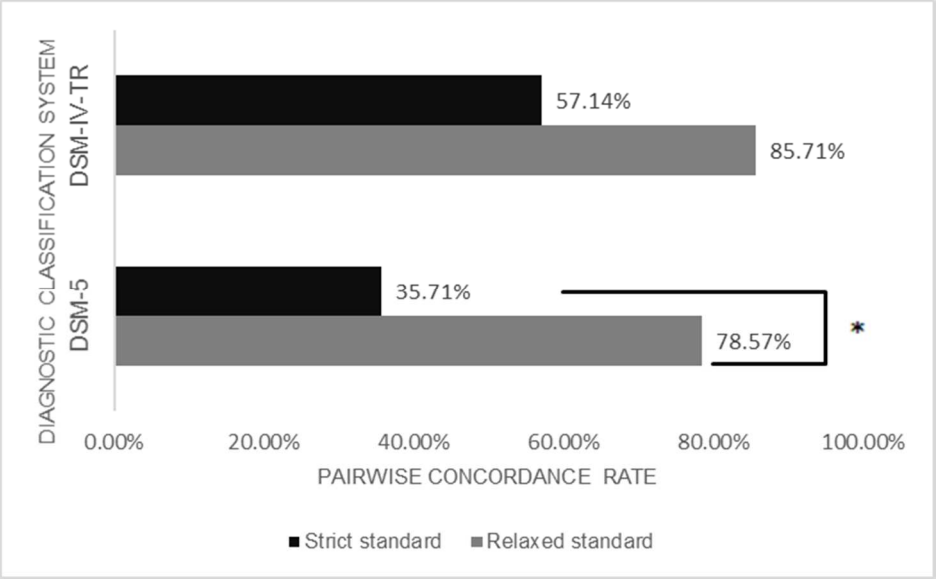

Figure 1. Pairwise concordance rates for monozygotic twin pairs according to diagnostic criteria for *DSM-IV-TR* pervasive developmental disorders and *DSM-5* autism spectrum disorder.

References

Abell, F., Happe, F., & Frith, U. (2000). Do triangles play tricks? Attribution of mental states to animated shapes in normal and abnormal development. *Cognitive Development, 15,* 1-6.

American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*, *3: DSM-III*. Washington, DC: American Psychiatric Association; 1980.

American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, Revised Third Edition.* Washington, DC: American Psychiatric Association; 1987.

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

American Psychiatric Association. *Diagnostic and statistical manual of mental disorders fourth edition (text revision): DSM-IV-TR*. Washington, DC: American Psychiatric Association; 2000.

American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (5th ed.): DSM-5*. Washington, DC: American Psychiatric Publishing; 2013.

Ankenman, K., Elgin, J., Sullivan, K., Vincent, L., & Bernier, R. (2014). Nonverbal and verbal cognitive discrepancy profiles in autism spectrum disorders: influence of age and gender. *American Journal on Intellectual and Developmental Disabilities, 119*, 84-99. doi: 10.1352/1944-7558-119.1.84

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*, 63-77. doi: 10.1017/S0033291700028099

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "reading the mind in the eyes" test revised version: A study with normal adults and adults with Asperger

syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 42,* 241-251. doi: 10.1111/1469-7610.00715

Bishop, D. V. (1989). Autism, Asperger's syndrome and semantic-pragmatic disorder: where are the boundaries? *British Journal of Disorders of Communication, 24,* 107-121.

Bishop, S. L., Richler, J., & Lord, C. (2006). Association between restricted and repetitive behaviors and nonverbal IQ in children with autism spectrum disorders. *Child Neuropsychology, 12* (4–5), 247–67. doi: 10.1080/09297040600630288

Bolte, S., Dziobek, I., & Poustka, F. (2009). Brief report: the level and nature of autistic intelligence revisited. *Journal of Autism and Developmental Disorders, 39*, 678-682. doi: 10.1007/s10803-008-0667-2

Brook, S. L., & Bowler, D. M. (1992). Autism by another name? Semantic and pragmatic impairments in children. *Journal of Autism and Developmental Disorders, 22*(1), 61-81. doi: 10.1007/BF01046403

Centers for Disease Control and Prevention, 2016. Prevalence and characteristics of autism spectrum disorder among children aged 8 years—Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2012. *MMWR Surveill Summaries*. 2016; *65*: 1-23. doi: 10.15585/mmwr.ss6503a1

Charman, T., Pickles, A., Simonoff, E., Chandler, S., Loucas, T., & Baird, G. (2011). IQ in children with autism spectrum disorders: data from the Special Needs and Autism Project (SNAP). *Psychological Medicine, 41,* 619-627. doi: 10.1017/S0033291710000991

Cohen, J. (1988*). Statistical power analysis for the behavioral sciences* (*2nd ed*.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Courchesne, V., Meilleur, A.-A. S., Poulin-Lord, M.-P., Dawson, M., & Soulieres, I. (2015). Autistic children at risk of being underestimated: school-based pilot study of a strength-informed assessment. *Molecular Autism, 6*(1), 12.

Damarla, S. R., Keller, T. A., Kana, R. K., Cherkassky, V. L., Williams, D. L., Minshew, N. J., & Just, M. A. (2010). Cortical underconnectivity coupled with preserved visuospatial cognition in autism: evidence from an fMRI study of an embedded figures task. *Autism Research, 3,* 273-279. doi: 10.1002/aur.153

Dawson, G., Estes, A., Munson, J., Schellenberg, G., Bernier, R., & Abbott, R. (2007). Quantitative assessment of autism symptom-related traits in probands and parents: Broader Phenotype Autism Symptom Scale. *Journal of Autism and Developmental Disorders, 37*, 523-536. doi: 10.1007/s10803-006-0182-2

Dawson, M., Soulieres, I., Gernsbacher, M. A., & Mottron, L. (2007). The level and nature of autistic intelligence. *Psychological Science, 18*, 657-662. doi: 10.1111/j.1467-9280.2007.01954.x

de Bildt, A., Sytema, S., Ketelaars, C., Kraijer, D., Mulder, E., Volkmar, F., & Minderaa, R. (2004). Interrelationship between Autism Diagnostic Observation Schedule-Generic (ADOS-G), Autism Diagnostic Interview-Revised (ADI-R), and the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR) classification in children and adolescents with mental retardation. *Journal of Autism and Developmental Disorders, 34,* 129-137. doi: 10.1023/B:JADD.0000022604.22374.5f

de Bruin, E. I., Verheij, F., & Ferdinand, R. F. (2006). WISC-R subtest but no overall VIQ-PIQ difference in Dutch children with PDD-NOS. *Journal of Abnormal Child Psychology, 34*, 254-262. doi: 10.1007/s10802-005-9018-3

Deng, W., Zou, X., Deng, H., Li, J., Tang, C., Wang, X., & Guo, X. (2015). The relationship among genetic heritability, environmental effects, and autism spectrum disorders. *Journal of Child Neurology, 30,* 1794-1799. doi: 10.1177/0883073815580645

Ehlers, S., Nyden, A., Gillberg, C., Sandberg, A. D., Dahlgren, S.-O., Hjelmquist, E., & Oden, A. (1997). Asperger syndrome, autism and attention disorders: a comparative study of the cognitive profiles of 120 children. *Journal of Child Psychology and Psychiatry, 38*, 207-217. doi: 10.1111/j.1469-7610.1997.tb01855.x

Elliott, C. D. (1997). The differential ability scales. *Contemporary intellectual assessment: Theories, tests, and issues*, 183-208.

Firkowska-Mankiewicz, A. (2011). Adult careers: does childhood IQ predict later life outcome? *Journal of Policy and Practice in Intellectual Disabilities, 8*, 1-9. doi: 10.1111/j.1741-1130.2011.00281.x

Folstein, S., & Rutter, M. (1977). Infantile autism: a genetic study of 21 twin pairs. *Journal of Child Psychology and Psychiatry, 18,* 297-321. doi: 10.1111/j.1469-7610.1977.tb00443.x

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

Fowler, R. L. (1987). Power and robustness in product-moment correlation. *Applied Psychological Measurement, 11*(4), 419-428.

Frazier, T. W., Thompson, L., Youngstrom, E. A., Law, P., Hardan, A. Y., Eng, C., & Morris, N. (2014). A twin study of heritable and shared environmental contributions to autism. *Journal of Autism and Developmental Disorders, 44*, 2013-2025. doi: 10.1007/s10803-014-2081-2

Gathercole, S. E., Willis, C. S., Baddeley, A. D., & Emslie, H. (1994). The children's test of nonword repetition: A test of phonological working memory. *Memory, 2,* 103-127. doi: 10.1080/09658219408258940

Gaugler, T., Klei, L., Sanders, S. J., Bodea, C. A., Goldberg, A. P., … & Buxbaum, J. D. (2014). Most genetic risk for autism resides with common variation. *Nature Genetics, 46,* 881-885. doi: 10.1038/ng.3039

Ghaziuddin, M. (2010). Brief report: should the DSM-V drop Asperger syndrome? *Journal of Autism and Developmental Disabilities, 40*, 1146-1148. doi: 10.1007/s10803-010-0969-z

Golden, C. J. (1980). *The Luria-Nebraska neuropsychological battery: manual*.

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. doi: 10.1007/s10803-008-0674-3

Grant, D. A., & Berg, E. A. (1948). A behavioral analysis of the degree of reinforcement and ease of shifting to new responses in a Weigl-type card sorting problem. *Journal of Experimental Psychology, 38,* 404-411. doi: 10.1037/h0059831

Gray, K. M., Tonge, B. J., & Sweeney, D. J. (2008). Using the Autism Diagnostic Interview-Revised and the Autism Diagnostic Observation Schedule with young children with developmental delay: evaluating diagnostic validity. *Journal of Autism and Developmental Disorders, 38,* 657-667. doi: 10.1007/s10803-007-0432-y

Hallmayer, J., Cleveland, S., Torres, A., Phillips, J., Cohen, B., Torigoe, T., Miller, J. … & 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

Heflin, L. J., & Simpson, R. L. (1998). Interventions for children and youth with autism: prudent choices in a world of exaggerated claims and empty promises. Part 1: intervention and treatment option review. *Focus on Autism and Other Developmental Disabilities, 13,* 194-211.

Holwerda, A., van der Klink, J. J. L., Groothoff, J. W., & Brouwer, S. (2012). Predictors for work participation in individuals with an autism spectrum disorder: a systematic review. *Journal of Occupational Rehabilitation, 22*, 333-352. doi: 10.1007/s10926-011-9347-8

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. *The American Journal of Psychiatry, 169*, 1056-1064. doi: 10.1176/appi.ajp.2012.12020276

Huerta, M., & Lord, C. (2012). Diagnostic evaluation of autism spectrum disorders. *Pediatric Clinics of North America, 59,* 103-111. doi: 10.1016/j.pcl.2011.10.018

Individuals with Disabilities Education Act, 20 U.S.C. § 1400 (2004).

Iovannone, R., Dunlap, G., Huber, H., & Kincaid, D. (2003). Effective educational practices for students with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities, 18*(3), 150-165.

JASP Team (2016). JASP (Version 0.8.0.0) [Computer software].

Jolliffe, T., & Baron-Cohen, S. (1997). Are people with autism and Asperger syndrome faster than normal on the embedded figures test? *Journal of Child Psychology and Psychiatry, 38,* 527-534. doi: 10.1111/j.1469-7610.1997.tb01539.x

Kanne, S. M., Gerber, A. J., Quirmbach, L. M., Sparrow, S. S., Cicchetti, D. V., & Saulnier, C. A. (2011). The role of adaptive behavior in autism spectrum disorders: implications for

functional outcome. *Journal of Autism and Developmental Disorders, 41*, 1007-1018. doi: 10.1007/s10803-010-1126-4

Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child, 2*, 217-250.

Kates, W. R., Burnette, C. P., Eliez, S., Strunge, L. A., Kaplan, D., Landa, R., Reiss, A. L., & Pearlson, G. D. (2004). Neuroanatomic variation in monozygotic twin pairs discordant for the narrow phenotype for autism. *American Journal of Psychiatry, 161*, 539-546. doi: 10.1176/appi.ajp.161.3.539

Kates, W. R., Mostofsky, S. H., Zimmerman, A. W., Mazzocco, M. M. M., Landa, R., Warsofsky, I. L., Kaufmann, W. E., & Reiss, A. L. (1998). Neuroanatomical and neurocognitive differences in a pair of monozygous twins discordant for strictly defined autism. *Annals of Neurology, 43,* 782-791. doi: 10.1002/ana.410430613

Kaufman, A. S., & Kaufman, N. L. (1983). *K-ABC--Kaufman assessment battery for children: Administration and scoring manual*. American Guidance Service.

Kaufman, A. S., & Kaufman, N. L. (1985). *Kaufman Test of Educational Achievement.* Circle Pines, MN: American Guidance Service.

Kim, Y. S., Fombonne, E., Koh, Y.-J., Kim, S.-J., Cheon, K.-A., & Leventhal, B. L. (2014). A comparison of DSM-IV pervasive developmental disorder and DSM-5 autism spectrum disorder prevalence in an epidemiological sample. *Journal of the American Academy of Child and Adolescent Psychiatry, 53*(5), 500-508. doi: 10.1016/j.jaac.2013.12.021

Klei, L., Sanders, S. J., Murtha, M. T., Hus, V., Lowe, J. K., Willsey, A. J. … & Devlin, B. (2012). Common genetic variants, acting additively, are a major source of risk for autism. *Molecular Autism, 3*(9), 1-13. doi: 10.1186/2040-2392-3-9

Korkman, M., Kirk, U., & Kemp, S. (1998). *NEPSY: A developmental neuropsychological assessment*. Psychological Corporation.

Korkman, M., Kirk, U., & Kemp, S. (2007). *NEPSY-II: A developmental neuropsychological assessment*. San Antonio, TX: The Psychological Corporation.

Kulage, K. M., Smaldone, A. M., & Cohn, E. G. (2014). How will DSM-5 affect autism diagnosis? A systematic literature review and meta-analysis. *Journal of Autism and Developmental Disorders, 44*, 1918-1932. doi: 10.1007/s10803-014-2065-2

LaMalfa, G., Lassi, G., Bertelli, M., Salvini, R., & Placidi, G. F. (2004). Autism and intellectual disability: A study of prevalence on a sample of the Italian population. *Journal of Intellectual Disability Research*, *48*, 262–267. doi: 10.1111/j.1365-2788.2003.00567.x

Lecavalier, L., Aman, M. G., Scahill, L., McDougle, C. J., McCracken, J. T., Vitiello, B., … & Kau, A. S. M. (2006). Validity of the Autism Diagnostic Interview-Revised. *American Journal on Mental Retardation, 111*(3), 199-215.

Le Couteur, A., Bailey, A., Goode, S., Pickles, A., Robertson, S., Gottesman, I., & Rutter, M. (1996). A broader phenotype of autism: the clinical spectrum in twins. *Journal of Child Psychology and Psychiatry, 37*(7), 785-801. doi: 10.1111/j.1469-7610.1996.tb01475.x

Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian cognitive modeling: A practical course.* Cambridge University Press.

Lincoln, A. J., Allen, M. H., & Kilman, A. (1995). The assessment and interpretation of intellectual abilities in people with autism. In E. Schopler & G. B. Mesibov (Eds.), Learning and cognition in autism (pp. 89-118). New York: Plenum Press.

Lord, C., Corsello, C., & Grzadzinski, R. (2014). Diagnostic instruments in autistic spectrum disorders. *Handbook of Autism and Pervasive Developmental Disorders, Fourth Edition.*

Lord, C., Leventhal, B. L., & Cook, E. H., Jr. (2001). Quantifying the phenotype in autism spectrum disorders. *American Journal of Medical Genetics Part B, 105*, 36-38.

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Jr., Leventhal, B. L., DiLavore, P. C., 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. doi: 10.1023/A:1005592401947

Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. (2012). *Autism Diagnostic Observation Schedule, 2nd Edition (ADOS-2)*. Torrance: Western Psychological Services.

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 24*, 659-685. doi: 10.1007/BF02172145

Lundqvist, D., Flykt, A., & Ohman, A. (1998). The Karolinska Directed Emotional Faces—KDEF. Stockholm, Sweden: Department of Clinical Neuroscience, Psychology section, Karolinska Institute.

Luria, A. R. (1976). *The working brain: An introduction to neuropsychology*. Basic Books.

Matson, J. L., & Shoemaker, M. (2009). Intellectual disability and its relationship to autism spectrum disorders. *Research in Developmental Disabilities, 30*(6), 1107–14. doi: 10.1016/j.ridd.2009.06.003

Mattila, M.-L., Kielinen, M., Linna, S.-L., Jussila, K., Ebeling, H., Bloigu, R., Joseph, R. M., & Moilanen, I. (2011). Autism spectrum disorders according to DSM-IV-TR and

comparison with DSM-5 draft criteria: an epidemiological study. *Journal of the American Academy of Child and Adolescent Psychiatry, 50*(6), 583-592. doi: 10.1016/j.jaac.2011.04.001

Mayes, S. D., & Calhoun, S. L. (2003). Analysis of WISC-III, Stanford-Binet: IV and academic achievement test scores in children with autism. *Journal of Autism and Developmental Disorders, 33,* 329-341. doi: 10.1023/A:1024462719081

McGue, M. (1992). When assessing twin concordance, use the probandwise not the pairwise rate. *Schizophrenia Bulletin, 18*(2), 171-176. doi: 10.1093/schbul/18.2.171

McPartland, J. C., Reichow, B., & Volkmar, F. R. (2012). Sensitivity and specificity of proposed DSM-5 diagnostic criteria for autism spectrum disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 51*(4), 368-383. doi: 10.1016/j.jaac.2012.01.007

Minshew, N. J., Goldstein, G., & Siegel, D. J. (1997). Neuropsychologic functioning in autism: profile of a complex information processing disorder. *Journal of the International Neuropsychological Society, 3*, 303-316.

Mitchell, S. R., Reiss, A. L., Tatusko, D. H., Ikuta, I., Kazmerski, D. B., Botti, J.-A. C., Burnette, C. P., & Kates, W. R. (2009). Neuroanatomic alterations and social and communication deficits in monozygotic twins discordant for autism disorder. *American Journal of Psychiatry, 166,* 917-925. doi: 10.1176/appi.ajp.2009.08101538

Nader, A.-M., Jelenic, P., & Soulières, I. (2015). Discrepancy between WISC-III and WISC-IV cognitive profile in autism spectrum: what does it reveal about autistic cognition? *PLoS ONE, 10*(12), e0144645. doi:10.1371/journal.pone.0144645

Narzisi, A., Muratori, F., Calderoni, S., Fabbro, F., & Urgesi, C. (2013). Neuropsychological profile in high functioning autism spectrum disorders. *Journal of Autism and Developmental Disorders, 43*, 1895-1909. doi: 10.1007/s10803-012-1736-0

Ozonoff, S., Young, G. S., Carter, A., Messinger, D., Yirmiya, N., Zwaigenbaum, L. …& Stone, W. L. (2011). Recurrence risk for autism spectrum disorders: a baby siblings research consortium study. *Pediatrics*, 128(3), e488-e495. doi: 10.1542/peds.2010-2825

Pennington, B. F., Filipek, P. A., Lefly, D., Chhabildas, N., Kennedy, D. N., Simon, J. H., …& DeFries, J. C. (2000). A twin MRI study of size variations in the human brain. *Journal of Cognitive Neuroscience, 12*(1), 223-232. doi: 10.1162/089892900561850

Pisula, E., & Ziegart-Sadowska, K. (2015). Broader autism phenotype in siblings of children with ASD—a review. *International Journal of Molecular Sciences, 16*(6), 13217-13258.

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. *American Journal of Psychiatry, 154*, 185-190. doi: 10.1176/ajp.154.2.185

Piven, J., & Palmer, P. (1999). Psychiatric disorder and the broad autism phenotype: evidence from a family study of multiple-incidence autism families. *American Journal of Psychiatry, 156*, 557-563.

Posthuma, D., de Geus, E. J. C., Neale, M. C., Hulshoff Pol, H. E., Baare, W. E. C., Kahn, R. S., & Boomsma, D. (2000). Multivariate genetic analysis of brain structure in an extended twin design. *Behavior Genetics, 30*(4), 311-319. doi: 10.1023/A:1026501501434

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Raven, J. C. (1998). *Raven's progressive matrices.* Oxford: Oxford Psychologists Press.

Reinvall, O., Voutilainen, A., Kujala, T., & Korkman, M. (2013). Neurocognitive functioning in adolescents with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 43*, 1367-1379. doi: 10.1007/s10803-012-1692-8

Reitan, R. M., & Wolfson, D. (1993). *Halstead-Reitan Neuropsychological Test Battery: Theory and clinical interpretation.* Tucscon, AZ: Neuropsychology Press.

Ritvo, R. A., Ritvo, E. R., Guthrie, D., & Ritvo, M. J. (2008). Clinical evidence that Asperger's disorder is a mild form of autism. *Comprehensive Psychiatry, 49*(1), 1-5. doi: 10.1016/j.comppsych.2007.06.010

Rommelse, N., Langerak, I., van der Meer, J., de Bruijn, Y., Staal, W., Oerlemans, A., & Buitelaar, J. (2015). Intelligence may moderate the cognitive profile of patients with ASD. *PLoS ONE, 10*(10), e0138698. doi:10.1371/journal.pone.0138698

Ronald, A., Happé, F., Bolton, P., Butcher, L. M., Price, T. S., Wheelwright, S., Baron-Cohen, S., & Plomin, R. (2006). Genetic heterogeneity between the three components of the autism spectrum: a twin study. *Journal of the American Academy of Child and Adolescent Psychiatry, 45*(6), 691-699. doi: 10.1097/01.chi.0000215325.13058.9d

Ronald, A., & Hoekstra, R. A. (2011). Autism spectrum disorders and autistic traits: a decade of new twin studies. *American Journal of Medical Genetics Part B, 156,* 255–274. doi: 10.1002/ajmg.b.31159

Rosenberg, R. E., Law, J. K., Yenokyan, G., McGready, J., Kauffman, W. E., & Law, P. A. (2009). Characteristics and concordance of autism spectrum disorders among 277 twin pairs. *Archives of Pediatrics and Adolescent Medicine, 163*(10), 907-914. doi: 10.1001/archpediatrics.2009.98

Rubia, K., Russell, T., Overmeyer, S., Brammer, M. J., Bullmore, E. T., Sharma, T., … & Taylor, E. (2001). Mapping motor inhibition: conjunctive brain activations across different versions of go/no-go and stop tasks. *Neuroimage, 13*(2), 250-261. doi: 10.1006/nimg.2000.0685

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

Shah, A., & Frith, U. (1983). An islet of ability in autistic children: A research note. *Journal of Child Psychology and Psychiatry, 24,* 613-620. doi: 10.1111/j.1469-7610.1983.tb00137.x

Shumway, S., Farmer, C., Thurm, A., Joseph, L., Black, D., & Golden, C. (2012). The ADOS calibrated severity score: Relationship to phenotypic variables and stability over time. *Autism, 5*, 267-276. doi: 10.1002/aur.1238

Siegel, D. J., Minshew, N. J., & Goldstein, G. (1996). Wechsler IQ profiles in diagnosis of high-functioning autism. *Journal of Autism and Developmental Disorders, 26,* 389-406. doi: 10.1007/BF02172825

Spiker, D., Lotspeich, L. J., Dimiceli, S., Myers, R. M., & Risch, N. (2002). Behavioral phenotypic variation in autism multiplex families: evidence for a continuous severity gradient. *American Journal of Medical Genetics, 114,* 129-136. doi: 10.1002/ajmg.10188

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, 30*, 405-416. doi: 10.1111/j.1469-7610.1989.tb00254.x

Swineford, L. B., Thurm, A., Baird, G., Wetherby, A. M., & Swedo, S. (2014). Social (pragmatic) communication disorder: a research review of this new DSM-5 diagnostic category. *Journal of Neurodevelopmental Disorders, 6*, 41. doi: 10.1186/1866-1955-6-41

Szatmari, P., Bryson, S. E., Boyle, M. H., Streiner, D. L., & Duku, E. (2003). Predictors of outcome among high functioning children with autism and Asperger syndrome. *Journal of Child Psychology and Psychiatry, 44*(4), 520-528. doi: 10.1111/1469-7610.00141

Tadevosyan-Leyfer, O., Dowd, M., Mankoski, R., Winklosky, B., Putnam, S., McGrath, L., … & Folstein, S. E. (2003). A principal components analysis of the Autism Diagnostic Interview-Revised. *Journal of the American Academy of Child and Adolescent Psychiatry, 42*, 864-872. doi: 10.1097/01.CHI.0000046870.56865.90

Tager-Flusberg, H. (1981). Sentence comprehension in autistic children. *Applied Psycholinguistics*, *2*(1), 5-24. doi: 10.1017/S014271640000062X

Tager-Flusberg, H., & Joseph, R. M. (2003). Identifying neurocognitive phenotypes in autism. *Philosophical Transactions of the Royal Society B Biological Sciences, 358*(1430), 303-314. doi: 10.1098/rstb.2002.1198

Thorndike, R. L., Hagen, E. P., & Sattler, J. M. (1986). *Stanford-Binet intelligence scale*. Riverside Publishing Company.

Thorndike, R. L., & Terman, L. M. (1973). *Stanford-Binet Intelligence Scale: Form LM: L972 Norm Tables*. Houghton Mifflin.

Tick, B., Bolton, P., Happé, F., Rutter, M., & Rijsdijk, F. (2016). Heritability of autism spectrum disorders: a meta-analysis of twin studies. *Journal of Child Psychology and Psychiatry, 57*(5), 585-595. doi: 10.1111/jcpp.12499

Tryon, P. A., Mayes, S. D., Rhodes, R. L., & Waldo, M. (2006). Can Asperger's disorder be differentiated from autism using DSM-IV criteria? *Focus on Autism and Other Developmental Disabilities, 21*(1), 2-6. doi: 10.1177/10883576060210010101

Wagenmakers, E. J. (2007). A practical solution to the pervasive problems of p values. *Psychonomic Bulletin and Review, 14*(5), 779-804. doi: 10.3758/BF03194105

Wechsler, D. (1967). *Wechsler preschool and primary scale of intelligence*. New York: Psychological Corporation.

Wechsler, D. (1974). *Wechsler intelligence scale for children-revised.* Psychological Corporation.

Wechsler, D. (1981). *WAIS-R manual: Wechsler adult intelligence scale-revised.* Psychological Corporation.

Wechsler, D. (1991). *WISC-III: Wechsler intelligence scale for children: Manual*. Psychological Corporation.

Wechsler, D. (1997). *WAIS-III: Administration and scoring manual: Wechsler adult intelligence scale*. Psychological Corporation.

Wechsler, D. (1999). *WASI (Wechsler Adult Scale-revised).* Psychological Corporation.

Wechsler, D. (2003). *Wechsler intelligence scale for children-WISC-IV.* Psychological Corporation.

Wing, L., Gould, J., & Gillberg, C. (2011). Autism spectrum disorders in the DSM-V: better or worse than the DSM-IV? *Research in Developmental Disabilities, 32*, 768-773.

Witkin, H. A., Oltman, P. K., Raskin, E., & Karp, S. A. (1971). *A manual for the embedded figures tests.* Palo Alto, CA: Consulting Psychologist Press.

Williams, D. L., Goldstein, G., & Minshew, N. J. (2006). Neuropsychologic functioning in children with autism: further evidence for disordered complex information-processing. *Child Neuropsychology, 12*, 278-298. doi: 10.1080/09297040600681190

Wilson, C. E., Happe, F., Wheelwright, S. J., Ecker, C., Lombardo, M. V., Johnston, P., …& Murphy, D. G. M. (2014). The neuropsychology of male adults with high-functioning autism or Asperger syndrome. *Autism Research, 7*, 568-581. doi: 10.1002/aur.1394

Youngstrom, E.A., Glutting, J. J., & Watkins, M. (2003). Stanford-Binet Intelligence Scale: Fourth Edition (SB4): Evaluating the empirical bases for interpretations. In C. R. Reynolds & R. W. Kamphaus (Eds.), *Handbook of psychological and educational assessment: Intelligence, aptitude, and achievement (2nd ed.)* (pp. 217-242). New York: Guilford.

Elizabeth P. McKernan

Vita

Contact:

430 Huntington Hall
Syracuse, NY 13244

Education:

Syracuse University 2015 - current
Department of Psychology, School Psychology Ph.D. Program

Villanova University 2010 - 2014
Department of Psychology/Department of Education and Counseling, B.A.