The Pennsylvania State University

The Graduate School

College of Education

THE USE OF COMPUTERIZED COGNITIVE ASSESSMENT WITH CHILDREN WITH AUTISM

A Dissertation in

School Psychology

by

Emily H. Perlman

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 2016

The dissertation of Emily H Perlman was reviewed and approved* by the following:

Barbara A. Schaefer
Associate Professor of Education
Dissertation Advisor
Chair of Committee
Director of Training/Professor in Charge for Graduate Programs in School Psychology

Krista Wilkinson
Professor, Department of Communication Sciences and Disorders

Shirley Woika
Associate Professor of Education, Director of Field and Clinical Training

Pamela Wolfe
Associate Professor of Education

*Signatures on file in the Graduate School

ABSTRACT

Autism spectrum disorder (ASD) is a rapidly increasing category of diagnosis in the United States today; however, little is definitively known about the cognitive abilities of individuals with ASD despite frequent study of this area. Results of the most recent studies have suggested that previously accepted idea about cognitive abilities for individuals with autism may be inaccurate. One reason for these mixed finding may be that the levels of communication and social skills vary greatly for individuals with ASD, and these variations can have powerful effects on the scores obtained on measures of intellectual abilities. One method of controlling for these specific variables may be through computerized testing, which research findings have demonstrated to be associated with higher scores on tests of executive functioning and working memory for individuals with autism but not for neurotypical individuals. Computer-based testing controls for the social and verbal aspects of the testing process, which are factors that may be affecting scores for individuals with autism, but not for others.

The current study used three measures to discriminate the causes of score discrepancies: a verbal measure, a nonverbal measure that incorporates a social component, and a computerized test that has no social component and only a minimal verbal component. The scores were analyzed to find which one best predicts academic achievement and is the optimal measure of cognitive ability.

The results indicated that no significant difference exists between scores obtained from each IQ test or in each test’s ability to predict either math or reading achievement; however, qualitative analysis indicated that differences may exists between scores of students with IQ scores below 70, but not for those with IQ scores above this level. The results also suggested that the level of a student’s communication skills is a significant component of what is measured by the IQ scores obtained from a verbal test, but not for those scores obtained from nonverbal or computerized measures. Future research needs to be done with a larger sample size to determine if differences in scores and the constructs they measure exist for individuals with autism who earn scores below 70 on common intelligence measures, as these students may be the most likely to have cognitive abilities underestimated.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

AKNOWLEDGEMENTS

I would first like to thank Dr. Schaefer as my excellent dissertation advisor. She allowed me the freedom to follow my interests and find a course of research that held a strong interest for me. She has continued to support my abilities and ideas throughout this long process, and for that, I am truly grateful. I also need to thank my committee members, Dr. Woika, Dr. Wolfe, and Dr. Wilkinson, without whom this dissertation would not have happened. They were always there to provide me ways to explore my research that I had never considered before, and they helped make my dissertation what it is today.

I would like to thank my parents (Dr. and Mr. Perlman) and my sister for all their help throughout my entire graduate career. Without their love and support- in ways both big and small- I never would have made it this far. They have always believed in my talents and have supported me in accomplishing any goal I set out for myself. I also want to thank my cohort for being there for me through graduate school and beyond. I cannot imagine having finished this program without their help, love, energy, support, and encouragement. Dr. Sarah Wollersheim Shervey and Dr. Brianne Mintern, you are the best cohort I could imagine ever having and, as I always say: Cohort is forever. Thank you to my Dearest Schmoo who has listened patiently to all my doubts, complaints, and exasperations about my dissertation, but who never let me quit. Your belief in me and pride in my accomplishments mean the world to me.

I need to give a heartfelt thanks to all the participating families in my study. They selflessly gave their time and energy to help me achieve my goal, and for that, I am grateful. I would especially like to thank Donna Budway who went out of her way to help me recruit as many participants as possible. I really appreciate the openness and willingness to help of the entire autism community, and I hope my research has provided them something in return.

Chapter One: Introduction

One of the fastest growing categories of psychiatric diagnosis in the United States today is autism spectrum disorder (ASD[1]). In the past, this classification has included autism, Asperger's syndrome (AS), Rett's disorder, childhood disintegrative disorder (CDD), and pervasive developmental disorder-not otherwise specified (PDD-NOS). However, recent changes to the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5; American Psychiatric Association [APA], 2013), a manual that is commonly used in diagnosing mental disorders in the US, have resulted in the creation of one general classification category (ASD) for individuals with similar symptoms. This spectrum of disorders is characterized by symptoms in the areas of social communication and interactions, as well as by symptoms of stereotyped behaviors, interests and activities. Individuals with ASD typically display deficits in social and communication skills, which may include delays in spoken language, inability to sustain a conversation, and a lack of age-appropriate relationships, as well as often-evidencing stereotypical or restricted patterns of interests or behaviors (APA, 2013). The severity of the deficits in each area varies by individual and can range from pervasive to mild. These issues, however, are likely to affect all areas of an individual's life and interfere with activities of daily living. Autism is commonly cited as having a high rate of comorbidity with mental retardation (APA, 2013), or what is now referred to as intellectual disability; however, cited rates vary from study to study (Charman, Pickles, et al., 2011; Edelson, Schubert, & Edelson, 1998). Little else is definitively known about the cognitive abilities for individuals with ASD despite extensive study of this area. Many past research findings have been challenged by the results of more recent studies.

[1] See Appendix for a glossary of abbreviations

Previous Findings on Cognitive Abilities of Individuals with ASD

It is difficult to determine what facts about the cognitive abilities of individuals with ASD have clear evidence to support them. For example, autism is generally accepted to co-occur often with intellectual disability. In the 1970s, the comorbidity rate was reported to be as high as 86% (Edelson, 2005); however, results of more recent research have challenged this finding. In some studies, rates of intellectual disability in a population of individuals with autism have been found to be as low at 19% (Edelson, Edelson, & Jung, 1998). Many of the other generally accepted findings about the intellectual abilities of individuals with ASD have also been challenged by newer findings. For example, research results previously suggested that individuals with ASD displayed uneven subtest profiles (de Bruin, Verheij, & Ferdinand, 2006; Mayes & Calhoun, 2003b; Ozonoff, South, & Miller, 2000), tended to score higher on the Performance IQ (PIQ) than on the Verbal IQ (VIQ; Coolican, Bryson, & Zwaigenbaum, 2008; Mayes & Calhoun, 2003b), and increased their IQ scores over time (Joseph, Tager-Flusberg, & Lord, 2002; Mayes & Calhoun, 2003a). However, in all of these areas, findings that are more recent have been mixed with many of the most recent studies suggesting that previously accepted patterns of cognitive abilities for individuals with autism are unfounded.

In the past few decades, research findings about the cognitive abilities of children with autism have grown tremendously and have been useful in determining appropriate interventions and educational strategies for this population. However, the unique challenges presented when working with individuals with ASD can often make data gathering for evidence-based decision-making a difficult process. Problem behaviors, issues of motivation or attention, and communication difficulties are common occurrences with this population and can all interfere with the assessment process that is necessary to make proper educational decisions (Koegel,

Koegel, & Smith, 1997). These challenges can introduce error into scores, which in turn lowers score reliability and validity. Such error makes drawing generalizations and conclusions about the skills and abilities of individuals with autism difficult. Thus, methods that produce scores with more evidence of reliability and validity must be developed before accurate conclusions can be drawn.

Challenges to Measuring Cognitive Ability

Social and communication issues are common to those diagnosed with autism and can affect the accuracy of cognitive ability scores for these individuals. Therefore, these issues must be addressed specifically in testing situations in order to obtain accurate test scores. For example, addressing particular interfering behaviors can increase motivation and attention during testing, leading to higher scores on tests (Koegel et al., 1997). When these behaviors are not addressed, scores may greatly underestimate a child’s true abilities.

Poor communication between the examiner and test taker can also interfere with an examiner’s ability to obtain an accurate estimate of cognitive ability. Research, however, has suggested that nonverbal measures of IQ lead to much higher and more accurate scores (Bölte, Dziobek, & Poustka, 2009; Mottron, 2004). Thus, the choice of measurement instrument is particularly important when assessing a child with autism to ensure that the scores on the test only reflect cognitive ability and not difficulties with communication skills.

Finally, deficits in social skills pose a challenge to obtaining an accurate estimate of cognitive ability. Research has been mixed as to whether measures of intelligence are correlated with social skills (Black, Wallace, Sokoloff & Kenworthy, 2009; Joseph et al., 2002). The cause of these mixed results may be due to the measures used. Some cognitive ability measures require much more social interaction with the examiner than others do, which may interfere with the ability of an individual with autism to display optimal cognitive functioning.

For tests of executive functioning (Ozonoff, 1995) and working memory (Nakahachi et al., 2006; Ozonoff & Strayer, 2001) administered by computer, individuals with autism tend to score comparably to neurotypical individuals, but they score below these control groups when taking the paper-and-pencil (and verbal) versions. Computer-based testing removes both the social and verbal aspects from the testing process, which may be the cause of the higher scores for individuals with autism and not the control group. These findings suggest that the social and verbal components inherent to most typical assessments may lead to lower scores for children with autism based on deficits in these functional areas. If so, these components must be isolated from the testing situation to ensure the accuracy of cognitive scores.

It is unclear from the three studies above (Nakahachi et al., 2006; Ozonoff, 1995; Ozonoff & Strayer, 2001) whether gains in computer-based test scores are due to controlling the verbal component, the social component, or both simultaneously. Scores for nonverbal cognitive measures may be commensurate with those of computerized assessments, thus obviating the need for the application of computer-based assessments. Alternatively, computer-based tests may control for specific variables that are important to hold constant in the assessment process for individuals with autism, in which case computer-based tests would lead to higher scores with greater predictive accuracy. Accurately measuring cognitive ability of individuals with autism is essential because scores may be predictive of future outcome. For example, cognitive ability scores may predict academic achievement, educational placement, and general future functioning (Arick et al., 2003; Gillberg, & Steffenburg, 1987; Howlin, Goode, Hutton, & Rutter, 2004; Mayes & Calhoun, 2008).

Project Significance and Aims

Given that research suggests a large difference between verbal and nonverbal measures of intelligence scores of children with autism, this study attempts to examine how verbal and social skills are related to IQ scores for individuals with autism. It has been suggested that both the verbal and social components may interfere with the ability of children with autism to display fully their cognitive potential on a verbally mediated ability measure (which includes both verbal and social aspects). Research has also suggested that verbal versus computerized testing produce differences in working memory and executive functioning scores, which may be related to these verbal and/or social components. This study attempts to unravel the factors relating to the discrepancy in these scores with regard to cognitive ability.

The current study uses three measures to discriminate the potential sources of score discrepancies. Scores from a verbal measure will be compared with those from a nonverbal measure that still incorporates a social component; both of these scores will be compared with a computerized test that has no social and a minimal verbal component. Any difference in scores on these measures is associated with the task demands of the included social or verbal aspects. The comparison of scores from the three measures should help determine which type of test scores constitutes an optimal test of the cognitive ability of a child with autism.

Research Questions

The goal of this study is to determine whether scores from a computerized, nonverbal, or verbal measure of cognitive ability are best for predicting achievement for children with autism. To that end, four research questions are addressed:

1. Is there a statistically significant difference in scores on three types of measures of cognitive ability for children with autism: computer based, nonverbal, and verbal?

2. Are there systematic differences between the scores obtained by these measures?
3. Are there significant differences in the ability of the scores to predict academic achievement for children with autism?
4. Are the levels of social and communication skills for children with autism predictive of their IQ scores for certain cognitive ability measures?

Chapter Two: Literature Review

Cognitive Ability in Children with ASD

The cognitive abilities of individuals with autism have long been studied. It has been commonly accepted that high rates of comorbidity with intellectual disability exist in individuals with ASD, with reported rates as high as 86% (Edelson, 2005). Further, it has been commonly cited that individuals with ASD have uneven subtest profiles with high scores on pattern copying subtests, like Block Design, and low scores on real-world knowledge subtests such as Comprehension (de Bruin et al. 2006; Mayes & Calhoun, 2003b; Ozonoff et al., 2000). These patterns supposedly lead to much higher PIQ scores than VIQ scores (Coolican et al., 2008; Mayes & Calhoun, 2003b). Claims have also been made that this difference diminishes over time and that overall IQ scores for children with autism have a tendency to increase (Joseph et al. 2002; Mayes & Calhoun, 2003a). Some researchers have suggested that higher IQ scores in children with ASD are equated with increased social skills or a better working memory (Charman, Jones, et al., 2011; Joseph, et al., 2002). All of these claims, however, have been called into question by the results of more recent studies, leaving very little certainty with regard to the cognitive abilities of individuals with autism.

Rates of Intellectual Disability

One often-documented fact about children with ASD is the high rate of intellectual disability that often accompanies this diagnosis. Although reported rates vary widely by study, a frequently estimated percentage of children with ASD who have an intellectual disability (IQ < 70) is currently cited to be around 50% (Charman, Jones, et al., 2011). However, it is difficult to be certain of the validity of this estimate or any other. Edelson (2006) recently reviewed the claims made in the literature about the percentages of individuals with autism who have a comorbid diagnosis of intellectual disability. She found 223 claims made between 1937 and 2003, and only 26% (*n* = 58) were based on empirical data. Of those claims that were non-empirical, less than half could be traced back to empirical data. Further, two-thirds of the empirical claims referred to articles published 25–45 years ago. Many of these studies used subtest scores, developmental scales, or adaptive measures to estimate IQ and few considered how the effects of autism may have interfered with the testing process or affected the scores. Based on the history of research in this area, it is difficult to determine an accurate estimate of the percentage of children with ASD who are cognitively functioning at a level of intellectual disability.

Recent evidence from empirical studies has suggested a wide range of intellectual disability rates in children with autism. Charman, Pickles, et al. (2011) conducted a study to examine the rates and severity of intellectual disability in 225 children with autism ages 9–14 years. Their results suggested that 55% of children in their sample had an intellectual disability and less than 20% had a moderate to severe disability. However, Edelson, Schubert, and Edelson (1998) also recently tested 258 participants with ASD, ages 4–41 years, using the Test of Nonverbal Intelligence: Second Edition (TONI-2; Brown, Sherbenou, & Johnsen, 1990). The results indicated a mean IQ score of 89, with only 19% of the participants falling into the range of intellectual disability. The results of this study may have been affected by the nonstandard procedure used when administering the TONI-2, as well as the selection effects of the time-intensive study procedure. However, the atypically low rates of intellectual disability found in this study also may be due to the nonverbal nature of the TONI-2.

The Edelson (2006) study raises the issue of whether any empirical research study can accurately capture average cognitive abilities of the population of children with ASD. Mottron

(2004) noted that many of the studies examining cognitive abilities of children with autism are cross-sectional studies of children with high-functioning autism. Their average Full Scale IQ (FSIQ) score is approximately 85 and their average chronological age is around 14 years. Mottron further noted that it is common to disqualify those with low developmental or chronological ages from research on autism. For example, Edelson disqualified approximately one-third of otherwise eligible participants due to their being "untestable," which may help to explain why the percentage of participants who fell within the intellectual disability range was much lower than other studies (Charman, Jones, et al., 2011; Charman, Pickles, et al., 2011). Further, the definition of autism spectrum disorder has been expanded over time to include individuals who are higher functioning than those previously included in this diagnosis. The inclusion of a more varied population in the classification of individuals with ASD is likely to increase the overall average IQ score of this population. Thus, as current research stands, it is unlikely that the results will be representative of all individuals with autism, a fact that limits the accuracy of the estimates of average cognitive functioning.

Further, this limit applies to the predictive validity of IQ scores as a measure for later cognitive functioning. A child with autism whose IQ score has been assessed to be in the range of intellectual disability may well fall in a different range of cognitive functioning if assessed later or with an alternate measure. Lord and Schopler (1989) reported that Developmental Quotients (DQ) or PIQ scores of 50 or below measured early in life reliably predicted the presence of an intellectual disability later. However, they noted that although early IQ scores can often be an accurate predictor of later intellectual functioning, factors such as the age of the child and the test administered must be taken into account when considering the predictive ability of the scores for students with ASD. For example, Lord and Schopler noted that the Bayley Scales of Infant and Toddler Development–Second Edition (Bayley-2; Bayley, 1993) a test of early development, certain social behaviors, such as smiling, help yield the DQ. Such behaviors are not necessarily an accurate predictor of cognitive functioning in a child with autism, and they may lead to underestimates of cognitive abilities in this population. In this way, early intervention with students who have autism can lead to increase in DQ scores that measure these skills, which may lead to differing rates of intellectual disability in children with autism depending on at what age students are tested. Measures that are nonverbal, untimed, and involve abstract reasoning may provide a more accurate measure of functioning in children with ASD, one in which many children do not score in the range of intellectual disability (Edelson, 2005). Thus, the measured rates of intellectual disability in children with autism may depend greatly on how and when these children are tested, and with what measure.

Subtest Profiles

Another often-purported facet of cognitive ability in individuals with autism has been the uneven profile of subtest scores. It is often reported that individuals with ASD perform well in nonverbal visual-spatial tasks such Block Design and poorly on real-world knowledge subtests such as Comprehension (de Bruin et al. 2006; Mayes & Calhoun, 2003b; Ozonoff et al., 2000). For example, Coolican, Bryson, and Zwaigenbaum (2008) examined cognitive abilities of 63 children aged 3–16 years with autism, AS, or PDD-NOS using the Stanford Binet-Fifth Edition (SB-V; Roid, 2003). They found that for the nonverbal subtests, mean scores on Fluid Reasoning, Quantitative Reasoning, and Visual Spatial Processing were higher than the Knowledge mean score at a statistically significant level. More recently, Nader, Courchesne, Dawson, and Soulières (2014) examined the Wechsler Intelligence Scale for Children–Fourth Edition (WISC-IV; Wechsler, 2003) cognitive profile of 25 children with autism and 22 typically developing children between the ages of 6 and 16. The results indicated that for the individuals with ASD, the Perceptual Reasoning Index (PRI) was higher at a statistically significant rate than the FSIQ, Verbal Comprehension Index (VCI), Working Memory Index (WMI), and Processing Speed Index (PSI). This pattern did not hold for the typically developing control group, who had similar FSIQ, VCI, PRI, and PSI scores, but a significantly lower WMI score.

Minshew, Turner, and Goldstein (2005) also recently studied the cognitive profiles of 113 adults and 102 children between the ages of 8 and 55 with high-functioning autism (HFA) who were given Wechsler-series intelligence tests. The authors noted that for the sample with HFA results indicated a peak in the scores for the Block Design subtests and a low point on the Comprehension subtest. They also reported that a group of typically developing controls did not display the same uneven profile. However, the authors failed to mention that although Block Design was one of the subtests on which the group of individuals with autism scored highly, the verbal subtests of Information and Similarities displayed the same peaks. Similarly, the subtest of Digit Span, which does not involve real-world knowledge, had a mean score equally as low as the Comprehension mean score. Thus, the subtest profile of individuals with ASD may not be as clear-cut as it is often presented to be. Additionally, other factors may have contributed to the earlier findings of uneven subtest profiles in samples of individuals with autism. Ozonoff, Goodlin-Jones, and Solomon (2005) noted that many studies that found results suggesting uneven subtest profiles used earlier editions of current cognitive assessments and that few studies have examined the most recent Wechsler series or Stanford-Binet intelligence tests. Further, uneven test profiles do not display any validity in diagnosis or prediction of academic achievement. Ozonoff et al. (2005) caution that profiles from cognitive tests should not be used to confirm a diagnosis of ASD or to differentially diagnosis subtypes of disorders on the spectrum. Research has also suggested that the overall FSIQ score is the best predictor of all areas of academic achievement for children with ASD and that using subtest scores or index scores did not significantly add to the predictive power (Mayes & Calhoun, 2008). Although an uneven profile of subtest scores for individuals who have autism may or may not be the case, this profile is irrelevant for differential diagnosis and prediction of achievement.

Verbal and Nonverbal Discrepancies

Related to the concept of uneven subtest profiles is the perception that for a child with ASD, PIQ scores are often much higher than VIQ scores. Research in this area has been mixed as to the veracity of this concept. The results of a recent study of the cognitive abilities of children with ASD, AS, or PDD-NOS using the SB-V suggested that the nonverbal IQ scores and verbal IQ scores were different at a statistically significant level (Coolican et al., 2008). In another study (Mayes & Calhoun, 2003b), young children with ASD ages 3–7 years were administered the Stanford-Binet-Fourth Edition (SB-IV; Thorndike, Hagen, & Sattler, 1986). The results of this study indicated that average nonverbal IQ scores were greater at a statistically significant level than average verbal IQ scores. This finding held for both the group of children with IQ scores above 80 and the group of children with IQ scores below 80.

The results of a recent meta-analysis further indicated that the discrepancy between PIQ and VIQ may vary based on diagnosis. In a review of 52 studies published between 1994 and 2012, Chiang, Tsai, Kuen Cheung, Brown, and Li (2014) found that both children and adults with Asperger's Disorder had a higher VIQ than PIQ, but that individuals with HFA did not demonstrate this discrepancy. However, the findings of a study (Minshew et al., 2005), which used Wechsler-series tests, suggested that only 10% of the 215 adult and child participants with ASD (n = 21) met their criteria of a significant difference between PIQ and VIQ scores, which was determined in the study to be a difference of 15.

Other studies have also found little evidence to support the veracity of a meaningful discrepancy between verbal and nonverbal or performance IQ scores. A review by Ozonoff, et al, (2005) noted that past research has suggested that most individuals with autism do not display evidence of a difference in VIQ and PIQ scores that is greater than 12. The results of a recent empirical study of children ages 9–14 years with ASD by Charman, Pickles, et al. (2011) did not support the idea that average PIQ scores are higher than average VIQ scores. These findings indicated no statistically significant difference between these two scores. The most common cognitive profile for children in this study (Charman, Pickles, et al., 2011) was for the PIQ score to be within 12 points of the VIQ score.

One possibility that may account for the differences in the results among these studies may be the age of the participants. It has been suggested nonverbal abilities may be much higher than verbal abilities in younger children but that the difference diminishes over time. In one study, Joseph, Tager-Flusberg, and Lord (2002) examined a sample of 120 children ages 3–13 years with autism using the Differential Abilities Scale (DAS; Elliott 2007). Their findings suggested that those in the preschool-age group had a statistically significant difference in favor of nonverbal scores over verbal scores; however, no difference between these two scores in the school-age group was detected. Further analysis (Joseph et al., 2002) indicated no difference between preschool-age and school-age groups on nonverbal IQ scores, but preschool-age verbal IQ scores were lower than school-age verbal IQ scores at a statistically significant level. The authors suggest that this pattern is consistent with non-verbal discrepancies lessening over time through an increase in verbal IQ scores; however, the fact that the study was cross sectional and that the two groups who were administered slightly different subtests of the DAS based on age may have confounded these results. Mayes and Calhoun (2003b) also found no difference between PIQ and VIQ scores for older children ages 6–15 years in their study, unlike differences found for younger children. However, the older children were tested using the Wechsler Intelligence Scale for Children–Third Edition (WISC-III; Wechsler, 1991) and not the SB-IV, which may account for results more similar to those of Minshew, et al. (2005), which also used the Wechsler tests.

Increasing IQ Scores

An increase in IQ scores over time for children with autism has been suggested by many researchers, although many more researchers maintain that they remain stable. A meta-analysis (Begovac, Begovac, Majić, & Vidović, 2009) indicated that 19 of the included 23 studies that examined the IQ scores of children with autism indicated no support for change in IQ scores over time; however, a small minority still suggested an increase. For example, Mayes and Calhoun (2003a) found in a study of 164 children ages 3–15 years that the average IQ score at 3 years old was 53 whereas the average IQ score for children ages 8 years and older was 91. The findings also suggested, similarly to Joseph et al. (2002), that verbal IQ scores increased more than nonverbal IQ scores and no difference appeared between these two scores in older children (Mayes & Calhoun, 2003a). The results of this study also indicated that this increase in IQ scores began sooner for those who began with higher IQ scores. Based on these results, the authors concluded that IQ scores for children with autism increase over time, specifically based on increases in verbal ability. Mayes and Calhoun (2003a) also noted that 33% of children with ASD experience a significant increase in IQ scores over time and that these significant increases are frequently reported for young children who have received extensive early intervention services. Sheinkopf and Siegel (1998) also noted increases in IQ scores in their study of 11 children with autism who had undergone Lovaas treatment for 6 months and a group of matched, waitlist controls. The results of this study suggested no difference in IQ scores at the pre-test; however, 6 months later, the treatment group had an average IQ score 25 points higher than the control groups, which was at a statistically significant level. The authors (Sheinkopf & Siegel, 1998) attributed this change to the effects of the intensive treatment regime on cognitive functioning.

All of the studies that have purported to find increases in IQ over time (i.e., Mayes & Calhoun, 2003a, 2003b; Sheinkopf & Siegel, 1998) have failed to address the difference in measures used to estimate cognitive ability over time. Mayes and Calhoun (2003b), for example, measured the cognitive functioning of their preschool-age group with the SB-IV and the cognitive functioning of their school-age group with the WISC-III. This change in test was unnecessary because the SB-IV can be used with both older and younger children. Although both these tests claim to measure the construct of intelligence, is it unclear whether one or the other measures more items involving culture of social dimensions, which may put a child with autism at a disadvantage. The other two studies (Mayes & Calhoun, 2003a; Sheinkopf & Siegel, 1998) also involved changes in measures that provide clearly discrepant scores for children with autism.

Magiati and Howlin (2001) examined the differences in scores on two different tests of developmental level in young children with autism, ages 27–58 months. Both the Bayley-2 and the Merrill-Palmer Scale (M-P; Stutsman, 1948) were administered to the participants to yield a DQ as an estimate of cognitive ability. The results revealed a statistically significant difference between scores on the two different tests with the mean score on the Bayley-2 being 23 points below the mean score on the M-P. The authors suggested that this difference may be due to the types of task on each test with the Bayley-2 including items involving social responsivity and with the M-P focusing on visual-spatial tasks (Magiati & Howlin, 2001). Lord and Schopler (1989) also noted that the Bayley-2 was consistently likely to underestimate both the M-P and the Leiter International Performance Scale (Leiter; Arthur, 1952), a measure of nonverbal intelligence, by 10 or more points – perhaps in part because it measures behaviors like smiling and paying attention to the examiner. Because these types of skills contribute to the DQ for this test, it is likely to be an underestimate of cognitive ability for children with autism. Further, when a child is first assessed with a measure that is likely to underestimate cognitive ability and later tested with a measure that does not underestimate, it may appear that cognitive ability has increased over time. Two of the studies mentioned above (Mayes & Calhoun, 2003a; Sheinkopf & Siegel, 1998) used a Bayley-2 as an initial measure and other tests to measure cognitive ability for at least some of their participants. Thus, the findings of many studies, which may be interpreted to suggest that IQ score changes over time, may be attributable to changes in instrumentation rather than actual change in intelligence.

Cognitive Abilities and Social Skills

A relationship between cognitive abilities and social skills in individuals with autism has been suggested; however, research in this area has been unclear. A study of children with autism by Joseph et al. (2002) found that symptoms related to social skills deficits were not associated with overall cognitive ability in the preschool-age group, but with verbal-performance difference in the school-age group. These results do not clearly indicate whether cognitive ability is correlated with social skills. Black, Wallace, Sokoloff and Kenworthy (2009) also found that the correlation between the verbal-nonverbal discrepancy and social skills was statistically significant for one measure but not another. Charman, Jones, et al. (2011) found no empirical evidence to support the idea of different phenotypes of social skills related to individuals with ASD who have high or low IQ scores. Charman, Pickles, et al. (2011) found that verbal-performance discrepancies were not related to symptoms and specifically that children who had greater verbal abilities did not display any less social impairment than those without, suggesting that a meaningful cognitive phenotype for IQ level as related to social skills may not exist. Charman, Jones, et al. (2011) noted that IQ scores are strongly correlated with everyday memory tasks and that these task are associated with autism symptom severity, as measured by the Autism Diagnostic Observation Schedule-General (ADOS-G; Lord, Rutter, & DiLavore, 1999). Lower levels on the everyday memory task were related to more severe social and communication deficits. Conversely, Charman, Jones, et al. (2011) noted that the directionality of this association is undetermined and this deficit was not directly correlated with IQ scores.

In an empirical study of 777 children ages 1–17 years that included both HFA (IQ $\geq$ 80) and low-functioning autism (LFA; IQ < 80), the Checklist for Autism Spectrum Disorder (CASD; Mayes & Calhoun, 1999) was administered to participants to determine differences between these two groups (Mayes & Calhoun, 2011). The results suggested that both groups had a majority of the 30 symptoms listed and that the symptom profiles of the two groups were similar. Based on the results of this study, no meaningful difference in the symptoms of social impairment displayed by individuals with ASD who have either high or low IQ scores was present. Further, it is possible that the relationship between social skills and cognitive ability scores may be an artifact of the tests used, which may inadvertently measure social skills for children with ASD.

In general, research findings on cognitive abilities in individuals with autism remain unclear. The rate of intellectual disability seems to be lower than previously claimed, but wide variability in estimates of this rate still appear. Specific subtest profiles for individuals with autism also seem to be an uncertain construct, as does the idea of a frequent discrepancy in verbal and nonverbal abilities. Further, research has often suggested that IQ in individuals with ASD increases from the preschool- to school-age years, but this is likely due to different instruments being used to estimate IQ at different points of measurement. Finally, social impairment has not been clearly linked to cognitive ability, although Joseph et al. (2002) has suggested this to be the case. Overall, the research findings in the areas of cognitive measurement of individuals with autism leave many uncertainties.

The Importance of Accurate Cognitive Estimates

The goal of IQ scores is to measure intelligence with as little error as possible. However, determining the accuracy of scores for children with autism is difficult. In one study, parents were more likely to agree with or underestimate their child's IQ score the higher it was and more likely to overestimate it the lower it was (Geiger, Smith, & Creaghead, 2002). However, it is difficult to determine whether it was the parents who were more accurate—as they are likely to see more behaviors over time—or the examiners who were accurate when using a standardized measure of cognitive ability.

Correlation with Outcome Measures

Accurate IQ scores for a specific population are only important if they can usefully predict other important criteria, which they often can for typically developing children as a whole. For children with autism, IQ scores have been suggested to be a useful predictor of outcome. In a longitudinal study of 23 children diagnosed with autism, those who had an IQ score greater than 50 measured during the preschool years also had better outcomes on average than those with scores below 50 using the outcome determined by a subjective scale of five levels ranging from “very poor outcome” to “good outcome” (Gillberg, & Steffenburg, 1987).

A similar study following 68 individuals with ASD into adulthood found similar results (Howlin et al., 2004). These results suggested that the outcome of only 4 individuals of the 23 on the group with initial IQ scores below 70 fell within the “Very Good” or “Fair” ranges and that the rest fell within the “Poor” or “Very Poor” ranges. For the 45 individuals with IQ scores above 70, 24 fell within the “Fair” to “Very Good” ranges. More recently, a study followed 85 individuals who had been diagnosed with ASD in early childhood (Anderson, Liang, & Lord, 2014). The results indicated that cognitive ability scores obtained at age 2 could predict intellectual functioning at age 19 with 85% accuracy. Further, individuals with IQ scores above 70 who participated in early interventions of at least twenty hours per week were more likely to have obtained a “Very Positive Outcome” by age 19. A study in Hong Kong of 64 adults between the ages of 18 and 28 who had been identified with ASD as children revealed similar results (Poon Mak, 2012). A regression analysis indicated that those individuals with a higher FSIQ during childhood were more likely to have a better outcome as an adult when interviewed and rated on a ten-point scale ranging from “Near Normal Functioning” to “Poor.” The findings from all these studies, though subjective in their measuring of functioning, indicate that cognitive ability measures at a young age may be able to predict later quality of life.

Other studies of the predictive accuracy of early IQ scores also indicated support for its usefulness over a shorter length of time. Arick et al. (2003) examined the progress of 67 preschool students ages 2–6 years with ASD who were engaging in the first 16 months of a 3-year intervention program. The results of the study suggested that the gain in language skills over 16 months was significantly related to initial DQ score and later measures of IQ score. Although the results are preliminary, they suggest that IQ scores for children with ASD have the possibility to be a relatively accurate predictor of the potential to benefit from instruction in a particular domain, in this case in the area of language learning. Harris and Handleman (2000) also examined the predictiveness of early IQ for placement in educational settings. They looked at the outcomes of 27 children with autism ages 10–14 years who had been administered IQ tests between the ages of 2 and 5 years. Of the children who originally scored above 80 (n = 14), 11 were in regular education settings and three were in special education settings 8 years later. Of those with IQ scores below 76 (n = 13), all were in special education settings, thus suggesting that early IQ scores can be predictive of later educational outcomes for children with ASD.

Correlation with Diagnosis

In addition to outcomes, IQ scores may be used for differential diagnosis in children with ASD. In one study of 126 Japanese children with PDD-NOS and autism, a statistically significant difference in estimates of cognitive ability at age 5 years was found, even after controlling for DQ score at age 2 years, on which the PDD-NOS group scored higher by 14.4 points (Takeda, Koyama, & Kurita, 2007). This difference in scores suggested a qualitative difference in the types of cognitive abilities of those with PDD-NOS and those with autism.

Many researchers, however, have cautioned against the use of cognitive ability scores as a tool for differential diagnosis. Lord and Schopler (1989) noted that narrowly categorizing children at a young age based on DQ or early IQ scores is inappropriate based on potential exceptions to the stability of these scores. Although Chiang et al. (2014) found means differences in FSIQ between individuals with AS and HFA, they caution that subgroups may exist within each group that demonstrate similar cognitive functioning. Kanai et al. (2011) also noted differences in mean FSIQ scores in a sample of 122 adults with either AS, HFA, or PDD-NOS; however, they noted that all three groups shared similar cognitive characteristics. Other researchers (Ozonoff et al., 2005) have noted that none of the instruments currently available provide technically adequate differential diagnosis for those with HFA, such as PDD-NOS and those with autism. Further, with the recent changes to the way ASD is diagnosed in the DSM-V (APA, 2013) the use of cognitive measure for differential diagnosis between subtypes of ASD may no longer be a meaningful goal.

Correlation with Achievement

Another area in which IQ may be a useful predictor is academic achievement. It is common practice in schools to use estimated cognitive ability to predict the expected level of academic achievement. In order to determine whether children with autism are performing up to their expected levels, it is important to measure cognitive ability accurately. Empirical research has suggested a relatively strong correlation exists between IQ and achievement for children with ASD. In one study (Mayes & Calhoun, 2003b), a sample of children with autism were administered either the Wechsler Individual Achievement Test (WIAT; Psychological Corporation, 1992) or subtests of the Woodcock-Johnson Tests of Achievement-Revised (WJ-R: Ach; Woodcock & Johnson, 1989) to measure academic performance and either the SB-IV or WISC-III to estimate cognitive ability. The results suggested that for the younger group who was administered the SB-IV and WJ-R: Ach, no statistically significant difference between IQ scores and reading, math, and writing scores emerged. For older children, administered the WISC-III and WIAT, math, reading, writing, and spelling scores did not differ from IQ scores at a statistically significant rate. The exceptions to this trend were that for the low-IQ score group of older children, reading decoding exceeded predicted IQ scores and for the high-IQ score group of older children, Written Expression scores were lower than IQ scores, both at a statistically significant level. In the sample (Mayes & Calhoun, 2003b), approximately half of the participants earned scores in reading decoding and spelling within the Average range (≥ 80). These findings suggested that IQ scores can serve as a valid predictor of achievement in many academic areas for children with autism. These results seem to hold true for both low- and high-IQ score groups (Mayes & Calhoun, 2003a; 2003b).

FSIQ scores seem to be the best predictor of achievement in all academic areas for children with autism. In one study of 54 children with autism 6–14 years of age, stepwise linear regression was used to determine the best predictor of academic achievement (Mayes & Calhoun, 2008). The results suggested that FSIQ scores were the best predictor of word reading, reading comprehension, mathematics, and written expression performance. They accounted for 41–64% (M = 52%) of the variance of the scores. Adding index scores or subtest scores to the prediction equation did not meaningfully contribute to the predictive capacity of the equation (Mayes & Calhoun, 2008). Further, the results of another study (Gwaltney, 2014) indicated that FSIQ scores accounted for a significant amount of the variance in achievement for students with HFA, and even more so than for a group of typically developing control students. In this study, both achievement and IQ scores were obtained for 37 students with HFA and 38 typically developing students between the ages of 8 and 16. For students with HFA, IQ accounted for 54% of the variance in reading comprehension scores and 56% of the variance in math problem solving scores. Thus, the results of these studies suggest that FSIQ scores can be a useful tool for practitioners to predict academic achievement in many skills areas for children with ASD.

The strong predictive validity of IQ scores makes them particularly useful in an academic setting. However, research has suggested that although FSIQ scores are the best predictors of achievement, they may not always maintain an adequate level of predictive accuracy. In a study of 100 adolescents with ASD in both mainstream educational settings and special educational placements, an abbreviated measure of cognitive ability was used to predict achievement scores in the areas of reading and writing (Jones et al., 2009). The results indicated that 30.3% of students (*n* = 30) earned a score in either reading or math that was significantly above their predicted achievement in that area and 16.2% (*n* = 16) displayed a similarly negative discrepancy in either of these two areas. These findings suggested that although FSIQ may be the best predictor of achievement at present, a meaningful margin of error in the predictive validity of scores may exist. Further, the results of Jones et al. (2009) indicated that it is most common for FSIQ scores to underestimate performance in a particular academic area, rather than overestimate it.

Assessing Children with Autism

Obtaining accurate IQ scores for children with autism is important for many reasons; however, obtaining these scores can be difficult. Although no measure is appropriate in all situations for assessing a child with autism, some cognitive tests have been recommended over others. The Leiter International Performance Scales–Revised (Leiter-R; Roid & Miller, 1997) and the DAS have both been suggested due to their range in both chronological and intellectual ability (Ozonoff et al., 2005). For younger children, the Bayley Scales of Infant and Toddler Development, Third Edition (Bayley-III; Bayley, 2005) and the Mullen Scales of Early Learning (Mullen, 1995) are also important because they can provide a DQ score for low-functioning children who are older than the test norms. Ozonoff et al. (2005) also recommend either the Wechsler series or a Stanford-Binet test for children who are verbal. However, even though these recommendations may somewhat narrow the field of assessments from which to choose, many factors still need to be accounted for when choosing an appropriate test for a child with ASD.

Test Factors Related to Score Accuracy

Different tests may yield very different scores for the same individual, thus calling into question the accuracy of cognitive ability scores. For example, in a study of 510 children with autism ages 9–15 years, average IQ scores as measured by the Raven's Standard Progressive Matrices (SPM; Raven, Raven, & Court, 2003; *M* = 88.3) were higher than IQ scores as measured by the WISC-III (*M* = 75.6; Charman, Pickles et al., 2011). Other studies have found similar results when comparing scores on these two tests for samples of individuals with autism (Bölte et al., 2009; Dawson, Soulières, Gernsbacher, & Mottron, 2007; Morsanyi & Holyoak, 2010). This pattern does not, however, seem to hold for samples of typically developing individuals. For example, in one study of 38 children with autism ages 6–17 years compared to 24 typical developing children, the findings suggested that although the children with autism scored higher on the SPM than on the WISC-III at a statistically significant level, the same did not apply for the typically developing control children (Dawson et al., 2007). Nader et al. (2014) found a similar pattern when comparing scores on the SPM and WISC-IV for 25 individuals with autism and 22 typically developing students all between the ages of 6 and 16. The results indicated that the group with autism scored much lower on the WISC-IV FSIQ than the SPM. Further, the results indicated overall lower WISC-IV FSIQ scores for students with autism than typically developing control students, but no difference in scores between the groups on the SPM, indicating that these two tests may be assessing different constructs in each population.

Evidence of this discrepancy is also evident in a study of 48 German individuals with autism as compared with 28 individuals with psychiatric disorder and 25 neurotypical controls all tested with the SPM and the German version of Wechsler-series tests (Bölte et al., 2009). However, results of this study also suggested that although SPM scores were higher than scores on the Wechsler series tests, the differences was only statistically significant for individuals with IQ scores less than 85. The results of another study (Mottron, 2004) of 26 French individuals ages 6–39 years with autism and AD suggested that although the SPM scores were higher than those obtained from French versions of Wechsler-series measures, the differences in scores were not statistically significant. The results of these studies (Bölte et al., 2009; Mottron, 2004) suggest that when individuals with autism are assessed with the SPM, they score relatively higher than when assessed with a Wechsler series-test; however, the reason for this difference remains unclear.

Explanations for Discrepancies in IQ Test Scores

One possible explanation for difference in cognitive ability estimated by different cognitive measures is based on the theory that individuals with autism display uneven subtest profiles. Those subtests on which individuals with autism perform highly may influence one test (e.g., the SPM) more than another test may (e.g., a Wechsler series test), or these uneven peaks and troughs in abilities may correspond with the types of skills each test measures (Mottron, 2004). However, based on previous research (Minshew et al., 2005) these uneven profiles may not exist or may be a test artifact. Blair (2006) suggests that fluid reasoning skills, which is the main ability measured by the SPM, should be an area of weakness, not strength, for individuals with autism based on its close relation to executive functioning. Therefore, the discrepancies in scores typically found between the SPM and Wechsler-series tests may not be due solely to the differential abilities measured by the different tests.

For individuals with autism, the demands of the test rather than intelligence may be what affect the scores so much. SPM is a test that is largely presented visually and nonverbally. The decreased social demands of this test, as compared with the many verbal-laden subtests of the

Wechsler series scales may lead to the higher SPM scores (Wallace, Anderson, & Happé, 2009). Both communication and social skills are known areas of deficits for individuals with ASD, so requiring the use of these types of skills when assessing cognitive ability may put children with autism at a disadvantage during these tests. The large difference in scores on the SPM and Wechsler-series scales suggests that using a test like SPM, which has fewer social and language demands, may yield more accurate measurements of cognitive ability (Bölte et al., 2009).

Some support for this theory can be found in confirmatory factor analysis of the structure of Wechsler series assessments. Goldstein et al. (2008) looked at the scores on the 11 overlapping subtests of two editions of Wechsler series adult and children's intelligence scales in a sample of 137 children and 117 adults with autism. The results suggested that the best-fitting model was a four-factor model comprised of VIQ, PIQ, Freedom from Distractibility, and Social Competence factors. This model was an improvement over the three-factor model at a statistically significant level. These results suggest that for individuals with autism, Wechsler series tests may be measuring social competency skills—a known area of difficulty for this population— and that this may lead to lower scores on these measures when compared with tests that have fewer socially-laden subtests. In newest versions of the Wechsler scales, the two subtests that depended most on the Social Competence factor were eliminated or classified as Supplementary. Thus, the discrepancy in scores on newer versions of the Wechsler tests and the SPM may be much lower. However, as mentioned above, research by Nader et al. (2014) does not suggest that this is the case.

Further evidence of language affecting cognitive scores can be found in a study of the correlation between scores on the Leiter- R and Kaufman Brief Intelligence Test, Second Edition (KBIT-2; Kaufman & Kaufman, 2004). Scattone, Raggio, and May (2012) examined the concurrent validity between these two assessments for a group students with ASD and a group with language impairments. Of the 53 participants in the study, who ranged in age from 4 to 13, 32 were diagnosed with autism. Overall, no significant difference was found between Leiter-R and KBIT-2 Nonverbal IQ scores for participants with autism; however, KBIT-2 Verbal IQ scores were significantly below these two scores for this group. Further, of the 32 participants with autism in the study, 44% (n = 14) had score differences between the Leiter-R and KBIT-2 that were greater than 10 points with 28% (n = 9) scoring higher on the Leiter-R and 16% (n = 5) scoring higher on the KBIT-2. The authors suggest that their results indicate that individuals with autism perform better when administered tests limiting both verbal directions and social interaction with an examiner.

The examination of early assessment measures yields additional information about discrepant scores across different measures. As noted previously, the Bayley-2 is likely to yield lower scores for children with autism compared with other measures based on the socially laden items within the scale (Lord & Schopler, 1989; Magiati & Howlin, 2001). In addition, the Bayley-2 and other early assessments, such as the Psychoeducational Profile-Revised (PEP-R; Schopler, Reichler, Bashford, Lansing, & Marcus, 1990) are designed to yield a DQ score and not an IQ score. Research has suggested that the DQ score yielded by the PEP-R is highly correlated with SB-IV scores and that it may be easier to administer to a child with autism (Delmolino, 2006); however, correlations in this study were below .80 (r = .73), which make the test inappropriate for use in individual or group decision-making in an educational setting (Sattler, 2008). Further, although DQ scores are often used as a proxy for IQ scores for young children, these scores do not measure the same constructs (Takeda, et al., 2007). As Lord and Schopler (1989) note, what any test measures is an important aspect that must be taken into account when interpreting scores and using them to predict future abilities or achievement. Thus, both what the test was designed to measure and what it may unintentionally measure, such as social skills and language, must be considered when choosing an appropriate measure of cognitive ability for a child with autism.

One problem with cognitive ability measures is that they are norm-referenced tests. These tests—and particularly those that yield DQ scores—are designed to assess a population of typically developing individuals rather than individuals with autism, who may develop skills in a different sequence (Magiati & Howlin, 2001). Although a strong correlation between two measures may be present, one test may be more appropriate than the other may be for certain populations. In one study (Edelson, 2005) of 35 children ages 4–18 years, both the Test of Nonverbal Intelligence-Third Edition (TONI-3; Brown, Sherbenou, & Johnsen, 1997) and the Analogic Reasoning (AR) subtest of the Universal Nonverbal Intelligence Test (UNIT; Bracken, & McCallum, 1998) were administered. Although the results initially suggested that the scores from both tests were highly correlated, further analysis indicated that those who completed more items on the AR subtest, which assessed real-world knowledge, obtained scores that were on average, 8 points lower than scores obtained on the TONI-3. The author suggested that this discrepancy in scores may have been due to a deficit in real-world knowledge caused by a lack of typical socialization (Edelson, 2005). However, in a study of 23 British children between the ages of 11 and 16 years with ASD who were administered picture and scene analogic reasoning tests involving real-world knowledge, no difference in scores between the children with ASD and a sample of typically developing children was present (Morsanyi & Holyoak, 2010). These findings suggested that a lack of real-world knowledge may negatively affect cognitive scores for children with autism and that other important factors must be considered when choosing a cognitive ability measure.

The results of these studies support the hypothesis that many subtests or even single items may be inappropriate for children with autism. This problem, as well as other factors, such as the social or linguistic demands of the tests, must be taken into account when determining the appropriateness of the measure for a child with autism. The accuracy of the score obtained may be determined by what facets the test actually measures, which is possibly different from what it was designed to measure. As Edelson (2005) noted, IQ scores for children with ASD may be dependent on the nature of the test rather than on the intelligence of a child.

Behavioral Factors Related to Score Accuracy

Many factors related the typical functioning of children with ASD must be considered when assessing this population. A diagnosis of ASD is made based on deficits in social communication skills as well as the presence of stereotyped and restricted behaviors or interests. These symptoms set apart a child with autism from one without this diagnosis and affect all aspects of a child's life—including the typical testing process—and can introduce error into test scores.

Verbal Ability

Many cognitive tests specifically aim to measure this construct; however, even nonverbal tests may inadvertently assess it. In a study of 39 Taiwanese children administered the TONI-2, age and verbal ability were the two most important predictors of scores despite the fact that the TONI-2 tests nonverbal intelligence (Edelson, Schubert, et al., 1998). Based on the results of this study, verbal ability seemed to be correlated with cognitive ability in individuals with ASD. Edelson, Shubert, et al. (1998) found similar results with a sample of 258 individuals with autism ages 4–41 years. Again, the results of a multiple regression indicated that age and verbal ability were the predictors that most contributed to the variability on scores on the TONI-2. These results also suggest a strong correlation between cognitive ability as measured by a nonverbal test and language skills. However, Edelson (2005) reported that deficits in language may be unrelated to cognitive functioning for children with ASD. Thus, these studies suggest that rather than a correlation between verbal and cognitive ability in individuals with autism, cognitive ability measures may inadvertently measure an individual's verbal skills. Further, even a nonverbal measure may not be enough to ensure the exclusion of extraneous factors included in estimates of cognitive ability for children with autism.

Further evidence indicating that adaptive skills such as communication and social skills may influence cognitive scores can be found in a study examining verbal, nonverbal, and FSIQ scores from both the WISC-IV and the SB-V. Baum (2013) found that in a study of 40 children between the ages of 10 and 16, FSIQ scores from both tests were correlated a statistically significant level with communication, social skills, and activities of daily living as measured by the Vineland Adaptive Behavior Scale. This correlation held true for both the SB-V Verbal and Nonverbal Indexes, as well as for the WISC-IV VCI; however, only the score on the Vineland Activities of Daily Living Composite was significantly correlated with scores on the WISC-IV PRI (Baum, 2013). These results indicate that FSIQ for individuals with ASD may be unduly influenced by language, social and adaptive skills. Additionally, some evidence suggests that measures of cognitive abilities that minimize language requirements, such as the WISC-IV PRI, may be less influenced by these factors. Thus, for individuals with ASD, it may be possible to limit the influence of verbal ability when measuring cognitive skills and, further, doing so may allow these individuals to demonstrate more ability than may otherwise be displayed.

Working Memory and Processing Speed

Deficits in working memory may be related to the accuracy of IQ scores for individuals; however, the empirical evidence for this is mixed. Some research suggests that working memory skills of children with autism are similar to typically developing individuals, whereas findings from other studies have suggested poor working memory skills in this population (Nakahachi et al., 2006; Ozonoff & Strayer, 2001). The results of a recent empirical study suggest that task demands rather than memory deficits lower scores on measures of working memory. In a study of 25 children and adolescents with autism ages 7 to 18 years compared with 15 individuals with Tourette's Syndrome and 15 neurotypical controls, three types of working memory tasks were administered (Ozonoff & Strayer, 2001). On all tasks, no difference appeared among the three groups on working memory scores, which suggests that a working memory deficit affecting cognitive ability scores was not present. Further, in another study of 16 Japanese individuals with autism and AD and 28 neurotypical controls, participants were administered the working memory subtests of the Japanese versions of the Wechsler Adult Intelligence Scale-Revised (WAIS-R; Wechsler, 1981) and the Advanced Trail Marking Test (ATMT; Reitan, 1956), a test of working memory administered on a touchscreen monitor (Nakahachi et al., 2006). A statistically significant difference between the group with autism and the control group on the Digit Symbol subtest of the WAIS-R was detected, but not on the fixed or random tasks on the ATMT or on the Digit Span subtest of the WAIS-R. For both of these empirical studies (Nakahachi et al., 2006; Ozonoff & Strayer, 2001), tasks were presented on computer monitors, so the unimpaired working memory skills shown by the participants with autism in these studies may indicate that social or linguistic demands of non-computerized tasks may have a negative effect on scores in this area, and possibly others. Task demands and not a working memory deficit may be what influence lower scores on these tests.

More recent versions of the WAIS-R and other Wechsler series scales have included the Digit Symbol subtest as a measure of processing speed rather than working memory. Thus, the results of Nakahachi et al. (2006) suggest that the discrepancy in Digit Symbol subtest scores of individuals with ASD and the control group may be due to a deficit in processing speed for individuals with ASD. However, the authors (Nakahachi et al. 2006) noted that no differences in reaction time on the ATMT tasks were observed between groups, indicating no difference in visual-motor coordination between groups. Further research (Wallace et al., 2009) of 23 children with ASD and a typically developing matched control group also indicated that processing speed deficits do not affect cognitive ability scores. In this study, inspection time, which is the time required to make a simple perceptual judgment, was examined and the results showed no difference between groups on this measure. The findings of Wallace et al. indicated that individuals with autism may not have deficits in processing speed and that any observed lower scores on a measure of processing speed may be an artifact of task demands and not ability.

Behavior, Motivation, and Attention

Behavior difficulties for children with autism often interfere with the standardized testing process. Ozonoff et al. (2005) noted that factors such as social difficulties, stereotyped language usage, off-task behaviors, and distractibility complicate the assessment process. In addition, motivation of the child can greatly affect the process and the resulting scores. In one study (Koegel et al., 1997), six preschool- and elementary-age children with autism were assessed under different conditions. In one condition, testing conditions were standard, and in the other, conditions were modified to accommodate each individual child’s unique behavior. For example, one participant engaged in “incessant perseverative stereotypical self-talking” (p. 238), which was dealt with by allowing breaks for self-talk that were contingent upon responding (Koegel et al., 1997). On average, higher test scores were obtained under the motivation/attention condition than under standardized testing conditions. These results were evident across tests of receptive vocabulary, receptive language, verbal intelligence, and nonverbal intelligence (Koegel et al., 1997). Based on these findings, it seems evident to individual behaviors must be accounted for when testing a child with autism so that optimal test scores are obtained.

Further, Ozonoff et al. (2005) noted that when assessing cognitive ability, most children should not be considered untestable and that considering them so reflects instead on the inexperience of the clinician or unavailability of an appropriate test. However, no "most appropriate" measure of cognitive ability for children with ASD has been defined (Delmolino, 2006). Many factors must be considered when choosing an appropriate test, including factors relating to both the test and the child being tested. Considerations, such as the chronological and mental age of the child, the range of standard scores, and types of abilities measured all must be taken into account (Ozonoff et al., 2005). These factors and many others can affect the accuracy of scores obtained on a cognitive ability measure when assessing a child with autism. However, many of these factors can be controlled via careful selection of a cognitive assessment measure. Taking into account the particular demands of the test, such as social and linguistic demands, as well as the constructs the test is designed to measure, can lead to IQ scores that have increased accuracy and predictive validity.

Computer Testing of Children with ASD

Potential issues related to both test characteristics and child characteristics for children with autism may be difficult to address; however, innovative methods of targeting these issues may enable examiners to neutralize many potential threats to the validity of cognitive ability estimates. Preliminary research has suggested that utilizing computers to administer tests may be an effective means of addressing many issues simultaneously, including social and linguistic demands of testing, and motivation problems.

Current Research

Using computer testing for children with ASD has mainly been conducted in the area of executive functioning. Specifically, the Wisconsin Card-Sorting Task (WCST; Heaton, Chelune, Talley, Kay, & Curtiss, 1993) has been adapted and studied with samples of individuals with autism. In this task, examinees are required to sort cards into piles and receive feedback until, through trial and error, they discover the correct rule for sorting. Throughout the administration, the rules are changed and the examinee must adapt to the change by repeating the trial-and-error process until a new rule is uncovered. Although the test yields many measures of performance, the most common are number or percentage of perseverative errors, total categories successfully sorted, and the number or percentage of errors (Ozonoff, 1995). Research has suggested that scores on the standard WCST tasks have high levels of reliability for a population with autism; however, this population is likely to score lower on number of categories successfully completed and higher or perseverative responses and overall errors than a typically developing sample (Ozonoff, 1995). Research involving a computerized version of the WCST found an exception to this trend. In a study of 24 children with autism and 24 typically developing controls matched on age and IQ score, half of the participants in each group were randomly assigned to complete the standard WCST and half were assigned to complete the computerized version of the WCST (Ozonoff, 1995). The results of the study indicated that although statistically significant differences between the individuals with autism and the controls were found on all three commonly used scoring systems when the standard WCST was administered, no differences in scores between the two groups emerged when they were tested using the computerized version of the test. Within-group analysis indicated that individuals with autism tended to perform better on the computerized version, although this difference did not reach statistical significance and that this trend did not appear in the control group (Ozonoff, 1995).

Researchers have suggested that the computerized version may reduce social and verbal demands, which in turn leads to fewer errors in performance (Strauss, Spreen, & Hunter, 2000). Because verbal and social demands are inherent in the original design of WCST and similar measures, it is important to ensure that these variables can be controlled for to determine how much they contribute to scores on this task, especially in the scores of individuals with autism. By using a computer-based testing format, the level of verbal and social skills of an individual are easier to regulate and can be generally consistent across testing situations. Doing so controls a potentially large amount of variation introduced into the scores, which can help determine how much these factors influence the total score.

Ozonoff et al. (2005) suggested that the use of a computerized version of a test is useful in determining the full extent of a child's capabilities in performance. Although the WCST is not designed to measure intelligence, it does measure other aspects of cognitive functioning. And, as previously mentioned, studies in the area of working memory, which is often included as a component of cognitive ability, have revealed similar results. In two studies (Nakahachi et al., 2006; Ozonoff & Strayer, 2001), no difference was evident in the scores of individuals with autism and neurotypical individuals on computer-administered tasks although some research had suggested a deficit in processing in this area. It is possible that the positive effects of eliminating the social component on tests of executive functioning and working memory may apply to tests of cognitive ability and may have similar positive effects on the scores, as suggested by Scattone et al. (2012).

In addition to addressing the social demands of a cognitive test, using a computerized measure can aid in potential confounds in testing a child with autism: motivation and attention. Evidence in favor of the use of computers to increase motivation for children with ASD can be seen in a study of reading interventions. In this study (Williams, Wright, Callaghan, & Coughlan, 2002), eight children with autism ages 3–6 years were engaged in either a personal reading instruction intervention or computerized version of the same program. Children in the computer group averaged more time-on-task minutes (M = 9.9) than the personal instructions group (M = 2.8). The results suggested that the computer distracted the children from self-stimulatory behaviors and that resistance to computer use was minimal. These findings (Williams et al., 2002) suggest that computers may be motivating for children with autism to use, and their use may increase attention to the task. Thus, using computers may be one useful method of decreasing problems with motivation and attention

Computerized assessment may capitalize on the screening out of unnecessary sensory information. Difficulties processing sensory information may make the minimal distractions, consistency, and predictability of computer use comparatively easy to the sensory processing demands of a standard testing situation (Williams et al., 2002). Therefore, the reduction in sensory demands may be the reason that scores from computer tests are often higher than scores obtained under standard testing conditions.

Issues with Computerized Testing

Computer testing for children with autism may solve some problems that typically occur, but it may also introduce new difficulties. For example, scores on computerized assessments may be affected by anxiety or unfamiliarity with a computer (Ozonoff, 1995; Schulenberg & Yutrzenka, 2004). However, if these factors were to affect scores with this population, the scores would likely be lower, not higher. Further, research with a computerized version of the Raven's Standard Progressive Matrices (SPM-C) with college students revealed no difference between scores on a measure of anxiety between those taking the computerized version and those taking the standard version (Williams, & McCord, 2006). However, it is the responsibility of the clinician to ensure that steps have been taken to allay any anxiety an examinee may have about computers, as well as to guarantee that examinees have adequate knowledge and physical ability to use a computer properly (Schulenberg & Yutrzenka, 2004). Further, the Joint Committee on Testing Practices (JCTP; 2004) recommends that examiners allow time for the test takers to become familiar with the format of the test and any equipment used. Failing to make certain that these criteria are met may increase the risk that reliable and valid scores will not be obtained.

Alternately, it has been suggested that the computerized version of a test may be easier than the standard version due to reduction of verbal and social demands. However, research has suggested that higher scores on the computerized version of a test do not hold for all populations, just individuals with autism (Ozonoff, 1995). The fact that scores are higher only for this population suggested that the format of the test may inadvertently be affected by areas of deficit common to individuals with autism, such as verbal and social skills. Further, reducing these demands may not make the test easier, but may make it more equitable for this population and may make the scores obtained more accurate in what they measure.

Another potential problem when using computerized assessments is equating scores on that measure with scores on the original version (Schulenberg & Yutrzenka, 2004). Differences in item functioning, factor structure, and constructs measured by the scale can introduce much error into the scores obtained from a computerized assessment (Strauss et al., 2000). These scores may not reflect the same construct or may reflect a much lower level of reliability than scores obtained from the original. Prior to the use of a computerized assessment, the equivalence for scores obtained on the computerized version and the paper-and-pencil version should be established and documented (American Psychological Association, 1986). Thus, research on both the reliability and validity of scores is necessary for all computerized assessments before they are used in a real-world setting.

Some evidence has suggested that scores obtained from computerized test may retain a similar level of reliability and factor structure to those obtained on the original. Strauss, Spreen, and Hunter (2000) noted that research on the SPM-C using scores obtained from 190 Australian children ranging in age from 10 to 14 years indicated that the computerized version had better item functioning than the standard version. Further, research on the SPM examined the scores of 50 college students who were administered either the computerized or standard version (Williams & McCord, 2006). The findings revealed no statistically significant differences in means or standard deviations between the two groups and a relatively high test-retest reliability coefficient for the computerized version of the test ($r = .95$). The results of these studies indicate that computerized version of a cognitive assessment, particularly the SPM, may maintain score equivalence, as well as a high level of reliability when compared to the original version. However, it is the responsibility of the clinician to guarantee that the psychometric properties of the scores obtained from any computerized assessment used have been properly examined and maintain the same standards.

It is also the responsibility of the clinician to ensure that the computerized test follows other guidelines suggested by the American Psychological Association (1986). For example, computerized assessments must provide a similar amount of feedback and allow examinees a similar amount of control as the original version. This guideline reflects the need for examinees to be able to verify their responses and potentially change it, if necessary. It is also important to note that computerized scoring procedures must be monitored and verified to ensure that proper scores are being obtained (American Psychological Association, 1986). As Harvey and Carleson (2003, p. 95) noted, "Computers are simply tools programmed by people." Computers and programs are as fallible as the individuals who programmed them and, as such, are subject to the same type of scrutiny in order to avoid error.

Maintaining ethical standards of practice when implementing computer-based testing both in clinical practice and in research is important (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999). Schulenberg and Yutrzenka (2004) recommend giving the client a chance to refuse computerized testing and to engage in the standard method of assessment if desired. Ensuring clinician competence with the technology used is also essential. Psychologists should not practice outside their area of competence and computer testing is no exception (Harvey & Carlson, 2003). Clinicians must make certain that they are familiar with both the hardware and software they use for computer-based testing to ensure they are working as intended. Finally, confidentiality must be maintained, as it is for any time of assessment or research situation, but which may be more difficult when using a computer (Schulenberg & Yutrzenka, 2004). In general, many pitfalls must be avoided when using a computer-based measure, but such tests have the potential to be effective measures of cognitive ability.

Although many potential drawbacks to the use of computerized assessments may be present, it is important to note that their use can be vital, especially for students with autism. Guidelines for fair testing practices (JCTP, 2004) note that examiners need to select tests that have appropriate modifications for students with disabilities who need accommodation. Computerized testing may be a good way to provide the appropriate accommodations necessary for students with autism in order to obtain the most accurate representation of their cognitive ability. Currently, little research has been conducted in the area of computerized intelligence assessments for children with autism. JCTP guidelines (2004) recommend that psychometric evidence for the scores from computer-based assessment with diverse groups be evaluated prior to use. Thus, investigating the reliability and validity of scores from specific computer-based tests among students with autism is essential if they are to be utilized in practice.

The Current Study

Rationale

The current study aims to examine how best to address the challenges associated with cognitively assessing children with ASD. Research has suggested that factors related to the test itself and factors related to the child being assessed must be accounted for when administering an IQ test to a child with autism.

Important factors include those related to the types of skills that the test is designed to measure (Mottron, 2004), such as verbal ability or fluid reasoning, and as well as subtests or items may inadvertently measure, such as real-world knowledge (Edelson, 2005), verbal skills (Edelson, Edelson, et al., 1998; Edelson, Schubert, et al., 1998), or social skills (Bölte et al., 2009; Magiati & Howlin, 2001). However, considering all these factors when choosing a test may be a complex process. For example, choosing to administer a nonverbal test may reduce the language demands on a child, but many nonverbal measures included pantomimed instructions, which may keep or increase the social demands of the test. Few tests can discard both social and verbal demands while ensuring that all items do not inadvertently measure culturally related skills such as real-world knowledge or social skills.

It may be possible to control the social and verbal demands of the testing environment using computerized testing. Utilizing a computer-based measure may lead to more accurate scores due to the controlling of social, verbal, or sensory interactions (Ozonoff, 1995). With the consistent and predictable computer format for testing, much of the variance in scores of cognitive ability for children with autism can be controlled. Although a computerized version of a cognitive ability measure may not necessarily demonstrate equivalence of scores, one such measure that maintains some preliminary evidence for equivalent reliability and factor structure of scores is the SPM-C (Strauss et al., 2000; Williams & McCord, 2006). Further, scores from the original version of this measure have been suggested to be superior to scores from other measures due to the abstract nature of the items and the minimization of verbal and social cues present (Bölte et al., 2009; Wallace et al., 2009). The computerized version of the SPM had the added advantage of being more intrinsically motivating for children with autism to complete than the original non-computerized version (Williams et al., 2002). Thus, the computerized version of the SPM may be one method of dealing with many of the challenges, both test- and child-related, involved in cognitively assessing a child with autism.

The accuracy of the scores obtained for a child with ASD on a cognitive measure is important for many reasons. The scores may not be useful for differential diagnosis (Ozonoff, 1995); however, they can be good predictors of academic achievement (Mayes & Calhoun, 2003a; 2003b) and of general measures of outcome (Gillberg, & Steffenburg, 1987). Prediction of academic achievement is especially important for school-age children given that performance well below the predicted level can indicate a problem or interference in the schooling of a child.

Although estimates of cognitive ability using current tests have a somewhat high correlation with achievement, more accurate cognitive scores may, in turn, provide more accurate predictions of achievement.

Hypotheses

The goal of the current study was to examine the differences in cognitive ability scores between a (a) computerized, (b) language-free, and (c) verbal measure of cognitive ability. The concurrent validity of each of the scores was examined based on their ability to predict academic achievement as measured by standardized reading and math assessments. Three hypotheses were posited in response to the research questions:

Hypothesis 1:

Based on the results of previous research (Bölte et al., 2009; Dawson et al., 2007; Morsanyi & Holyoak, 2010), it was hypothesized that a statistically significant difference between scores obtained for a sample of children with autism on different measures of cognitive abilities would be found.

Hypothesis 2:

With regard to the predictive ability of the scores, it was hypothesized that the computer-based measure would account for the most variance in academic performance and that the scores from the other measures would not add to this predictive ability.

Hypothesis 3:

It was hypothesized that social and communication skills would predict cognitive ability scores obtained on a verbal measure, but would not predict those obtained on a computerized test of cognitive ability. It was also hypothesized that social but not communication skills would predict the scores on a nonverbal cognitive measure.

Chapter Three: Method

Participants

The participants in the study were 12 children and youth between the ages of 8 and 16 (*M* = 12.00, *SD* = 3.00) years who currently reside in Virginia. A majority of the sample were boys (75.00% [*n* = 8]). They ranged in grade from third to tenth. For the participating children, 25.0% were in Grade 3 (*n* = 3), 25.0% were in Grade 4 (*n* = 3), 8.3% were in Grade 5 (*n* = 1), 8.3% were in Grade 7 (*n* = 1), 8.3% were in Grade 8 and 25.0% were in Grade 10 (*n* = 3). All children had an educational diagnosis of autism and were not labeled as ID by their schools. However, it should be noted that although no students were reported to have educational labels of ID, 42% of participants earned K-BIT-2 FSIQ scores that would have placed their intellectual functioning in the ID range. Further, to participate, students were required to have at least partial sight and hearing to be included in the study in order to be able to complete all assessments. They were also required to have minimal gross motor capabilities in order to be able to accurately point to items on the tests.

The level of education obtained by the parents of participants was reported to be some college, but no degree for 8.3% of mothers (*n* = 1) and 0.0% of fathers, bachelor's degree for 50.0% of mothers (*n* = 6) and 58.3% of fathers (*n* = 7), master's degree for 16.7 % of mothers (*n* = 2) and 8.3 % of fathers (*n* = 1), and doctorate/professional degree for 25.0% of mothers (*n* = 3) and 33.3% of fathers (*n* = 4). With regards to student ethnicity, 16.7% (*n* = 2) of participating children were reported to be of Hispanic origin, 75.0 % (*n* = 9) were reported not be of Hispanic origin and 8.3% (n = 1) preferred not to answer. Based of parent report of a child's race, 16.7% (*n* = 2) of the participating children were Black or African American, 58.3% (*n* = 7) were White, and 25.0% (*n* = 3) were of mixed race.

Measures

Raven's Standard Progressive Matrices. The Raven's Standard Progressive Matrices (SPM; Raven, Raven, & Court, 2003) was designed to measure the educative ability component of general intelligence. The test is comprised of five sets of 12 visual puzzles. Each puzzle exhibits serial changes in two dimensions with one part of the puzzle missing. The examinee must choose the correct missing piece from an array of eight options. As the items progress they become increasingly more difficult. This test provides a single summary score obtained by adding the number of items answered correctly. This score is then converted into a percentile representing the percentage of the reference group who performed better than or equal to an individual.

Reliability evidence from the 1979 British standardization sample indicated internal consistency ranging from .97 to .99 with various socio-economic groups. Similar levels of internal consistency were noted among US and South African samples (α = .97 – 1.00). Test-retest reliability ranged from .83 to .93 for different age groups, with the highest correlation for individuals under the age of 30. Research has also indicated that correlations between the SPM and the WISC-R and SB-IV range from .54 to .86 with an English-speaking sample of children and adolescents.

Computerized Raven's Standard Progressive Matrices. Computerized versions of the SPM (SPM-C) have been developed for many different computer platforms. As with the SPM, participants must chose the missing piece of a puzzle from an array of eight choices. Examinees make their choice by moving the computer cursor to click on the option they have chosen. They are then asked to confirm their choice by clicking the next arrow. This test provides a single summary score obtained by adding the number of items answered correctly. This score is then converted into a percentile representing the percentage of the reference group who performed better than or equal to an individual. Research has suggested that the test-retest reliability coefficient for the two tests is .88. Research has reported no statistically significant difference between scores obtained on the SPM and the SPM-C (Williams & McCord, 2006). Scores on the SPM-C also correlate highly with Wechsler-series nonverbal reasoning subtests (*r* = .81 with Matrix Reasoning) and measures of overall cognitive ability (*r* = .74-.84 with FSIQ scores).

Kaufman Brief Intelligence Test, Second Edition. The Kaufman Brief Intelligence Test, Second Edition (KBIT-2; Kaufman & Kaufman, 2004) is a brief intelligence test designed for individuals ages 4 to 90 years, which measures both verbal and nonverbal intelligence. Verbal Intelligence is measured through two subtests (Verbal Knowledge and Riddles) and nonverbal intelligence is measured through the Matrices subtest. The Verbal Knowledge requires examinee to identify the correctly described picture from an array of six images, whereas the Riddles subtest requires examinee to correct identify the described item from clues. For the Matrices subtest, the examinee must choose the correct missing piece of an image from an array. Scores from all three tests combine to yield an estimate of FSIQ. Scores on the Verbal Scale, Nonverbal Scale, and FSIQ are presented as standard scores with a mean of 100 and a standard deviation of 15. The KBIT-2 was normed on a sample of 2,120 children and adults that was representative of US Census data.

Test-retest reliabilities with an average of a four-week interval range from .88 to .93 for the Verbal scale (*M* = .91), from .76 to .89 for the Nonverbal Scale (*M* = .83), and from .88 to .92 for the IQ composite (*M* = .90). The mean split-half reliability coefficient for the IQ Composite scores of all age groups, based on the Spearman-Brown formula, is .93. KBIT-2 IQ Composite scores correlate highly with IQ scores from the KBIT (*r* = .80), the WISC-III (*r* = .76), and the WISC-IV (*r* = .77). Group differences were also found on KBIT-2 FSIQ scores for special populations as compared to a control group, with the control group having higher scores: learning disabilities (difference of 9.4), children with ADHD (difference of 7.9), individuals with mental retardation (difference of 33.4), and students with Traumatic Brain Injury (difference of 25.5; Kaufman & Kaufman, 2004).

Gilliam Autism Rating Scale- Second Edition. The Gilliam Autism Rating Scale-Second Edition (GARS-2; Gilliam, 2006) is a rating scale used to help clinicians identify and diagnose children with autism. The rating scale has two forms: a parent scale and a teacher scale. The GARS-2 is designed to be completed for individuals ages 3 to 22. The scale consists of 42 items grouped into three subscales of Social Interaction, Communication, and Stereotyped Behaviors, which are based on the classification system of the DSM-IV-TR (APA, 2000). Raw scores for these scales can be converted into standard score with a mean of 10 and a standard deviation of 3. The GARS-2 also yields an overall Autism Index score, which is represented as a standard score with a mean of 100 and a standard deviation of 15. Ratings are made on a frequency basis and the scale can be completed in 5-10 minutes.

The norms for the GARS-2 are based on a representative sample of 1,107 individuals with autism from 48 states. Strong reliability evidence has been reported for both test-retest (*r* = .70 - .90) and internal consistency (*r* = .84 - .88) of scores on the subtests and the Autism Index (*r* = .94; .88, respectively). Research has also suggested that the subscales are representative of the characteristics of autism and that the Autism Index scores strongly correlate with overall rating scores of other autism rating scales such as the Autism Behavior Checklist (*r* = .64). GARS-2 scores have also demonstrated differential diagnosis ability between individuals with autism (*N* = 100) and those without disabilities (*N* = 49), with intellectual disabilities (*N* = 72), or with multiple disabilities (*N* = 76). Sensitivity and specificity for distinguishing between these groups range from .84 to 1.00. Correlations among GARS-2 subtest are between .46 and .59 (Gilliam, 2006).

Basic Achievement Skills Inventory (BASI). The BASI Surveys (Bardos, 2004) are computer-based, multi-level, norm-referenced achievement test. The Survey version tests of both reading (Verbal Skills) and mathematics (Math Skills) performance. The test requires students to answer multiple-choice questions and provides norm-referenced scores for individuals ages 8 to 80. Each examinee must complete each section within 25 minutes. The BASI Surveys in math and reading each yields scaled scores, percentile ranks, and grade equivalent scores which are calculated by the computer program. The Verbal Skills section accesses the specific skills of Vocabulary, Language Mechanics, and Reading Comprehension. The Math Skills section assesses the areas of Math Computation and Math Application. The BASI Survey was normed on a sample of 5,000 children and adults in sample representative of 2000 US Census data. Scores are reported as standard scores with a mean of 100 and a standard deviation of 15.

The test-retest reliability estimates for BASI Survey scores range from .23 to .63. For children ages 8 – 18, internal consistency ranges by grade from .73 to .91 for Math Skills and from .75 to .90 for Verbal Skills. BASI Verbal Skills scores have relatively high correlations with scaled scores on The Iowa Tests of Educational Development Reading Total scale (*r* = .56). Scores on the BASI Math Skills section also display evidence of criterion-related validity based on correlation by grade with scaled scores on The Iowa Test Development Quantitative Thinking Advanced Skills scale (*r* = .41; Bardos, 2004).

Virginia Standards of Learning Assessments in Reading and Math. The Virginia

Standards of Learning (SOL; Virginia Department of Education, 2008) assessments are administered to students in the state of Virginia at the end of the school year from Grade 3 to Grade 11. The Virginia SOL assessments are standards-based tests designed to measure student performance on Virginia‘s content standards, in the areas of reading, writing, mathematics, science, and history/social science. The reading and math tests consist of multiple-choice items. These tests are administered during the school day in either an online or a paper-and-pencil format over a three-week period in May of each year. Only one test is administered to students per school day. Scores range from 0 to 600 with a score of 400 or above representing a "proficient" score.

Internal consistency scores range from .84 to .89 for the online reading SOL assessment and from .82 to .90 for the paper-and-pencil reading assessment. Internal consistency scores range from .87 to .91 for the online math SOL test and from .87 to .93 for the paper-and-pencil math test. Correlations between the SOL reading assessment and Stanford Achievement Test (SAT) 9 Reading scores were .76 - .79 in Grades 3 and 5, .80 - .81 in Grade 8, and .57 - .62 in Grade 11. Correlations between the SOL math assessment and SAT 9 Math scores were .72 in Grade 3, .76 in Grade 5, .82 in Grade 8, and .71 in Grade 11 (Virginia Department of Education, 2008). Despite gaining parent authorization to review SOL scores, only 4 of 12 students had this score available. Subsequently, no further evaluation using these scores was possible

Procedure

Participants were recruited via email solicitations to support groups of parents with children on the autism spectrum or via flyers at offices of psychologists specializing in work with students on the autism spectrum. Parents who were interested in the study contacted the principal investigator of the study through phone or email. Participants were screened for the study through a phone interview with the parent to determine if they qualified to be included in the study. Participating youth were required to have an educational diagnosis of autism and no education diagnosis of Intellectual Disability based on parent report. Data from students who participated in the study and were determined to earn IQ scores that fell in the ID range (< 70) were included in the study. Participating students were also required to have a minimum level of verbal, motor, hearing, and vision skills in order to be able to complete the assessments. Testing sessions were then scheduled with families that qualified. Most families opted to complete testing sessions in their homes with two participants opting to engage in the testing session at a private room in their local library.

Three tests of cognitive ability were administered to all participants: the computerized SPM-C (a nonverbal and non-social measure), the SPM (a nonverbal, but socially interactive test), and the KBIT-2 (a social and verbal test). All tests were administered by a doctoral-level graduate student with experience administering assessments to children with autism. Testing took place on two different days that were approximately one to two weeks apart. Each test took about 15-30 minutes to complete and two tests were completed during each testing day. The SPM and the SPM-C were administered on different days in order to allow the most time to pass between administrations as a means of diffusing practice effects. Participants were randomly assigned so that half were administered the SPM and BASI on the first day and the other half were administered the SPM-C and KBIT-2 on the first testing day to counterbalance practice effects.

All participating youth were asked to sign their assent to participate in the study and parents were required to sign consent for their child. The examiner then completed testing with the participants in a location in their house or in a public space that was generally quiet and free from distractions. All standard administration guidelines and procedures were followed for the tests. Once testing had been completed, the examiner allowed the child to pick a small prize as a gift for their cooperation such as a scented pencil or puzzle eraser. This procedure was repeated on the next testing day with the two tests that the participant had not yet completed. Parents received $10 for allowing their child to participate in both days of testing and were entered into a drawing for a $100 gift card upon the study's completion.

Parents were asked to sign a form to allow their child's school to share their child's SOL scores with the study. A survey of demographic information was also administered to parents, which included questions related to race, ethnicity, and educational attainment of parents or guardians. The GARS-2 Parent Form was administered to parents about each participating child in order to determine levels of functioning in the areas of communication and social skills.

Chapter Four: Results

Descriptive Statistics

The means, standard deviations, skewness, kurtosis, and ranges of the KBIT-2 FSIQ scores, SPM scores, SPM-C scores, GARS-2 scores, BASI-V scores, and BASI-M scores were calculated. Descriptive statistics for these scores are presented in Table 1, and correlations between these variables are presented in Table 2.

Table 1

Summary of Means, Standard Deviations, Skewness, and Kurtosis for KBIT-2 FSIQ, SPM, SPM-C, GARS-2, and Reading and Math Achievement Scores (N = 12)

	M	*SD*	Skew	Kurtosis	Range
KBIT-2 IQ score	76.00	32.35	0.03	-1.81	40 – 124
SPM Score	83.03	29.48	-0.43	-1.02	32 – 120
SPM-C score	74.94	24.71	-0.14	-1.40	35 – 107
BASI Math Score	92.92	18.07	-0.21	-0.43	60 – 121
BASI Verbal Score	97.75	18.53	-0.25	-1.31	70 – 123
GARS-2 Communication Score	7.42	3.00	0.37	-0.99	4 – 13
GARS-2 Social Interaction Score	8.08	2.54	1.17	1.46	5 – 4
GARS-2 Stereotyped Behavior Score	8.08	2.23	0.46	-1.07	5 – 2
GARS-2 Autism Index Score	86.33	13.03	0.28	-1.07	70 – 109

Note. BASI = Basic Academic Skills Inventory; GARS-2 = Gilliam Autism Rating Scale, Second Edition; SPM-C = Raven's Computerized Standard Progressive Matrices; SPM = Raven's Standard Progressive Matrices; KBIT-2 - Kaufman Brief Intelligence Test, Second Edition; KBIT-2, SPM, SPM-C, BASI Math, BASI Verbal and GARS-2 Autism Index scores all have a mean of 100 and a standard deviation of 15. GARS-2 Communication, GARS-2 Social Interaction, and GARS-2 Stereotyped Behavior Scores all have a mean of 10 and a standard deviation of 3.

Table 2

Intercorrelations Between for KBIT-2 FSIQ, SPM, SPM-C, GARS-2, and Reading and Math Achievement Scores (N = 12)

	KBIT-2 IQ	SPM	SPM-C	BASI Math	BASI Verbal	GARS-2 Comm.	GARS-2 SI	GARS-2 SB
SPM Score	.78**							
SPM-C score	.71*	.96**						
BASI Math Score	.57	.60*	.53					
BASI Verbal Score	.88**	.73**	.69*	.69*				
GARS-2 Comm. Score	-.51	-.31	-.29	-.43	-.64*			
GARS-2 Social Interaction Score	.25	.06	-.05	-.20	.10	.29		
GARS-2 Stereotyped Behavior Score	.03	.13	.02	-.06	-.07	.47	.50	
GARS-2 Autism Index Score	-.12	-.07	-.15	-.31	-.29	.79**	.76**	.80**

Note. KBIT-2 = Kaufman Brief Intelligence Test, Second Edition; SPM = Raven's Standard Progressive Matrices; SPM-C = Raven's Computerized Standard Progressive Matrices; BASI = Basic Academic Skills Inventory; GARS-2 = Gilliam Autism Rating Scale, Second Edition; Comm. = Communication; SI = Social Interaction; SB = Stereotyped Behavior

* $p < .05$. ** $p < .01$.

Analyses of Variance

A two-way repeated-measure ANOVA was performed to compare the standard scores of participants on each of the three cognitive ability tests they completed across different orders of testing. The assumptions of normality, linearity, homogeneity of variance, and sphericity were tested. Histograms and p-p plots of the data were visually examined to determine normality. Skew and kurtosis values were also calculated, examined, and determined to be within normal limits. Mauchley's Test of Sphericity was used to test the assumption of sphericity, and the results indicated that this assumption was violated, χ^2 (2) = 9.04, p = .01. To correct for this violation, the Greenhouse-Geisser correction was applied to the degrees of freedom, as the Greenhouse Geisser ε was determined to be less than .75 (ε = .61). Results of the test of the main effect of order are presented in Table 3, and the results of the 3 x 2 repeated measures ANOVA are reported in Table 4.

Table 3

Summary of Two-Way Analysis of Variance for Test Presentation Order on KBIT-2, SPM, and SPM-C Scores

	M	*SD*	*SS*	*df*	*MS*	*F*	*p*	Partial η^2
Order 1	94.04	20.36						
Order 2	66.52	28.12						
Intercept			225577.94	1	225577.94	129.85	.00	.93
Test Order			6625.39	1	6625.39	3.81	.08	.28
Error			17372.13	10	1737.21			

Note: *N* = 12. SPM-C = Raven’s Computerized Standard Progressive Matrices; SPM = Raven’s Standard Progressive Matrices; KBIT-2 = Kaufman Brief Intelligence Test, Second Edition

Table 4

Summary of 3 x 2 Repeated Measures Analysis of Variance on KBIT-2, SPM, and SPM-C Scores

	M	*SD*	*SS*	*df*	*MS*	*F*	*p*	Partial η^2
SPM								
Order 1	100.47	16.13						
Order 2	70.57	31.38						
SPM-C								
Order 1	91.85	16.66						
Order 2	62.85	22.94						
KBIT-2								
Order 1	89.80	28.87						
Order 2	66.14	33.03						
Test Type			482.74	1.22	394.37	1.30	.29	.12
Test Type × Order			66.39	1.22	54.23	.18	.73	.02
Error			3723.02	12.24	304.15			

Note: *N = 12*. SPM-C = Raven's Computerized Standard Progressive Matrices; SPM = Raven's Standard Progressive Matrices; KBIT-2 = Kaufman Brief Intelligence Test, Second Edition

No statistically significant main effect emerged for testing order, $F(1, 10) = 3.81$, $p = .08$, $\eta^2 = .28$. Neither a main effect for test type, $F(1.22, 12.24) = 1.30$, $p = .29$, $\eta^2 = .12$, nor an interaction between order and test type, $F(1.22, 12.24) = .18$, $p = .73$, $\eta^2 = .02$, were evident. As no significant differences were found, no post hoc tests were completed and no effect sizes were calculated for any group differences.

Correlation

For reading achievement, the correlations between BASI-V scores and each of the IQ scores were compared using the Steiger's *z* to determine predictive utility. To account for multiple comparisons, a Bonferroni correction was used to adjust the *p*-value to .017. Comparisons between the three correlations revealed that no correlation was different from any other at a statistically significance level (KBIT-2 and SPM: $t(9) = 1.34$, $p = .09$; KBIT-2 and SPM-C: $t(9) = -1.46$, $p =.07$; SPM-C and PRM-S $t(9) = 0.61$, $p = .27$).

For the math achievement, Steiger's *z* was again used to compare the correlations between each IQ score and the BASI-M scores. These comparisons revealed no statistically significant differences for the correlations between KBIT-2 and SPM scores, $t(9) = -.17$, $p = .43$, KBIT-2 and SPM-C scores, $t(9) = .19$, $p = .42$, or SPM and SPM-C scores, $t(9) = .90$, $p = .18$, at the .017 level.

Multiple Regression

Finally, three multiple regression models were applied with each type of IQ score as a criterion variable. Although 30 participants per variable is generally recommended for this type of analysis, VanVoorhis and Morgan (2007) suggest that when using 6 or more predictors, the absolute minimum of participants should be 10. Further explanations of the limits that sample size impose of the results of this study is presented in the discussion section. In each model, GARS-2 Social Interaction scores were entered in Step 1, and GARS-2 Communication scores were entered in Step 2. For the model regressing these variable on the KBIT-2, Step 1 was not statistically significant, $F(1\ 10) = .68$, $p = .42$; however, the addition of the GARS-2 Communication score in Step 2, $F(1, 9) = 5.90$, $p = .04$, was statistically significant. For KBIT-2 scores, GARS-2 Social Interaction scores accounted for 6% of the variance and GARS-2 Communication scores accounted for an additional 37% of the variance. Overall, 43% of the variance in KBIT-2 IQ scores could be accounted for by these two variables combined. A summary of the regression analyses is displayed in Table 5.

Table 5

Hierarchical Regression Analysis Summary for GARS-2 Scores Predicting KBIT-2 Scores (N = 12)

Predictor	B	$SE\ B$	β	t	ΔR^2	R^2	R^2_{adj}
Step 1							
GARS-2 Social Interaction	3.22	3.90	.25	.83	.06	.06	-.03
Step 2							
GARS-2 Communication	-6.87	2.83	-.64	-2.43*	.37	.43	.31

Note. GARS-2 = Gilliam Autism Rating Scale, Second Edition; KBIT-2 - Kaufman Brief Intelligence Test, Second Edition

*$p < .05$.

For the model regressing SPM on the same variables, the addition of variables during both Step 1, $F(1,10) = .04$, $p = .85$, and Step 2, $F(1,9) = 1.21$, $p = .30$, was not statistically significant. Neither variable in the model was statistically significant on an individual basis.

GARS-2 Social Interaction scores accounted for 0% of the variance and GARS-2 Communication scores accounted for 12% of the variance. Overall, approximately 12% of the variance in SPM scores could be accounted for by these variables combined. A summary of the regression analyses is displayed in Table 6.

Table 6

Hierarchical Regression Analysis Summary for GARS-2 Scores Predicting SPM Scores (N = 12)

Predictor	*B*	*SE B*	β	*t*	ΔR^2	R^2	R^2_{adj}
Step 1							
GARS-2 Social Interaction	.70	3.67	.06	.19	.00	.00	-.10
Step 2							
GARS-2 Communication	-3.54	3.21	-.36	-1.10	.12	.12	-.07

Note. GARS-2 = Gilliam Autism Rating Scale, Second Edition; SPM = Raven's Standard Progressive Matrices

*$p < .05$.

For the same model regressed on SPM-C scores, neither the variable added during Step 1, $F(1,10) = .03$, $p = .87$, nor Step 2, $F(1, 9) = .81$, $p = .39$, were statistically significant. GARS-2 Social Interaction scores accounted for 0% of the variance and GARS-2 Communication scores accounted for 8% of the variance, leading to 8% of variance in SPM-C scores accounted for overall. Table 7 displays a summary of this regression analyses.

Table 7

Hierarchical Regression Analysis Summary for GARS-2 Scores Predicting SPM-C Scores (N = 12)

Predictor	*B*	*SE B*	β	*t*	ΔR^2	R^2	R^2_{adj}
Step 1							
GARS-2 Social Interaction	-.51	3.07	-.05	-.16	.00	.00	-.10
Step 2							
GARS-2 Communication	-2.47	2.75	-.30	-.90	.08	.09	-.12

Note. GARS-2 = Gilliam Autism Rating Scale, Second Edition; SPM-C = Raven's Computerized Standard Progressive Matrices

*p < .05.

Qualitative analysis

All participating students were reported by parents not to have an educational label of ID; however, as the cognitive abilities of 42% of participating students fell in the cognitive ability range typically required for this designation, a comparison was completed between the data for those with IQ scores below 70 and those with IQ scores above 70. The groups were determined based on scores the participants obtained on the KBIT-2. Due to the small sample size of the current study, further statistical analysis comparing groups based on higher- (IQ > 70) and lower-IQ (IQ < 70) score could not be completed; however, this data was graphed and visually analyzed. In addition, the line of best fit for each group on the three scatter plots was calculated and compared. The results are presented in Figure 1-3.

The slopes of the lines best fitting the scatter plot of SPM and SPM-C scores were similar for the lower-FSIQ group (*r* = .94) and the higher-FSIQ group (*r* = .87). However, slopes for the lines of best fit for the scatter plot of the SMP and the KBIT-2 scores varied by group. The slope for the lower-FSIQ group (*r* = .05) was close to 0, whereas the slope for the higher-IQ group was much closer to 1 (*r* = .88). The same pattern was also noted for the slopes of the lines for the scatter plot including SPM-C and KBIT-2 scores (lower-FSIQ *r* = .06; higher-FSIQ *r* = .84).

Figure 1. *Scatter Plot of SPM and SPM-C Scores for both Higher-FSIQ and Lower-FSIQ Groups*. Circles represent the higher-FSIQ group and exes represent the lower-FSIQ group. The formula for the line of best fit for the higher-IQ group is $y = .94x - 6.67$. The formula for the line of best fit for the lower-IQ group is $y = .87x + 6.37$.

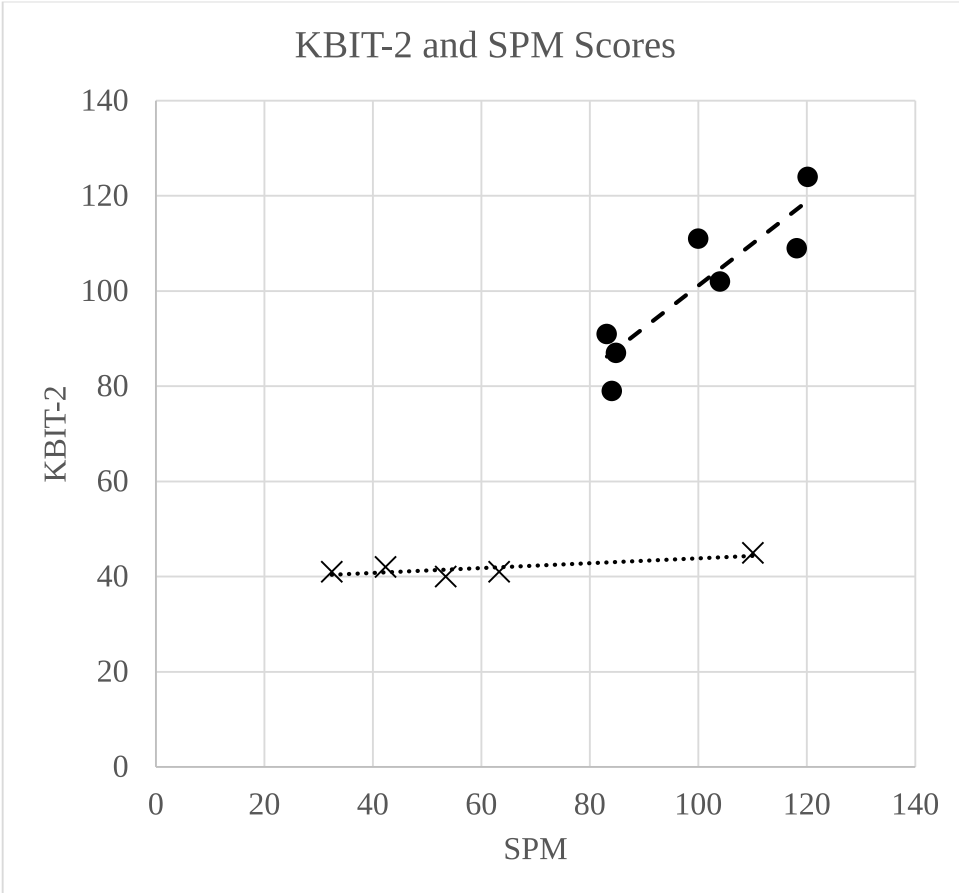

Figure 2. *Scatter Plot of SPM and KBIT-2 Scores for both Higher-FSIQ and Lower-FSIQ Groups*. Circles represent the higher-FSIQ group and exes represent the lower-FSIQ group. The formula for the line of best fit for the higher-IQ group is $y = .88x + 12.74$. The formula for the line of best fit for the lower-IQ group is $y = .05x + 38.68$.

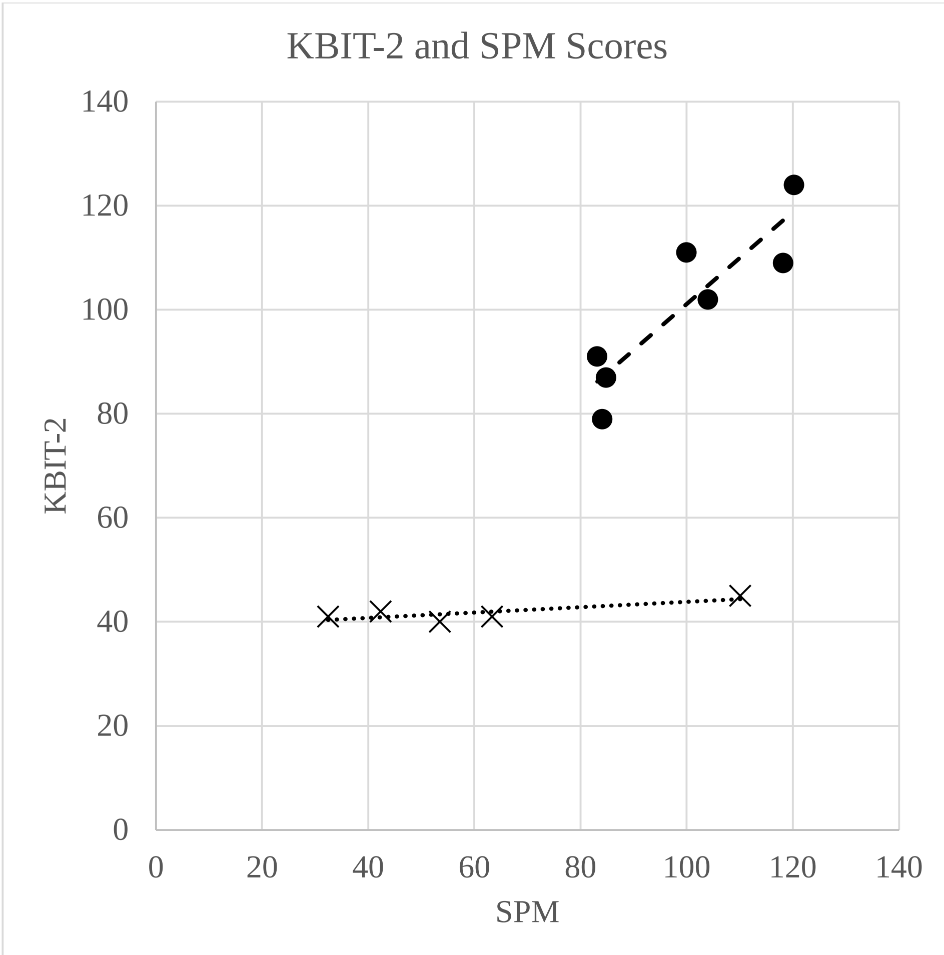

Figure 3. *Scatter Plot of SPM-C and KBIT-2 Scores for both Higher-FSIQ and Lower-FSIQ Groups*. Circles represent the higher-FSIQ group and exes represent the lower-FSIQ group. The formula for the line of best fit for the higher-IQ group is $y = .87x + 27.43$. The formula for the line of best fit for the lower-IQ group is $y = .06x + 38.42$.

Chapter Five: Discussion

The purpose of the current study current study was to determine what type of assessment best measures cognitive skills for individuals with ASD and best predicts their academic achievement. Further, this study aimed to determine if any of the three types of assessments administered (i.e., verbal, nonverbal, and computerized) was inadvertently measuring constructs other than intelligence for this population. Overall, the results indicated that no significant difference exists between scores obtained from each IQ test or in each test's ability to predict either math or reading achievement. However, the results suggested that the level of a student's communication skills is a statistically significant component of what is measured by the IQ scores obtained from a verbal test, but not for those scores obtained from nonverbal or computerized measures.

Hypothesis 1

Descriptive statistics indicated that SPM-C scores were the lowest and SPM scores were the highest; however, the ANOVA results found no statistically significant difference between the IQ scores obtained by the KBIT-2, the SPM, and the SPM-C. All of these scores were correlated at a statistically significant level ($p < .05$). These results indicate that the first hypothesis of the current study, which posited that a statistically significant difference between scores obtained on different measures of cognitive abilities would be found, is not supported. All three tests, when administered to a sample of students with ASD, provided similar estimates of cognitive ability. Further, the high level of correlation among these scores suggests that they are likely measuring similar constructs within the population of the current study.

These results are similar to those found by Minshew et al. (2005) in which 90% of the 215 children and adults assessed had no significant difference in scores on verbal and nonverbal indexes. These findings also align with prior work by Charman, Pickles et al. (2011) and Ozonoff et al. (2005), both of whom found that for individuals with autism, verbal and nonverbal scores are generally within twelve points of each other. The results of the current study provide further evidence that a difference between the verbal and nonverbal cognitive skills does not generally exist for individuals with ASD.

The results of the current study, which indicate that no statistically significant difference exists between scores from different IQ tests, should be interpreted with caution as they may be due to the small sample size that was obtained. The difference of approximately eight points between the SPM and SPM-C scores and of approximately seven points between the SPM and KBIT-2 scores may have been statistically significant given a sample size larger than the twelve participants in the current study. More research is needed to determine the significance of any differences in scores obtained by each type of test with this population.

Further, when the data was divided into higher-IQ and lower-IQ groups, the results of the visual analysis of the scatter plots of IQ scores and comparison of the slopes indicated that scores may be measuring different constructs based on the ability level of the participant. When comparing the KBIT-2 scores to the scores obtained by the two nonverbal measures, the slopes indicated that with the higher-IQ group the scores on the measures co-varied very closely; however, the slopes of the lower-IQ groups indicated that although scores on the nonverbal measures varied greatly, suggesting a wide range of ability levels, the score on the KBIT-2 remained at a similarly low level. Thus, this data is suggestive that for individuals on the autism spectrum who have IQs measured to be below 70 on standard intelligence tests, cognitive abilities may exist that are not being accurately captured by these scores. As no statistical analysis of this trend was completed, further research with a larger sample size is needed to adequately test this hypothesis and do further comparison between scores of students at different ability levels.

Hypothesis 2

The results indicated that all three obtained IQ scores were correlated at a statistically significant level ($p < .05$) with the BASI-V score, but only the SPM was correlated at this level with the BASI-M score. However, the differences in correlations between the BASI-M score and each IQ score were not different at a statistically significant level. Although the results indicate that SPM scores are correlated with math scores at a statistically significant level and that KBIT-2 and SPM-C scores are not, no practical significance to the difference in these correlations exists. Thus, the results indicate that all three IQ scores would be able to predict math achievement with similar levels of accuracy.

Although all three IQ score were correlated with the BASI-V score at a statistically significant level, when predicting reading scores, no statistically significant difference in IQ score intercorrelations was present. Thus, all three IQ scores have equal practical significance in their accuracy with predicting reading achievement. The results are in line with previous research (Gwaltney, 2014; Mayes & Calhoun, 2003b; 2008) that typical tests of cognitive ability can have good predictive validity for reading and math achievement; however, they fail to support the hypothesis that scores from the SPM-C would predict achievement better than SPM and KBIT-2 scores.

The results of the current study, which suggest that all three IQ tests provide scores that are equally predictive for a student’s academic achievement, may indicate a benefit for practitioners. These results suggest that either a verbal, a nonverbal, or a computerized assessment will provide scores with similar predictive utility. In such a case, a practitioner would have flexibility in choosing which assessment would be the best fit for the testing situation and the student's needs. Nonverbal assessments may be necessary for students with low communication skills or for those who are English language learners. Similarly, computerized assessments may allow for better cooperation and motivation with some students (Williams et al., 2002). For the cases in which either a nonverbal or a computerized test is utilized, the results of the current study indicate that the scores provided will likely have similar predictive utility to that of a traditional verbal cognitive assessment.

It should also be noted that previous research indicated that levels of variance accounted for by IQ scores when predicting math (56%) and reading (54%) achievement were similar to each other (Gwaltney, 2014). In the current study, the IQ scores predicted between 53% and 60% of the variance in math scores and between 69% and 88% of the variance in reading achievement, demonstrating higher predictive validity for reading than had been previously demonstrated. Gwaltney's study (2014) was conducted solely with students with HFA and the difference in results suggest that IQ scores may be more predictive of academics for individuals with ASD who are not as high-functioning, such as the participants in the current study; however, further research is needed to determine difference in predictive validity for different populations.

Hypothesis 3

Regressing IQ scores on GARS-2 scores did not account for a statistically significant percentage of variance in either SPM or SPM-C scores; however, the inclusion of GARS-2 Communication scores in the second step of the hierarchical regression after GARS-2 Social Interaction scores was statistically significant (p = .04) for KBIT-2 scores. Further, the variance accounted for in these scores by this addition ($\Delta R^2 = .37$) represents a large effect size. The results of the current study lend some support to the third hypothesis that KBIT-2 scores may be measuring constructs other than intelligence.

Communication skills may be one potential variable reflected in KBIT-2 scores, but not in SPM and SPM-C scores. This may occur because the KBIT-2 requires some verbal responses or it may also be due to the design of the test. Indeed, one KBIT-2 subtest requires oral answers -- only one word in length -- so individuals with greater communication difficulties may have found this subtest more challenging. Further, two subtests of the KBIT-2 require understanding of receptive language, which again may be a struggle for students with lower communication skills. For students with greater communication difficulties, the KBIT-2 may have been measuring skills in expressive and receptive language and not solely cognitive ability. However, the GARS-2 Communication scale generally measures pragmatic communication skills, such as use of gestures, pronouns, and eye contact, as well as measuring atypical communication behaviors, such as echolalia, atonality, and babbling. These types of pragmatic skills are not necessarily essential for completing the verbal subtests of the KBIT-2, so difficulties in this area should not necessarily affect scores. Thus, the fact that the inclusion of GARS-2 Communication scores added a statistically significant amount of predictive validity to the model predicting KBIT-2 scores may be due to the fact that this type of IQ test measures constructs it was not intended to when used with a population with ASD. Further research is needed to determine the underlying factors related to this finding.

Alternatively, the results may be due to KBIT-2 scores being predictive of communications skills instead of the inverse. The scores on all three types of IQ tests administered did not demonstrate any difference in predicting academic achievement. Although KBIT-2 scores were correlated at a higher level (r = .88) with BASI-V scores than either SPM score (r = .73) or SPM-C scores (r = .69), this difference did not reach statistical significance. Notably, the statistically significant negative correlation between BASI-V scores and GARS-2 scores (r = -.64) suggests that higher communication abilities play an important role in being able to learn language arts skills in the classroom. Thus, those who have lower skills on tests of verbal cognitive ability may find it more difficult to develop appropriate communication skills, which may lead to lower GARS-2 scores.

Further, the results of the current study may reflect the likelihood that those with lower cognitive skills also may have more trouble learning necessary adaptive skills, such as functional communication and study skills, to be successful in a classroom. Arick et al.'s (2003) research suggested as much, with students with lower levels of cognitive skills increasing language skills to a lesser degree during an intervention. Thus, the results indicating that GARS-2 Communication skills are predictive of KBIT-2 IQ scores and significantly correlated with BASI-V scores may be due to the overall lower level of cognitive ability in the sampled participants. This situation would lead to both low IQ scores obtained on the KBIT-2 (M = 76.00) and more difficulties in developing communication skills. Low cognitive ability and low communication skills also likely lead to more difficulty in learning language arts skills and lower achievement in reading, as suggested by the correlation between GARS-2 Communication scores and BASI-V scores in the current study.

The results indicating the connection between communication skills and academic learning suggest that a focus on developing communication skills is important for students with autism. It may be that implementing early interventions programs focusing on the development of pragmatic language skills may eventually lead to more positive academic outcomes for this population. The results of the current study emphasize the importance of communication interventions for students with ASD, though more research as to the effects of this type of intervention is needed.

Limitations

Several limitations to the current study can be identified. Primarily, this study was designed to pilot the hypothesis that computer-based cognitive tests had greater predictive validity than standard assessments or nonverbal assessments. Based on the a priori power analysis, a sample size of 14 would be needed to find statistical significance between groups. The current sample size of 12 is not a large enough sample to determine statistical significance based on this power analysis.

Further, in an attempt to utilize recent SOL scores as an outcome measure, the grade level of participants was restricted to Grades 3 to 10. Thus, the results of the current study can only be generalized to students in those grades and of a similar age range. However, data on SOL scores was unavailable for a majority of participants for various reasons. For example, some students were determined by their schools to be too low functioning to take the SOL tests or were administered alternative assessments. Future studies may benefit from gathering this data and using it for an achievement measure, as it may better assess the knowledge a student is taught in a classroom throughout the school year.

The study also excluded participants that were identified as ID and included only those with an educational diagnosis of autism. As noted previously, it is unclear what percentage of the population that has an autism spectrum disorder also has ID; however, it can be acknowledged that these disorders co-occur to some extent. Thus, excluding students with ID may have omitted a potentially substantial portion of the childhood population diagnosed with autism. It is unknown whether the effects found in the current study would generalize to these excluded individuals and future research is needed to determine this. Additionally, requiring an educational diagnosis of autism likely excluded some of the higher functioning students on the autism spectrum. An educational diagnosis of autism generally requires both a medical diagnosis, as well as an educational need for special education services. Many students with HFA may have the medical diagnosis, but may perform adequately in a regular education setting; thus, they may not receive an educational diagnosis of autism because they do not need specialized instruction. Requiring this educational classification likely excluded many students with HFA, which limits the generalizability of the results in this population.

Another limitation of the current study is the use of the computer-based BASI scores to estimate academic achievement. As the test is administered using a computer, students who are less familiar with technology are at a disadvantage (Ozonoff, 1995; Schulenberg & Yutrzenka, 2004). These scores may have reflected difficulty using a computer as opposed to true academic achievement. In addition, the BASI is a survey of academic skills that does not accurately differentiate between specific academic areas. For example, a student may have strong skills in vocabulary, but weaker skills in reading comprehension; however, he or she only receives one over score for verbal skills on the BASI. Thus, this score may not accurately reflect the full range of academic skills, which limits the results of the study.

Further, new challenges when utilizing computerized assessments may have influenced the results. During the current study, each student appeared comfortable and interested in using the computer; however, it is plausible that for certain participants the sensory components of the test, such as the noise made when clicking the computer buttons or the feel of the keyboard, may have been a distraction. Students with autism are more likely to be sensitive to sensory stimulation, which can make using a computer motivating to use, but does not guarantee an individual will be focused on the assigned task. The specific computer used, including hardware and software, or the child's familiarity with technology, may make the experience of a computerized assessment different for each child. His or her reaction to this experience may introduce extraneous variables into the data and negatively affect the results.

Additionally, the use of the Raven's SPM and KBIT-2 limit the validity of the cognitive ability scores obtained for participants and the effects found during the study. The norms for the Raven's SPM with a US sample are from 1987. Norms from this year are likely to overestimate cognitive ability due to the Flynn effect (Flynn, 2007), leading to the standard that tests be re-normed at least every ten years (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999). Obtaining scores on the Raven's SPM and KBIT-2, which has been more recently normed, may not have resulted in an equal comparison, and may have overestimated the beneficial effects of using the Raven's SPM with a population on the autism spectrum. Further, the KBIT-2 is an abbreviated measure of cognitive ability. The reliability of the measure, although high, is not adequate for individual classification. The lower reliability scores limit the certainty of the effect sizes obtained in this study.

The KBIT-2 purports to measure both fluid and crystallized intelligence, whereas the SPM and SPM-C only measure fluid intelligence. Recent analysis has suggested that most research has not clearly revealed a difference between verbal and nonverbal IQ scores (Ozonoff et al., 2005), especially with older children and adolescents (Mayes & Calhoun, 2003b), suggesting that most students with autism do not have large differences between their fluid and crystallized intelligence abilities. However, it is possible that scores found in the current study may be due to ability differences in these two types of cognitive skills and rather than differences in how cognitive skills were measured.

Further, differences in the types of matrix tests in the three types of cognitive test used in this study may have affected the skills measured by each assessment. The KBIT-2 includes items that require visual processing and items that require fluid reasoning skills. For example, individuals must identify which image is missing from a larger design and identify objects that go together conceptually (i.e., a pillow goes with a bed). Both the SPM and SPM-C include items only requiring visual processing (i.e., examinees must find the missing part of a picture). While all tasks measure nonverbal cognitive ability, the tasks may measure a different set of specific skills, which may limit the comparability of the scores.

Another limitation of the current study is that most participating families opted to complete the testing sessions in their homes. While doing so facilitated participation and eliminated the travel burden for both parents and children, it also made the testing environments variable for each participant. At times, various factors, such as noise from other parts of the home or distracting objects in the room may have affected the focus and motivation of the participants. The difference in environment introduced extraneous variables into the data, which may have affected the results.

Furthermore, maintaining the precise protocol of limiting verbal and social interactions with the participating children when they were completing the SPM-C was not always possible. During some testing sessions, certain participants needed more encouragement or direction than did others in order for the participant to complete all tests. Whether these “extra” interactions with the examiner positively or negatively affected their scores, either by providing more motivation or by distracting them with an extraneous social interaction is unknown. As with the factors of the home environment, it is likely that this extraneous variable affected the scores on the tests and, in turn, limited the overall validity of the results of the study.

Future Directions

The current study was designed to determine how different types of tests measure cognitive skills and predict achievement for individuals with autism. Future research must be done with a larger sample to confirm the effects identified and determine the statistical and practical significance. Further, employing a more representative sample, one that is reflective of current national, regional, or state census data, would allow the results of the study to be generalized to a larger subset of the population. Given the relative homogeneity of the sample in terms of race, ethnicity, and parental education, future research should include students in lower SES levels and those of minority backgrounds to determine if the effects hold for these populations as well.

Secondly, it would be valuable to replicate this work using measures with stronger technical specifications in future studies on this topic. For example, the Raven's SPM norms have not been updated recently, which may mean that they overestimate ability due to the Flynn effect (Flynn, 2007). A more appropriate measure may be the Naglieri Nonverbal Assessment Test, Second Edition (NNAT-2; Naglieri, 2007), which has been designed for administration to children on the computer and has been normed within the past ten years. The use of this test in the current study was considered; however, at the time of the study the NNAT-2 was unavailable for individual testing and only available to be administered through school wide testing procedures.

Further, instead of using the abbreviated KBIT-2 measure, a full-scale cognitive test could be administered to obtain a more accurate assessment of cognitive ability. Similarly, in future studies the BASI, which is a short survey or general math and reading skills, could be replaced with a more comprehensive test of academic performance. Gathering data on state-level assessments would also be beneficial, as these measure skills taught directly in the classroom throughout the academic year. Further, the GARS-2 may not be the most effective measure of the intensity of behaviors associated with autism because it was normed on a sample of students who were already diagnosed on the autism spectrum. Because of this fact, some atypical behaviors may be scored in a typically functioning range due to being compared to a sample in which such behaviors are the norm. A measure such as the Autism Spectrum Rating Scale (Goldstein & Naglieri, 2003) may have more validity in measuring the severity of behaviors in future studies. Addressing technical issues with the measures used in the current study will help ensure the validity of the results of future studies in this area.

A third important area of expansion for future studies would be to include participants who are in grades lower than third and higher than tenth. Younger students may be more likely to display behaviors disruptive to the testing session, leading to lower cognitive ability scores, which potentially underestimate their ability. This population would especially benefit from a cognitive measure that increased motivation. Increasingly, very young students are learning to use technology and may find a test that utilizes it much more engaging than they would a typical cognitive assessment. Further, it would be interesting to see if the results of the study hold true for older students, who may have more developed social skills, so may be less distracted by social interaction with an examiner or less motivated by computer use. In order to determine if the effects of this study hold for these age groups, further research needs to be done with these populations.

Expanding research to include students who may have been diagnosed with Intellectual Disability (ID) will be important. These students especially are at risk of having their abilities underestimated as a result of typical intelligence test scores. For example, students with ASD who score in the ID range may include those who display the most disruptive behaviors during the testing process or experience the most difficulty interacting with the examiner in the testing situation. Thus, it is likely that if a manner of testing could be found to circumvent these difficulties, the biggest gains would be found with the population, leading to larger effect sizes. Further, trends in the data from the current study suggest that the difference in constructs measured by different types of ability measures may occur only with individuals scoring below 70 on verbal intelligence tests. Future studies are needed to determine if test type is more important for students based on factors such as their level of cognitive functioning or the level of disruptive behaviors they display.

Further research should also be conducted on the overall use of computerized assessment, especially with a population on the autism spectrum. Many assessments, both academic and cognitive, are now designed to be completed by students on computers. Thus far, limited research has been done on the effects of this medium of test administration on scores and student's behaviors in this population. As mentioned previously, computers may be motivating to use, but may also represent a distraction. Certain students may be focused on the sound the computer makes or the enjoyment of clicking a button instead of the task in which they are engaged.

Previous research (Nakahachi et al., 2006; Ozonoff, 1995; Ozonoff & Strayer, 2001) has suggested that computerized assessments have the potential to be useful in minimizing or eliminating some difficulties that students with autism experience during a typical testing situation. However, as the results of the current study suggest, computerized testing may not always provide the best results for this population or any other. As the fields of education and psychology move forward with computerized testing, it is essential to ensure that such testing can provide the most accurate, valuable, and useful scores for all students.

References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing.* Washington, DC: American Educational Research Association.

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

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

American Psychological Association. (1986). *Guidelines for computer-based tests and interpretations*. Washington, DC: Author.

Anderson, D. K., Liang, J. W., & Lord, C. (2014). Predicting young adult outcome among more and less cognitively able individuals with autism spectrum disorders. *Journal of Child Psychology and Psychiatry, 55*, 485-494. doi:10.1111/jcpp.12178

Arick, J. R., Young, H. E., Falco, R. A., Loos, L. M., Krug, D. A., Gense, M. H., & Johnson, S. B. (2003). Designing an outcome study to monitor the progress of students with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities, 18*, 75-87. doi:10.1177/108835760301800201

Arthur, G. (1952). *The Arthur adaptation of the Leiter International Performance Scale.* Chicago, IL: The Psychological Service Center Press.

Bardos, A. N. (2004). *Basic Achievement Skills Inventory: Manual for the BASI survey test*. Minneapolis, MN: Pearson.

Baum, K. T. (2013). *Measurement of intelligence in children and adolescents with autism spectrum disorder: Factors affecting performance*. (doctoral dissertation, University of

Cincinnati, 2011). Retrieved from http://search.proquest.com/docview/1494002080?accountid=13158

Bayley, N. (1993). *Bayley Scales of Infant Development* (2nd ed.). San Antonio, TX: Psychological Corporation.

Bayley, N. (2005). *Bayley Scales of Infant Development* (3rd ed.). San Antonio, TX: Psychological Corporation.

Begovac, I., Begovac, B., Majić, G., & Vidović, V. (2009). Longitudinal studies of IQ stability in children with childhood autism—Literature survey. *Psychiatria Danubina, 21*, 310-319. Retrieved from http://search.proquest.com/docview/816387533?accountid=13158

Black, D. O., Wallace, G. L., Sokoloff, J. L., & Kenworthy, L. (2009). Brief report: IQ split predicts social symptoms and communication abilities in high-functioning children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 39*, 1613-1619. doi:10.1007/s10803-009-0795-3

Blair, C. (2006). How similar are fluid cognition and general intelligence? A developmental neuroscience perspective on fluid cognition as an aspect of human cognitive ability. *Behavioral and Brain Sciences, 29,* 109–125.

Bölte, 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

Bracken, B. A., & McCallum, R. S. (1998). *Universal Nonverbal Intelligence Test*. Itasca, IL: Riverside.

Brown, L., Sherbenou, R. J., & Johnsen, S. K. (1990). *Manual for the Test of Nonverbal Intelligence* (2nd ed.). Austin, TX: PRO-ED.

Brown, L., Sherbenou, R. J., & Johnsen, S. K. (1997). *Manual for the Test of Nonverbal Intelligence* (3rd ed.). Austin, TX: PRO-ED.

Charman, T., Jones, C. R. G., Pickles, A., Simonoff, E., Baird, G., & Happé, F. (2011). Defining the cognitive phenotype of autism. *Brain Research, 1380*, 10-21. doi:10.1016/j.brainres.2010.10.075

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

Chiang, H., Tsai, L. Y., Cheung, Y. K., Brown, A., & Li, H. (2014). A meta-analysis of differences in IQ profiles between individuals with Asperger's disorder and high-functioning autism. *Journal of Autism and Developmental Disorders, 44*, 1577-1596. doi:10.1007/s10803-013-2025-2

Coolican, J., Bryson, S. E., & Zwaigenbaum, L. (2008). Brief Report: Data on the Stanford–Binet Intelligence Scales (5th Edition) in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders,* 38, 190–197. doi:10.1007/s10803-007-0368-2

Dawson, M., Soulières, 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 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,* 263-271. doi:10.1007/s10802-006-9033-z

Delmolino, L. M. (2006). Brief report: Use of DQ for estimating cognitive ability in young children with autism. *Journal of Autism and Developmental Disorders, 36,* 959–963.

doi:10.1007/s10803-006-0133-y

Edelson, M. G. (2005). A car goes in the garage like a can of peas goes in the refrigerator: Do deficits in real-world knowledge affect the assessment of intelligence in individuals with autism? *Focus on Autism and Other Developmental Disabilities, 20,* 2-9. doi:10.1177/10883576050200010101

Edelson, M. G. (2006). Are the majority of children with autism mentally retarded? A systematic evaluation of the data. *Focus on Autism and Other Developmental Disabilities, 21,* 66-83. doi:10.1177/10883576060210020301

Edelson, M. G., Edelson, S. M., & Jung, S. (1998). Assessing the intelligence of individuals with autism: A cross-cultural replication of the usefulness of the TONI. *Focus on Autism and Other Developmental Disabilities, 13,* 221-227. doi:10.1177/108835769801300403

Edelson, M. G., Schubert, D. T., & Edelson, S. M. (1998). Factors predicting intelligence scores on the TONI in individuals with autism. *Focus on Autism and Other Developmental Disabilities, 13*, 17-26. doi:10.1177/108835769801300102

Elliott, C. D. (2007). *Differential Ability Scales* (2nd ed.). San Antonio, TX: Harcourt Assessment.

Geiger, D. M., Smith, D. T., & Creaghead, N. A. (2002). Parent and professional agreement on cognitive level of children with autism. *Journal of Autism and Developmental Disorders, 32,* 307-312. doi:10.1023/A:1016382819186

Gillberg, C., & Steffenburg, S. (1987). Outcome and prognostic factors in infantile autism and similar conditions: A population-based study of 46 cases followed through puberty. *Journal of Autism and Developmental Disorders, 17,* 273-287. doi:10.1007/BF01495061

Gilliam, J. E. (2006). *Gilliam Autism Rating Scale* (2nd ed.). Austin, TX: PRO-ED.

Goldstein, G., Allen, D. N., Minshew, N. J., Williams, D. L., Volkmar, F., Klin, A., & Schultz, R.

T. (2008). The structure of intelligence in children and adults with high functioning autism. *Neuropsychology, 22,* 301-312. doi:10.1037/0894-4105.22.3.301

Goldstein, S., & Naglieri, J. A. (2009). *Autism Spectrum Rating Scales (ASRS) Technical Manual.* Tonawanda, NY: Multi-Health Systems, Inc.

Gwaltney, M. V. (2014). *Effects of IQ, memory and ADHD symptoms on achievement in students with high functioning autism.* (doctoral dissertation, University of California at Davis, 2014). Retrieved from http://search.proquest.com/docview/1501839677?accountid=13158

Harris, S. L., & Handleman, J. S. (2000). Age and IQ at intake as predictors of placement for young children with autism: A four- to six-year follow-up. *Journal of Autism and Developmental Disorders, 30*, 137-142. doi:10.1023/A:1005459606120

Harvey, V. S., & Carlson, J. F. (2003). Ethical and professional issues with computer-related technology. *School Psychology Review, 32*, 92-107.

Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G. G., & Curtiss, G. (1993). *Wisconsin Card Sorting Test manual: Revised and expanded.* Odessa, FL: Psychological Assessment Resources.

Howlin, P., Goode, S., Hutton, J., & Rutter, M. (2004). Adult outcome for children with autism. *Journal of Child Psychology and Psychiatry, 45,* 212-229. doi:10.1111/j.1469-7610.2004.00215.x

Joint Committee on Testing Practices. (2004). *Code of fair testing practices in education* (pp. 1065-1067). Washington, DC: Author.

Jones, C. R. G., Happé, F., Golden, H., Marsden, A. J. S., Tregay, J., Simonoff, E, Pickles, A., Baird, G., & Charman, T. (2009). Reading and arithmetic in adolescents with autism spectrum disorders: Peaks and dips in attainment. *Neuropsychology, 23,* 718-728. doi:10.1037/a0016360

Joseph, R. M., Tager-Flusberg, H., & Lord, C. (2002). Cognitive profiles and social-communicative functioning in children with autism spectrum disorder. *Journal of Child Psychology and Psychiatry, 43,* 807-821. doi:10.1111/1469-7610.00092

Kanai, C., Tani, M., Hashimoto, R., Yamada, T., Ota, H., Watanabe, H., . . . Kato, N. (2012). Cognitive profiles of adults with Asperger's disorder, high-functioning autism, and pervasive developmental disorder not otherwise specified based on the WAIS-III. *Research in Autism Spectrum Disorders, 6*, 58-64. doi:10.1016/j.rasd.2011.09.004

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Brief Intelligence Test* (2nd ed.). Circle Pines, MN: AGS.

Koegel, L. K., Koegel, R. L., & Smith, A. (1997). Variables related to differences in standardized test outcomes for children with autism. *Journal of Autism and Developmental Disorders, 27*, 233-243. doi:10.1023/A:1025894213424

Lord, C., & Schopler, E. (1989). The role of age at assessment, developmental level, and test in the stability of intelligence scores in young autistic children. *Journal of Autism and Developmental Disorders, 19*, 483-499. doi:10.1007/BF02212853

Lord, C., Rutter, M., & DiLavore, P. (1999). *Autism Diagnostic Observation Schedule-Generic* (ADOS-G). Torrance, CA: The Western Psychological Corporation.

Magiati, I., & Howlin, P. (2001). Monitoring the progress of preschool children with autism enrolled in early intervention programmes: Problems in cognitive assessment. *Autism, 5,* 399. doi:10.1177/1362361301005004005

Mayes, S. D., & Calhoun, S. L. (1999), Symptoms of autism in young children and correspondence with the DSM, *Infants and Young Children, 12* , 90–97.

Mayes, S. D., & Calhoun, S. L. (2003a). Ability profiles in children with autism: Influence of age and IQ. *Autism, 7*, 65-80. doi:10.1177/1362361303007001006

Mayes, S. D., & Calhoun, S. L. (2003b). 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

Mayes, S. D., & Calhoun, S. L. (2008). WISC-IV and WIAT-II profiles in children with high-functioning autism. *Journal of Autism and Developmental Disorders, 38,* 428-439. doi:10.1007/s10803-007-0410-4

Mayes, S. D., & Calhoun, S. L. (2011) Impact of IQ, age, SES, gender, and race on autistic symptoms. *Research in Autism Spectrum Disorders*, *5*, 749-757 doi:10.1016/ j.rasd.2010**.** 09.002

Minshew, N. J., Turner, C. A., & Goldstein, G. (2005). The application of short forms of the Wechsler intelligence scales in adults and children with high functioning autism. *Journal of Autism and Developmental Disorders, 35,* 45-52. doi:10.1007/s10803-004-1030-x

Morsanyi, K., & Holyoak, K. J. (2010). Analogical reasoning ability in autistic and typically developing children. *Developmental Science, 13*, 578-587. doi:10.1111/j.1467-7687.2009.00915.x

Mottron, L. (2004). Matching strategies in cognitive research with individuals with high-functioning autism: Current practices, instrument biases, and recommendations. *Journal of Autism and Developmental Disorders, 34,* 19-27. doi:10.1023/B:JADD. 0000018070.88380.83

Mullen, E.M. (1995). *Mullen Scales of Early Learning*. Circle Pines, MN: American Guidance

Service

Nader, A., Courchesne, V., Dawson, M., & Soulières, I. (October 12, 2014). Does WISC-IV underestimate the intelligence of autistic children? *Journal of Autism and Developmental Disorders,* 1-8. doi:10.1007/s10803-014-2270-z

Naglieri, J. A. (2007). *Naglieri Nonverbal Ability Test* (2nd ed.). San Antonio, TX: The Psychological Corporation.

Nakahachi, T., Iwase, M., Takahashi, H., Honaga, E., Sekiyama, R., Ukai, S.,… Takeda, M. (2006). Discrepancy of performance among working memory-related tasks in autism spectrum disorders was caused by task characteristics, apart from working memory, which could interfere with task execution. *Psychiatry and Clinical Neurosciences, 60*, 312-318. doi:10.1111/j.1440-1819.2006.01507.x

Ozonoff, S. (1995). Reliability and validity of the Wisconsin Card Sorting Test in studies of autism. *Neuropsychology, 9*, 491-500. doi:10.1037/0894-4105.9.4.491

Ozonoff, S., Goodlin-Jones, B. L., & Solomon, M. (2005). Evidence-based assessment of autism spectrum disorders in children and adolescents. *Journal of Clinical Child and Adolescent Psychology, 34*, 523–540. doi:10.1207/s15374424jccp3403_8

Ozonoff, S., & Strayer, D. L. (2001). Further evidence of intact working memory in autism. *Journal of Autism and Developmental Disorders, 31,* 257-263. doi:10.1023/A:1010794902139

Ozonoff, S., South, M., & Miller, J. N. (2000). DSM-IV-defined Asperger syndrome: Cognitive, behavioral and early history differentiation from high-functioning autism. *Autism, 4,* 29-46. doi:10.1177/1362361300004001003

Poon Mak, S. M. (2012). *Adult outcome of children with autism with normal intelligence*.

(doctoral dissertation, The Chinese University of Hong Kong, 2009). Retrieved from http://search.proquest.com/docview/1025845030?accountid=13158

Psychological Corporation. (1992). *Wechsler Individual Achievement Test.* San Antonio, TX: Psychological Corporation.

Raven, J., Raven, J. C., & Court, J. H. (2003). *Manual for Raven's Progressive Matrices and Vocabulary scales. Section 1: General overview*. San Antonio, TX: Harcourt Assessment.

Reitan, R. (1956) *Trail Making Test: Manual for Administration, Scoring and Interpretation.* Bloomington, IN: Indiana University

Roid, G. H. (2003). *Stanford Binet Intelligence Scales* (5th ed.). Itasca, IL: Riverside Publishing.

Roid, G., & Miller, L. (1997). *Leiter International Test of Intelligence– Revised.* Chicago, IL: Stoelting.

Sattler, J. M. (2008). *Assessment of Children: Cognitive Foundations* (5th ed.). San Diego, CA: Author.

Scattone, D., Raggio, D. J., & May, W. (2012). Brief report: Concurrent validity of the Leiter-R and KBIT-2 scales of nonverbal intelligence for children with autism and language impairments. *Journal of Autism and Developmental Disorders, 42*, 2486-2490. doi:10.1007/s10803-012-1495-y

Schopler, E., Reichler, R. J., Bashford, A., Lansing, M. D., & Marcus, L. (1990). *Psychoeducational Profile Revised, Volume 1* (PEP-R). Austin, TX: Pro-Ed.

Schulenberg, S. E., & Yutrzenka, B. A. (2004). Ethical issues in the use of computerized assessment. *Computers in Human Behavior, 20*, 477-490. doi:10.1016/j.chb.2003.10.006

Sheinkopf, S. J., & Siegel, B. (1998). Home-based behavioral treatment of young children with autism. *Journal of Autism and Developmental Disorders, 28*, 15-23.

Strauss, E., Spreen, O., & Hunter, M. (2000). Implications of test revisions for research. *Psychological Assessment*, *12*, 237-244. doi:10.1037/1040-3590.12.3.237

10.1037//1040-3590.12.3.237

Stutsman, R. (1948). *Merrill-Palmer Scale of Mental Tests*. Los Angeles, CA: Western Psychological Services.

Takeda, T., Koyama, T., & Kurita, H. (2007). Comparison of developmental/intellectual changes between autistic disorder and pervasive developmental disorder not otherwise specified in preschool years. *Psychiatry and Clinical Neurosciences*, *61*, 684-686. doi:10.1111/j.1440-1819.2007.01740.x

Thorndike, R. L., Hagen, E. P., & Sattler, J. M. (1986). *Stanford-Binet Intelligence Scale* (4th ed.). Itasca, IL: Riverside Publishing.

VanVoorhis, C. R. W. & Morgan, B. L. (2007). Understanding power and using rules of thumb to determine sample sizes. *Tutorials in Quantitative Methods for Psychology, 3,* 43-50.

Virginia Department of Education. (2008). *Virginia standards of learning assessments technical report: 2008 – 2009 Administration cycle*. Richmond, VA: Author.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale-Revised*. Oxford, England: Psychological Corporation.

Wechsler, D. (1991). *Wechsler Intelligence Scale for Children manual* (3rd ed.). New York, NY: The Psychological Corporation.

Wechsler, D. (2003). *Wechsler Intelligence Scale for Children manual* (4th ed.). Minneapolis, MN: Pearson.

Wallace, G. L., Anderson, M., & Happé, F. (2009). Brief report: Information processing speed is intact in autism but not correlated with measured intelligence. *Journal of Autism and*

Developmental Disorders, 39, 809–814. doi:10.1007/s10803-008-0684-1

Williams, J. E., & McCord, D. M. (2006). Equivalence of standard and computerized versions of the Raven Progressive Matrices Test. *Computers in Human Behavior, 22*, 791-800. doi:10.1016/j.chb.2004.03.005

Williams, C., Wright, B., Callaghan, G., & Coughlan, B. (2002). Do children with autism learn to read more readily by computer assisted instruction or traditional book methods?: A pilot study. *Autism, 6*, 71-91. doi:10.1177/1362361302006001006

Woodcock, R. W., & Johnson, M. B. (1989). *Woodcock-Johnson Psycho-educational Battery* (Rev. ed.). Allen, TX: DLM Teaching Resources.

Appendix

Glossary

ADOS-G - Autism Diagnostic Observation Schedule-General

APA - American Psychiatric Association

AR - Analogic Reasoning

AS - Asperger’s Syndrome

ASD - Autism Spectrum Disorder

ATMT - Advanced Trail Marking Test

BASI-M - Basic Skill Achievement Inventory, Math assessment

BASI-V - Basic Skill Achievement Inventory, Verbal assessment

Bayley-2 - Bayley Scales of Infant and Toddler Development–Second Edition

Bayley-III - Bayley Scales of Infant and Toddler Development, Third Edition

CASD - Checklist for Autism Spectrum Disorder

CDD - Childhood Disintegrative Disorder

DAS - Differential Abilities Scale

DQ - Developmental Quotient

FSIQ - Full Scale IQ

GARS-2 - Gilliam Autism Rating Scale- Second Edition

HFA - high-functioning autism

KBIT-2 - Kaufman Brief Intelligence Test, Second Edition

Leiter - Leiter International Performance Scale

Leiter-R - Leiter International Performance Scales–Revised

LFA - low-functioning autism

M-P - Merrill-Palmer Scale

Mullen - Mullen Scales of Early Learning

NNAT-2 - Naglieri Nonverbal Ability Test, Second Edition

PDD-NOS - Pervasive Developmental Disorder-Not Otherwise Specified

PEP-R - Psychoeducational Profile-Revised

PIQ - Performance IQ

PRI - Perceptual Reasoning Index

PSI - Processing Speed Index

SAT - Stanford Achievement Test

SB-IV - Stanford-Binet-Fourth Edition

SB-V - Stanford Binet-Fifth Edition

SPM - Raven’s Standard Progressive Matrices

SPM-C - Raven’s Standard Progressive Matrices (Computerized version)

TONI-2 - Test of Nonverbal Intelligence: Second Edition

TONI-3 - Test of Nonverbal Intelligence-Third Edition

UNIT - Universal Nonverbal Intelligence Test

VIQ - Verbal IQ

VCI - Verbal Comprehension Index

WAIS-R - Wechsler Adult Intelligence Scale-Revised

WCST - Wisconsin Card-Sorting Task

WIAT - Wechsler Individual Achievement Test

WISC-III - Wechsler Intelligence Scale for Children–Third Edition

WISC-IV - Wechsler Intelligence Scale for Children–Fourth Edition

WJ-R: Ach - Woodcock-Johnson Tests of Achievement-Revised

WMI - Working Memory Index

EMILY H. PERLMAN

11132 Forest Edge Dr., Reston VA 20190 | emilyhperlman@gmail.com | (908) 963-7774

EDUCATION

Ph.D. in School Psychology Anticipated graduation: August 2016
The Pennsylvania State University, University Park, PA
NASP-approved, APA-accredited program
Dissertation Title: *The Use of Computerized Cognitive Assessment with Children with Autism*

M.Ed. in School Psychology May 2010
The Pennsylvania State University, University Park, PA
NASP-approved, APA-accredited program
Thesis Title: *Reliability and Factor Structure of Scores from Mexican Children on a Spanish Translation of the Wechsler Abbreviated Intelligence Scale (WASI)*

B.S. in Psychology, minor in Criminal Justice May 2005
University of Richmond, Richmond, VA
Graduated magna cum laude, with honors

CREDENTIALS AND CERTIFICATES

Licensed School Psychologist
Virginia Board of Psychology
April 2013 – April 2018

Education Specialist I, School Psychologist K–12
Pennsylvania Department of Education
April 2012 – April 2018

PROFESSIONAL ASSOCIATIONS: MEMBERSHIP

Fairfax Association of School Psychologists *August 2013 - Present*
American Psychological Association, Division 16 Member *February 2009 – Present*
National Association of School Psychologists *October 2008 – Present*
Association of School Psychologists of Pennsylvania, Student Member *September 2008 – August 2012*

CINICAL EXPERIENCE

School Psychologist, Fairfax County Public Schools *August 2013 – Present*
Fairfax County, Reston, VA

- Conducted evaluations of students to determine eligibility for special education services in two elementary schools.
- Served as a member of Local Screening Committee, Positive Behavioral Support, and Responsive Instruction committees, generating academic and behavioral interventions to enable student learning.
- Provided one-on-one interventions and facilitated group counseling sessions for both special and regular education students from Kindergarten to sixth grade.
- Consulted with teachers and administrators about behavior and academic issues of specific students

PUBLICATIONS

Bellinger, J. M., Perlman, E. H., & DiPerna, J. C. (Winter, 2011). Social Skills Interventions for Individuals with Autism Spectrum Disorder. *School Psychology Forum: Research in Practice, 5, 141-159.*