Comparison of the *Leiter International Performance Scale-Revised* and the *Stanford-Binet Intelligence Scales, 5th Edition* in Children with Autism Spectrum Disorders

THESIS

Presented in Partial Fulfillment of the Requirements for the Degree Master of Arts in the
Graduate School of The Ohio State University

By

Sabrina Nicole Grondhuis

Graduate Program in Psychology

The Ohio State University

2010

Master's Examination Committee:

James Mulick, Advisor

Michael Aman

Luc Lecavalier

Copyright by
Sabrina Nicole Grondhuis
2010

Abstract

Both the *Leiter International Performance Scale-Revised (Leiter-R)* and the *Stanford-Binet Intelligence Scales, 5th Edition (SB5)* are standardized intelligence tests that can be used with a variety of special populations. The *Leiter-R* is a nonverbal assessment shown to be useful in evaluating clients with restricted language capabilities. Although the *Leiter-R* and more traditional *SB5* are frequently used to assess intellectual ability in children with autism spectrum disorders (ASDs), data about the relationship between both measures are not available for the most recent editions of the tools. A retrospective chart review of patient files from a major children's hospital was conducted, and data from 1,071 patients at risk of an intellectual or developmental disability were obtained. From that sample, analyses were conducted on 50 children with both the *Leiter-R* and the *SB5*. All participants received an ASD diagnosis and were between the ages of three and twelve years. The final sample contained 41 males (82%) and 9 females (18%), similar to the ratio found in the general population for ASDs. Age ranged from 36 to 131 months ($M = 66.7$, $SD = 24.1$). ASD diagnoses were AD ($n = 26$; 52%) and non-AD ($n = 24$; 48%). Mixed method ANOVAs and *t*-tests found that this population received significantly higher scores on the *Leiter-R* than the *SB5* when evaluated based on the full sample. Analyses also showed a significant effect for diagnostic category, level of functioning, and age of the child. No differences were found between the two genders.

These analyses strongly suggest that the *Leiter-R* and the *SB5* may not be equivalent measures of intellectual functioning in children with ASDs. Because these children regularly scored between one and two standard deviations higher on the *Leiter-R* than *SB5*, it seems prudent to assume that a nonverbal measure like the *Leiter-R* may be an appropriate assessment for estimating the upper bounds for an ASD population.

Alternatively, they may also suggest that the *Leiter-R*, by virtue of being a nonverbal intelligence test, assesses areas known to be strengths for those with ASDs. These findings are in line with previous work that utilized earlier editions of the measures.

This document is dedicated to my parents.

Acknowledgments

I wish to thank my committee for all of their thoughtful suggestions and guidance throughout the course of this project.

I would also like to thank the staff of the Nationwide Children's Hospital Child Development Center for their patience and enthusiasm for the project, as well as the patients that were evaluated there.

I thank Cristan Farmer for her valuable insight into the world of academia and her recommendations for this thesis. Lastly, I would like to thank Erin Kellogg for being a sounding board and cheerleader for progress; the hours spent on walks talking through complex ideas helped add clarity and clarification where needed.

Vita

January 22, 1986Born, Buffalo Grove, IL
2008.....B.A Psychology, Univ of California, Irvine
2008-2009University Fellow
2009 to presentGraduate Teaching Associate, Department
of Psychology, The Ohio State University

Fields of Study

Major Field: Psychology

Table of Contents

Comparison of the <i>Leiter International Performance Scale-Revised</i> and the <i>Stanford-Binet Intelligence Scales, 5th Edition</i> in Children with Autism Spectrum Disorders	1
THESIS	1
Abstract	ii
Acknowledgments	v
Vita	vi
Fields of Study	vi
Table of Contents	vii
List of Tables	viii
Chapter 1: Introduction	1
Chapter 2: Methods	14
Chapter 3: Results	16
Chapter 4: Discussion	20
References	35

List of Tables

Table 1. Sample Characteristics.....	28
Table 2. Summary of <i>t</i> -test Results on Leiter-R and SB5 Comparisons.....	29
Table 3. Analysis of Variance for the Effect of Diagnostic Category and Test.....	30
Table 4. Analysis of Variance for the Effect of Age Group and Test.....	31
Table 5. Analysis of Variance for the Effect of Level of Functioning on the <i>Leiter-R</i> and Test.....	32
Table 6. Analysis of Variance for the Effect of Level of Functioning on the <i>SB5</i> and Test.....	33
Table 7. Analysis of Variance for the Effect of Gender and Test.....	34

Chapter 1: Introduction

Autism Spectrum Disorders (ASDs) are characterized by deficits in both social interaction and communication, as well as by restricted and repetitive patterns of behavior, and are a type of Pervasive Developmental Disorder (PDD) (American Psychiatric Association, 2000). The conditions included under the DSM-IV PDD umbrella include Autistic Disorder (AD), Asperger's Disorder, Pervasive Developmental Disorder – Not Otherwise Specified (PDD-NOS), Rett's Disorder, and Childhood Disintegrative Disorder (American Psychiatric Association, 2000). All five of these disorders share the three core deficit areas, although diagnoses differ based on the severity of impairment and the symptom presentation (Bryson & Smith, 1998). Children meet criteria for an AD diagnosis when they possess (a) at least two different impairments in social interaction, (b) at least one impairment in communication, and (c) at least one restricted, repetitive, or stereotyped behavior, interest, or activity. Additionally, there must be evidence that the onset of symptoms was at, or before, three years of age.

Asperger's Disorder has similar diagnostic criteria, although communication delays are exclusionary. The criteria for PDD-NOS include the impairments in social interaction, but the diagnosis requires either impairments in communication OR restricted, repetitive, or stereotyped behavior, interest, or activities (American Psychiatric Association, 2000). These three disorders (AD, Asperger's Disorder, and PDD-NOS)

make up the ASDs. Childhood Disintegrative Disorder and Rett's Disorder are less frequently diagnosed and are considered by some to be distinctive disorders due to their respective etiologies (Fombonne, 2003). Consequently, these latter disorders will be omitted from the remainder of this discussion.

Fombonne (2007), a respected epidemiologist, published a review of the recent epidemiology literature and found that, although the prevalence of ASDs has been on the rise for the past three decades, the incidence, or the number of new cases of the disorder, has not changed substantially. Current data indicate that ASDs occur in 60 out of 10,000 people (1 in 167; Fombonne, 2005), although other respected researchers believe that the figure might, in fact, be considerably lower (e.g., Gillberg & Wing, 1999). Substantial debate surrounds the question of whether the incidence of ASDs is on the rise or if the increase is a result of relaxed diagnostic criteria or inclusion of other medical or neuropsychological disorders within the ASD category. Many in the field agree that the rise is a function of relaxed diagnostic criteria (Fombonne, 2005).

Intellectual disability (ID) is characterized by deficits in adaptive functioning, an intelligence quotient (IQ) that falls at least two standard deviations below average (i.e., below 70), and onset prior to adulthood (American Psychiatric Association, 2000). Although ID is not a core symptom of ASD, there is an established relationship between the two constructs. Seventy percent of children with AD have a co-morbid diagnosis of ID (Bolte & Poustka, 2002; Fombonne, 2005). It has been suggested that many children who would have been diagnosed with ID in the past would today receive an ASD diagnosis, even with identical symptomatology (Lecavalier et al., in press).

Although ID is common in the ASD population, measuring intelligence in those with both ID and ASD is uniquely challenging. Children with these diagnoses frequently have behavioral difficulties (Matson & Shoemaker, 2009), diminished attention spans (Matson, Dempsey, LoVullo, & Wilkins, 2008), and significant language limitations. Each of these may impair performance on IQ tests and on the child's overall ability to learn (Tsatsanis, Dartnall, Cicchetti, Sparrow, Klin & Volkmar, 2003).

Although some controversy surrounds the specific meaning of Ψ intelligence, Φ Ψ cognitive ability, Φ and Ψ cognitive functioning, Φ these terms will be used interchangeably throughout this thesis. An intelligence test is usually given because a clinician is interested in global cognitive functioning (Halperin & McKay, 1998). IQ then, is an estimate of intellectual ability based on normed-referenced responses from similarly-aged individuals (Lilienfeld, Lynn, Namy, & Woolf, 2006).

Ways to Measure Intelligence

Instruments for assessing cognitive functioning are valuable tools at the very heart of the assessment process for clinicians who seek to understand individual children with ASD. IQ tests are part of the foundation for most psychological assessments. There are many tests from which to choose. Traditionally-structured tests like the *Stanford-Binet Intelligence Scale, 5th Edition* (SB5; Roid, 2003) or the *Wechsler Intelligence Scale for Children, 4th Edition* (WISC; Wechsler, 2003) are the most common. This is likely due to their systematic and empirical construction and to renorming nearly every decade of late (obtaining norms is a prohibitively expensive exercise). These tests require clinicians to administer a series of questions or tasks that usually are arranged in order of increasing

difficulty within subtests. Items are presented verbally, and answer formats are verbal or involve the completion of an activity (e.g., block design). Some nonverbal intelligence tests like the *Leiter International Performance Scale – Revised* (*Leiter-R*; Roid & Miller, 1997), give directions through pantomime and require nonverbal (motor-based) responses. Some experts argue that these nonverbal tests level the playing field by allowing clinicians to obtain an estimate of a patient's true intelligence without the influence of language, which may depress an individual score if he or she has a language disability (Tsatsanis et al., 2003).

It is certainly true that language is an important aspect of adaptive functioning in the real world, and researchers have found that cognitive performance is highly influenced by verbal ability (Lennen, Lamb, Dunagan, & Hall, 2010). Despite the importance of language in learning and intelligence, neither language disorder nor ASDs alone are a guarantee of reduced cognitive capacity (Chakrabarti & Fombonne, 2005). Language restrictions and behaviors associated with ASDs may affect the validity of a highly-structured traditional verbally-based assessment. Therefore, a nonverbal measure of intelligence may offer avenues for understanding test items and provide a means for responding that is not dependent on verbal ability. The use of nonverbal tests in addition to traditional measures may result in a range of scores in which at least some aspects of the child's abilities may lie.

Traditional, verbally-based cognitive assessments are the undisputed gold standard in the field of intelligence testing (Sattler, 2001). IQ scores are regarded as a snapshot estimate of the individual's level of cognitive functioning at the time of the

testing (DeThorne & Schaefer, 2004). However, there are substantial subgroups in the general population who would be unable to utilize tools optimally that require verbal responses. Thus, nonverbal intelligence tests allow those with language complications to be evaluated, and clinicians can choose an assessment tool that is hoped to be more appropriate for the patient.

DeThorne and Schaefer (2004) supported the view that Nonverbal IQ is a fallacy, and that no such nonverbal ability actually exists. DeThorne and Schaefer (2004) argued that these nonverbal tests do measure cognitive abilities, just those that are expressed nonverbally. These tests use pictures and figures instead of words and numbers, but are specifically designed to assess similar cognitive abilities as verbal intelligence measures (Roid & Miller, 1997). Nonverbal IQ, in this sense, is then not synonymous with performance IQ as it was often described during the 20th century (DeThorne & Schaefer, 2004).

The *Leiter-R* manual explicitly claims that the test is evaluating the general ability to perform complex nonverbal mental manipulations related to conceptualization, inductive reasoning and visualization (Roid & Miller, 1997). An absence of verbal subtests may give nonverbal intelligence measures a restricted profile for crystallized intelligence, or accumulated knowledge of the world over time, which is typically the construct associated with academic achievement (Roid & Miller, 1997). This is unquestionably a large limitation, but it does appear that nonverbal tests tap into an individual's fluid intelligence, which is the ability to manipulate information in novel ways, and the ability to process stimuli visually. Thus, if a person is unable to process

verbal items or the psychometrician deems it inappropriate to administer a verbal IQ test, a nonverbal test may be used instead with some obvious value. The data derived from a nonverbal measure, although less complete in scope, could be helpful for making diagnostic and treatment decisions and for determining in some sense, and the range of a person's intellectual capacity.

Nonverbal measures of cognitive functioning have been used for over two decades in children with ASD in an attempt to work around at least some of the characteristic communication impediments commonly observed in ASD (Tsatsanis et al., 2003). It is important to note that the *Leiter-R* is not the only nonverbal assessment tool available. The *Ravens Progressive Matrices* (Raven, 1998) test is a nonverbal problem solving and reasoning task that includes matrix patterns that increase in difficulty. Some tests, like the *SB5*, include a separate nonverbal score that can be calculated independent of the Full Scale or Abbreviated scores. However, these tests are not truly nonverbal because they require spoken directions (Lennen et al., 2010). Studies have shown that many children with ASDs perform significantly better on nonverbal measures of intelligence than on verbal cognitive assessments. Nevertheless, it is important to investigate the relationship of nonverbal tests with verbal measures to have fuller understanding of where deficits lie and to be able to predict performance on the array of academic and practical tasks that such children will be asked to do (Banach, Thompson, Szatmari, Goldberg, Tuff, Zwaigenbaum & Mahoney, 2009; Maltz, 1981). Additionally, at least one study demonstrated that nonverbal IQs are more stable over time for an ASD population, and thus might be a better long-term predictor of cognitive ability level

(Bishop, Richler, & Lord, 2006). Interestingly, this finding was not found in individuals with typical development (Hopkins & Bracht, 1975; Strand, 2004).

Properties of Leiter International Performance Scale-Revised

The *Leiter-R* is the most current version of Russell Leiter's nonverbal intelligence scale, originally developed in 1929. It was designed to assess nonverbal cognitive ability in people with a variety of language complications. It is often used in populations with hearing and motor impairments, brain injuries, lack of English proficiency, and ASDs (Bradley-Johnson, 1998; Sparrow & Davis, 2000). The *Leiter-R* is divided into two batteries that include 10 subscales each. The Visualization and Reasoning Battery includes the following subscales: Matching, Sequential Order, Picture Context, Figure Ground, Design Analogies, Form Completion, Repeated Patterns, Classification, Paper Folding, and Figure Rotation. The Visualization and Reasoning battery is used to obtain a composite IQ. The Attention and Memory Battery subscales includes Forward Memory, Backward Memory, Attention Divided, Delayed Pairs, Associated Pairs, Immediate Recognition, Attention Sustained, Visual Coding, Spatial Memory, and Delayed Recognition. The Attention and Memory subscales are not considered in the Full Scale composite score, but they are used to evaluate deficits in attention or memory domains. Subtests chosen for administration are determined by the age of the person being evaluated. The *Leiter-R* offers normative data for people aged 2 years 9 months to 20 years 11 months based on a representative sample of 1,719 subjects (Roid, Pomplun, & Martin, 2009).

The *Leiter-R* appears to be psychometrically sound, with reliability estimates for composite and IQ scores ranging from alphas of .88 to .90 for the brief screener and .91 to .93 for the full-scale administration (Roid et al., 2009). Concurrent validity was evaluated with respect to the *Universal Nonverbal Intelligence Test (UNIT)*; Bracken & McCallum, 1998); Full Scale IQs from the two tests were moderately correlated in a typical population ($r = .72$; Hooper & Bell, 2006). Available data suggest a moderate correlation between the *Leiter* and the *WISC-III* (Wechsler, 1991) when evaluated with a bilingual population ($r = .76$; Cathers-Schiffman & Thompson, 2007).

A comparison of the *Leiter-R* and *WISC-III* was also included in the administration manual. The population consisted of 126 children who were either typically developing, had cognitive delays, were gifted, or were bilingual. Roid and Miller (1997) reported a mean *Leiter-R* score of 95.8 ($SD = 23.1$) and a mean score on the *WISC-III* that was 98.7 ($SD = 25.7$; Roid & Miller, 1997). The 2.9 point difference between the two scales seems negligible from a practical standpoint and suggests that the *Leiter-R* is equivalent to traditionally-structured verbally-presented intelligence tests. Nevertheless, this is the only comparison that I could locate to compare the absolute scores from the *Leiter-R* and another intelligence measure that included typical children. It seems possible to argue that more direct comparisons of the *Leiter-R* to traditional tests in typical populations should be a necessary precondition for standardization of the test. Multiple reviews of the *Leiter-R* have been written, and their authors have consistently concluded that the test is well established and that it is suitable for use in people who are

unable to complete traditional types of cognitive testing (Campbell, Brown, Cavanagh, Vess, & Segall, 2008; Kuschner, Bennetto, & Yost, 2007; Tsatsanis et al., 2003).

The *Leiter International Performance Scale* (Leiter, 1980) was evaluated for use with children with ASD, and was found to yield a good indication of cognitive and educational potential for children who were otherwise unable to complete traditional intelligence assessments (Shah & Holmes, 1985). The subsequent revision (*Leiter-R*) included a significant redesign of the materials and administration procedures, resulting in a substantially different test. The Leiter and *Leiter-R* are highly correlated ($r = .87$) and do not produce significantly different IQ scores. However, the sample sizes used in these analyses of consistency were small (Tsatsanis et al., 2003). Therefore, further evaluation of the *Leiter-R*, particularly against other well-established cognitive achievement tests, is warranted.

It is important to note that while the *Leiter-R* supplies normative data for a number of specialty groups, like those with intellectual or learning disabilities, no norms exist for ASD populations. The literature reports that persons with ASDs have distinctive intellectual patterns that include characteristic strengths. These strengths often appear on visual-spatial tasks. Weaknesses often appear on tasks such as verbal reasoning (Carpentieri & Morgan, 1994). Clinicians may be able to draw more specific conclusions from scores if they are able to use an assessment tool that was normed for people with the same disability. This seems especially true if people with that disability tend to perform in a predictable pattern.

Properties of *Stanford-Binet Intelligence Scales, 5th Edition*

The *SB5* is a descendent of the very first intelligence measure ever created. It tests general intellectual ability, and the Fifth Edition affords the ability to report Full Scale, Abbreviated, Nonverbal, and Verbal IQ scores (Roid, 2003). One of the updated features of the *SB5* is the nonverbal component, in that close to half of the responses no longer require verbal *answers* (Coolican, Bryson & Zwaigenbaum, 2008). The Full Scale IQ is derived from the verbal and nonverbal scales, each with five subtests bearing the same names: Fluid Reasoning, Knowledge, Quantitative Reasoning, Visual Spatial Processing, and Working Memory.

Even though the *SB5* has made advances by including a Nonverbal section, it is not universally accepted as adequate. Lennen and colleagues (2010) evaluated the *SB5* for an ASD population and used verbal ability as a covariate. They found that the score produced on the Nonverbal section was still influenced a great deal by the child's level of language. They ultimately concluded that the *SB5* Nonverbal score was YunderestimatingΦ the cognitive ability of children with AD, and that fully nonverbal measures might be able to give a more YaccurateΦ representation of IQ.

The *SB* is among the most popular intelligence measures, and it has been subject to much evaluation, even within the field of intellectual and developmental disabilities. Silverman, Mizejeski, Ryan, Zigman, Krinsky-McHale, and Urv (2010) studied the performance of adults with ID on the *Wechsler Adult Intelligence Scale-III* (Wechsler, 1997) and *SB5*. Their findings revealed that performance on the full scale *WAIS* ($M = 58.1$, $SD = 9.80$, range = 39-88) was, on average, 16.7 points higher than on the *SB5* ($M = 41.3$, $SD = 9.33$, range = 23-67), greater than one standard deviation. This study was far

from perfect, as tests were not administered at the same time point. Also, these analyses utilized an older population; the median birth year was 1947. The Fifth Edition of the *SB* is particularly well-represented in the field despite its more recent appearance. However, it has not yet been extensively studied in comparison to the *Leiter-R*.

Past Research Comparing the Leiter to Other Measures

There have been no published comparisons between the *Leiter-R* and the *SB5*, probably due to the relative newness of the *SB5*. In the past, the Wechsler tests have been more frequently used than the *SB* in comparisons with the *Leiter-R*. Nevertheless, studies comparing the two measures for a population with ASD are limited. Shah and Holmes (1985) compared the original *Leiter* and the *WISC-R* (Wechsler, 1976) in a small sample of 18 children with AD. The authors reported no significant differences in mean performance on the two Full Scale IQ scores (*Leiter* $M = 69$, $SD = 18$; *WISC-R* $M = 56$, $SD = 12$); they observed a moderate correlation between the measures ($r = .74$, 55% shared variance). The original *Leiter* and the *SB-LM* were shown to correlate significantly ($r = .77$, 59.3% shared variance) at a moderate magnitude in a sample of 60 typically-developing children (Reeve, French, & Hunter, 1983). Some investigators have observed that children can receive very different scores on the *Leiter-R* and *SB5*. In the Reeve et al. (1983) study, the group mean scores were 105.9 ($SD = 16.74$) and 106.9 ($SD = 16.68$) for the *Leiter* and *SB*, respectively. Half of the participants were within 7 points on the two intelligence measures, although some experienced a discrepancy of 34 points (once with the *SB* higher and once with it lower). This is an interesting finding in light of the fact

that this and other studies of previous versions of the *Leiter* and *SB* were highly correlated (Pearson's $r = .92$; Bessent, 1950).

While studies exist using previous editions of the *Leiter* and *SB*, the *Leiter-R* and *SB5* have not been compared in a population of children diagnosed with ASD. These two measures are very popular (Sattler, 2001) and they are frequently used for diagnostic purposes with ASD clients. Thus, a thorough comparison of child performance on the two measures seems warranted.

Objectives of Current Study

The lacuna in the literature regarding the relationship between the *Leiter-R* and the *SB5* for children with ASD needs to be addressed. In the current study, I aimed first to evaluate the correspondence of performance on the *Leiter-R* and *SBV* instruments in a sample of children with ASD. My second major goal was to investigate whether the magnitude of this relationship varied with (a) the age of the child, (b) ASD diagnosis (AD or non-AD), (c) level of functioning (i.e., IQ), and (d) gender. Intelligence measures are most stable when administered after the age of four (Lord & Schopler, 1989; Mayes & Calhoun, 2003). However, the movement towards early intervention (and test-author published normative data at a floor of two years of age) necessitates further evaluation at these younger ages. Therefore, I included younger children (three to twelve years of age) in my sample.

More children than ever before are now evaluated for ASDs, and cognitive development testing is usually a major evaluation component (Bolte & Poustka, 2002). Important decisions about education, services, and treatment plans are often constructed

based upon diagnostic assessments (Bolte & Poustka, 2002; Howlin, 1998). Therefore, it is important to understand fully the impact of ASDs on the validity of traditional cognitive measures. This is especially true if either of the tests produces an inaccurate estimate of the child's cognitive functioning.

Based on existing literature, I expected that in this sample of children with ASDs the *Leiter-R* would produce a higher IQ score than the *SB5*, regardless of diagnosis, age, sex, or level of functioning. The deficits in communication as a core diagnostic criterion simply makes it seem more likely that a nonverbal assessment measure may provide an elevated score in comparison to the verbal measure that does not compensate for language problems. Therefore, my hypothesis for this thesis was that children with ASD, in general, would perform significantly higher on the *Leiter-R* than on the *SB5*. I predicted that this would be the case for children with both AD and non-AD. I made no specific predications for age, level of functioning, and gender subgroups.

Chapter 2: Methods

Participants

The current study used an existing database from a large Midwestern children's hospital. A retrospective chart review was conducted using data from patients seen between 2004 and 2009 ($N = 1,071$) at a specialty clinic for developmental disabilities. Children were assessed with standardized tests by either trained psychometricians, advanced graduate students in a doctoral program, or post-doctoral interns. Inclusion criteria included the following: (a) a clinical ASD diagnosis, (b) Full Scale test results available for both the *Leiter-R* and *SB5*, and (c) ages between three and twelve years, inclusive. There were no exclusion criteria for this study in place to restrict participation in this study. I saw this as desirable, as it would potentially enable a full range of scores, hopefully with minimal attenuation to the ranges.

This resulted in a sample of fifty children (see Table 1 for demographic information). The Full Scale inclusion criterion was deemed important because the Abbreviated version of the *SB5* misrepresented the cognitive abilities of children with ASD in roughly 27% of cases (Coolican et al., 2008). The final sample contained 41 males (82%) and 9 females (18%), a similar gender ratio to that found in general for individuals with ASDs. Age ranged from 36 to 131 months ($M = 66.67$, $SD = 24.12$). The ethnic composition was as follows: 33 (66%) participants were Caucasian, 8 (16%) were

African American, 3 (6%) were biracial, and 1 (2%) was Asian American, with five participants (10%) declining to state an ethnic category. ASD diagnoses were AD ($n = 26$; 52%) and non-AD ($n = 24$; 48%), which comprised children who had either Asperger's Disorder ($n = 3$) or PDD-NOS ($n = 21$).

Procedure

Groups were compared for obtained IQ on both measures using paired sample t -tests. This was deemed the most appropriate method because scores on both intelligence tests came from the same individual at the same time point.

Groups were compared to one another using a mixed method analysis of variance (ANOVA). This allowed the two intelligence tests to be a within-group factor since all group members took both assessments. The between-groups factor was the additional variable under investigation (e.g., diagnostic category, age grouping, gender). These were between-groups subject factors, because a child is only able to belong to a single group at a time (e.g., a child cannot have both AD and PDD-NOS).

Chapter 3: Results

Comparisons Using The Full Sample

For the whole sample, the mean score on the *Leiter-R* was 88.20 ($SD = 26.26$, range 41 to 145) and the mean on the *SB5* was 66.42 ($SD = 21.02$, range 40 to 114). Paired samples t -tests were used to compare the *Leiter-R* and the *SB5*. The difference between these two tests was statistically significant for the full sample [$t(49) = 10.465$, $p < .001$] (see Table 2). The two tests had only a moderate correlation (Intraclass correlation coefficient, $r = .57$).

Differences Between Diagnostic Categories

A mixed method 2 (*Leiter-R/SB5*) x 2 (ASD Diagnostic Category) ANOVA, with *Leiter-R/SB5* (Test) score as the within-subjects factor and diagnostic category as the between-subjects factor, was used to examine differences between diagnostic groups. A significant main effect for Test was found [$F(1,48) = 108.91$, $p < .001$] (see Table 3) that demonstrated significantly higher scores on the *Leiter-R*. There was also a statistically significant main effect for Diagnostic Category [$F(1,48) = 7.8$, $p < .007$]. The interaction effect (Test by Diagnostic Category) was not significant [$F(1,48) = .537$, $p = .467$].

Differences within the diagnostic categories were evaluated with paired sample t -tests (see Table 2). Children with AD had a *Leiter-R* mean score of 79.42 ($SD = 24.99$) and a *SB5* mean score of 59.12 ($SD = 17.51$) [$t(25) = 7.150$, $p < .001$]. Differences were

also observed within the non-AD group. Those children had a mean score on the *Leiter-R* of 97.71 ($SD = 24.68$) and *SB5* mean score of 74.33 ($SD = 21.96$) [$t(23) = 7.58, p < .001$].

Differences Between Age Groups

The data set was divided into two age groups by approximate median split at 60 months. Data were evaluated with a mixed model 2 (*Leiter-R/SB5*) x 2 (below 60 months/60 months or above) ANOVA, with the intelligence Test scores as the within-subjects factor and the Age category as the between-subjects factor. A significant main effect of Test was found [$F(1,48) = 126.99, p < .001$] (see Table 4), which demonstrated that *Leiter-R* scores were significantly higher than *SB5*. A significant main effect was also found for Age that indicated that older children performed more poorly than the younger group [$F(1,48) = 7.79, p < .007$]. Mean scores and standard deviations for the two age groups appear in Table 2. The analyses also revealed a significant interaction between Test and Age [$F(1,48) = 8.82, p < .005$].

Differences Between Levels of Functioning

The sample was divided into high and low Levels of Functioning by Test, by splitting the groups at an IQ of 70, which is the cutoff for ID. For the *Leiter-R*, 11 children (22%) had scores below 70, while 39 (78%) had scores of 70 points or above. A 2 (*Leiter-R/SB5*) x 2 (Level of Functioning on *Leiter-R*) mixed model ANOVA was conducted. There was a statistically-significant main effect of Test, with elevated scores for the *Leiter-R* [$F(1,48) = 58.94, p < .001$](see Table 5). There also was a significant main effect of Functional Level, with the high-functioning group scoring significantly better than the lower-functioning group [$F(1,48) = 44.65, p < .001$]. Finally, there was a

significant interaction between Test and Functional Level [$F(1,48) = 17.08, p < .001$], with higher-functioning children demonstrating a larger discrepancy between the two measures than the lower functioning group (see Table 2 and 5).

When split by *SB5* score, a larger number of children stood *below* 70 IQ points ($n = 32, 64\%$), than above the 70-point cutoff ($n = 18, 36\%$). Once again, the findings were analyzed as a function of Functional Level (2 levels, low vs. high) and Test (*Leiter-R* vs. *SB5*). Main effects were significant for Test, which again showed higher scores from the *Leiter-R* than for the *SB5* [$F(1,48) = 97.57, p < .001$] (see Table 6) and for Functional Level [$F(1,48) = 75.31, p < .001$]. Naturally, the IQ split resulted in better scores within the high-functioning group. The interaction term was not significant [$F(1,48) = .07, p = .797$]. The means for *SB5*-defined functional groups appear in Table 2, while ANOVAs appear in Table 6.

Comparisons Based on Gender

A mixed method 2 (*Leiter-R/SB5*) x 2 (male/female) ANOVA was performed. As predicted, there was a significant main effect due to Test, based upon the elevated *Leiter-R* [$F(1,48) = 59.34, p < .001$] (please see Table 7). However, there was no main effect for Gender [$F(1,48) = 1.18, p = .282$] and no interaction effect [$F(1,48) = .18, p = .675$].

Paired samples *t*-tests were used to evaluate the categories individually. Boys had mean *Leiter-R* scores of 86.78 ($SD = 27.26$) and mean *SB5* scores of 64.59 ($SD = 20.68$) [$t(40) = 9.50, p < .001$]. For girls, the mean *Leiter-R* score was 94.67 ($SD = 21.24$) and the mean *SB5* score was 74.78 ($SD = 21.71$) [$t(8) = 4.19, p < .003$].

***Leiter-R* Compared to *SB5* Subsections**

The Nonverbal domain of the *SB5* was compared to the *Leiter-R* full-scale composite score. The mean *SB5* nonverbal score was 72.02 ($SD = 22.36$), whereas the mean for the *Leiter-R* was 88.20 ($SD = 26.26$; see Table 2). A paired samples *t*-test showed that the *Leiter-R* and the *SBV* Nonverbal scores were significantly different from one another [$t(49) = 7.758, p < .001$], although they were moderately correlated (Intraclass correlation coefficient, $r = .67$).

The Verbal section of the *SB5* was also compared to the *Leiter-R* total score to determine how much disadvantage resulted from a verbal-only assessment. There were only 48 sets of scores that included the *Leiter-R* and the Verbal section of the *SB5*, so the *Leiter-R* mean changed slightly to 88.38 ($SD = 26.17$). The verbal *SB5* resulted in an average score of 63.81 ($SD = 19.47$) which was significantly lower than the *Leiter-R*, [$t(47) = 10.243, p < .001$] (see Table 2). The two test scores had a moderate correlation (Intraclass correlation coefficient, $r = .47$).

Chapter 4: Discussion

Comparisons Using Full Sample

On average, scores from the *Leiter-R* were 21.78 points higher than those derived from the *SB5*. This is both highly statistically significant ($t(49) = 10.465, p < .001$) and clinically significant. This difference is greater than one standard deviation (both IQ tests have a mean of 100 and a standard deviation of 15), and could likely alter the functional level (borderline, mild, moderate, severe, etc.) in which the score falls. This may be because the *Leiter-R* does not penalize children with language deficits in the same way that the *SB5* does, which requires verbal answers. This finding is in line with my original hypothesis and with a body of evidence using earlier versions of the tests previously mentioned (i.e., Reeve et al., 1983). A difference of this magnitude would certainly prompt clinicians to make different predictions for future learning and educational success.

Differences Between Diagnostic Categories

ANOVAs demonstrated significant main effects for both the IQ tests and ASD diagnoses. When the relationship between the two tests was also evaluated within the individual diagnostic categories by paired samples *t*-tests, it revealed that both the AD and non-AD groups scored significantly higher on the *Leiter-R* than the *SB5*.

These results demonstrate the merit of using a nonverbal measure such as the *Leiter-R* when evaluating children suspected of having an ASD in general, regardless of the specific diagnosis in question. The finding that all children with ASDs, on average, received higher scores on the *Leiter-R* than *SB5* could be an indicator that the *Leiter-R* may produce a cognitive profile more indicative of their upper level of functioning. An alternate explanation is that the two tests are evaluating different cognitive constructs, and additionally that the *Leiter-R* may be tapping into aspects of intelligence that are particular strengths for the ASD population.

Differences Between Age Groups

Both the *Leiter-R* and the *SB5* Full Scale IQs were negatively correlated with age. Past research has demonstrated that youngsters with ID do not learn as quickly as their peers (Matson & Shoemaker, 2009). Thus, even if a child is maintaining a consistent level of cognitive gain, IQ scores may still decrease because learning is not occurring at the same rate as for typical children. Evidence for this phenomenon was shown in the lower average scores for the older children relative to the younger children.

The sample split at 60 months made sense because it corresponded to the age that commonly splits preschool children and children in elementary school. Further, this was an age dichotomy that has precedence in the literature (Bishop et al., 2006), and much practical application in North America. Also interesting was the significant interaction effect between age and test. The magnitude of differences between the *Leiter-R* and *SB5* depended upon age group; the younger children showed more test disparity (27.52 points) than the older children (16.04 points). Considered another way, children younger than

five years of age showed nearly two standard deviations of discrepancy between IQ tests. Conversely, children over the age of five years averaged one standard deviation difference. This suggests that the implications of using verbal vs. nonverbal IQ tests are even more important for younger children than for older children. The large difference between the two age groups could be explained by discrepant language abilities. Children with AD have delayed communication skills and some never develop functional language (Bryson & Smith, 1998). It seems reasonable to assume that the younger group of children studied was likely to have larger language deficits, leading to significantly lower verbal IQ scores relative to nonverbal scores.

Differences Between Levels of Functioning

The sample was divided by a score of 70 on both the *Leiter-R* and *SB5*. There was a main effect of IQ in both analyses, indicating that the main effect of Test (*Leiter-R* vs. *SB5*) was significantly different regardless of which test is used to determine cognitive function. The more interesting results lay in the interactions between level of functioning and test. This interaction term was significant for the groups split based on *Leiter-R* score, but not for the *SBV* split. The lower-functioning group had an average difference of 7.73 points between tests, while the higher-functioning group had a much larger difference of 25.74. This result may be in part due to a Yfloor effect. Children with an IQ less than 70 only had a 30 point range in which to vary (basal of 40 up to 70), while the higher-functioning group can achieve scores up to 150 (a range of 80).

The *Leiter-R* likely results in a more accurate upper estimate of IQ, so a split based on *Leiter-R* score should be the more accurate characterization of the sample. If the

Leiter-R placed the child in the ID range for IQ, there is a greater likelihood that the child actually belonged there, and therefore would have similar scores on the *SB5*. The range of scores that is possible outside the realm of ID is much greater (71 to 150 for normal intelligence compared to 40 to 70 for ID) and, therefore, more variation would be expected. This likely does not hold true when the *SB5* makes the distinction between low and high functioning. This probably is because the *SB5* required the child to use language. In the presence of an ASD, underestimating ability level would leave greater room for variation in the low-functioning category.

Comparisons Based on Gender

ASD is reported to occur at a 4.3:1 male-to-female ratio in the general population (Fombonne, 2003), which is close to the 4.5:1 male to female ratio that was found in this sample. Analyses were conducted to explore whether the differences seen between the two intelligence tests held true for both genders. Both males and females demonstrated the characteristic pattern of higher scores on the *Leiter-R* than on the *SB5*, but no significant interaction emerged. As males and females performed similarly on the assessments of cognitive functioning, it seems that gender does not necessarily need to be considered when selecting tests of cognitive functioning or when interpreting such tests.

***Leiter-R* Compared to *SB5* Subsections**

As previously stated, although the *SB5* has a nonverbal portion, some grasp of language is necessary to understand the directions that are given to the individual being tested. Perhaps for this reason, the nonverbal *SBV* and the *Leiter-R* were not equivalent. Although the discrepancy between the two tests was reduced from 21.78 (using the full

scale *SB5*) to 16.18 (using the nonverbal section), a difference persisted that was both statistically and clinically significant. This replicated the findings of Lennen and colleagues (2010). It suggests that even the nonverbal portion of the *SB5* results in lower estimates of IQ in children with AD, as compared to fully nonverbal intelligence tests like the *Leiter-R*.

Limitations

The current study has several limitations. The data used were collected up to five years prior to the project, and it was not possible to control what tests were given to the patients. This drastically reduced the number of cases that qualified for inclusion in these analyses, and it left questions regarding whether children who received alternative batteries of tests differed in a systematic way from the youngsters who were evaluated here. No records were available to investigate how choices regarding test administration were made, and without such a record there is no way to determine whether the sample was biased. It is possible that the sample was biased towards children with lower IQs who were deemed to need a nonverbal instrument such as the *Leiter-R*. Conversely, the sample may have been more homogeneous than average for this clinical population because children who were unable to complete the Full Scale administration of the two assessments were otherwise excluded. Nevertheless, if either of these biases were in effect, then the sample may not have been truly representative of the typical array of children who present at specialty clinics that assess children for ASDs. Additionally, the measures were administered by a variety of individuals, so there is no way to guarantee

absolute standardization. However, this could also be seen as a strength because this is exactly the type of circumstance that occurs in real-world settings.

Conclusions

These analyses strongly suggest that the *Leiter-R* and the *SB5* are not equivalent measures of intellectual functioning in children with ASDs. Because these children regularly score between one and two standard deviations higher on the *Leiter-R* than *SB5*, it seems prudent to assume that one explanation is that these measures are representing both the upper and lower limits of ability level for these children. As advocates for this population, it is natural and appropriate to aim for that upper limit, but for some children, achieving that apparent level of functioning may not be possible. Intelligence tests are valuable for their prescriptive and predictive functions, but due to the limited research in this area, we need more data before we know how these measures perform in those two capacities.

The analyses conducted do not answer questions about whether or not these two intelligence tests are assessing different constructs and what the data imply for the long-term prognosis of these children with ASDs. The observed intraclass correlation coefficient between the two measures was modest to say the least ($r = .57$ for the full sample analysis). This is indicative of only moderate concordance between the assessments, and represents statistical evidence that they are not analogous cognitive measurements. The lack of agreement between the *Leiter-R* and *SB5* likely arises from the different roles language is playing in these two IQ tests.

This document already discussed the unique cognitive profile that is demonstrated by children with ASDs. That profile includes characteristic strengths (visual-spatial tasks) and weaknesses (verbal reasoning). It possible the discrepancies observed were due to the *Leiter-R* highlighting those strengths and minimizing the weaknesses by virtue of being a nonverbal measure. When viewed through this light, the *Leiter-R* would understandably produce higher scores for this population than cognitive assessments like the *SB5* that are presented verbally and included some of those categories of intelligences that are more difficult for those with ASDs.

The literature acknowledged that language plays an important role in both academic success as well as later ability to function in life (DeThorne & Schaefer, 2004). Regardless of how high children score on the *Leiter-R* when compared to the *SB5*, if they are within the range of ID for the verbal measure they are likely still going to struggle. We currently live in a social world that uses language for even the most mundane tasks, but it is important to recognize that it is not the only way to be assessed, or to carry out tasks. The magnitude of the discrepancy between the measures should make clinicians cautious about using only verbal tests for this population, regardless of the fact that we do not fully understand what the discrepancy means. More research is needed to flesh out the long-term implications that these findings have for children with ASD.

Inaccurately inflating or underestimating a child's cognitive ability would be a grave disservice, particularly since clinical recommendations are so frequently used to make decisions regarding educational placement and treatment. Programs designed for ability levels not consistent with child potential may not provide opportunities for

growth, and could impede the child's chances of learning skills to live as productive a life as possible.

Although understanding the nature of a child's performance on a traditional intelligence test that requires use of language can be helpful, including a nonverbal measure like the *Leiter-R* in a battery of tests may provide a more well-rounded picture of the child's ability. Given the limited scope of the current literature in this area, dual administration of the *SB5* and *Leiter-R* would afford researchers opportunities to make more complete evaluations of these measures in the future that would hopefully answer some of the questions that are still pending.

I recognize that in life, there are situations that arise that may alter the preferred outcome of an intelligence test. This could be the case when allocation of specialized services is involved, and a low IQ could allow a child to receive more assistance. Particularly in sensitive instances, it is important for clinicians to recognize what pattern of results can be expected from the administration of tests available. This comprehensive evaluation will help clinicians select tests that would allow them more accurately to tailor their recommendations for education, services, and treatment.

	<i>N</i>	Percent
<i>Gender</i>		
Male	41	82%
Female	9	18%
<i>Ethnicity</i>		
Caucasian	33	66%
African American	6	12%
Biracial	3	6%
Asian	1	2%
Unknown	5	10%
<i>Diagnosis</i>		
Autistic Disorder	26	52%
Non-Autistic Disorder	24	48%
<i>Age in Months</i> <i>M(SD)Range</i>		
Full Sample	66.67 (24.12) 36 – 131	

Table 1. Sample Characteristics

Comparison	<i>Leiter-R, M(SD)</i>	<i>SB5 M(SD)</i>	Difference	<i>t</i> -test	<i>p</i> value
All Subjects (n = 50)	88.20 (26.26)	66.42 (21.02)	21.78	10.47	< .001
<i>Diagnostic Category</i>					
Autistic Dis (n = 26)	79.42 (24.99)	59.12 (17.51)	20.30	7.15	< .001
Non-AD (n = 24)	97.71 (24.68)	74.33 (21.96)	23.38	7.58	< .001
<i>Age</i>					
< 60 months (n = 25)	99.44 (19.21)	71.92 (18.52)	27.52	9.93	< .001
> 60 months (n = 25)	76.96 (22.27)	60.92 (22.27)	16.04	5.95	< .001
<i>Level of Functioning</i>					
Low <i>Leiter-R</i> (n = 11)	51.91 (9.62)	44.18 (5.08)	7.73	2.80	< .019
High <i>Leiter-R</i> (n = 32)	98.44 (19.39)	72.69 (19.47)	25.75	11.86	< .001
Low <i>SB5</i> (n = 32)	75.28 (20.84)	53.09 (8.96)	22.19	8.12	< .001
High <i>SB5</i> (n = 18)	111.17 (17.98)	90.11 (14.07)	21.06	6.52	< .001
<i>Gender</i>					
Males (n = 41)	86.78 (27.26)	64.59 (20.68)	22.19	9.50	< .001
Females (n = 9)	94.67 (21.24)	74.78 (21.71)	19.89	4.19	< .003
<i>SB5 Sections</i>					
<i>SB5</i> Nonverbal (n = 50)	88.20 (26.26)	72.02 (22.36)	16.18	7.76	< .001
<i>SB5</i> Verbal (n = 48)	88.38 (26.17)	63.81 (19.47)	24.56	10.24	< .001

Table 2. Summary of *t*-test Results on Leiter-R and SB5 Comparisons

Source	df	F	<i>p</i>
Between Subjects			
Diagnosis	1	7.80	< .007
Error	48		
Within Subjects			
Tests	1	108.91	< .001
Tests x Diagnosis	1	.537	.467
Error	48		

Table 3. Analysis of Variance for the Effect of Diagnostic Category and Test

Source	df	F	<i>p</i>
Between Subjects			
Age Group	1	7.79	< .007
Error	48		
Within Subjects			
Tests (T)	1	126.99	< .001
Test x Age	1	8.82	< .005
Group Error	48		

Table 4. Analysis of Variance for the Effect of Age Group and Test

Source	df	F	<i>p</i>
Between Subjects			
Functioning Level	1	44.65	< .001
Error	48		
Within Subjects			
Tests	1	58.94	< .001
Tests x Functioning	1	17.08	< .001
Error	48		

Table 5. Analysis of Variance for the Effect of Level of Functioning on the *Leiter-R* and Test

Source	df	F	<i>p</i>
Between Subjects			
Functioning Level	1	75.31	< .001
Error	48		
Within Subjects			
Tests	1	97.57	< .001
Tests x Functioning	1	.067	.797
Error	48		

Table 6. Analysis of Variance for the Effect of Level of Functioning on the *SB5* and Test

Source	df	F	<i>p</i>
Between Subjects			
Gender	1	1.18	.282
Error	48		
Within Subjects			
Tests	1	59.34	< .001
Tests X Gender	1	.18	.675
Error	48		

Table 7. Analysis of Variance for the Effect of Gender and Test

References

- American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders*, fourth edition text revision, Washington, DC: APA.
- Banach, R., Thompson, A., Szatmari, P., Goldberg, J., Tuff, L., Zwaigenbaum, L., Mahoney, W. (2009). Brief report: Relationship between non-verbal IQ and gender in autism. *Journal of Autism and Developmental Disorders*, 39, 188–193.
- Bessent, T.E. (1950). A note on the validity of the Leiter International Performance Scale. *Journal of Consulting Psychology*, 234.
- Bishop, S.L., Richler, J., & Lord, C. (2006). Association between restricted and repetitive behaviors and nonverbal IQ in children with autism spectrum disorders. *Child Neuropsychology*, 12, 247-267.
- Bolte, S. & Poustka, F. (2002). The relation between general cognitive level and adaptive behavior domains in individuals with autism with and without co-morbid mental retardation. *Child Psychiatry and Human Development*, 33(2), 165-172.
- Bracken, B.A. & McCallum, R.S. (1998). *Universal Nonverbal Intelligence Test examiner's manual*. Itsaka, IL; Riverside.
- Bradley-Johnson, S. (1998). Test Reviews: Leiter International Performance Scale-Revised. *Psychology in the Schools*, 35(2), 195-197.
- Bryson, S.E. & Smith, I.M. (1998). Epidemiology of autism: Prevalence, associated characteristics, and implications for research and service delivery. *Mental Retardation and Developmental Disabilities Research Reviews*, 4, 97-103.
- Campbell, J.M, Brown, R.T, Cavanagh, S.E., Vess, S.F., & Seggall, M.J. (2008). Evidence-based assessment of cognitive functioning in pediatric psychology. *Journal of Pediatric Psychology*, 33(9), 999-1014.
- Carpentieri, S.C. & Morgan, S.B. (1994). Brief report: A comparison of patterns of cognitive functioning of autistic and nonautistic retarded children on the Stanford-Binet-fourth edition. *Journal of Autism and Developmental Disorders*, 24(2), 215–223.

- Cathers-Schiffman, T.A. & Thompson, M.S. (2007). Assessment of English- and Spanish-speaking students with the WISC-III and Leiter-R. *Journal of Psychoeducational Assessment*, 25(1), 41-52.
- Chakrabarti, S. & Fombonne, E. (2005). Pervasive developmental disorder in preschool children: Confirmation of high prevalence. *American Journal of Psychiatry*, 126(6), 1133–1141.
- Coolican, J., Bryson, S.E., Zwaigenbaum, L. (2008). Brief report: Data on the Stanford-Binet Intelligence Scales (5th ed) in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 38, 190-197.
- DeThorne, L.S. & Schaefer, B.A. (2004) A guide to child nonverbal IQ measures. *American Journal of Speech-Language Pathology*, 13, 275-290.
- Faul, F., Erdfelder, E., Lang, A.-G. & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175-191.
- Fombonne, E. (2003). Epidemiological surveys of autism and other pervasive developmental disorders: An update. *Journal of Autism and Developmental Disorders*, 33(4), 365-381.
- Fombonne, E. (2005). Epidemiology of autistic disorder and other pervasive developmental disorders. *Journal of Clinical Psychiatry*, 66(10), 3-8.
- Fombonne, E. (2007). Epidemiological surveys of pervasive developmental disorders. In F.R. Volkmar (Ed.), *Autism and Pervasive Developmental Disorders*, Second Edition (pp. 33-68). New York: Cambridge University Press.
- Gillberg, C. & Wing, L. Autism: not an extremely rare disorder. *Acta Psychiatrica Scandinavica*, 99, 399–406.
- Halprin, J.M & McKay, K.E. (1998). Psychological testing for child and adolescent psychiatrists: A review of the past 10 years. *Journal of American Academy of Child and Adolescent Psychiatry*, 37(6), 575–584.
- Hooper, V.S. & Bell, S.M. (2006). Concurrent validity of the Universal Nonverbal Intelligence Test and the Leiter International Performance Scale-Revised. *Psychology in the Schools*, 43(2), 143-148.
- Hopkins, K.D. & Bracht, G.H. (1975). Ten-year stability of verbal and nonverbal IQ scores. *American Educational Research Journal*, 12(4), 469-477.

- Howlin, P. (1998). Autism and intellectual disability: Diagnostic and treatment issues. *Journal of the Royal Society of Medicine*, 93, 351-355.
- Kuschner, E.S., Bennetto, L., Yost, K. (2007). Patterns of nonverbal cognitive functioning in young children with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 37, 795-807.
- Lecavalier, L., Snow, A.V., & Norris, M. (in press). Autism Spectrum Disorders and Intellectual Disability. In Matson, J.L. & Sturmey, P. (Eds.), *International Handbook of Autism and Pervasive Developmental Disorders*. Springer.
- Leiter, R.G. (1980). *Leiter International Performance Scale, instruction manual*. Chicago: Stoelting.
- Lennen, D.T., Lamb, G.D., Dunagan, B.J., & Hall, T.A. (2010). Verbal prowess equals higher IQ: Implications for evaluating autism. *Research in Autism Spectrum Disorders*, 4, 95-101.
- Lilienfeld, S.O., Lynn, S.J., Namy, L.L., & Woolf, N.J. (2006). *Psychology: From Inquiry to Understanding*. New York: Pearson Custom Publishing.
- 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(4), 483-499.
- Maltz, A. (1981). Comparison of cognitive deficits among autistic and retarded children on the Arthur Adaptation of the Leiter International Performance Scales. *Journal of Autism and Developmental Disorders*, 11(4), 413-426.
- Matson, J.L., Dempsey, T., LoVullo, S.V., Wilkins, J. (2008). The effects of intellectual functioning on the range of core symptoms of autism spectrum disorders. *Research in Developmental Disabilities*, 29, 341-350.
- Matson, J.L. & Shoemaker, M. (2009). Intellectual disability and its relationship to autism spectrum disorders. *Research in Developmental Disabilities*, 30, 1107-1114.
- Mayes, S.D. & Calhoun, S.L. (2003). Analysis of WISC-III, Stanford-Binet:IV, and academic achievement test scores in children with autism. *Journal of Autism and Developmental Disorders*, 33(3), 329-341.
- Preacher, K. J. (2002). Calculation for the test of the difference between two independent correlation coefficients [Computer software]. Available from <http://www.quantpsy.org>.

- Raven, J.C. (1998). *Manual for Raven's Progressive Matrices and Vocabulary Scales*. Oxford: Oxford Psychologists Press.
- Reeve, R.R, French, J.L, & Hunter, M. (1983). Brief report: A validation of the Leiter International Performance Scale with kindergarten children. *Journal of Consulting and Clinical Psychology*, 31(3), 458-459.
- Roid, G.H. (2003). *Stanford Binet intelligence scales* (5th ed.). Itasca, IL; Riverside Publishing.
- Roid, G.H. & Miller, L.J. (1995, 1997). *Leiter International Performance Scale-Revised*. Wood Dale, IL: Stoelting Col.
- Roid, G.H., Pomplun, M., & Martin, J.J. (2009). Nonverbal intellectual and cognitive assessment with the Leiter International Performance Scale-Revised (Leiter-R). *Practitioner's Guide to Assessing Intelligence and Achievement*. Naglieri, J.A. & Goldstein, S.
- Sattler, J.M. (2001). *Assessment of Children: Cognitive Applications*, 4th Ed. La Mesa CA: Jerome M. Sattler, Publisher, Inc.
- Shah, A. & Holmes, N. (1985). Brief report: The use of the Leiter International Performance Scale with autistic children. *Journal of Autism and Developmental Disorders*, 15(2), 195-203.
- Silverman, W., Mizejeski, C., Ryan, R., Zigman, W., Krinsky-McHale, S., & Urv, T. (2010). Stanford-Binet and WAIS IQ differences and their implications for adults with intellectual disability (aka mental retardation). *Intelligence*, 38, 242-248.
- Sparrow, S.S. & Davis, S.M. (2000). Recent advances in the assessment of intelligence and cognition. *Journal of Child Psychology and Psychiatry*, 41(1), 117–131.
- Strand, S. (2004). Consistency in reasoning test scores over time. *British Journal of Educational Psychology*, 74, 617-631.
- Tsastsanis, K.D., Dartnell, N., Cicchetti, D., Sparrow, S.S., Klin, A., & Volkmar, F.R. (2003). Concurrent validity and classification accuracy of the Leiter and Leiter-R in low-functioning children with autism. *Journal of Autism and Developmental Disorders*, 33(1), 23-30.
- Wechsler, D. (1976). *Wechsler Intelligence Scale for Children: Revised Manual*. Windsor, England: N.F.E.R Publishing Company.

Wechsler, D. (1991) *Manual for the Wechsler Intelligence Scale for Children* (3rd Ed.) San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1997). *Manual for the Wechsler Adult Intelligence Scale* (3rd Ed.) New York: The Psychological Corporation.

Wechsler, D. (2003). *Wechsler Intelligence Scale for Children-Fourth Edition (WISC-IV)*. San Antonio, TX; The Psychological Corporation.