Underlying Factor Structures of the Stanford-Binet Intelligence Scales – Fifth Edition

A Thesis

Submitted to the Faculty

of

Drexel University

by

Danielle Chase

in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

November 2005

Dedication

This dissertation is dedicated to my parents, Fred and Renee Chase.

Acknowledgements

Thank you to Doug Chute, Mitzie Grant, Joe McCarty, Steve Platek, and John Kounios, my committee members.

Thank you to Renee Chase, Fred Chase, Jesse Chase, and Charlotte Weiss, my family who read for me, cooked for me, supported me, and prayed for me while I went through this process.

Thank you to Katie Wiggins, Tara Gray, and Kim Larson, my best friends who cheered me on through the big diss.

Table of Contents

List of Tables

Abstract

Underlying Factor Structures of the Stanford-Binet Intelligence Scales – Fifth Edition

Danielle Chase

Douglas L. Chute, Ph.D.

The majority of commonly used intelligence measures provide information about the underlying factor structure of a test by including both exploratory, as well as confirmatory, factor analyses in their manuals (Wechsler, 2002; Wechsler, 2003). However, the manual of the Stanford-Binet Intelligence Scales – Fifth Edition (SB5) does not include an exploratory factor analysis (EFA). The initial concern is that EFA and confirmatory factor analysis (CFA) are supposed to be used together when constructing a new test instrument (Gorsuch, 1983). The second greatest concern is that CFA only confirms one factor structure, while EFA can be used as an exploratory measure to find all possible factor structures.

To ensure that findings of the CFA of the SB5 were not sample-specific an EFA was conducted on the data. The hypothesis of this study was that an EFA of the SB5 would yield a different underlying structure than the five-factor model used by Roid, which in fact it did. The factor structure found was dichotomous and named General Knowledge and Ordering/Sequencing of Information. The variables that loaded on factor one all reflected an individual's fund of knowledge. Therefore, this first factor was named General Knowledge (GK). The variables that loaded onto the second factor had one major quality in common, which was that they required the ability to order and sequence information. Hence, this second factor was named Ordering/Sequencing of Information (OSI). The EFA of the SB5 data indicated that a

different factor structure underlies the measure. This information is very useful to clinicians and may guide interpretation of the SB5 in practice. These findings also confirmed the necessity of EFAs when constructing a new test battery, even when the battery is based on a pre-established theory.

CHAPTER ONE: BACKGROUND AND LITERATURE SURVEY

Intelligence

A definition of intelligence that most experts would accept includes the constructs of goal directed behaviors that are adaptable across environments (Williams, 1996). In two studies that asked experts to define intelligence there were themes common to both definitions. The first common theme was focused on the individual learning from experience, and the second on the individual's ability to adapt to the environment (Williams, 1996).

Intellectual Development

Though developmental theories have changed fantastically over the last century, one of the leaders in developmental theory is still Jean Piaget. Piaget is a major initiator of the field who articulated a system for the development of intellectual concepts. Many modern developmental theorists have found fault in Piagetian methodologies, primarily because the majority of subjects studied were his own children. The majority of these modern theorists did not agree with the developmental, age-related milestones that Piaget put forward in the first half of the 20^{th} century. However, the underlying developmental sequence that Piaget offered is one that continues to be widely accepted by the field of psychology in general (Diessner & Tiegs, 2001).

Piaget's 1962 publication of his lectures on the stages of child development, in the Bulletin of the Menninger Clinic, explored others' theories of intelligence before offering his own. He described Claparède's definition of intelligence as "an adaptation to new situations," and Karl Buhler's "as an act of immediate comprehension; this is to say, an insight" (Piaget, 1962, p 120). However, Piaget did not subscribe to either definition. Piaget refuted Buhler's intelligence as an insight theory, citing the example of a mathematician. Piaget maintained "that when a mathematician solves a problem, he ends by having insight, but up to that moment he feels, or gropes for, his way; and to say that the trial-and-error behavior is intelligent and the intelligence starts only when he finds the solution to the problem, seems a very narrow definition" (Piaget, 1962, p 120). Piaget maintained that the trial and error behavior that was described in Claparède's definition was apparent at every level of intelligence, including hypothesis testing at the most superior level (Piaget, 1962).

Piaget stated that both Claparède's and Bhuler's definitions related intelligence to "static conditions," and that he felt intelligence was a fluid condition. Piaget offered the notion that intelligence is "a compensation for an external disturbance," such that when an external disturbance occurs, an individual's intelligence is found in how he compensates for the disturbance (Piaget, 1962, p 120). Supporting his notion of intelligence, Piaget posited that compensation for external disturbances was achieved by initiation of an activity, which is not a static condition but a fluid one.

Piaget observed that intelligence does not appear instantly as an entity separate from the processes that preceded it. Rather, it develops contiguously with the interplay of innate and acquired processes upon which intelligence is dependent. Therefore, in order to understand intelligence, the underlying processes on which it is dependent must first be explored. The most fundamental processes, reflexes and reactions, posturing, and vocalizations, were originally innate mechanisms; however, when a relationship was established with the outside world, automatization became systemization and, in turn, behavior. Each resulting behavior was contingent upon the preceding organically evolved episode, and so on and so forth.

Fundamental Piagetian ideas, such as behavior resulting from the adaptation of a biological mechanism to one's environment, contribute to the current conceptualization of the nature-nurture debate in that there is a biological substrate to all behavior that is expressed via the environmental experiences of the organism. Today, equal weight is given to both processes. (Diessner & Tiegs, 2001). Ideally, as development continues behaviors will adapt to serve comprehensively the individual in his environment. (Diessner & Tiegs, 2001)

Intellectual Theory

Though Plato may have been the first to theorize about the construct of intelligence, it was Aristotle who first made the delineation between "excellence of intelligence and excellence of character," or between two intelligences, intellectual and moral (Tigner & Tigner, 2000, 168). Aristotle wrote of a triarchic theory of intelligence, not so dissimilar from Sternberg's more current triarchic theory, and included theoretical, practical, and productive intelligences. Sternberg's correlates were analytical, practical, and creative intelligences, but that will be discussed later.

Theoretical intelligence is a combination of the use of inductive and deductive processes to arrive at the understanding of a construct. Practical intelligence, according to Aristotle, is intellectual virtue and explains the role of intellect in ethical behavior (Tigner & Tigner, 2000). Practical intelligence is not only the understanding of the best course of action for a particular situation, but is also the execution of the highest behavioral standards. Aristotle's third component of intelligence is productive intelligence, which he identified as the capacity to make art, which he stated involved a "true course of reasoning" (Tigner & Tigner, 2000, 173).

Similar to Aristotle, though not exactly, is Sternberg's triarchic theory of intelligence which includes analytical, practical, and creative intelligences. Sternberg included in his conceptualization of analytical intelligence, one's ability to analyze, compare, and evaluate what is needed in order to make appropriate decisions (Tigner & Tigner, 2000). Furthermore, it includes an individual's ability to monitor and evaluate one's own performance. Sternberg's practical intelligence is reflective of an individual's ability to generalize a basic fund of knowledge across situations and experiences. Creative intelligence, in Sternberg's theory, is an individual's ability to deal with new, as well as, recurring situations. This intelligence is less structured than the others and not only allows for an individual to come up with novel ideas, but to defend those ideas in the face of controversy. Though Sternberg presents three different intelligences, his work reflects an understanding of a general concept of intelligence, similar to Spearman's "g" (Sternberg, 2000).

Spearman, in 1904, put forth the concept of a "g" factor, or an overall general intelligence, based on the positive correlations between cognitive tests (Duncan, Seitz, Kolodny, Bor, Herzog, Ahmed, Newell, & Emslie, 2000). He used a factor analysis of many cognitive measures in order to suggest that the main underlying component of these measures was an overall intelligence, or "g" (Spearman, 1904; Duncan et al, 2000). At the other end of the spectrum is Gardener's notion of multiple intelligences, which suggested that there are eight intelligences, to which we are biologically predisposed and are influenced by the environment (Kezar, 2001).

Piaget's theory of intelligence was diarchic in that there was operative intelligence and learning. Operative intelligence refers to highly integrated and generalized sets of actions that are adaptive in nature (Schonfeld, 1986). This is not dissimilar from the Cattell-Horn model's fluid intelligence. Learning, which is not dissimilar from the Cattell-Horn model's crystallized intelligence, is knowledge that is "a function of environmental data" (Schonfeld, 1986, 205). Genevan theorists maintain that learning is moderated by operative level, that is, one's amount of environmental knowledge is a function of the individual's ability to adapt.

Another multiarchical theory of intelligence is the Cattell-Horn-Carroll theory. In 1993, John "Jack" Carroll conducted a thorough investigation of the psychometric properties of human cognition. His study concluded that the Cattell-Horn *Gf*-*Gc* theory was, as far as psychometric theories of intelligence go, the most empirically grounded. Cattell's theory, that of the two intelligences, was labeled *Gf*-*Gc* for it's delineation of fluid *Gf* and crystallized *Gc* intelligences. He continued that the bridge between theory and practice was very well encapsulated in the Cattell-

Horn-Carroll (CHC) cognitive theory. The CHC theory is an amalgamation of the Cattell-Horn *Gf-Gc* theory and the three step theory put forth by Carroll (Alfonso, Flanagan, & Radwan, 2005; Roid, 2003a).

By 1991, Horn had developed a more expansive factor model of intelligence which included 9-10 *Gf-Gc* abilities: Fluid Intelligence (*Gf*), Crystallized Intelligence (*Gc*), Short-Term Acquisition and Retrieval, Visual Intelligence, Auditory Intelligence, Long-Term Storage and Retrieval, Cognitive Processing Speed, Correct Decision Speed, Quantitative Knowledge, and lastly, Comprehension and Expression of Reading and Writing Skills (Alfonso, Flanagan, & Radwan, 2005; Roid, 2003a). Carroll then integrated the *Gf-Gc* model into his own three-tiered model of human cognitive abilities.

The least restrictive level of Carroll's theory was an overall *g*, or general intelligence (Roid, 2003a). The next tier of Carroll's theory consisted of eight abilities representing the steadfast characteristics inherent to human control of behavior within any specific domain. These abilities were: Fluid Intelligence, Crystallized Intelligence, General Memory and Learning, Broad Visual Perception, Broad Auditory Perception, Broad Retrieval Ability, Broad Cognitive Speediness, Reaction Time/Decision Speed. The third and most broad tier of Carroll's model encompassed 69, focused abilities that are reflected in the eight general abilities, and moreover by the one general factor of intelligence, *g* (Alfonso, Flanagan, & Radwan, 2005; Roid, 2003a).

Originally, intelligence tests measured verbal and nonverbal functioning, and offered an overall estimate of cognitive functioning based primarily on these two constructs (Alfonso, Flanagan, & Radwan, 2005). However, Carroll and the CHC theory allow for a more comprehensive cognitive picture. By integrating this logic into the SB5, it allowed for the test to assess more than simply verbal and nonverbal functioning. Because children oftentimes have subtle deficits, the more comprehensive the test the more likely the clinician is to understand the child. Increased understanding of the various aspects of a cognitive profile allow for more accurate diagnosis and also for the clinician to provide more efficient and appropriate recommendations for the child (Alfonso, Flanagan, & Radwan, 2005; Roid, 2003a).

Assessment

Use of Intelligence Tests

Today, intelligence tests are most commonly administered on an individual basis and used to guide decision making with regard to exceptionality, eligibility, and educational placement (Salvia & Yssledyke, 1995). The aim of assessment is to gain insight into an individual that will aide in the decision making process with regard to screening, problem solving, diagnosis, therapy, rehabilitation, and progress evaluation. Screening evaluations are relatively brief and are generally used to identify an individual’s eligibility for certain programs and to gauge the necessity for a complete battery (DSM-IV-TR, 2000).

Development of the Intelligence Test

Before the emergence of psychology as a discipline, Jean Esquirol noticed that there seemed to be two distinct domains of mental impairment, mental incapacity and mental illness. Those affiliated with the former, to whom he referred to as “idiots,” never developed their intellectual abilities (Sattler, 2001, 129). Those affiliated with the latter, to whom he referred to as “mentally-deranged persons,” did in fact develop intellectual abilities but then lost them over time (Sattler, 2001, 129). After identifying these two groups, Esquirol attempted to develop a scientific method to differentiate between them. He first focused on physical measurements, and then incorporated speech patterns into his analyses. Esquirol’s initial descriptions of verbal patterns associated with various levels of “idiocy” are regarded as the earliest form of intelligence testing (Viney & King, 2003).

As psychology emerged as its own discipline in latter part of the nineteenth century, advances in the field of testing and measurement increased substantially. Sir Frances Galton, who was among the first to make a significant contribution to the field of psychometrics, is widely regarded as the patriarch of the testing movement. Galton is credited with the development of many statistical concepts, such as regression to the mean and correlation; Galton’s contributions opened the door to the study of intelligence (Francher, 1985). Following Galton’s development of statistical concepts geared towards measuring intellectual capacity, Karl Pearson developed the product-moment correlation formula for linear correlation, the partial correlation coefficient, the phi coefficient, and the chi-square test for establishing the “goodness of fit” of a particular data set to the expected distribution (Horn, 1968).

Concurrent with the development of statistical measures to quantify individual differences with regard to physicality, behavior, and mentality, was the exploration of the concept of intelligence (Viney & King, 2003). Around the turn of the twentieth century, psychologists and their colleagues began to explore the area of intelligence testing. Previously, researchers and clinicians alike had focused their interests not on intelligence, but the broader category of individual differences. In the 1890s James McKeen Cattell, an American student of Galton's, brought the idea of intelligence testing to America; however, when Galton's test was unable to predict academic achievement, enthusiasm over the concept faded (Viney & King, 2003).

Across the ocean in late 1901, the French research psychologist Alfred Binet revealed to his colleagues his intention to measure intelligence using specially developed tests and measures. At the time, those interested in the measure of individual difference, specifically with regard to intelligence, had been focusing on elementary processes as measured by response time. However, while response time indicated how quickly one was able to complete a task, it offered no insight into the cognitive underpinnings of the thought process (Thorndike, Hagen, & Sattler, 1986a).

Binet and his colleague, Thèodore Simon, initiated the study of memory for numbers and designs, as well as the ability to solve spatial or conceptual problems, employing tasks that related directly to everyday tasks. Binet also collected data on tasks that children were typically able to solve; thus, providing a normative sample by which to measure others. His comparative approach gave way to the first functional intelligence test, the Binet-Simon scale, which was published in France in 1905 (Salvia & Ysseldyke, 1995). The Binet-Simon was later revised in France on two separate occasions and re-released in 1908 and 1911, respectively (Viney & King, 2003; Roid, 2003).

Though the Binet-Simon scale was not well received in France, Lewis Terman of Stanford University, took interest and bought the publishing rights from Binet for the sum of one dollar. After translation, adaptation, and normative processes, Terman published the Stanford-Binet Intelligence Scale in 1916 (Viney & King, 2003). The Stanford-Binet scales would be revised and republished numerous times in coming years. The version that is used today is the Stanford-Binet Intelligence Scale – Fifth Edition.

In the years after 1916, numerous intelligence tests were released, some of which claimed inherent specialized conveniences and applications. For example, Robert Yerkes and colleagues developed the Army Alpha test during World War I. The special convenience of this test was that it was intended for group administration, such as for screening incoming military personnel. Similarly, Yerkes and colleagues put forth the Army Beta test, which was designed for group administration to illiterate individuals (Viney & King, 2003). While the focus of the test items had become more academic in nature, the interest in prediction of functional ability in the field was still prevalent.

History of the Binet Scales

Alfred Binet (1857-1911) set out to develop a series of tasks designed to measure individual differences. The differences that he intended to delineate included a number of complex mental facilities, such as memory, imagery, imagination, attention, comprehension, aesthetic sentiment, moral sentiment, muscular strength, motor ability, and hand-eye coordination (Roid, 2003a). The original Binet Scale was commissioned by the French government due to the need for a reliable diagnostic system to identify children with mental retardation. Together with physician Theodore Simon, Binet created the Binet-Simon scale, which was published in 1905 (Thorndike, Hagen, & Sattler, 1986a).

The 1905 Binet-Simon scale differed greatly from the scale that we use today. The original scale consisted of 30 pass/fail items. The tasks were also different from today's items and required a combination of mental and physical strategies to complete each task. The major breakthrough of the Binet-Simon scale was the complexity of the tasks and the breadth of mental abilities measured. Furthermore, intelligence was finally able to be measured during a clinical interview, as opposed to in laboratories or by using physical measurements (Roid, 2003a; Thorndike, Hagen, & Sattler, 1986a).

Although the Binet-Simon scale is quite antiquated with regard to today's intelligence scale standards, many current day innovations were derived from this scale. The concepts of strict administration, age-graded norms, and a rank order of items ranging from least to most difficult, are but a few. Furthermore, the inclusion of age-graded norms provided for the first estimate of mental age (Roid, 2003a; Thorndike, Hagen, & Sattler, 1986a).

The first revision of the Binet scale was in 1908; however, the majority of the scale was left unchanged. By 1911, the scale was in its second revision and the age range had been extended through adulthood, as opposed to its previous use for the diagnosis of mental retardation in children. With the inclusion of adults, the scales needed to be rebalanced, which Binet did by including five items for each age level. The abilities targeted by the 1911 edition were language, auditory processing, visual processing, learning and memory, and problem solving (Roid, 2003a; Thorndike, Hagen, & Sattler, 1986a).

By 1912, Lewis M. Terman of Stanford University began revisions on the 1911 Binet scale. Terman's version was one of the premier editions published in the United States. However, because Terman had been attending closely to Binet's work over the years, he'd compiled a list of ways to improve upon his already advanced measure. Much of his focus was on the mental age equivalents. Terman noticed that the Binet equivalents tended to overestimate the mental ages of young children and underestimate those of older children. Adding and eliminating items, by 1916 Terman had tested more than 2,300 children using a modified version of the Binet-Simon scale. Terman's changes were published as the *Stanford Revision and Extension of the Binet-Simon Intelligence Scale.* After publication of his changes, a revised version of the Binet-Simon scale was published in 1916 and was entitled the *Stanford-Binet Intelligence Scale* (Roid, 2003a; Thorndike, Hagen, & Sattler, 1986a).

The advantages that the Stanford-Binet had over other intelligence scales of the time were many. The first, and seemingly most simplistic, was that the 1916 version was the most comprehensive revision of Binet's original scale. The second, and perhaps the most important, was that the standardization procedure used by Terman was the most rigorous of the time. The third advantage was the inclusion of an extensive manual, both for administration of the test as well as for use as a teaching aide for understanding the test. The final advantage was an adaptation of Stern's (1912) concept of an intelligence quotient. Stern put forth the notion that to derive an intelligence quotient, the mental age of the subject was to be divided by the chronological age. Terman incorporated this concept into the Stanford-Binet (Minton 1988).

Although standardization procedures for the Stanford-Binet were both novel and comprehensive for their time, Terman felt that they were still suboptimal. He also felt that some of the items lacked validity, that the floors and ceilings were inadequate, that the test did not span a large enough age range, and that coaching might have been a problem for readministration, as there was no alternate form of the test. By 1937 Terman had revised the Stanford Binet, with the help of Maud Merrill, into the *Revised Stanford-Binet Intelligence Scale*. The revision included two alternate forms, the L form and the M form, each with 129 items. Though the revision was a vast improvement, there were still some difficulties associated with the test. One difficulty was that the items were still graded on a pass/fail basis. Another was that the toys that were used as a part of the assessment battery were Germanmade and in the years leading up to, as well as following, World War II, these toys could not be replaced (Roid, 2003a; Thorndike, Hagen, & Sattler, 1986a).

The third revision of the Stanford-Binet came after Terman's death in 1960 and was constructed primarily by Merrill. The *Stanford-Binet Intelligence Scale, Form L-M*, was different from its predecessor in that it included a deviation intelligence quotient with a normative mean of 100 and a standard deviation of 16. This version also included the 142 most pertinent items from the two previous forms of the test (Roid, 2003a; Thorndike, Hagen, & Sattler, 1986a).

The fourth revision of the Stanford-Binet, the Stanford-Binet Intelligence Scales – Fourth Edition (SB4), retained much of the content of the Form L-M edition. The same age range was covered, many of the same items and tasks were retained, and the basal and ceiling procedures were quite similar. Though much of the fourth edition was the same as its predecessor, there were many new changes as well. In contrast with the developmental age format, similar items were now grouped together into point scales. The greatest advance of the fourth edition was that, like the fifth edition, the fourth edition was based on a hierarchical model of intelligence. The four main areas assessed were verbal reasoning, abstract/visual reasoning, quantitative reasoning, and short-term memory. The fourth edition not only provided an overall intelligence quotient, but composite scores as well. Furthermore, to establish a basal level, the Vocabulary subtest of the fourth edition was used as a routing subtest, along with the subject's chronological age (Roid, 2003a).

The SB4 differs significantly from the latest version of the test, the Stanford-Binet Intelligence Scales – Fifth Edition (SB5), with regard to theoretical structure, as the SB5 adheres rather strictly to the CHC theory and the SB4 adheres to a less strict psychometric design. There are five main differences between the SB4 and the SB5. The first is the addition of the fifth factor; the five factors of the SB5 are: Fluid Reasoning, Knowledge, Quantitative Reasoning, Visual-Spatial Processing, and Working Memory. The second difference between the editions is the way in which the authors responded to user feedback, in that they included more child-friendly materials; for example, toys and colorful manipulatives. The third addition to the SB5 was the enhancement of the nonverbal content. The nonverbal portion of the SB5 accounts for 50% of the test and ranges across all factors, which is unique to the SB5 among cognitive batteries. The fourth difference is the increased breadth of the scale, in that the ceilings are higher and the basals are lower. This change allows for better assessment of both the mentally retarded and the mentally gifted. The fifth major difference between the SB4 and the SB5 is the enhanced usefulness of the test. The SB5 was physically designed for easier administration and theoretically offers more information than did its predecessor. Included in the SB5 is a five factor verbal and nonverbal contrast, an abbreviated version, and a nonverbal form of the test.

Table 1.

Comparison of the Various Binet Editions

Edition	Structure	Abilities Measured
1916	Parallel Vocabulary Tests	General Intelligence
	Single Age Scale	
1937	Form L Vocabulary Test	General Intelligence
	Parallel Age Scales	
1960/1973	Vocabulary Test	General Intelligence
1986	Vocabulary Routing Test	General Intelligence
	Single Age Scale	Verbal Reasoning
		Abstract/Visual Reasoning
		Quantitative Reasoning
		Short-Term Memory
2003	Hybrid Structure	General Intelligence
	Verbal Routing Test	Knowledge
	Verbal and Nonverbal Age Scales	Fluid Reasoning
		Quantitative Reasoning
		Visual-Spatial Processing
		Working Memory
		Verbal IQ
		Nonverbal IQ

Stanford-Binet Intelligence Scales– Fifth Edition

Prior to its 2003 revision, the Stanford-Binet Intelligence Scale was last revised in 1986. The most current revision, the 2003 revision, is the SB5. The changes between the fourth and the fifth editions of the Binet scale included changes in the layout of the test, norming standards, and the underlying theoretical structure of the instrument. (Roid, 2003a).

The SB5 is used to assess intellectual ability in individuals between the ages of two and 89 years, is individually administered, and contains 10 subscales. The three areas assessed by the SB5 are: general cognitive functioning, verbal and nonverbal intelligence, and five CHC factors formed into groups along verbal/nonverbal measures. The five CHC factors that the SB5 measures are Fluid Reasoning, Knowledge, Quantitative Reasoning, Visual-Spatial Processing, and Working Memory. Together, the ten subtests yield an overall estimate of cognitive functioning, which is the Full Scale Intelligence Quotient (Roid, 2003a).

Composites and Testlets of the SB5

The SB5 is comprised of 5 composite scores each with a verbal and a nonverbal testlet, for a total of 10 testlets.

Fluid Reasoning

Fluid Reasoning, as defined by Roid, is “the ability to solve verbal and nonverbal problems using inductive or deductive reasoning.” The tasks required by this section of the SB5 assess the individual’s ability to determine the underlying relationships between pieces of novel information. The inductive reasoning component requires the individual to derive the general whole from its specific parts. Likewise, the deductive reasoning component requires that the individual draw a conclusion, implication, or specific example from a general piece of information about the topic (Roid, 2003b).

The Fluid Reasoning subtests within the verbal domain progress through Early Reasoning, Verbal Absurdities, and Verbal Analogies, with the Early Reasoning subtest being the most basic tool for evaluation of this subdomain. At the most basic level, the individual is required to sort and classify pictured objects. At a more moderate level, the individual is required to identify what is absurd or impossible about verbally presented sentences, which requires the individual to make generalizations about the information provided. At the highest level, the individual is required to reason verbally through a set of analogies, which requires verbal fluency, long-term storage of vocabulary meanings and variations, and verbal problem solving strategies (Roid, 2003b).

The Fluid Reasoning subtests within the nonverbal domain are Object-Series Matrices. Initially, the individual is required to match objects. These objects are then placed into a series, either repetitive or not, that the individual must complete. The last phase is similar to the classic matrix-reasoning measures that are common among intelligence testing. These subtests measure an individual's sequential and inductive reasoning abilities and the ability to solve novel figural problems (Roid, 2003b).

Knowledge

Knowledge, as defined by Roid, "is a person's accumulated fund of general information acquire at home, school, or work." This construct is often referred to as crystallized intelligence, as it involves learned material that has been stored in long-term memory (Roid, 2003b).

The Knowledge testlets within the verbal domain is Vocabulary. In keeping with the previous versions of the Binet, as well as most published cognitive batteries, the Vocabulary subtest was included. First, the subject is required to identify body parts, toy objects, and picture vocabulary. As the difficulty level increases, the subject must clearly define vocabulary words. At the upper levels, performance on this testlet is influenced by schooling (Roid, 2003b).

The Knowledge testlets within the nonverbal domain include Procedural Knowledge and Picture Absurdities. At the lowest end of the spectrum, the subject is required to communicate basic human needs using gesture. As the task demands increase, the subject is presented with impossible pictures in which he is required to point out what is odd or impossible about the scene. The Nonverbal Knowledge tasks tax an individual's basic level of common knowledge about people, nature, and physical laws of the universe. Though these are the main areas targeted, these testlets also require perception of detail, attention, concentration, geography, science, and inference skills (Roid, 2003b).

Quantitative Reasoning

Quantitative Reasoning, as defined by Roid, "is an individual's facility with numbers and numerical problem solving, whether with word problems or with pictured relationships." The items included on the SB5 Quantitative Reasoning testlets target problem solving abilities as opposed to rote mathematical knowledge (Roid, 2003b).

The Quantitative Reasoning testlets within the verbal domain measure an individual's ability to use a variety of mathematical skills. The lower levels of the testlets assess the individual's ability to count toys or pictured objects, as well as to use basic addition and subtraction skills. As the demands of the testlets increase, the individual is required to use geometric skills, measurement skills, and to complete word problems involving multiplication. At the highest level, there are multiple ways of arriving at the solution and in depth problem solving abilities are critical to success at this level (Roid, 2003b).

The Quantitative Reasoning testlets within the nonverbal domain have been carried over from the SB4; however, the focus of the testlets from the SB5 are on the reasoning behind the mathematical concepts, as opposed to the rote solving of mathematical items. Initially, basic concepts such as relative size, counting, and addition are targeted. As the testlets progress, items become more complex and abilities such as depicting figural series, functional relationships, linear transformations, and logical or algebraic transformations are taxed. In order to succeed on the higher level tasks, the subject must use problem solving strategies, persistence, and cognitive flexibility (Roid, 2003b).

Visual-Spatial Processing

Visual-Spatial Processing, as defined by Roid, “measures an individual’s ability to see patterns, relationships, spatial orientations, or the gestalt whole among diverse pieces of a visual display.” The items included on the SB5 include a Form Board and Form Patterns, which assess the individual’s ability to move pieces and shapes to form a proper whole (Roid, 2003b).

The Visual-Spatial Processing testlets within the verbal domain assess the individual’s ability to understand spatial concepts and relationships. The lower levels of the test include terms such as “ahead” and “behind,” and do not rely heavily upon expressive vocabulary. However, as the task demands increase, expressive vocabulary is needed to explain the complex relationships between geographic information (Roid, 2003b).

The Visual-Spatial Processing testlets within the nonverbal domain incorporate the Form Board activity from the SB4; however, tasks have been added in order to expand the evaluation of Nonverbal Visual-Spatial Processing tasks. Initially, shapes are matched and then inserted into forms. As the individual progresses, accurate duplication of patterns using the provided shapes is targeted. All levels within this testlet address visual construction abilities (Roid, 2003b).

Working Memory

Working Memory, as defined by Roid, “is a class of memory processes in which diverse information stored in short-term memory is inspected, sorted, or transformed.” The tasks of the SB5 Working Memory testlets require an individual to function well in all areas of working memory. For example, at the lower end of the testlets subjects simply repeat series of numbers or words; however, as the testlets increase in difficulty individuals are required to manipulate the presented information (Roid, 2003b).

The Working Memory testlets within the verbal domain begin with Memory for Sentences, which has long been a component of the Binet scales. As the testlets increase in difficulty, the individual is required not only to retain bits of information in working memory, but to manipulate these bits as well. The individual must filter out the irrelevant information and maintain focus on the pertinent. Oftentimes individuals are able to complete the rote memory sections but encounter difficulty when information manipulation is required (Roid, 2003b).

The Working Memory testlets within the nonverbal domain begin by assessing the individual's ability to hold fundamental, observable objects in short-term memory and progress into a rote memory block tapping task. However, towards the higher end of the testlets, the information presented becomes less concrete and more complex. Furthermore, the information must be manipulated, which places both memory, organizational, and visual-spatial demands on the individual (Roid, 2003b).

Table 2.

The Stanford-Binet Intelligence Scales – Fifth Edition

	Domain	
Factor Index	Nonverbal	Verbal
Fluid Reasoning	*Activity: Object-Series/Matrices* Requires the ability to solve novel figural problems and identify sequences of pictured objects or matrix-type figural and geometric patterns	*Activities: Early Reasoning, Verbal Absurdities, Verbal Analogies* Requires the ability to analyze and explain, using deductive and inductive reasoning, problems involving cause-effect connections in pictures, classification of objects, absurd statements, and interrelationships among words
Knowledge	*Activity: Procedural Knowledge, Picture Absurdities* Requires knowledge about common signals, actions, and objects and the ability to identify absurd or missing details in pictorial material	*Activity: Vocabulary* Requires the ability to apply accumulated knowledge of concepts and language and to identify and define increasingly difficult words
Quantitative Reason.	*Activity: Nonverbal Quantitative Reasoning* Requires the ability to solve increasingly difficult premathematic, arithmetic, algebraic, or functional concepts and relationships depicted in illustrations	*Activity: Verbal Quantitative Reasoning* Requires the ability to solve increasingly difficult mathematical tasks involving basic numerical concepts, counting, and word problems

Visual-Spatial Proc.	*Activity: Form Board, Form Patterns* Requires the ability to visualize and solve spatial and figural problems presented as "puzzles" or complete patterns by moving plastic pieces into place	*Activity: Position & Direction* Requires the ability to identify common objects and pictures using common visual-spatial terms such as "behind" and "farthest left," explain spatial directions for reaching a pictured destination, or indicate direction and position in relation to a reference point
Working Memory	*Activity: Delayed Response, Block Span* Requires the ability to sort visual information in short-term memory and to demonstrate short-term and working memory skills for tapping sequences of blocks	*Activity: Memory for Sentences, Last Word* Requires the ability to demonstrate short-term and working memory for words and sentences and to store, sort, and recall verbal information in short-term memory

Factor Analyses of the SB5

The SB5 manual does not include an exploratory factor analysis (EFA) regarding instrument development; however, a confirmatory factor analysis (CFA) is included. The CFA that was conducted on the SB5 was not done to confirm the five-factor model the test was built around, but to confirm the presence of verbal and nonverbal domains, which were found to be present (Roid, 2003a).

The rationale behind the inclusion of the CFA and the exclusion of the EFA was that "the SB5 factor structure was established a priori guided by contemporary intelligence theory (e.g., Cattell-Horn-Carroll)" (Roid, 2003a). Because the SB5 was constructed using the CHC five-factor model, EFA was not needed. The major problem that was found with this explanation are that both CFA and EFA can, and should, be used together to compliment each other while developing or revising an instrument. Furthermore, CFA only confirms one factor structure, while EFA can be used as an exploratory measure to find all possible factor structures. To ensure that the findings of the CFA for the SB5 were not sample specific, both an EFA and a

CFA should have been conducted on the data. In the case of the SB5, EFA was only used to ensure that the test was a suitable measure of overall cognitive ability, which it is (Roid, 2003a).

The rationale provided by the SB5 manual for the exclusion of an EFA was that because the SB5 was, a priori, based on the CHC theory, they used a CFA instead. Roid explained that because there is no model-data-fit index inherent to EFA, save possibly the chi-square of the maximum likelihood method, that EFA should not have been done in this case. However, this appears rather suspect in that the author wanted the data to fit a particular theory so he used the analysis that would yield the results he was looking for. Furthermore, Roid himself admits that "traditional EFA procedures, and especially those using default settings, are not expected to reproduce the five-factor model accurately" (Roid, 2003b).

A cornerstone of the science-practitioner model is that studies be replicable in order to be considered reliable and valid. If the author himself states that the study is not replicable, how is it possible that we consider the SB5 to be reliable and valid across the many levels that the test is used for? Furthermore, what implications does this have in clinical practice? If the SB5 does not in fact accurately reflect the five factors that it claims to, is the information that clinicians provide to their patients, based on findings from the SB5, an accurate representation of the patient's functioning?

Dombrowski, DiStefano & Noonan, stated in the *Communiqué*, the Newspaper of the National Association of School Psychologists, that "both EFA and CFA used in an exploratory manner should be replicated on independent samples of data to ensure that the findings are not sample specific" (Dombrowski, DiStefano & Noonan, 2004). If in fact the EFA is replicated and the findings do not mirror those of Roid's, it will largely limit the overall generalizability of the SB5. The EFA that was conducted on the normative data of the SB5 only indicated that the SB5 is a strong measure of "*g*," most likely due to the inclusion of memory and cognitively complex items (Roid, 2003b). Furthermore, based on an EFA of the correlations between the SB5's 20 half scales, clear pictures were not presented of either the CHC five-factor model or of the verbal-nonverbal dichotomy, which is a two-factor model. However, when the subscales were separated into verbal and nonverbal, the five-factor model was able to be seen clearly using an EFA of only the verbal subtests. This finding indicates "that the nonverbal subscales are in need of additional testing and refinement to ensure that these dimensions can be uniquely and accurately measured in the company of verbal dimensions" (Dombrowski, DiStefano & Noonan, 2004).

Hypotheses

The findings of this study are important to the clinician in that interpretation stems from the results, and if the results are based on a factor structure that does not best account for the data, interpretation might be skewed. If this is the case, the information gleaned from the SB5 may not best describe the individual. If in fact a new factor structure is discovered, one that better fits and explains SB5 data, it will change the way in which the SB5 is both used and interpreted. The alternative potential underlying factor structures that we have reason to expect to find are:

1) A single factor, presumably *g*.
2) A verbal/nonverbal dichotomy.
3) A three-factor model. Save the visual spatial and quantitative reasoning subtests, the three remaining verbal subtest will load on one factor and the three remaining nonverbal subtests will load on another factor. The third factor will be in the four subtests that comprise visual spatial and quantitative reasoning abilities; these subtests will load on the same factor. These four subtests will have their own factor loading because they measure similar constructs and involve similar abilities, such as measurement and problem solving abilities.

If the findings of our data analyses mirror those of Roid, and the SB5 does in fact adhere to the five-factor model (or to a verbal/nonverbal dichotomy with the five scales attributed to each), the outcomes will indicate that Roid's findings were not sample specific and that the interpretation of the SB5 according to its manual is generalizable to the population as a whole. Though the above question of the underlying factors of the SB5 is interesting, and could possibly change the way in which clinicians interpret SB5 findings, it does not speak to the applications of the SB5 in clinical settings. An interesting question that applies directly to the clinical use of the SB5 is whether or not the SB5 can predict specific diagnoses. Furthermore, are there patterns within the scores on the SB5 typical to specific diagnosis, and if so, how does this relate to intelligence? How does gender play into the relationship of the SB5 to intelligence?

It was hypothesized that, in small subgroup correlations, the verbal factor of the SB5 (should one be found) would correlate significantly with the standard scores of the verbally based adaptive, educational, and neuropsychological measures (see table 3). Furthermore, it was hypothesized that the nonverbal factor of the SB5 (should one be found) would correlate significantly with other nonverbal measures (see table 3). These hypotheses were offered as it was presumed that the anticipated correlated measures were based on similar constructs; such constructs would provide convergent validity for the discovered factor scores.

CHAPTER TWO: APPARATUS AND TEST PROCEDURE

Participants

An initial subject pool of 133 cases was available; however, included in the analyses of this study were only 78 cases. Initially, the subject pool was restricted to 107, as these were the only cases with complete SB5 data available. The excluded 26 cases did not have complete SB5 data, as these cases may have referenced a previous administration of the SB5 but did not provide a complete data set for analysis. Because a clinical sample was used, the mentally retarded population was overrepresented in our sample (FSIQ < 70). Because this study aimed to generalize its findings to the population, and because Roid's normative group "was nationally represented and matched to percentages of the stratification variables identified in U.S. Census Bureau (2001)," the 24 cases that fell in the mentally retarded range were excluded (Roid, 2003). Similarly, the two cases that fell in the Very Superior range were also excluded (FSIQ > 130). Of note, both the MR and the 2 gifted subjects were considered for group membership; however; the remaining 78 cases, which were more reflective of normal functioning, had complete SB5 scores and were included in the final analyses. The rational of selection was based upon the statistical convention that two standard deviations from a mean of a measure reflect approximately 96% of the population (Micceri, 1989). The SB5 has a mean of 100 and a standard deviation of 15; therefore, all subjects whose Full Scale IQ scores ranged outside 70-130 were excluded.

The data of 78 cases were included in this analysis with the Full Scale Intelligence Quotients ranging from a standard score of 70 (mean 100, standard deviation 15) to a standard score of 130, with a mean of 94.7 and a standard deviation of 11.3. 52 males and 26 females were included in the study. The minimum age was 2 and the maximum age was 19; the mean age was 8.2 with a standard deviation of 3.6. 47 subjects had an Attention-Deficit/Hyperactivity Disorder (ADHD), 60 subjects had an Anxiety Disorder (ADG), 12 subjects had Mood Disorders, 4 subjects had a Bipolar Disorder, 9 subjects had a Depressive Disorder, and 22 subjects had a conduct disorder, and 33 subjects had comorbid ADHD and ADG (see Appendix B). All data were taken from individual scores on the SB5 and the above mentioned academic and neuropsychological tests. Cases with missing SB5 data were omitted; the resultant number of cases was 78. All data from the 78 cases were used in both the factor analysis and the post hoc analyses. Data used in the post-hoc correlation analyses were included on a case-by-case basis (see table 3).

All data were collected during Multidisciplinary Evaluations (MDEs) done on an outpatient basis, as well as stand-alone outpatient evaluations, at the Franciscan Hospital for Children in Boston, Massachusetts. MDSs are comprised of numerous evaluations, including psychological, occupational, and speech/language. Psychological evaluators included Danielle Chase, M.S., Ph.D. Intern; Joseph C. McCarty, Ph.D., NCSP; Allen Brown, Ph.D./J.D., Chris Lopes, M.A., Ph.D Intern; James Nguyen, M.A., Ph.D. Intern; Sarah Fournier, M.A., Psy.D. Practicum Student; and Jennifer Markey, M.A., Psy.D Practicum Student.

Participants were referred for MDEs for a variety of reasons. Referrals came primarily from psychiatrists and neurologists. The majority of referral questions were to confirm a DSM diagnosis, assess the need for medicinal management, and to quantify and evaluate cognitive, emotional, and behavioral difficulties. All MDEs provided recommendations that addressed referral concerns. Because of the pathologies associated with the participants, diagnoses were coded in order to track the variance that may stem from the inclusion of various pathologies, as well as for the purpose of tracking the predictive power of the SB5 testlets regarding said diagnosis. Similarly, gender was also coded to study any variance that may result from using a heterogeneous population as well as enabling the determination as to if there is a difference in the factor findings of the SB5 for females then for males, or vice versa.

Inclusion Criteria

- Subject is included in the listing of examinees by one of the aforementioned MDE examiners; therefore, ages will range from two years, ten months – nineteen years, nine months
- Subject was evaluated for MDE on or after July 14, 2004
- Subject's record included a set of SB5 scale scores

Exclusion Criteria

- Subject was evaluated prior to July 14, 2004 or after August 2005
- Subject was not evaluated by the above listed MDE examiners
- Subject was not assessed using a full administration of the SB5

Power Analysis

The commonly accepted number of subjects per variable for a factor analysis is at least five (Gorsuch, 1983). As published by the authors (Gorsuch, 1983), the scoring structure of the SB5 consists of 10 subtests yielding a need for a minimum of 50 subjects; 78 complete SB5 cases were available for this analysis ensuring that an adequate subject-to-variable ratio was involved.

Procedure/Data Analysis

As noted above, data were collected by MDE evaluators at FHC in Boston, MA, over the course of thirteen and a half months. These data were collected in accordance with the SB5 administration and scoring procedures. The scale scores were entered into a data set and then analyzed using a factor analysis (Roid, 2003b). All identifying information was removed from the data set and included were the SB5 scaled and composite scores. The data was cleaned and proper entry was ensured.

Initially, the means and standard deviations of the scaled scores were compared with those that Roid used in his stability sample in order to approximate the generalizability of this study. One hypothesis of the study was that there would be only one factor that represents an overall g; therefore, it continues that all of the variables are interrelated and highly correlated. When it is known that the variables are correlated, an Oblimin rotation is used in order to minimize variance. In the event that the correlation analysis indicates that the factors are uncorrelated, the Varimax rotation will be used.

An exploratory factor analysis was then conducted to examine the underlying factor structure of the SB5, based on the 10 scales that are included in the measure. The principal components method of extraction was used (Joreskog & Lawley, 1968; Pedhazur, 1982). A variety of criteria were used to determine the number of factors to retain and analyze. They included Cattell's (1966) scree test, a weighted, reduced correlation matrix, and the interpretability of the Oblimin rotation of factors. Cattell's Scree Plot method retains its variables based on where the plot levels off and the eigenvalue method retains its variables that have factor loadings higher than one (Morrison, 1990). All analyses used the Statistical Package for the Social Sciences – Edition 10 (SPSS) factor analysis software. The exploratory factor analysis was then re-run, splitting the male and female cases to evaluate the hypothesis that the gender of the subjects will influence the underlying factor structure of the SB5. Furthermore, post hoc regression analyses were conducted to address the question of whether or not the SB5 can predict specific diagnoses. The diagnostic categories were ADHD, Anxiety-NOS, Mood Disorder, Depressive Disorder, Oppositional Defiant Disorder (ODD), and ADG.

To ensure that our data were valid and that the factors measured the constructs that we indicated that they did, the correlations between the SB5 and various other measures were explored. The following table contains the number of cases per test used in these analyses.

Table 3.

Correlations Between Verbal and Nonverbal Measures and Respective Factors

Test	Number of Cases
Stanford-Binet Intelligence Scales – Fifth Edition (SB5)	78
Peabody Individual Achievement Test (PIAT) Math*	11
Oral and Written Language Skills (OWLS) Written Expression**	13
Diagnostic Assessment of Reading (DAR) Spelling**	10
Kauffman Educational Achievement Tests (K-TEA) Reading Comprehension**	12
Diagnostic Assessment of Reading (DAR) Word Identification**	9
Vineland Communication**	41
Vineland Daily Living Skills**	41
Vineland Socialization**	41

*Correlations with anticipated "verbal" factors
**Correlations with anticipated "nonverbal" factors

As noted in the table above, post hoc correlations were run to evaluate the interrelatedness of the found verbal factor and the standard scores of previously established measures of verbal information. Similarly, correlations between the found nonverbal factor and the standard scores of previously established measures of nonverbal information were run.

Factor analysis, regression analyses, and correlations are types of parametric statistics. Inherent to parametric statistics are certain assumptions, such as a normal distribution, an interval level of measurement, and homogeneity of variance. Taken from the multiple general linear hypothesis are many of the assumptions of factor analysis, such as linear relationships, interval data, untruncated variables, exclusion of irrelevant variables, lack of mulitcollinearity, and multivariate significance testing (DeVellis, 1991). "However, it has long been established that moderate violations of parametric assumptions have little or no effect on substantive conclusions in most instances" (Cohen, 1969: 266-267).

Because of the large sample size that was used, we have ensured that the amount of error associated with each individual case would be minimized to a roughly equivalent level. It is hoped then that the latent variables involved exerted nearly the same amount of influence on all items. To ensure sound measurement, descriptive statistics were run to guarantee that the data was relatively free of coding errors and that there were no gross outliers. In turn, this would ensure that there was no skew or kurtosis. To ensure normality, boxplot tests were run. A principal components analysis was used. The “principal components analysis is a statistical technique that linearly transforms an original set of variables into a substantially smaller set of uncorrelated variables that represents most of the information in the original set of variables” (Dunteman, 1989, 1).

CHAPTER THREE: DATA

Initial Analyses; n = 107

Before analyses took place, all data were “cleaned.” Frequency scores were tabulated and outliers were corrected if initially entered incorrectly. Furthermore, the comparability of our sample’s values to those anticipated by Roid’s normative values was evaluated in order to see if our findings could be generalized. The means and the standard deviations of the initial 107 subjects’ SB5 scaled scores were compared with the expected means and standard deviations of the Roid’s test (see Table 4.).

The sample means for each subdomain were compared with the population mean that was set by Roid in the process of “smoothing” his national SB5 data set to the normal curve. Originally the p-value was set at .05; however, all obtained t-values were greater than the critical t-value, indicating that our sample was significantly different from Roid’s expected sample. When a more conservative p-value was used (.005), nine of the ten subdomains yeilded a t-value obtained from the analysis that was greater than the critical t-value. This indicated that the means of the nine sample groups were significantly different than the means of Roid’s expected groups, and hence that our sample was a poor candidate for generalization to the general population. There was a single subdomain, Verbal Fluid Reasoning, that yeilded a t-value obtained less than the critical value of t, indicating that the mean of this sample group was not significantly different than the mean of Roid’s group at the more conservative standard.

Table 4.

Comparison of the Sample to the Normative Sample Expected by Roid

Subtests	Chase n=107 Mean	SD	Expected n=107 Mean	SD	t-obtained df=212
Nonverbal					
Fluid Reasoning	8.90	4.48	10	3	2.110**
Knowledge	7.86	3.66	10	3	4.678
Quant. Reasoning	7.57	3.79	10	3	5.211
Visual-Spatial	8.01	3.88	10	3	4.197
Working Memory	7.59	3.93	10	3	5.042
Verbal					
Fluid Reasoning	7.08	4.04	10	3	6.002
Knowledge	7.94	3.36	10	3	4.731
Quant. Reasoning	7.57	3.59	10	3	5.373
Visual-Spatial	8.46	4.06	10	3	3.155
Working Memory	7.23	3.79	10	3	5.928

t-critical is 1.960 when alpha is set at .05; t-critical is 2.576 when alpha is set at .005
*nonsignificant at .05
**nonsignificant at .005

Though one group was not significantly different from Roid's, the majority were. This suggested that there were issues with considering the full sample of 107 for analysis. A further sign of this arose with the determinant of the correlation matrix, which speaks to the fitness of this sample for factor analysis (SPSS, 1999). The determinant closely approximated zero (determinant = 1.140E-04), indicating that a factor analysis of this data was not appropriate as the variables were too strongly intercorrelated. In the full sample of 107, the variance of the variables overlapped to such a degree that there was insufficient distinction between the variables to support reliably the presence of even a single factor.

Further examination the data indicated that 27 subjects were in the mentally retarded (MR) range, and that the majority of these subjects had earned scaled scores of one on all or most subtests of the SB5. This in turn decreased the overall variance by creating such a strong resemblance among the variables that they mirrored one another as if they were all measuring the same construct. Therefore we can infer that there was not enough variance among the scores to comprise an underlying factor structure of the data. In other words, the many instances of scaled scores of one across most or all of the variables skewed their distributions in similar ways, inflating the resemblance between them. Hence, not surprisingly, the inclusion of individuals with MR prevented the sample variance from configuring in a way that resembled the general population (Roid's scores).

Because the intent of the study focused on generalizability, two standard deviations from the mean was chosen as an exclusion cut-off within the sample of 107, as this would then leave a subsample that was representative of 96% of the general population (Micceri, 1989). This led to the exclusion of the data of the 27 MR subjects, as well as the two cases that were more than two standard deviations above the mean. Left for analysis was a subsample of 78 cases with IQ scores within the general range of typical functioning set by the +/- two standard deviation inclusion principal (Micceri, 1989).

Subsample Analysis; n = 78

After slimming the sample to 78, the mean scores from each subdomain were then compared with those predicted by Roid (see Table 5.).

Five of the ten subdomain mean scores (including his routing scales - Nonverbal Fluid Reasoning and Verbal Knowledge - which he considered to be the core tasks of the measure) were not significantly different than Roid's at the .05 level. At the .005 level, eight of the ten subtests were not significantly different from Roid's, indicating that the sample population grossly resembled the normative population that Roid used. The two subtests that were significantly different than Roid's were Verbal Fluid Reasoning and Verbal Working Memory, though these were not far removed from the higher cut-off level. A reason why a clinical sample would differ from a typically-functioning one on these particular scores is that they may successfully track compromised ability to draw verbal deductions and pay attention to overheard material (respectively) due to the interference of the diagnostic conditions involved.

Table 5.

Comparison of the Subsample to the Normative Sample Expected by Roid

	Chase	Expected		t-obtained	
	n=78	n=78		df=154	
Subtests	Mean	SD	Mean	SD	
Nonverbal					
Fluid Reasoning	10.68	2.84	10	3	1.453* **
Knowledge	9.37	2.23	10	3	1.456* **
Quant. Reasoning	8.82	2.75	10	3	2.561**
Visual-Spatial	9.47	2.64	10	3	1.172* **
Working Memory	9.15	2.91	10	3	2.021**
Verbal					
Fluid Reasoning	8.50	3.04	10	3	3.102
Knowledge	9.29	2.22	10	3	1.681* **
Quant. Reasoning	8.90	2.23	10	3	2.599* **
Visual-Spatial	9.78	3.28	10	3	.4371**
Working Memory	8.47	3.12	10	3	3.112

t-critical is 1.960 when alpha is set at .05; t-critical is 2.576 when alpha is set at .005
*nonsignificant at .05
**nonsignificant at .005

CHAPTER FOUR: RESULTS

Appropriateness for Factor Analysis

When the factor analysis was run on the remaining 78 cases, a two factor model was found. Unlike the preliminary factor analysis, the analysis of 78 cases was in fact valid because the determinant of the correlation matrix was over two million, indicating that there was a large amount of variability among the variables (determinant = 24,528,135.201). In addition, the Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) and Bartlett's Test of Sphericity were also run. According to the KMO, which indicates that the variables are measuring a common factor as the KMO approaches one, the amount of common variance shared by the variables is in the "meritorious" range (KMO = .838)

Table 6.

KMO Values

KMO Value	Degree of Common Variance
0.90-1.0	Marvelous
0.80-0.89	Meritorious
0.70-0.79	Middling
0.60-0.69	Mediocre
0.50-0.59	Miserable
0.00-0.49	Don't Factor

Tabachnick & Fidell, 1996

Because the KMO was in the meritorious range, the factors to be extracted are expected to account for a substantial amount of the sample's variance, indicating that it is acceptable to proceed with the factor analysis (Harman, 1976).

Bartlett's Test of Sphericity determines the difference between the target correlation matrix and the Identity Matrix is significant via an F-test. The Identity

Matrix is hypothesized as one kind of worst-case, unfactorable matrix in which there is no shared variance between the variables; i.e., each variable correlates only with itself. Because there is no variance in the Identity Matrix, if a factor analysis were conducted upon it, there would be as many factors as there were variables because there would be no way to improve upon the variables themselves as components. For the current study, Bartlett's test was significant at the $p < .000$ level, indicating that the study's matrix is significantly different from the Identity Matrix; therefore, it is factorable, as there is variance within the matrix (chi-square = 206.661; df = 45; $p < .000$). Bartlett's test indicated that that the correlation between variables did not approximate the Identity Matrix and that the variables were hence sufficiently correlated to support the use of a factor analysis (Harman, 1976).

Factor Analysis

Principal component analysis (PCA) involves a mathematical procedure that transforms a number of (possibly) correlated variables into a (smaller) number of uncorrelated variables called principal components. The first principal component accounts for as much of the variability in the data as possible, and each succeeding component accounts for as much of the remaining variability as possible (SPSS, 1999). Though the hypotheses of this study proposed Oblimin rotation (as it was hypothesized that the variables were all intercorrelated), usually associated with PCA is the Varimax rotation, which tries to maximize the variance of a factor. Thus, each factor has a small number of large loadings and a large number of small loadings. By maximizing the variance, it is easy to see which variables load on which factors.

Varimax also configures the data to form orthogonal (uncorrelated) factors by maximizing the unique variance within each factor and minimizing the amount of variance shared between them. To decrease the amount of standard error of the factor loadings, the Kaiser normalization was used. The Kaiser normalization functions to adjust the values and to provide the simplest possible structure (SPSS, 1999). The factor loadings were realized in three iterations, indicating that the underlying structure that was found by this study is a strong series of associations. The low number of iterations it took to realize the factor loadings provides compelling evidence that the factor structure that was found is a natural fit to the data.

Factor Rotation

As it had been hypothesized that the variables would be significantly correlated, the data was initially factor analyzed using an Oblimin rotation. However, the two resultant factors proved not to be significantly related, correlating at $r = .415$; this entails that less than 20 percent of the variance of the factors were shared. Furthermore, altogether, the factors from the Oblimin rotation only accounted for five percent of the total variance among the variables, a paltry amount that indicated that this sort of rotation (and hence the hypothesized correlated-variables-model) was a poor fit to the sample variance.

For this reason, an uncorrelated factors model was tried via the Varimax rotation. When the analysis was run the two factors explained approximately 50 percent of the variance of the variables. Prior to the Varimax rotation, factor one explained 38 percent of the variance and factor two explained ten percent of the variance. However, after rotation, each factor accounted for a little more than 24 percent of the variance. Hence, the uncorrelated-variables model proved much more robust in explaining sample variance.

Factor Extraction

As confirmation that a two-factor model was appropriate, eigenvalues and the Scree Plot were considered. The eigenvalue is the sum of the squared loading values of the variables onto a given factor. To be worth considering, its value must exceed that of one variable's squared variance, or 1.0 (Lawley & Mazwell, 1971). Based on the practice that factors should be extracted if the eigenvalue is higher than one, two factors were extracted from this data (Morrison, 1990). Furthermore, Cattell's Scree Plot flattened at a low level after two peaks, indicating that any factor identified after the second peak would explain less variance than one variable (Morrison, 1990).

Communalities

As for how well the variables contributed to the factors, the communalities listed below are a measure of the proportion of variance that each item has in common with other items, i.e. how much of the variance is attributable to the factors. The proportion of variance that is unique to each item is then the respective item's total variance minus the communality. The initial communalities were 1.0, which means that 100 percent of the variance within a given variable is explained by itself. This is altered by segmenting out amounts related to factors. The extracted communalities are the percent of variance in each variable explained by only the factors that were extracted. Because there should be fewer factors than variables, this number should be less than 1.0.

The variable that contributed the most to the factors was Nonverbal Fluid Reasoning, with a shared variance due to the underlying factors that was approximately 70 percent. Verbal Quantitative Reasoning shared close to 70 of its variance with the other variables, and Verbal Fluid Reasoning shared close to 64 percent. The three variables that contributed the least to the factors were Nonverbal Knowledge (16%), Nonverbal Quantitative Reasoning (35%), and Verbal Knowledge (38%).

Table 7.

Rescaled Communalities

Subtest	Initial	Extraction
Verbal		
Fluid Reasoning	1.000	0.639
Knowledge	1.000	0.377
Quantitative Reasoning	1.000	0.694
Visual-Spatial	1.000	0.489
Working Memory	1.000	0.457
Nonverbal		
Fluid Reasoning	1.000	0.708
Knowledge	1.000	0.165
Quantitative Reasoning	1.000	0.346
Visual-Spatial	1.000	0.511
Working Memory	1.000	0.558

Factor Loadings

The resultant factor structure that was found was a two factor model, with five variables loading on the first factor, four on the second factor, and one split between the two factors. The cut-off point that was used to decide on which factor the variable loaded onto was 0.405 (Feinstein, Fallon, Petkva, & Liebowitz, 2003).

Table 8.

Rotated Component Matrix Extraction Method: Principal Components Analysis.

	Rescaled Component	Rescaled Component
Subtest	Factor One	Factor Two
Verbal		
Fluid Reasoning	**0.797**	0.063
Visual-Spatial	**0.657**	0.240
Knowledge	**0.603**	0.115
Quant. Reasoning	**0.559**	**0.617**
Working Memory	0.302	**0.605**
Nonverbal		
Quant. Reasoning	**0.483**	0.335
Knowledge	**0.324**	0.244
Fluid Reasoning	-0.115	**0.833**
Working Memory	0.359	**0.655**
Visual-Spatial	0.404	**0.590**

Rotation Method: Varimax with Kaiser Normalization. N = 78. Bold face denotes major factor membership.

On the first factor, the variables that loaded were Verbal Fluid Reasoning, Verbal Visual-Spatial Reasoning, Verbal Knowledge, Nonverbal Quantitative Reasoning, and Nonverbal Knowledge. The variables that loaded on the second factor were Nonverbal Fluid Reasoning, Nonverbal Working Memory, Verbal Working Memory, and Nonverbal Visual-Spatial Reasoning. Verbal Quantitative Reasoning was split relatively evenly between the two factors, with factor one accounting for approximately 31 percent of its variance and factor two accounting for approximately 36 percent of its variance.

Table 9.

Factor Membership

General Knowledge	**Ordering/Sequencing Information**
Verbal Fluid Reasoning	Nonverbal Fluid Reasoning
Verbal Visual-Spatial Reasoning	Nonverbal Working Memory
Verbal Knowledge	Verbal Working Memory
Nonverbal Quantitative Reasoning	Nonverbal Visual-Spatial Reasoning
Nonverbal Knowledge	*Verbal Quantitative Reasoning*
Verbal Quantitative Reasoning	

The first factor was named General Knowledge (GK) and the second factor Ordering/Sequencing Information (OSI). Of note, Nonverbal Knowledge failed to reach the cut-off of .405, as its loading on factor one was .324, while its loading on factor two was .244. However, it was included on factor one (as opposed to being cut entirely) as GK was associated with ten percent of the variance of Nonverbal Knowledge while OSI only accounted for approximately five percent of its variance, suggesting that it was relatively more affiliated with the first factor. The generally weak performance of this variable was well-predicted by the communalities of this assessment.

Post-hoc Analyses

Post-hoc correlation analyses (see appendix A) were conducted to assess the validity of our factors. Factor one, GK, was significantly correlated with Vineland Adaptive Behavior Scales (VABS) Communication (r = .000), VABS Daily Living Skills (r = .014), DAR Word Identification (r = .029), and K-TEA Reading

Comprehension (*r* = .019). Both OWLS Written Expression and PIAT Math approached (*r* = .055) significance when correlated with GK. Factor two, OSI, was significantly correlated with VABS Communication (*r* = .033), OWLS Written Expression (*r* = .023), PIAT Math, and DAR Spelling (*r* = .010). Both DAR Word Identification (*r* = .060) and K-TEA Reading Comprehension (*r* = .058) approached significance with regards to their correlations with OSI.

Discriminant Analyses

Originally, post hoc regression analyses were proposed; however, we chose to use post-hoc discriminant analysis as were not trying to make predictions of values based on a continuous variable but rather were attempting to predict group membership. Though we used a different analysis than originally proposed, “it is unlikely that the two methods will give markedly different results” as they share similar statistical purposes and formulae (Press & Wilson, 1978, 705).

Initially, group membership was based simply on whether or not one was diagnosed with Attention-Deficit/Hyperactivity Disorder (ADHD), Anxiety Disorder NOS, Depression, Bipolar Disorder, or Oppositional Defiant Disorder (ODD). Then, Bipolar and Depression were combined into a Mood Disorders group, while a more broad Anxiety Disorders Group (ADG) was created to include those diagnosed with Anxiety NOS, Posttraumatic Stress Disorder (PTSD), and Generalized Anxiety Disorder (GAD). It was observed that the intersection of ADHD group and the ADG was fairly sizeable (n = 33) and this dual diagnosis group (ADHD/ADG) was tested as well. See Appendix B. for a comprehensive list of diagnostic groupings.

Unfortunately, apparently there was not enough variance to explain more than one variable for any of the groups (which is somewhat expectable with binary group designations – with zero for nongroup membership and one for group membership).

Table 10

Discriminant Analyses

Diagnosis	**SB5 Scale**	**f-value**	**sig.**	**df1**	**df2**	**%**
ODD	Verbal Working Memory	12.145	0.001	1	76	74
ADG	Verbal Working Memory	5.889	0.018	1	76	74
ADHD	Verbal Working Memory	8.116	0.006	1	76	62
Anxiety-NOS	Verbal Fluid Reasoning	10.076	0.002	1	76	69
Depression	Verbal Visual-Spatial	4.665	0.034	1	76	89
Mood	Nonverbal Vis.-Spatial	4.050	0.048	1	76	85

ADHD group membership was correctly predicted by scores on Verbal Working Memory 62 percent of the time. Because the correlation between Verbal Working Memory and ADHD was negative (r = -.311; p=.006), as ADHD is present, scores on Verbal Working Memory go down. Verbal Working Memory also predicted membership in the ADG and ODD groups. Interestingly, membership in either group was accurately predicted by Verbal Working Memory 74 percent of the time, and both of the correlations between ADG and ODD groups and Verbal Working Memory were positive (r =.268; p =.018; and r =.371; p =.001, respectively). This indicated that as ADG and ODD scores become positive, so do scores on Verbal Working Memory.

When Verbal Visual Spatial Processing scores were used to predict group membership in the Depression group, they were able to precisely predict membership

89 percent of the time. Furthermore, Depression was significantly correlated with Verbal Visual Spatial Processing (*r* =.240; p =.034), indicating that as Depression increases, so does Verbal Visual Spatial Processing. For the Mood group, Nonverbal Visual-Spatial Processing was able to correctly predict membership 85 percent of the time and the correlation between them was positive (*r* =.225; p =.048), indicating that as one increases so does the other. The last variable that was able to predict group membership was Verbal Fluid Reasoning, which was able to truthfully predict membership in the Anxiety – NOS group 69 percent of the time. The Anxiety – NOS group was positively correlated with Verbal Fluid Reasoning (*r* =.342; p =.002), indicating that as anxiety increases so do Verbal Fluid Reasoning scores.

Table 11

Correlation Analysis; n = 78

Subtest	ODD	ADHD	ADG	Depress.	Mood	Anx-NOS
Verbal Visual-Spatial						
Correlation	0.104	0.262**	*0.159	0.240*	0.167	0.107
Significance	0.366	0.020	0.163	0.034	0.145	0.351
Verbal Working Memory						
Correlation	0.371**	*-0.311**	*0.268*	0.120	0.076	0.236*
Significance	0.001	0.006	0.018	0.296	0.510	0.038
Nonverbal Visual-Spatial						
Correlation	0.217	-0.087	0.164	0.209	0.225*	0.212
Significance	0.056	0.451	0.152	0.066	0.048	0.062
Verbal Fluid Reasoning						
Correlation	0.243*	-0.129	0.243*	0.055	0.124	0.342***
Significance	0.032	0.259	0.032	0.631	0.297	0.002

**Correlation is significant at the 0.01 level (2-tailed)
*Correlation is significant at the 0.05 level (2-tailed)

Because three of the groups were significantly correlated with Verbal Working Memory, it was possible that ODD, ADG, and ADHD were intercorrelated. To address this possibility, the intercorrelations between the three groups were examined, as were their sample distributions. The three demographic/diagnostic variables failed to intercorrelate (see table 12) suggesting some functional independence. Then, when the cross-hatched distribution grid of the three diagnoses was considered (see table 13) it was fairly clear that each condition was distinct. Despite fair overlap between the ADHD and ADG groups, a greater number remained separate and hence unique to even group, helping the groups retain their individuality. Therefore, it is possible that Verbal Working Memory operates differently for each.

Table 12.

Correlation Matrix of Diagnostic Categories that are Significantly Predicted by Verbal Working Memory

		ADG	ADHD	ODD
ADG	Pearson corr.	1.000		
	2-tail sig			
ADHD	Pearson corr.	-0.080	1.000	
	2-tail sig	0.489		
ODD	Pearson corr.	-0.005	-0.015	1.000
	2-tail sig	0.966	0.897	

**denotes significant correlations at the .01 level

Table 13.

Sample distribution by diagnosis, unique and comorbid

	ADG	ADHD	ODD
ADG	unique: n = 57		
ADHD	comorbid: n = 33	unique: n = 47	
ODD	comorbid: n = 33	comorbid: n = 16	unique n: = 22

Number comorbid for membership in all three groups = 10

A second discriminant analysis of gender and SB5 scale scores was run in an attempt to predict gender from SB5 scale scores. The discriminant analysis indicated that none of the SB5 subscales were able to predict membership in either gender group. Further analyses were conducted, in the form of simple t-tests, in an attempt to compare the group means of male and female across SB5 subscales. It was discovered that males and females did not perform significantly differently on any of the SB5 subscales.

CHAPTER FIVE: CONCLUSIONS AND RECOMMENDATIONS

Communalities

Communalities are the sum of the squared factor loadings for each variable. Therefore, it is the percent of variance in a given variable explained by all of the factors. Most of the variables (five) contributed roughly half of their variance. Nonverbal and Verbal Fluid Reasoning were the best contributors, each at just over two thirds of their variance. Both Verbal and Nonverbal Fluid Reasoning tasks involve drawing deductions from the data that is available; with Verbal Fluid Reasoning drawing upon what is known, and Nonverbal Fluid Reasoning drawing on what is seen. It seems that extrapolating from what you know and from what you see are both foundational to intelligence.

Furthermore, Nonverbal Quantitative Reasoning and Verbal Knowledge were among the poorest contributors, each at just over a third of their variance. So what about these two subtests are not relating to either factor? Verbal Knowledge requires extensive, descriptive verbalization, a skill which in and of itself is neither a General Knowledge nor a Sequencing ability per se, though it is a medium for conveying information. Nonverbal Quantitative Reasoning did not seem to relate to General Knowledge, but did have something of an Ordering/Sequencing component; however, this was apparently not predominant enough to relate more strongly to this scale. The visually-inspected intuitive comparisons between amounts involved may well represent unique content.

Nonverbal Knowledge was by far the worst contributor at only 16 percent. Nonverbal Knowledge may have been the worst predictor because it because it is a task that hence requires an individual to evaluate real world scenes and then make a deductive leap as to what is wrong with them. It is the only task that involves such depictions in such volume and one of the only two tasks (with Verbal Knowledge) that hence requires extensive verbal description, making the intuitive deductions unique in the battery.

Factor Analysis

Factor One

The variables that loaded on factor one all reflected an individual's fund of knowledge. Therefore, this first factor was named General Knowledge (GK). However, the reader should proceed with care here; typically, when clinicians refer to a "fund of knowledge," they are usually referring to verbal information. However, this study found that this information repertoire is not always strictly verbal and that some of the variables that aligned themselves with the GK factor were in fact "nonverbal" variables by Roid's classification. This may have happened because most of the SB5's nonverbal tasks involve some degree of verbal prowess, either in comprehending the spoken directions or in mediating the content of replies. This is probably why the expected pure verbal/nonverbal dichotomy was not found by factor analysis. However, there may also have been pull for these variables based upon the nonverbal intellectual content involved being attracted to the informational content of the verbal scales.

An examination of the contents of GK is in order. Verbal Fluid Reasoning requires the child to describe scenes, to understand absurdities, and to comprehend abstract relationships between words; all of which require the individual to have a fund of information about real-world circumstances. This is intriguing since Roid stated in the SB5 manual that fluid reasoning subtests were supposed to be free of all academic knowledge (Roid, 2003a, 41). Perhaps this may indeed function as a "common sense," deductive measure for adults as Roid intended, but a child nonetheless needs some subject mastery of facts to navigate this, and it hence fits well on this factor.

The Verbal Visual-Spatial subtest requires that the examinee understand directional relationships. In order to complete the tasks, mastery of verbal directional terms and prepositions is needed. The subtest of Verbal Knowledge begins with identifying body parts and continues until the individual is providing definitions of words, which of course, demands a general fund of knowledge in order to answer. As with VIQ, the verbal information dynamic of these scales is intuitively obvious as an explanation of factor membership here.

This was not the case with the next variable to land on GK: Nonverbal Quantitative Reasoning. This subtest requires that the examinee understand verbal directions pertaining to amount and measurement. Thus, although the subtest is considered by Roid to be nonverbal due to the concepts involved, a verbal repertoire is required to understand the instructions and in turn to answer the questions. It was unclear whether the fact based content or the verbal comprehension of directions and item content needed is the reason for this variable's attraction to this factor. The last variable that loaded onto GK was Nonverbal Knowledge. This subtest requires the individual to identify body parts and to verbalize what is absurd or incongruous about pictures – again involving a verbal component. This furthermore related to the General Knowledge because in order to be able to identify incongruity in pictures, the subject must first have an understanding of what a normal picture should look like. Of note, however, the Nonverbal Knowledge variable's factor loadings on GK were paltry at best, and indicated that only approximately ten percent of the variance was due to GK. Though this amount of variance accounted for by GK was rather minimal, factor two accounted for even less, at about six percent. Hence, nonverbal knowledge likely belongs here, though as a minimal contributor. Again, perhaps the visual deductive leaps needed (beyond mere common-sensical knowledge about the world) were unique and not easily related to either factor here.

Factor Two

The variables that loaded onto the second factor had one major quality in common, which was that they required the ability to order and sequence information. Hence, this second factor was named Ordering/Sequencing of Information (OSI). Nonverbal Fluid Reasoning is dependent upon the ability to match objects and to determine the sequence of objects – first in linear rows and later in completing two-dimensional (up/down, right/left) grids of symbols and shapes. Nonverbal Working Memory is perhaps the task most closely linked with the sequencing concept in that the individual must retap the blocks that the examiner has tapped. At lower levels, the sequence must be exactly the same, which is similar to the other subtests on this factor in requiring the individual put words or objects into an order correctly. Higher levels require manipulation of the order according to simple rules, moving it from a short term memory task to a working memory task. Similarly, Verbal Working Memory requires the individual initially to repeat simple sentences, and later to retain targeted pieces of information taken from a series of verbal statements in a sequential order. Nonverbal Visual-Spatial Processing abilities were measured by the SB5 tasks where the individual must put pieces of a picture into a certain pattern based on a pictorial model order, which is akin to sequencing in that the individual must actively guide the placement and organization of the geometric shaped component pieces.

Though the OSI factor did not mirror the traditional NVIQ construct, there were many similarities. Three of the four subtests that loaded onto the OSI factor were deemed nonverbal by Roid: Nonverbal Fluid Reasoning, Nonverbal Working Memory, and Nonverbal Visual-Spatial. The last variable that loaded onto the OSI factor was Verbal Working Memory. Because of the majority of variables that loaded onto the OSI factor were nonverbal, the authors of this study felt as though comparisons with previously established measures of NVIQ were relevant. However, the two are not interchangeable. OSI is distinct with a focus on ordering that transcends the nonverbal/verbal distinction in component material, while NVIQ is intended as a more broad-reaching construct that includes examining pictorial scenes (as in Nonverbal Knowledge), and reading and math applications (as in Nonverbal Quantitative Reasoning).

<u>Split Variable</u>

There was one variable that was split between the factors: Verbal Quantitative Reasoning. Verbal Quantitative Reasoning initially requires the individual to count, and then progresses from simple mathematical equations to more complex word problems. This subtest requires both an individual's fund of knowledge (math functions and learned real-world relationships between objects) as well as the ability to sequence (in counting and in deducing the next item to arise in series of tiles with numbers of symbols on each determined by progressions based upon mathematical functions). The subtest progresses to its most sophisticated extreme of complex word problems which require both verbal comprehension as well as sequencing abilities in mentally grouping the order of facts.

Post-Hoc Correlation Analyses

Correlation analyses were run to examine how GK and OSI relate to pre-established verbal and nonverbal measures. The rationale behind doing this was that if positive correlations were found, results would clarify what role each of the newfound factors might play in each area relative to the more conventionally-defined verbal and nonverbal functions.

GK's correlation with the VABS Communication score was significant and indicated that these two sets of scores shared an underlying dynamic that drew the two together (see Appendix A). As the VABS Communication score is a measure of verbal communication, we can infer that GK reflects a related body of knowledge; however, the abilities underlying both maybe reflected differently on the two measures, as GK is task-based and VABS is derived from parent ratings. It is likely that the verbal skills that are assessed by the VABS Communication score reflect the fund of knowledge that is the foundation of GK. Similarly, the DAR Word

Identification and K-TEA Reading Comprehension measures were also significantly correlated with GK. This lends support to GK playing an active role in some traditionally designated verbal skill domains.

However, in a somewhat less straightforward way, the VABS Daily Living Skills score was also found to be significantly correlated with GK, indicating again that there may be a nonverbal component to GK as well. However, it must be noted that the VABS Daily Living Skills domain involves procedural knowledge and could hence also involve a verbal component of rehearsed self-cueing. Regardless of whether nonverbal actions or verbally-learned guidelines are involved, the resultant knowledge base needed to perform the skills is reflected in the significant correlation between GK and the VABS Daily Living domain skills domain score.

What is interesting is that although the majority of the subtests that loaded onto OSI were "nonverbal," the majority of significant correlations that were realized with this factor were with previously established verbal measures. The verbal measures that correlated significantly with OSI were VABS Communication, OWLS Written Expression, and DAR Spelling. Though the latter two are traditionally considered to be verbal measures, the ability to sequence and order is needed to perform these tasks proficiently. That is, spelling is a series of letters, while writing entails organization of words (grammar, ordering ideas, etc.). A spoken equivalent to the latter element may be involved in VABS Communication which also includes an academic skills subscale that taps reading and writing skills as well. The only traditionally nonverbal measure that correlated significantly with OSI was the PIAT

Math, likely because a mastery of numerical order is needed in order to complete both simplistic as well as more complex mathematical problems.

Discriminant Analysis

For the three diagnostic groupings for which Verbal Working Memory was a significant predictor, the question remained as to whether these three groups overlapped so significantly that they all related to the variable the same way, or whether they were so distinct that Verbal Working Memory related with each distinctly. Refer to the results section for a complete illustration.

Discriminant analyses of the ADHD, ADG, and ODD groups yielded the findings that, of the ten SB5 subscales, Verbal Working Memory was the best predictor of membership for each. It was also established that, despite a healthy overlap between the ADHD and ADG subgroups in this sample, these three diagnostic groups were distinct (see tables 9 and 10) and may relate to Verbal Working Memory in three distinct ways.

ODD and ADG correlated positively with Verbal Working Memory, indicating that as they increased, so did scores on Verbal Working Memory. However, ADHD had a negative correlation with Verbal Working Memory, which means that as ADHD increased, Verbal Working Memory decreased. The positive correlation between ADG and Verbal Working Memory may be reflective of the hypervigilance that often accompanies anxiety. The positive correlation between ODD and Verbal Working Memory may be reflective of the verbal and memory skills that are needed to argue with another individual, in that the individual needs to listen more to be actively engaged in an ongoing argument; listening is a large part of Verbal Working Memory. The negative correlation between ADHD and Verbal Working Memory is rather obvious, as when attentional behaviors increase, working memory decreases due to the individual's inability to attend to the information.

The Anxiety–NOS group was significantly correlated with Verbal Fluid Reasoning, indicating that as anxiety increased so did Verbal Fluid Reasoning scores. This correlation may be again reflective of the hypervigilance that is associated with anxiety facilitating a tendency to be over-ready to draw conclusions from, and relationships between, stimuli via verbal reasoning. Depression was significantly correlated with Verbal Visual-Spatial Processing, indicating that as depression increases, so does Verbal Visual-Spatial Processing. The Mood group was significantly correlated with Nonverbal Visual-Spatial Processing, also a positive correlation. The reason for these associations with the Visual-Spatial Processing spectrum may be reflective of artistic intelligence. Some theorists claim that mood disorders are statistically overrepresented in those with artistic talents, which in turn would be associated with this type of intelligence (Barron & Harrington, 1981). If this relationship was true of even a few of our sample subjects, it may have inflated these relationships, which may in turn be symptomatic of over-fit to our clinical sample.

An Example

The following profile is a hypothetical example of the differences in interpretation of scores based initially on Roid's factor analysis and secondly on the factor analysis of this study.

The below table (14) is the SB5 profile of an almost 6 year old male, "Jack." Adhering to Roid's factor analysis, the interpretation of this profile is as follows:

> The SB5 is a measure that estimates an individual's current level of cognitive functioning. It yielded both an overall intelligence quotient as well as measures of verbal and nonverbal abilities. The Average score on the SB5 is 90-109, and Low Average is 80-89.
>
> Jack was administered the SB5, which yielded overall scores varying between the Average and the Low Average ranges, with verbal abilities better developed than nonverbal abilities. Jack's verbal scores tended to be Average, without much variability, which suggested that he has a good potential for academic learning in most subjects in school. However, he did have both relative strengths and weaknesses within the verbal domain.
>
> Jack's greatest relative strength was found in working memory, indicating that Jack is able to retain verbal information and then mentally manipulate it. Because his Verbal Working Memory is far superior to his Nonverbal Working Memory, we can infer that Jack has more of a proclivity for retaining simple verbal information than nonverbal information. This finding is consistent with the pattern of his overall IQ scores, and may suggest visual inattention.
>
> Both within the nonverbal domain and across domains, Nonverbal Quantitative Reasoning was Jack's greatest strength. In that one area, Jack demonstrated a High Average grasp of nonverbal concepts of quantity and comparative relationships. By contrast, though still just below the Average range, Jack's Verbal Quantitative Reasoning skills were a relative weakness for him and indicated that he is having at least minimal difficulty in the application of rudimentary math skills to complex word problems. However, Jack's greatest relative weaknesses, again both within the nonverbal domain and across the entire measure, were found on Fluid Reasoning and Working Memory tasks. This finding is consistent with poor mental flexibility and mental manipulation.

Table 14. Jack's profile with Roid's interpretation

Stanford-Binet Intelligence Scales – Fifth Edition (SB5)		
Scales	Standard Score	Percentile
Nonverbal		
Fluid Reasoning	4	2
Knowledge	10	50
Quantitative Reasoning	13	84
Visual-Spatial Processing	6	9
Working Memory	4	2
Verbal		
Fluid Reasoning	10	50
Knowledge	9	37
Quantitative Reasoning	7	16
Visual-Spatial Processing	9	37
Working Memory	12	75
Composite Scores		
Fluid Reasoning Composite	82	12
Knowledge Composite	97	42
Quantitative Reasoning Composite	100	50
Visual-Spatial Processing Composite	85	16
Working Memory Composite	89	23
Nonverbal Intelligence Quotient	83	13
Verbal Intelligence Quotient	96	39
Full Scale Intelligence Quotient	89	23

Adhering to the study's factor solution, interpretation of the below table (15) is as follows:

Table 15. Jack's profile with Chase's interpretation

Stanford-Binet Intelligence Scales – Fifth Edition (SB5)		
	Scaled Score	Percentile
Ordering/Sequencing Scales		
Fluid Reasoning (NV)	4	2
Working Memory (NV)	4	2
Working Memory (V)	9	37
Visual-Spatial Processing (NV)	9	37
Quantitative Reasoning (V)	7	16

Fund of Knowledge Scales		
Fluid Reasoning (V)	10	50
Visual-Spatial Processing (V)	9	37
Knowledge (V)	12	75
Quantitative Reasoning (NV)	13	84
Knowledge (NV)	10	50
Quantitative Reasoning (V)	7	16

Jack was administered a cognitive battery that evaluated his General Knowledge and Ordering/Sequencing abilities. General Knowledge refers to the information that one has learned from one's environment, which at the pediatric level is that which is gleaned from school. Ordering/Sequencing involves the ability to put concepts into a logical order, at the most basic level of which are activities such as counting and saying the alphabet. Overall, Jack's knowledge skills were greater relative strengths than were his sequencing abilities.

Within the General Knowledge domain, Jack's greatest relative strength was in Nonverbal Quantitative Reasoning, closely followed by Verbal Knowledge. At the most basic level, Jack demonstrated an understanding of stimuli magnitude, was able to make comparisons between two units of measure, and was able to identify simple words and pictures. At the most complex level, Jack was able to demonstrate a proclivity for understanding the values that are assigned to specific numbers, as well as to provide definitions for single words.

Within the Ordering/Sequencing domain, Jack's scores displayed more scatter between subtests than on the Knowledge subtests, indicating both relative strengths and weaknesses. Jack's greatest strength, with regard to sequencing abilities, was displayed on the Verbal Working Memory task. At the most basic level he was able to repeat strings of words in the proper sequence; and at the most difficult level, he was able to hold sequences of words in his working memory and then mentally manipulate them.

By contrast, Jack's greatest relative weaknesses within the Sequencing domain were Nonverbal Working Memory and Nonverbal Fluid Reasoning. These findings reflect difficulty with visual sequencing of objects and nonverbal pattern analysis. Despite having good general knowledge of the world, Jack is limited in how well he detects and anticipates logical orders that arise within it, perhaps relating to a visual-based inattention to such stimuli.

Both interpretations of the data set describe the child; however, what differs is the way in which the case was conceptualized. The interpretation using Roid's underlying factor structure relies heavily upon the verbal/nonverbal dichotomy. This is a problem because this dichotomy did not reveal itself in a pure form when factor analyzed by this study; therefore, the interpretation of the data set via Roid's theory may be based on a dichotomy that is not the most accurate approach to describing the child's needs and abilities. This would make great sense as it appears that almost all of Roid's nonverbal subtests have a verbal component. Though verbal and nonverbal abilities were reflected in this study's interpretation of the results, they were not strictly categorized as such and therefore allowed for the use of both verbal and nonverbal information within each identified domain (i.e. GK and OSI).

What is interesting about the model proposed by this study is that it offers a glimpse into the individual's fund of knowledge (GK) and then provides information about how the individual is using that information (OSI). As opposed to the work of Roid, who differentiates between verbal and nonverbal intelligence, this study differentiates between what information is known and then how that information maybe manipulated. Furthermore, both interpretations of Jack's scores indicated visual inattention. However, the Chase description is much more comprehensive at illustrating the dynamic at work among the scales as it breaks down visual inattention into manageable components, such as visual sequencing and nonverbal pattern analysis. Breaking down weaknesses into manageable parts is the first step in skill remediation.

Interpretation of the SB5: Then and Now

The SB5 provides global verbal and nonverbal scores as well as an overall full scale intelligence quotient. In addition to these scores are composite scores that address overall general abilities in five areas. These five areas correspond to the CHC five factor model by which the SB5 was designed. Included in the SB5 are ten subtests that can also be used for analysis and are divided among the five constructs into verbal and nonverbal processing for each. While all of this is well and good, it was discovered by this study that the five factor model on which the SB5 was constructed does not reliably hold true across clinical samples; and, as our sample of 78 approximated the general population, may well fail to hold true in nonclinical samples as well.

Certainly in the clinical sample that was used by this study, Roid's findings were not generalizable. This is of concern because clinicians and psychometricians tend to see more clinical samples of patients than non, including examiners who use the SB5. If Roid's factor structure is not generalizable, it is probable that the information that clinicians are providing to their patients, based on scores derived from this structure, is not what best describes the tested individual. In order for the clinician to convey the information in an understandable manner, the clinician must be well-informed of the interpretability of the measure. Roid's interpretation guide is based on the same five factor model (or, alternatively, two five sub areas being represented equally on two factors) that the test was based on, and if the five factor model is in fact not the true structure of the SB5, interpretative results will be skewed.

Due to the limitations of our clinical sample it is not clear that the factor structure that was discovered here is the only one that underlies the SB5, nor is it clear that it is the structure that best describes SB5 data, though there are some compelling signs in its favor (it explained at least half of the sample variance and converged in three iterations). What is clear is that Roid's factor model can be surpassed by another model, as the GK/OSI one presented here, and is hence not the best fit to the data of all individuals.

What does this tell us about intelligence?

Though the authors of this study acknowledge the presence of a general intelligence "g," the findings from the current study do not support a single factor model, but rather a dichotomous model of intelligence. However, the dichotomy uncovered by this study was not the traditional verbal/nonverbal breakdown of intelligence; neither was it the crystallized/fluid dichotomy per se. The dichotomy found by this study was a General Knowledge/Ordering-Sequential processing model. General Knowledge encompasses that which applies to a lifetime of learned knowledge. Order-based/Sequential processing involves placing information in a step-by step order, such as words in sentences, number series, or sequences of hand movements.

Though there are marked similarities between GK and VIQ, and OSI and NVIQ, many differences exist as well. Both GK and VIQ offer insight into the individual's fund of knowledge, VIQ does not parcel out the actual knowledge base from how the individual uses the learned information. GK focuses solely on the acquisition of information, while OSI focuses on the mental manipulations and applications of the learned material. Similarly, NVIQ provides for a nonverbal knowledge base but subsumes manipulation in there as well. By parceling out the manipulations and applications of information from the actual basis of knowledge, the potential of an individual is highlighted.

Limitations

The major limitation of this study was that a clinical sample was used. This is a limitation because the sample that Roid used to norm his test was nonclinical. In order for a true comparison between samples, demographic information should be matched. This was not the case in this study.

Furthermore, the study had an n large enough for analysis; however, the n was too small for many of the subgroup analyses that we wanted to run, for example, males versus females. Also, race and other demographic data, including SES and handedness, were not available for inclusion in this study.

Given our clinical sample and the inclusion of multiple calculations, it is possible that our model has been over-fit to our sample. To examine this limitation properly, the study would need to be replicated using either another clinical sample or a nonclinical sample. Another clinical sample could be used as it would answer our exact question of was the model over-fit to our data; however, if a nonclinical sample were used, it would answer the question of how generalizable is our model, or Roid's for that matter, to the general population.

Because the discriminant analyses were conducted primarily on categorical values it limited the sophistication of prediction, as categorical values tend to reduce variability. Therefore, it may have seemed that again, our model was over-fit to the data.

Conclusions and Future Directions

The main hypothesis of this study was that a factor structure other than Roid's would underlie the SB5 data. The specific hypotheses that were offered were:

1) A single factor, presumably *g*.
2) A verbal/nonverbal dichotomy.
3) A three-factor model. Save the visual spatial and quantitative reasoning subtests, the three remaining verbal subtest will load on one factor and the three remaining nonverbal subtests will load on another factor. The third factor will be in the four subtests that comprise visual spatial and quantitative reasoning abilities; these subtests will load on the same factor. These four subtests will have their own factor loading because they measure similar constructs and involve similar abilities, such as measurement and problem solving abilities.

Though our main hypothesis was confirmed by the uncovering of our two factor model, the specific hypotheses were not. Neither a single factor model, a verbal/nonverbal dichotomy, nor a three-factor model was found by this study. However, due to the limitations of this study it is not to say that future studies will replicate our findings.

It was also hypothesized that, in small subgroup correlations, the verbal factor of the SB5 (should one have been found) would correlate significantly with the standard scores of the verbally based adaptive, educational, and neuropsychological measures. Furthermore, it was hypothesized that the nonverbal factor of the SB5 (should one have been found) would correlate significantly with other nonverbal measures. However, because the factors did not come out as either purely verbal, or purely nonverbal, these hypotheses were not able to be taken to fruition.

This study uncovered an alternate factor structure than the five factor model put forth by Roid. As explained throughout, this is of interest because the interpretation of an intelligence measure is highly dependent upon the underlying factor structure of the model. Without a sound structure to underlie an intelligence measure, the findings and interpretability of the measure are compromised. This is of serious concern because of the widely used applications of many intelligence measures, specifically the SB5.

To ensure sound measure, the underlying structure of the model should be both replicable as well as generalizable. Furthermore, it should be generalizable across all populations for which the measure is intended. A future direction of this study would be to identify if Roid's five factor model holds up in a nonclinical population. Though the implications of this would speak to greater generalizability of the measure, it must be noted that the majority of individuals to whom intelligence measures are administered are in some type of clinical population. Therefore, generalizability across clinical populations is important and should also be considered a future goal of this study.

List of References

1. Alfonso, V. S., Flanagan D. P., & Radwan, S. (2005). The impact of the Cattell-Horn Carroll theory on test development and interpretation of cognitive and academic abilities. *Contemporary Intellectual Assessment, (2nd ed.). Theories, Tests, and Issues.* New York, NY: Guilford Publications.

2. American Psychiatric Association: *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision.* Washington, DC: American Psychiatric Association.

3. Atkinson L. (1990). Intellectual and adaptive functioning: some tables for interpreting the Vineland in combination with intelligence tests. *American Journal of Mental Retardation*, 95, 198-203.

4. Barron, F., & Harrington, D.M. (1981). Creativity, intelligence, and personality. *Annual Review of Psychology*, 32, 439-476.

5. Biederman, J., Newcorn, J., & Sprich, S. (1991). Comorbidity of attention deficit hyperactivity disorder with conduct, depressive, anxiety, and other disorders. *American Journal of Psychiatry*, 148, 564-577.

6. Binet A, Simon T. (1916). *The development of intelligence in children* (E.S. Kit, Translator). Baltimore: Williams & Wilkins.

7. 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 & Human Development*, 33, 165-72.

8. Burns, B., & Zeman, D. (1980). A comparison of laterality indices in college and retarded subjects. *The Journal of Psychology*, 104, 241-7.

9. Carpentieri, S., & Morgan S.B. (1996). Adaptive and intellectual functioning in autistic and nonautistic mentally retarded children. *Journal of Autism and Developmental Disorders*, 26, 611-620.

10. Carroll, J.B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.

11. Carroll J.B. (1997). Psychometrics, Intelligence, and Public Perception. *Intelligence,* 24, 25-52.

12. Carroll, J.B. (1997). The three-stratum theory of cognitive abilities. *Contemporary Intellectual Assessment: Theories, Tests and Issues.* New York: Guilford.

13. Carroll, J.B & Horn, J.L. (1981). On the scientific basis of ability testing. *American Psychologist*, 36, 1012-1020.

14. Cattell R.B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology.*

15. Chernick, M.R., & Friis R.H. (2003). *Introductory biostatistics for the health sciences.* John Wiley & Sons Inc.

16. Cohen, D. (1974). *Intelligence, what is it?* New York: M. Evans and Company.

17. Cohen, J. (1992). Quantitative methods in psychology: A power primer. *Psychological Bulletin*, 112, 155-9.

18. Crain, W. (2000). *Theories of development: Concepts and applications,* (4th ed.). New Jersey: Prentice Hall.

19. Darlington, R.B., Weinberg, S., & Walberg, H. (1973). Canonical variate analysis and related techniques. *Review of Educational Research*

20. Dawson, J.L. (1977). An anthropological perspective on the evolution and lateralization of the brain. *Annals of the New York Academy of Science*, 299, 424.

21. DeVellis, R.F. (1991). *Scale development: Theory and applications*. Newbury Park, CA: SAGE Publications.

22. Diessner, R., & Tiegs, J. (2001). *Notable selections in human development* (2nd ed.). Connecticut: McGraw Hill/Dushkin.

23. Dombrowski, DiStefano, & Noonan. (2004). *The Stanford-Binet Intelligence Scales* (5th ed.). The Communiqué.

24. Duncan, J., Seitz, R.J., Kolodny, J., Bor, D., Herzog, H., Ahmed, A., Newell, F.N., Emslie, H. (2000). A neural basis for general intelligence. *Science*, 289, 457-460.

25. Dunteman, G.H. (1989). Principal components analysis. Newbury Park, CA: Sage Publications, Inc.

26. Fancher, R.E. (1985). *The intelligent men: Makers of the IQ controversy*. New York: W. W. Norton & Company, Inc.

27. Feinstein, S.B., Fallon, B.A., Petkova, E., Liebowitz, MR. (2003). Item-by-Item factor analysis of the Yale-Brown Obsessive Compulsive Scale Symptom Checklist. *Journal of Neuropsychiatry and Clinical Neuroscience*, 15, 187-193.

28. Gardner H. (1983). *Frames of mind*. New York: Basic Book Inc.

29. Gardner H. (1999). *Intelligence reframed.* New York: Basic Book Inc.

30. Gorsuch, R. L. (1983). *Factor analysis*. Hillsdale, NJ: Lawrence Erlbaum.

31. Gottfredson, L.S. (2002). Dissecting practical intelligence theory: Its claims and evidence. *Intelligence,* 30, 1-55.

32. Harman, H. (1976). *Modern factor analysis.* University of Chicago Press.

33. Horn, J.L. (1967). Intelligence-Why it grows, why it declines. *Trans-action*, 5, 23-31.

34. Horn, J.L. (1968). Organization of abilities and the development of intelligence. *Psychological Review*, 75, 242-259.

35. Horn, J.L. (1978). Human ability systems. *Life-Span development and behavior*, 1, 211-256. New York: Academic Press.

36. Horn, J.L. (1978). The nature and development of intellectual abilities. *Human Variation: The Biopsychology of Age, Race, and Sex.* New York: Academic Press.

37. Horn, J.L. (1979). Trends in the measurement of intelligence. *Intelligence*, 3, 229-239.

38. Horn, J.L. (1987). A context for understanding information processing studies of human abilities. *Speed of Information-Processing and Intelligence.* Norwood: Ablex.

39. Horn, J.L. (1998). A basis for research on age differences in cognitive capabilities. *Human Cognitive Abilities in Theory and Practice.* Mahwah: Erlbaum.

40. Horn, J.L., & Cattell, R.B. (1967). Age differences in fluid and crystallized intelligence. *Acta Psychologica*, 26, 107-129.

41. Horn, J.L., & Knapp, J.R. (1973). On the subjective character of the empirical base of Guilford's structure-of-intellect model. *Psychological Bulletin*, 80, 33-43.

42. Hull, J.T., & Thompson, J.C. (1980). Predictive adaptive functioning of mentally retarded person in community settings. *American Journal of Mental Deficiency*, 3, 253-261.

43. Jensen, A. (1980). *Bias in mental testing*. New York: The Free Press.

44. Jensen, A.R. (1998). *The G factor: The science of mental ability.* Westport: Praeger.

45. Jensen, A.R., (2002). Galton's legacy to research on intelligence. *Journal of Biosience*, 34, 145-172.

46. Joreskog, K.G., & Lawley, D.N. (1968). New methods in maximum likelihood factor analysis. *British Journal of Mathematical and Statistical Psychology,* 21, 85-96.

47. Kalat, J.W. (1998). *Biological psychology* (6th ed.). Brooks/Cole Publishing Company: Pacific Grove.

48. Kaplan, S.L., & Alfonso, V.C. (1997). Confirmatory factor analysis of the Stanford-Binet Intelligence Scale: Fourth Edition with preschoolers with developmental delays. *Journal of Psychoeducational Assessment*, 15, 226-236.

49. Kaufman, A.S., & Kaufman, N.L. (1983). *K-ABC: Kaufman assessment battery for children.* American Guidance Company: Circle Pines, MN.

50. Kezar, A. (2001). Theory of multiple intelligences: Implications for higher learning. *Innovative Higher Education*, 26(2), 141-154.

51. Kolb, B., & Wishaw, I.Q. (1996). Fundamental of Human Neuropsychology, (4th ed.). W.H. Freeman and Company.

52. Kolb, B. (1989). Brain development, plasticity, and behavior. *American Psychologist*, 44, 1203-12.

53. Kristjansdottir, R., Uvebrant, P., Wiklund, L.M. (2000). Clinical characteristics in children with white matter abnormalities. *European Journal of Pediatric Neurology,* 4, 17-26.

54. Lawley, D.N. & Maxwell, AE. (1971). *Factor analysis as a statistical method.* London: Butterworth and Co.

55. Micceri, T. (1989). The unicorn, the normal curve, and other improbable creatures. *Psychology Bulletin,* 105(1), 156-166.

56. Mohan, A., Singh, A.P., Mandal, M.K. (2001). Transfer and interference of motor skills in people with intellectual disability. *Journal of Intellectual Disability Research*, 45, 361-69.

57. Morrison, D.F. (1990). *Multivariate Statistical Methods*. New York: McGraw-Hill.

58. Neisser, U., Boodoo, G., Bouchard, T.J., Boykin, W.A., Brody, N., Ceci, S.J., Halpern, D.E., Loehlin, J.C., Perloff, R., Sternberg, R.J., Urbina, S. Intelligence: knowns and unknowns. (1996). *American Psychologist,* 51, 77-101.

59. Paquette, C., Tosoni, C., Lassonde, M., Peretz, I. (1996). Atypical hemispheric specialization in intellectual deficiency. *Brain and Language*, 52, 474-83.

60. Pedhazur, E. (1982). *Multiple regression in behavioral research: Explanation and prediction* (2nd ed.). New York: Holt, Rinehart, and Winston.

61. Pennington, B.F., & Smith, S.D. (1988). Genetic influences on learning disabilities: an update. *Journal of Consulting and Clinical Psychology*, 56, 817-23.

62. Pipe, M.E. (1988). Atypical laterality and retardation. *Psychological Bulletin*, 104, 343-47.

63. Press, S.J., & Wilson, S. (1978). Choosing between logistic regression and discriminant analysis. *Journal of the American Statistical Association*, 73, 699-705.

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

65. Roid, G.H. (2003b). *Stanford-Binet Intelligence Scales – Examiner's Manual* (5th ed.). Itasca, IL: Riverside Publishing,.

66. Roizen, N.J., Blondis, T.A., Irwin, M., & Stein, M. (1994). Adaptive functioning in children with Attention-Deficit Hyperactivity Disorder. *Archives of Pediatric and Adolescent Medicine*, 148, 1137-42.

67. Salvia, J., & Ysseldyke, J.E. (1995). *Assessment* (6th ed.). Boston: Houghton Mifflin Company.

68. Sattler, J.M. (2001) *Assessment of children: Cognitive applications* (4th ed.). San Diego: Jerome M. Sattler, Publisher, Inc.

69. Sternberg, R.J. (2000). The holey grail of general intelligence. *Science*, 289 (5478), 399-410.

70. Sternberg, R.J., & Kaufman, J.C. (1998). Human abilities. *Annual Review of Psychology*, 49, 479-502.

71. Taylor, F. (2005). Book reviews. *Annals of Science*, 60(2), 207-218.

72. Terman, L.M. (1911). The Binet-Simon Scale for measuring intelligence: Impressions gained by its application. *Psychological Clinic*, 5, 199-206.

73. Terman, L.M. (1916). *The measurement of intelligence*. Boston: Houghton Mifflin,.

74. Terman, L.M. (1921). A symposium: Intelligence and its measurement. *Journal of Educational Psychology*, 12, 127-133.

75. Thorndike, E.L. (1927). The measurement of intelligence. New York: Bureau of Publications, Teachers College, Columbia University.

76. Thorndike, R.L., & Hagen, E.P. (1977). *Measurement and evaluation in psychology and education* (4th ed.). New York: Wiley.

77. Thorndike, R.L., Hagen, E.P., & Sattler, J.M. (1986). *Guide for administering and scoring the Stanford-Binet Intelligence Scale: Fourth Edition*. Chicago: Riverside.

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

79. Thorndike, R.M. (1990). Would the real factors of the Stanford-Binet Fourth Edition please come forward? *Journal of Psychoeducational Assessment*, 8, 412-435.

80. Tabachnick, B.G., & Fidell, L.S. (1996). *Using multivariate statistics.* (3rd. Ed). New York: HarperCollins College Publishers,.

81. Viney, W., & King, D.B. (2003). *A history of psychology: Ideas and context.* Boston: Allyn and Bacon.

82. Vogel, S.A. (1990). Gender differences in intelligence, language, visual-motor abilities, and academic achievement in students with learning disabilities: a review of the literature. *Journal of Learning Disabilities*; 23(1), 44-52.

83. Williams, W.M. (1996). Consequences of how we define and assess intelligence. *Psychology, Public Policy, and Law*, 2(3/4), 506-535.

84. Wine, J. (1971). Test anxiety and direction of attention. *Psychology Bulletin*, 76(2), 92-104.

85. Zafeiriou, D.I. (2004). Primitive reflexes and postural reactions in the neurodevelopmental examination. *Pediatric Neurology*, 31, 1-8.

APPENDIX A: CORRELATION ANALYSES

Correlations

		Regression Factor Score One	Regression Factor Score Two
VABS Comm.	Pearson corr.	0.531*	0.334*
	2-tail sig	0.000	0.033
	N	41	41
VBAS DLS	Pearson corr.	0.318*	0.052
	2-tail sig	0.014	0.745
	N	41	41
VBAS Soc.	Pearson corr.	0.292	0.109
	2-tail sig	0.064	0.499
	N	41	41
Word ID (DAR)	Pearson corr.	0.718*	0.647
	2-tail sig	0.029	0.060
	N	9	9
Written Expr. (OWLS)	Pearson corr.	0.534	0.624*
	2-tail sig	0.055	0.023
	N	13	13
Math (PIAT)	Pearson corr.	0.591	0.833*
	2-tail sig	0.055	0.001
	N	11	11
Spelling (DAR)	Pearson corr.	0.514	0.764*
	2-tail sig	0.129	0.010
	N	10	10
Read Comp. (K-TEA)	Pearson corr.	0.662*	0.560
	2-tail sig	0.019	0.058
	N	12	12

*Correlation is significant at the .05 level (2-tailed)

APPENDIX B: DIAGNOSTIC CATEGORY

Diagnostic Category	Number of Cases
Total Anxiety	60
300.00	*36*
300.01	1
300.02	11
300.30	1
307.51	1
309.81	9
313.89	1
Total ADHD	47
314.00	11
314.01	32
314.90	4
Total ADHD+Anxiety	33
Total Bipolar Disorder	4
296.42	1
296.50	1
296.62	1
296.80	1
Total Depressive Dis.	9
296.30	1
296.32	2
300.40	5
311.0	1
Total Mood Disorders	14
296.90	1
Bipolar	4
Depressive	9
Total ODD	22
313.81	22

Vita

Danielle Chase
Philadelphia, Pennsylvania
United States of America

Education:

- Ph.D. Candidate, Clinical Psychology
  Drexel University, Philadelphia, PA (APA) Approved
  2005 (anticipated completion)
- M.S., Clinical Psychology, December 2002
  Drexel University, Philadelphia, PA
- B.A., Educational Psychology/Neuropsychology, May 2000
  Bucknell University, Lewisburg, PA

Clinical Experience:

- Franciscan Hospital for Children and Rehabilitation Center
  Brighton, MA, 2004-2005, APA Accredited Internship
  Supervisor: Mary Ann Mullin, Ph.D.
- Neuropsychological Private Practice
  Epilepsy Department, Thomas Jefferson University Hospital
  Philadelphia, PA, Summer 2004
  Supervisor: Joseph I. Tracy, Ph.D.
- St. Christopher's Hospital for Children
  Philadelphia, PA, 2003-2004
  Supervisor: Mitzie Grant, Ph.D.
- Drexel University Counseling Center
  Philadelphia, PA, 2003-2004
  Supervisor: Amy Henning, M.A.
- Heart Failure/Transplant Unit
  Hahnemann University Hospital, Philadelphia, PA, 2002-2004
  Supervisor: Ralph Petrucci, Ed.D.
- Eastern Pennsylvania Psychiatric Institute
  Philadelphia, PA, 2002-2004
  Supervisor: Dr. Ralph Petrucci, Ed.D.

Publications:

- Danielle Dilkes, Steven M. Platek. Syntax: The Evolutionary Stepchild? Behavioral and Brain Sciences, 27(4), 511-512, 2005.
- R. J. Petrucci, K. C. Truesdell, A. Carter, N. E. Goldstein, M. M. Russell, D. Dilkes, J. M. Fitzpatrick, C. E. Thomas, M. E. Keenan, L. A. Lazarus, N. D. Chiaravalloti, J. E. Trunzo, J. W. Verjans, E. C. Holmes, L. E. Samuels, J. Narula. Cognitive Dysfunction in End-Stage Heart Failure Patients Requiring Inotrope Infusion of Mechanical Cardiac Assist Devices: Another Parameter of End-Organ Damage? Submitted for publication.

Professional Presentations:

- 2002 National Academy of Neuropsychology
  Danielle Dilkes, Joseph I. Tracy, et.al. The Wada Procedure as a predictor of post-surgical seizure outcome in anterior temporal lobectomy patients.
- 2002 American Epilepsy Society
  Danielle Dilkes, Joseph I. Tracy, et.al. The Wada Procedure as a predictor of post-surgical seizure outcome in anterior temporal lobectomy patients.