To the Graduate Council:

I am submitting herewith a dissertation written by Connie Theriot England entitled "Predictive and Concurrent Validity of the Universal Nonverbal Intelligence Test and the TONI-II and Achievement." I recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Education.

R. Steve McCallum

R. Steve McCallum, Major Professor

We have read this dissertation
and recommend its acceptance:

Thomas W. George

Thomas W. George

Charles H. Hargis

Charles H. Hargis

R A Saudargas

Richard A. Saudargas

Accepted for the Council:

Cwminkel

Associate Vice Chancellor and
Dean of The Graduate School

PREDICTIVE AND CONCURRENT VALIDITY OF THE UNIVERSAL NONVERBAL INTELLIGENCE TEST AND THE TEST OF NONVERBAL INTELLIGENCE-II WITH ACHIEVEMENT

A Dissertation

Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Connie Theriot England

August, 1997

Dedicated to my mother

Evelyn D. Theriot

ACKNOWLEDGMENTS

I am most appreciative to the members of my committee for their time and effort devoted to this project. A special thanks goes to Dr. Steve McCallum, Chairperson of my committee, for his unrelenting support and guidance in directing this research.

A special thanks also goes to Mr. Joe Butt, School Psychologist at the Tennessee School for the Deaf, who enthusiastically participated in this project.

ABSTRACT

The primary purpose of this study is to investigate the correlation and predictive utility of three constructs presumed to underlie intelligence operationalized by the Universal Nonverbal Intelligence Test (UNIT; Bracken & McCallum, In Press) as memory and reasoning and operationalized by the Test of Nonverbal Intelligence-II as figural problem solving (TONI-II; Brown, Sherbenou & Johnsen, 1990). The criterion measure is achievement, as measured by the Math and Reading Comprehension subtests of the Peabody Individual Achievement Test - Revised (PIAT-R; Markwardt, 1989) and the Stanford Achievement Test - Eighth Edition (SAT-8; Madden & Gardner, 1989) for a population of hearing impaired students. An additional purpose is to investigate the effects of experiential deficits of hearing impaired children by comparing the scores from the UNIT and TONI-II between two subgroups of students with hearing impairments: those with hearing parents and those with non-hearing parents.

TABLE OF CONTENTS

LIST OF TABLES

Tables (cont.)

CHAPTER I

INTRODUCTION

Purpose

The primary purpose of this study is to investigate the predictive and concurrent utility of three constructs presumed to underlie intelligence operationalized by the Universal Nonverbal Intelligence Test (UNIT; Bracken & McCallum, In Press) as memory and reasoning and operationalized by the Test of Nonverbal Intelligence-II as figural problem solving (TONI-II; Brown, Sherbenou & Johnsen, 1990). The criterion measure is achievement, as measured by the Math and Reading Comprehension subtests of the Peabody Individual Achievement Test - Revised (PIAT-R; Markwardt, 1989) and the Stanford Achievement Test - Eighth Edition (SAT-8; Madden & Gardner, 1989) for a population of hearing impaired students. An additional purpose is to investigate the effects of experiential deficits of hearing impaired children by comparing the scores from the UNIT and TONI-II between two subgroups of students with hearing impairments: those with hearing parents and those with non-hearing parents.

Rationale

Use of Cognitive Constructs to Predict Academic Achievement for Students with Hearing Impairments

Intelligence tests, and more specifically the general ability factor measured by such tests, are important insofar as they provide predictive information about educational or occupational variables deemed important by society. In educational settings, measures of intellectual functioning are used primarily to predict academic performance or to aid in the identification of specific strengths or weaknesses in the student's learning ability. When an intelligence test provides useful information regarding future performance on a specific criterion, e.g., academic performance, the measure is said to have "predictive validity." Jensen (1980) reports that typical correlations between intellectual and achievement measures are around .50, with verbal indexes yielding higher correlations than nonverbal measures.

Assessment tools available to the professional diagnostician for use with those groups of students for whom traditional intelligence measures are contraindicated, i.e., those possessing limited English proficiency, or who are language, hearing, and/or motor impaired, are limited. An extensive review of the literature reveals two major problems associated with the intellectual assessment of persons with hearing impairments: few professionals have training in the assessment of persons with hearing impairments and a limited number of instruments which have criterion-related validity is available for use with this population.

In 1974, Levine conducted a survey of the psychological testing of persons with hearing impairments. His study revealed that only twenty percent

(20%) of the psychologists involved in the assessment of children with deafness were formally trained in the area of deafness. Of the psychologists that Levine surveyed their common concern was their inability to adequately communicate with their hearing impaired subjects.

Compounding the problem of having only a limited number of psychologists trained in the area of deafness, is the issue of inadequate testing materials for this population. "Researchers have usually preferred to use performance tests of intelligence with deaf children (Sullivan, 1982) or to take only the performance component of existing scales as an estimate of intelligence" (Brunich, 1981; Moores, 1982; Vernon, 1968). The most preferred tests of intelligence have been the WISC and WISC-R Performance Scales (Weschler, 1974), the Hiskey-Nebraska Test of Learning Aptitude (Hiskey, 1966) and the Leiter-International Performance Scale (Anderson & Stevens, 1970; Levine, 1974). Although the WISC-R does provide norms for persons with hearing impairments, and an administration adaptation standardized for use with persons with hearing impairments (Ray, 1979), of the three aforementioned tests, only the Hiskey-Nebraska was specifically developed for and standardized upon children with deafness. However, the norms are dated and it is not model based.

Furthermore, "despite the popularity of nonverbal IQ tests for psychoeducational assessment of deaf children, criterion-related validity is lacking" (Braden, 1989, p. 329). In a survey of published and unpublished works, Braden identified only one study reporting WISC-R PIQ-criterion correlations (Hiroshoren, Hurley & Kavale, 1979). In this research the correlations between WISC-R PIQ and Stanford Achievement Test grade equivalents ranged from .09 to .35, and were considerably less than the expected moderate range (rs of .50 to .70) of criterion-related validity coefficients found with other populations. In a study of the correlations between SAT Subtests and Nonverbal IQs only limited correlations were revealed and these were well below any practical significance. Braden suggests that achievement may be an inappropriate criterion for correlation of IQ scores because deafness results in a drastic reduction in incidental learning of language and culturally-specific information. Additionally, he reasons that the low correlation between PIQ and academic achievement may result from a lack in overlap among the psychological constructs tapped by tests currently in use with this population.

Several authors take the theoretical position that persons with hearing impairments are intellectually normal but suffer from an experientially deprived learning environment. Jensen (1985) and Braden (1984) argue that "deprived" deaf children do not conform to "Spearman's Supposition" that the magnitudes of the mean differences between certain groups of individuals is directly related to the loadings of those measures on the hypothetical general intelligence factor "g." Jensen asserts that "Although there are significant ability differences between normally hearing and congenitally deaf children, the differences are not positively related to "g", if anything, they are negatively related to the tests' _g_ loadings. The language deprivation caused by deafness evidently takes its toll mostly on common factors other than "g", or on tests' specificities (i.e., nonerror variance that is not shared with any other tests in the battery) (Jensen, 1985, pp. 204-206). Jensen further asserts that many verbal tests are merely a reflection of a person's ability to recall stored information. On the other hand, nonverbal tests require more complex manipulations of elements to produce a correct answer. The difference in "g" loading for verbal and nonverbal tests indicate that nonverbal tasks rely more heavily on reasoning and inference than verbal tasks. "The verbal-nonverbal distinction is much more important than the "g"-loadedness of items for groups in which English is a second language" (Jensen, p. 161).

Current nonverbal intelligence measures often lack one or more essential components necessary for a good assessment. "For example, a "good" nonverbal measure of intelligence should not require any receptive or expressive language for administration, should assess multiple facets of intelligence, should assess higher order cognitive processes, and should provide materials and activities that are inherently interesting and motivating" (McCallum & Bracken, in press). The Universal Nonverbal Intelligence Test (UNIT; Bracken & McCallum, 1992) was specifically designed to fulfill this need.

The UNIT has its theoretical basis in Jensen's (1980) two tiered, hierarchial conceptualization of intelligence. Jensen provides a classification scheme of mental abilities which include Level I Associative skills, encompassing rote learning and memory, and Level II Cognitive skills including reasoning and problem solving abilities. The UNIT is designed to measure higher order mental functioning, such as reasoning and judgement, as well as more basic memory skills. Not only does it measure two primary intellectual abilities - memory and reasoning, but it also provides subtests which tap both symbolic and non-symbolic abilities. The less symbolically laden subtests require less internal mediation. For example, three of the UNIT's subtests (Symbolic Memory, Object Memory, and Analogic Reasoning) "require the use of concrete and abstract symbols to conceptualize the environment" (McCallum & Bracken, in press). Nonsymbolic subtests (Spatial Memory, Cube Design, and Mazes) "require the ability to perceive and make meaningful judgements about the physical relationships within our environment" (McCallum & Bracken, in press). Criterion-related validity studies of the UNIT for use with a population of hearing impaired individuals is needed. Specifically, there is a need to determine the predictive validity of the constructs assessed by the UNIT: memory, reasoning, nonsymbolic, ability, symbolic ability. Another cognitive construct with particular relevance for individuals with hearing impairments is figural problem solving. This construct has been operationalized by the TONI-II.

Summary

Assessment of persons with hearing impairments has typically included the use of the performance section of the WISC-R, or use of various other "nonverbal" instruments, i.e. Leiter International Performance Scale, (Leiter, 1979), the Hiskey-Nebraska Test of Learning Aptitude (Hiskey, 1966), the Columbia Mental Maturity

Scale (Burgemeister, Blum, & Lorge, 1972), and the Raven Progressive Matrices (Raven, Court, & Raven, 1986). All of these instruments suffer from one or more of the following problems: outdated norms, narrow conception of intellectual assessment, poor standardization, or lack of a theoretical basis. There is a need for a model-driven, theoretically based, multidimensional nonverbal measure of intelligence standardized for use with this population. A new test devoted to measuring intelligence nonverbally will soon be available. This test is designed to assess four important cognitive constructs: memory, reasoning, symbolic and nonsymbolic abilities. Predictive validity of these constructs is particularly needed for persons with hearing impairments. In addition there is a need to assess the predictive relationship among those constructs and another particularly relevant construct for this population - figural problem solving, as assessed by the TONI-II.

A review of assumptions related to standardized intelligence tests follows. First, an examination of factors relating to a test's predictive validity is presented. Then an overview of the theoretical constructs associated with intelligence test construction will be given.

Literature Review

Assessment of Cognitive Constructs

Historically I.Q. scores have had little importance in and of themselves; consequently, their usefulness is indicated primarily by their ability to accurately predict a person's performance on a criterion of true importance, for example, academic achievement. "When the correlation between scores on a particular test and some external criterion of practical importance is well established in a certain population, the test can be used to predict performance on the criterion" (Jensen, p.141, 1981). In educational contexts, achievement test scores and classroom performance are important variables for consideration in prediction. A test's "predictive validity" is related to the accuracy with which a test score predicts performance on another criterion. Recent efforts to develop tests which measure cognitive constructs presumed to underlie certain academic abilities culminated in development of the Woodcock-Johnson Cognitive Battery. The development of this test has stimulated anew the hunt for relationships between cognitive constructs and achievement.

Cognitive constructs: Historical antecedents

Generally proponents of standardized measures of mental ability rely on the following implicit assumptions regarding the use of intelligence tests: (a) the tasks tap cognitive processes underlying intellectual behavior; (b) the tasks embody the attributes that require intellectual behavior; and, (c) the tasks reflect equivalent prior experiences. Theoretical specificity regarding the nature of the cognitive processes underlying intellectual behavior, however, remains an elusive construct.

Sir Francis Galton is regarded as the father of the testing movement (Shouksmith, 1970). He developed two important statistical concepts - regression to the mean and correlation. Although his primary assumption regarding intelligence, (i.e., because knowledge of the environment reaches us through the senses, those with the highest intelligence would have the best sensory discrimination skills), generally proved to be invalid, his statistical analysis procedures allowed the study of psychometrics to flourish.

Recognizing the practical applicability of analyzing individual differences in behavior (i.e., selection of people for training and for diagnostic evaluations), James Cattell, in 1890, coined the term "mental test." Although his tests generally measured sensory and motor abilities and had little predictive validity for educational achievement or any other aspect of intellectual functioning, they did introduce psychology to the practical aspects of mental ability testing.

Early discussions regarding the nature of intelligence can be traced to Spearman's (1927) The Ability of Man. For Spearman, intelligence is composed of two factors, a general factor (g) and a number of very specific factors (Spearman, 1923). The g factor, as Spearman postulated, accounts for the majority of differences between groups of people and this single largest independent factor will be present in any mental test, regardless of form, content, or sensorimotor ability. As a result of Spearman's study on the nature of black-white differences on intelligence test measures, he formulated an hypothesis which stated that the mean black-white difference "was most saturated with g" (Spearman, 1927, p.370), suggesting that population differences were not attributable merely to cultural and linguistic idiosyncracies.

One of the earliest theories of multiple intelligences is attributed to Thurstone (Sattler, 1993). For Thurstone, (1947) intelligence is a composite of seven "vectors of the mind," i.e., perceptual speed, verbal comprehension and frequency, memory, reasoning, spatial ability, and numbers. Guilford (1967) offers an extension of Thurstone's theory in his Structure of Intellect Model. For Guilford, intelligence is composed of 120 traits which fall along three dimensions to produce requisite task-specific factors. These three dimensions include: activities or operations performed, the material or content on which the operations are performed, and the product that is the result of the operations.

Another perspective on multiple processes which theoretically undergird intellectual functioning is offered by Gardner in his "multiple intelligences" theory. Gardner suggests that individuals have eight mental skills, (i.e., musical, bodily-kinesthetic, logical-mathematical, linguistic, spatial, interpersonal, intrapersonal and naturalistic) which are the building blocks guiding the behaviors. His theory emphasizes the relatively independent nature of these "intelligences."

One of the most comprehensive theories of multiple intelligences is presented in the Cattell-Horn model (Horn & Cattell, 1966; Horn, 1988). Their theory holds that intelligence consists of "g" and is further conceptualized as two types of abilities - fluid and crystallized. Fluid intelligence refers to essentially nonverbal, relatively culture-free mental efficiency, and crystallized intelligence refers to acquired skills and knowledge that are strongly dependent for their development on exposure to the culture (Sattler, 1992, p. 48). They postulate that it is through the exercise of fluid intelligence that crystallized intelligence is developed.

More recently, Sternberg (1985) postulated that intelligence is comprised of three components: (a) metacomponents of higher-order executive processes, (b) performance components, and, (c) knowledge-acquisition components. In his triarchic theory of intelligence, Sternberg proposes three interrelated subtheories: (a) a componential theory which focuses on the internal mental operation of the individual; (b) a contextual theory which focuses on the external world of the individual; and, (c) an experiential subtheory which relates both the internal and external worlds of the individual. For Sternberg, the basic unit of analysis for explaining individual differences in intellectual functioning is the cognitive component. However, the use of these cognitive processes are examined within the context of meaningful learning situations or on tasks that are novel or to which individuals respond with automaticity.

Das and his colleagues proposed an information-processing model of intelligence based in part on Luria's work in neuropsychology (Sattler, 1991). They suggested that the two primary modes of processing information are simultaneous (integrated) and successive (sequential). They view intelligence as the ability to use the information obtained through these modes to "plan and structure one's behavior with an end in view" (Das, 1973, p.27).

Jensen proposed a hierarchical model of intelligence which delineates "Level 1" associative abilities (i.e., rote learning and short-memory skills), and "Level 2" abilities which require mental manipulation of sensory input to solve problems. "The crucial distinctions between levels I and II are differences in the complexity of the transformation and the mental manipulations that occur between the presentation of a given mental task and the response" (Sattler, 1992, p.53).

Many of the aforementioned constructs have been operationalized in the development of intelligence testing. Alfred Binet and Theodore Simon (1895) were noteworthy forerunners in the development of intelligence scales. Although they regarded intelligence as a collection of faculties, they developed their series of tasks on an empirical criterion - how well the tests differentiated between older and younger children. The items were ranked in order of difficulty and were used to diagnose degrees of mental retardation. What they thought the tests were measuring was based solely on their opinion, or face validity.

The next major development in the field of intelligence testing was introduced by David Wechsler in 1939. His scales, WISC, WISC-R, etc., were designed to measure the "global nature of intelligence" which he considered to be part of the larger whole of personality (Sattler, 1992). As Wechsler stated, "Unlike all other factors [g] cannot be associated with any unique or single ability; g is involved in many different types of ability; it is in essence not an ability at all, but a property of the mind" (1958, p. 124). He suggested that intelligence is composed of qualitatively different abilities, but did not provide empirical referents for his definitional terms (Sattler, 1991).

Generally, most early tests of intelligence lacked strong theoretical underpinnings and retained many of the characteristics of the Binet-Simon Scale. One of the first intelligence tests which did provide a theoretical framework in its conceptualization of intelligence was the Kaufman's Assessment Battery of Children (K-ABC). The K-ABC was developed from theory and research in cognitive psychology and neuropsychology. The test's authors conceptualized intelligence as an individual's ability to process information and they categorized these processes as either simultaneous or sequential. Although both praise and criticism have surrounded the K-ABC since its publication, it is a noteworthy forerunner to new tests based on various theoretical approaches to measuring cognitive abilities.

Nonverbal Measures of Intellectual Constructs

As previously noted there is a higher correlation between academic achievement and verbally-related constructs of intelligence than between academic achievement and nonverbal measures. However, when use of ordinary intelligence tests are not suitable, performance measures may be a better indicator of the true intellectual abilities of the child.

Tests developed Specifically as Nonverbal Measures of Intelligence

One of the first performance measures was introduced by Itard who attempted to train, by means of colored shapes and forms, "The Wild Boy of Aveyron." He was followed by his pupil Seguin (1846), who devised a number of form-boards specifically for those challenged mentally.

One of the earlier, well-known nonverbal measures of intelligence was introduced by Pitner and Paterson in 1943. Their Performance Scale, although mainly used in the testing of hearing children when first introduced, was primarily designed with the deaf in view (Drever & Collins, 1946). The authors of the scale had three specific objectives: (a) to provide a multidimensional scale, utilizing a variety of test formats, so that a variety of factors entering into intelligence might be brought into play; (b) to provide novel situations within the test's format, thus eliminating the possibility that the test's solution would be familiar to the child; and, (c) to employ tests which required no verbal instructions for the hearing impaired child, and in some instances to give a practice test before the actual test was used.

Another early nonverbal measure developed by Drever and Collins (1946) is the Performance Tests of Intelligence: A Series of Non-linguistic Tests for Deaf and Normal Children. Their test consists of eight subtests: Block Designs, Knox Cubes, Dominoes, Size and Weight, Manikin and Profile, Form Boards, Cube Construction and Picture Completion. Test administration guidelines include both "sign" and "say" versions, so that the test may be used with either hearing or nonhearing subjects.

Within the past two to three decades there has been a growth in the number of tests designed to minimize the effects of language mediation in the measurement of intelligence. The following is a review of the current body of tests designed as nonverbal measures of intelligence.

Raven's Coloured Progressive Matrices Test. The RCPM, composed of 36 items, is an individually administered test of manifest reasoning. The RCPM items tap various perceptual and reasoning functions through a visual mode. The child has to recognize a missing part in a visual pattern and complete it from six alternatives given at the bottom of the page. Most of the items require perception (i.e., recognizing a Gestalt and comparing identical elements), spatial orientation, and recognition of shapes. The last five items require analogical thinking.

Raven's Standard Progressive Matrices Test. The RSPM was designed to measure mental ability as represented by Spearman's principle of "g" (Spearman, 1927) using abstract figural matrices. The test is intended to assess a person's ability to see the relationships between abstract figures, "conceive the nature of the figure completing each system of relations presented, and by so doing, develop a systematic method of reasoning" (Raven, Court & Raven, 1977, p.2). Like the RCPM, the RSPM, requires the individual to select the best option to fit in the missing location in the matrix based on the relationships he/she has uncovered through reasoning. The Raven Standard Progressive Matrices contains five sets of twelve matrices for a total of sixty items. The test has been normed in Europe and the United States. Although the Raven is a useful measure of non-verbal reasoning ability, it provides a measure of intelligence based on figural reasoning only. Vernon (1976) suggests that the Raven Progressive Matrices Test is a good second measure of intelligence which should be used to substantiate another more comprehensive test of intelligence.

Leiter International Performance Scale. The Leiter (Leiter, 1979) is a general intellectual scale designed to measure the cognitive functioning of individuals between the ages of two and eighteen, who are either communicatively disordered, hearing impaired, or non-English speaking. The Leiter can be administered with minimal verbal communication and requires no verbalizations on the part of the examinee. Criticisms of the test include: lack of a theoretical basis for conceptualization of results; uneven item difficulty levels; and, outdated normative data. Levine's review of twenty-five studies involving the Leiter Scales indicate wide differences, both higher and lower, between it and other tests (Levine, 1979).

Vernon suggests that the Leiter lacks sufficient validation but is useful in testing young deaf children. He states that since the Leiter has a mean of 95 and a standard deviation of 20 direct correspondences to other measures such as the Wechsler or Binet cannot be assumed (Vernon, 1983).

Columbia Mental Maturity Scale. The Columbia Mental Maturity Scale, Third Ed. (CMMS; Burgmeister, Blum & Lorge, 1972) is a test of general reasoning ability for ages three years, six months to nine years, eleven months. It is a pictorial classification test, originally developed for students with cerebral palsy. The test requires only limited motoric or language responses. Limitations of the CMMS is its lack of an underlying theoretical model for interpretation of results, outdated standardization data, and its limited focus on abilities assessed.

The CMMS may be inappropriate for use with hearing impaired populations who may have difficulty in understanding the concept of pointing to the "one that does not belong. For deaf children who clearly understand the direction, the test can be used for screening purposes" (Sattler, 1992, p. 315).

Goodenough Draw-A-Man Test. The Goodenough Draw-A-Man Test is a brief, nonverbal test of intelligence that can be administered either individually or to a group. The test requires the child to draw a person. Points are awarded for various body parts drawn by the child. "The Draw-A-Man Test is based on Goodenough's (1926) assumption that the intelligence of early school-age children can be estimated from their early drawings of a figure of a man" (Sattler, 1992, p. 311). The test is a relatively poor predictor of scores on other intelligence tests and correlation with achievement is also poor.

Vernon states that directions for the Goodenough are extremely difficult to give to young deaf children in a standardized manner. He also suggests that the scoring is less objective than is desirable, and may lead to relatively unreliable scores.

Nebraska Test of Learning Aptitude; Hiskey-Nebraska Test of Learning Aptitude. The Nebraska Test of Learning Aptitude is a nonverbal intelligence test that has been found to be useful in assessment of children with a variety of language handicaps. The test includes a scale totaling 124 individual parts. The parts of the scale include: Memory for colored sticks; Bead stringing; Pictorial associations; Block building; Memory for digits; Completion of drawings; Pictorial identification; Paper folding; Visual Attention span; Puzzle block; and, Pictorial analogies. Reliability and validity are satisfactory. It has parallel instructions for administering each subtest to deaf and to normal-hearing children. The normative data is seriously out of date and may not be representative of the hearing-impaired population. Vernon suggests that the Hiskey-Nebraska is basically a sound and useful test but he does not recommend the test for hearing impaired children between three and four years of age.

Naglieri's Matrix Analogies Tests: Matrix Analogies Test-Expanded and the Matrix Analogies Test-Short Form. Naglieri's developed the MAT-EF and the MAT-SF to meet the need for a well normed, well constructed, nonverbal test of intelligence (Naglieri, 1985). The MAT-SF is intended to be used as a group screening test and the MAT-EF was designed for individual use. The MAT-EF is divided into four groups of items based upon the problem solving requirements of the specific tasks. The Pattern Completion items require the subject to choose the option which most accurately completes the pattern. In the Reasoning by Analogy subtest the individual must determine how the change in one item is analogous to the change in the target item. Serial Reasoning measures the degree to which an individual is able to uncover the order in which items appear throughout the matrix. Spatial Visualization requires the subject to imagine how a figure would look when two or more designs are combined. The MAT-EF is reported to have significant correlation with the WISC-R Performance IQs for samples of normal, and hearing impaired students. The MAT-SF also shows significant correlations with achievement for hearing impaired children with the Stanford Achievement Test Reading and Spelling scores.

Test of Nonverbal Intelligence. The Test of Nonverbal Intelligence, Second Edition is a "highly standardized, psychometrically sound, norm referenced intelligence test with an administration and response format that eliminates language and reduces motoric and cultural factors" (TONI-II; Brown, Sherbenou, & Johnsen, 1990, p.1). It was developed as an appropriate nonverbal cognitive measure for persons aged 5 to 85 years. The TONI-II items require test subjects to solve problems by identifying relationships among abstract figures and then solve problems created by the manipulation of these relationships. The authors give theoretical justification for selecting problem solving as the core for TONI-II items by referring to Sternberg's contention that problem solving appears to be a general component or construct of intelligent behavior rather than a splinter skill or subcomponent (Sternberg, 1980, 1984). Major limitations of the TONI-2 are: reliance on a single subtest format as a measure of intelligence, inadequate data regarding its statistical properties, and predictive validity derivation of data from small sample sizes (Naglieri, 1990).

Universal Nonverbal Intelligence Test. The Universal Nonverbal Intelligence Test (UNIT, Bracken & McCallum, 1994) is a newly developed assessment of cognitive abilities which uses a nonverbally mediated method of administration and is designed to be used with those persons for whom traditional, verbally-laden instruments are inappropriate (i.e., non-English speaking, hearing impaired, communicatively disordered, or limited English proficiency). The UNIT is an individually administered measure of intellectual functioning in persons aged 5 to 17. Initial findings suggest that the UNIT is a psychometrically sound instrument.

Specific features of the UNIT which make it unique include: its multi-subtest battery design which assesses a wide range of abilities; its theoretical underpinnings conceptualized to measure memory and reasoning skills through both symbolic and non-symbolic tasks; and its reliance on an entirely pantomimed method of administration. The gestures selected for use in the UNIT are based on the commonality of such pantomimed movements used in the United States as well as in other countries. Since no formal sign language knowledge is required for administration, persons administering the test will not require additional skills in this area. Demonstration items, sample items, and checkpoint items (i.e., items on which examiner feedback are allowed) are included in each subtest to ensure maximum task comprehension. The UNIT provides for a Full Scale IQ, and Memory, Reasoning, Symbolic and Nonsymbolic functioning indexes.

The UNIT has its theoretical basis in the work of Arthur Jensen with emphasis on measurement of the "g" factor and categorization of mental abilities into Associative Level I abilities and Cognitive Level II abilities. The test consists of six subtests which were developed to measure intelligence via symbol-laden content and less symbolically loaded tasks. Since the UNIT is currently in the standardization stage of development, there is only limited data regarding its predictive capabilities with achievement. However, "data comparing UNIT scores with scores form the Woodcock-Johnson Cognitive Battery-Revised (Woodcock & Johnson, 1989), Extended Battery, provide support for good concurrent (Reed & McCallum, 1994) and predictive validity" (Williams & McCallum, 1994). Correlation coefficients between the UNIT global scores and the WJ-R Cognitive Ability Score (Extended) ranged from .60 (Memory Quotient and WJ-R) to .68 (UNIT Full Scale IQ and WJ-R). Coefficients showing the relationship between the UNIT and the end-of-year Comprehensive Test of Basic Skills (CTBS, 1989) are moderate and encouraging, particularly given the nonverbal nature of the UNIT and the verbal nature of academic achievement (range .37 to .56) (McCallum & Bracken, in press). Strong correlations ranging from .48 to .61 were found between the Symbolic Quotient and all areas of achievement. The Symbolic Quotient was more strongly related to achievement than the Nonsymbolic Quotient. Memory and Reasoning Quotient were about equally related to achievement. Analogic Reasoning was found to be the best predictor of reading achievement (.46).

With the exception of the MAT-EF, no predictive validity data are available for any of the instruments appropriate for use with a population of persons with hearing impairments.

Assessment of Cognitive Constructs for Persons with Hearing Impairments

Generally the controversy over the level of intelligence of persons with hearing impairments as compared to that of their hearing peers, has been laid to rest. Although early investigations of the intellectual abilities of this population suggested that deaf individuals were cognitively inferior to hearing individuals, perceptions regarding this position have been largely refuted. Deaf children usually obtain scores within the normal range on nonverbal intelligence tests (Meadow, 1975).

One of the earliest researchers in the field of intelligence testing for hearing impaired individuals is Pitner. He conducted a study of 2,172 children drawn from 26 residential schools for the deaf. Using a group intelligence test, he found a two to five year retardation educationally. Pitner reviewed all data to 1940 and suggested that persons with hearing impairments suffer a ten point inferiority in IQ scores when compared to their hearing peers. The origin of this deficiency, he attributed to both sensory deprivation and complications associated with deafness. As a result of these findings Pitner concluded that language laden tests such as the Binet-Simon Scale were "absolutely inadequate to test the mentality of deaf children," (Pitner & Patterson, 1943). Pitner and Paterson developed a performance scale which required practically no instructions other than natural gestures as a way to measure the cognitive abilities of persons with hearing impairments. They considered the non-language feature of their test as a requisite for measurement of any group of individuals who have been "completely shut off from hearing language." They contend that no kind of test involving reading or spoken language can be used with the hearing impaired population, since the measurement would simply reflect language scores rather than true cognition.

Myklebust, in 1953, took exception to Pitner's conclusion of hearing impairment causing a 10 point mental deficiency. As a result of his review of testing of hearing impaired individuals he reported that "during the past 15 years various workers have uniformly concluded that . . . deaf children are not generally inferior in intelligence as concluded by Pitner" (Myklebust, 1953, p.346). Myklebust did suggest, however, that although students with hearing impairments were quantitatively equal to their hearing peers, qualitative differences still existed in perceptual functioning and conceptual reasoning. Myklebust postulated that the absence in one of the senses results in a shift in cognitive structures, thus his theory has come to be known as the "Organismic Shift Theory." The net effect of such a shift accounts for the qualitative differences found in the intellectual functioning of hearing impaired individuals. Myklebust concluded that the deaf tended to be more concrete as compared to their hearing peers who were said to be more abstract.

> When auditory sensation are lacking or present only to a minimal degree, the nature of (the deaf person's) perceptions, imagery, symbols, and concepts is altered. The levels of symbolization are most affected; development of certain types of abstract behavior is impeded (Myklebust, 1960, p. 229).

Myklebust's theory appears to have its foundation in the Whorfian Hypothesis (Whorf, 1956), that language and conceptualization are inextricably connected; i.e., if one's language is different, then so must one's thought processes be different.

Although contemporary views are not generally supportive of the view of deaf individuals as being more concrete, some of the recent studies on information processing may lend support to the Organismic Shift Hypothesis. Deaf children who have an early, severe hearing loss may not have the lateral specialization of the brain for language that characterizes the hearing population (Kelly & Tomlinson-Keasey, 1977). If the deaf child is presumed to be processing information in the right hemisphere as a result of visual input rather than auditory input, and the hearing child is processing information in the left hemisphere, then tests which tap higher levels of language symbolization would be problematic to the deaf child regardless of other environmental influences because of his/her lack of development of the left hemisphere which has been deemed the language center of the brain. On the other hand, deaf children would be proficient on tasks requiring visual processing since these processes are presumed to be localized in the right hemisphere. Although such information processing research is only in its rudimentary stages, limited support for this thesis is provided in Todman and Cowdy's (1993) study using Pascual-Leone's (1970) compound stimulus visual information (CSVI) task. The results of their study found that deaf subjects significantly outperformed hearing subjects of the same age and nonverbal reasoning ability on tasks requiring visual action coding (i.e., unordered recall of simultaneously presented pairings of spatial stimuli and action responses.)

Another study which tends to support Myklebust's theory was done by Hiskey (1979). When Hiskey compared the performance of students aged 4-10 with hearing impairments and those with hearing on the Nebraska Test of Learning Aptitude, he found that students with hearing impairments scored slightly lower. The common difference was one half year. He suggests that the average intelligence quotient for the deaf is in the mid 90's when norms for hearing children are used. He concludes that hearing children have a decided advantage on items where immediate recall "was aided by vocalization, and the deaf excelled slightly on some items demanding visual perception (almost entirely) as the approach to analysis and performance (Hiskey, 1979, p. 338). Hiskey suggested that the qualitative differences in perceptual and conceptual functioning and reasoning may add support to Myklebust's Organismic Shift Theory. Yet, he cautions against taking these results too seriously since he suspects that the deaf child may have more problems understanding pantomimed instructions than the hearing child does the verbal direction of the examiner. He concludes that deaf children are likely to equal hearing children when language is not a primary factor and even to excel when visual perception is of paramount importance (Hiskey, 1979).

Levine posited that persons with hearing impairments possessed normal intelligence, but that differences in academic achievement or performance on intelligence tests is a result of experiential deficits rather than sensory loss per se. In his "Experiential Deficits Hypothesis," Levine asserts that qualitative differences between hearing and non-hearing individuals corresponds to the hearing impaired person's lack of experiences.

Support for this position is found in a study by Watt (1979); he studied the performance of deaf, hard of hearing, and hearing children on a series of tasks including conservation-quantitative thinking, horizontally-spatial thinking, and picture arrangement-social thinking. He summarized his findings in this way:

> From the tasks used the results have pointed to the fact that the development of deaf children's cognitive abilities is remarkably unaffected by the absence of verbal language. Secondly, valuable support has been given to the idea that if deaf children perform poorly on some intellectual tasks,

> this performance can be more adequately accounted for by experiential than linguistic deficiency. The suggestion that experiential deficiency more adequately explains performance is supported by the fact that the improvement of the deaf children on the conservation tasks took place without a concomitant improvement in the comprehension and usage of written language (Watts, 1979, p. 55).

The heterogeneity of the hearing impaired population is an important aspect to be considered whenever issues of intelligence are reported. The term "deaf" encompasses two distinct populations. A distinction may be based upon the hearing status of the child's parents. Deaf children of deaf parents constitute a different population than deaf children of hearing parents. The critical difference between these two groups is whether or not the deaf child had the opportunity to acquire a native language (Brill, 1974; Gallaudet research Institute, 1987; Rodda & Grove, 1987). Deaf children who grow up in homes with deaf parents are immersed in language much the same as hearing children (Hoffmeister, Will, 1980, Newport and Mei, 1985). Through direct exposure to language in the natural family environment, the deaf child does not suffer from the degree of "language deprivation" usually associated with deafness.

Deaf children of hearing parent's may suffer from Levine's notion of experiential deficits. This group of deaf children do not have ready access to the language around them and consequently suffer from a lack of much information which language conveys. Research employing deaf subjects conducted by Gallaudet Research Institute (1987) supports this conclusion:

> The justification for studying deaf children of deaf parents separate from deaf children of hearing parents is attributable to the fact children in the two categories typically are born into and develop within quite different linguistic, social and cognitive environments, even though they may attend the same school programs (p. 1).

To consider both these groups as a single population often reveals contradictory results (Vernon, 1968). Studies, however, separating these two groups reveal more consistent findings. "Deaf children of deaf parents outperform children with hearing parents on virtually any measure of intelligence or academic achievement investigated" (Isham & Kamin, 1993, p.43). Brill (1974) and Rodda & Grove (1987) report similar findings. Mean IQ scores of deaf children of deaf parents are significantly higher than those with hearing parents. An interesting finding of studies conducted by Gallaudet Research Institute (1987) on 65 deaf students using the Wechsler Intelligence Scale for Adults (WAIS) found that the mean Performance IQ for the group having deaf parents was 113.02 (SD=14.86). This study indicates that as a group deaf children of deaf parents score even higher than hearing norms. The reason for this finding is not clear; however, studying these two groups separately is now accepted standard procedure at Gallaudet University, a major research institute on the deaf. There is a need to determine whether this experiential deficit hypothesis will be supported when ability is determined between measures of intelligence obtained nonverbally. Because the UNIT affords a "Symbolic Quotient" and a "Nonsymbolic Quotient" it is possible to address the experiential deficit hypothesis. One would predict that the two groups will differ on both the Symbolic Quotient of the UNIT and the TONI-II scores because these areas probably reflect one's experience more than the UNIT Nonsymbolic Quotient.

The interpretation of deaf children's test results is difficult, because of communicative barriers and nonintellective factors that may have detrimental effects on cognitive performance. Deaf persons, for example, have been found to have a higher degree of impulsivity, egocentricity, dependency, and rigidity than hearing individuals (Altshuler, Deming, Vollenweider, Rainier & Tender, 1976; Harris, 1978; Levine, 1981). Deaf children are also more likely than hearing children to be low in motivation, independence, and initiative (Meadow & Dyssegard, 1983), to deny responsibility for the effects of their behavior, and to possess external loci of control (Bonder & Johns, 1977; Dillon, 1980; Harris, 1978; White, 1982). The influence of nonintellective factors on cognitive performance, particularly on intelligence test scores, has been intensively investigated on different groups varying in age, etiology of manifest problem, and level of cognitive functioning (e.g., Dweck & Elliot, 1983; Haywood, 1968; Zigler & Butterfield, 1968; Phillips, 1984; Tzuriel, Samuels, & Feuerstein, 1988).

Summary

Historically, the development of intellectual measures for persons with hearing impairments has lacked strong theoretical underpinnings making interpretation of their results tentative at best. Nonverbal measures have generally addressed only a narrow scope of abilities and the efficacy of their correlational and predictive validity is questionable.

Intellectual assessment of children with hearing impairments have generally supported the thesis that the "deaf" are intellectually normal. Two hypotheses accounting for differences in test scores found between hearing and nonhearing subjects are the Organismic Shift Hypothesis postulated by Myklebust and the Experiential Deficits Hypothesis postulated by Levine. Confounding the aforementioned problems, are the problems associated with a lack of trained professionals in the area of deafness, and the paucity of cognitive measures appropriate for use with this population.

Statement of Problem

Despite the widespread use of intelligence tests for the measurement of cognitive functioning of school-aged children, there exists a void in psychometrically sound nonverbal instruments for intelligence testing. For some populations, e.g., hearing impaired individuals, verbal measures are inappropriate. There is a need for a model-driven, theoretical based, multidimensional, nonverbal intelligence measure. One aspect of determining quality measure is the investigation of the predictive and concurrent utility of the constructs purported to be assessed by the instrument. The relative predictive and concurrent validity characteristics are unknown for constructs of memory, reasoning and figural objects for a hearing impaired population. These constructs are assessed by the UNIT and the TONI-II. In addition, because the UNIT provides measures of multiple constructs the relations between those constructs and achievement need to be clarified. For example, in a hearing population the UNIT Symbolic Quotient predicted better than the Nonsymbolic. Will this pattern hold for a hearing impaired population?

Those interested in determining the usefulness of new and sound instruments for the hard of hearing or deaf population are also interested in investigating the reason for differential responding on these test for two nonverbal populations of "deaf" children. That is, there is a need to explore the experiential deficits hypothesis. Such a study would require study of deaf children of hearing parents, and deaf children of deaf parents. These two sub-populations have been shown to differ significantly on both academic and intelligence measures. Studying these two groups as separate populations is warranted.

Research Questions and Hypotheses

The following research questions and hypotheses will be addressed by this study:

I. What is the relationship between the theoretical constructs of reasoning and memory operationalized by the UNIT, and figural problem solving, operationalized by the TONI-II, and achievement?

A. Nondirectional Hypothesis 1: The Reasoning Quotient of the UNIT, operationalized through the reasoning subtests, the Memory Quotient of the UNIT, operationalized through the memory subtests, and the Deviation Quotient of the TONI-II, operationalized through figural problem solving, will predict equally well achievement on the Stanford Achievement Test (Total Score).

B. Nondirectional Hypothesis 2: The Reasoning Quotient of the UNIT, operationalized through the reasoning subtests, the Memory Quotient of the UNIT, operationalized through the memory subtests, and the Deviation Quotient of the TONI-II, operationalized through figural problem solving, will predict equally well achievement on the Stanford Achievement Test (Reading Score) and the PIAT-R (Reading Score).

C. Nondirectional Hypothesis 3: The Reasoning

Quotient of the UNIT, operationalized through the reasoning subtests, the Memory Quotient of the UNIT, operationalized through the memory subtest, and the Deviation Quotient of the TONI-II, operationalized through figural problem solving, will predict equally well achievement on the Stanford Achievement Test (Math Score) and the PIAT-R (Math Score).

II. What is the relationship between the Symbolic and Nonsymbolic Quotient of the Unit and achievement?

A. Directional Hypothesis 1: The Symbolic Quotient operationalized through the symbolic subtests of the UNIT will predict achievement as measured by the SAT (Total Score) significantly better than the Non-Symbolic Quotient, as operationalized by the non-symbolic subtests of the Unit.

B. Directional Hypothesis 2: The Symbolic Quotient of the UNIT operationalized by the symbolic subtests will predict achievement as measured by the SAT (Reading Score) and the

PIAT-R (Reading Score) significantly better than the Non-Symbolic Quotient operationalized by the non-symbolic subtests of the UNIT.

C. Directional Hypothesis 3: The Symbolic Quotient of the UNIT operationalized by the symbolic subtests will predict achievement as measured by the SAT (Math Score) and the PIAT-R significantly better than the Nonsymbolic Quotient of the UNIT, as operationalized by the nonsymbolic subtests.

III. What is the relationship between experience, defined as having parents who share the same language structure as their children versus those who do not, and children's cognitive ability?

A. Directional Hypothesis 1: The mean Symbolic Quotient of the UNIT operationalized by the symbolic subtests for hearing impaired children of hearing impaired parents will be significantly better than the mean of the measure for hearing impaired children of hearing parents.

B. Nondirectional Hypothesis 2: The mean Non-Symbolic Quotient of the UNIT operationalized by the non-symbolic subtests for hearing impaired children from hearing impaired parents will be equal to the mean of hearing impaired children from hearing parents.

C. Directional Hypothesis 3: The mean Total Score of the TONI-II operationalized as figural problem solving of hearing impaired children from hearing impaired parents will be significantly higher than the mean of hearing impaired children from hearing parents.

CHAPTER II

METHOD

Participants

Forty-six students, aged five to seventeen, were selected as participants for this study; all attend the Tennessee School for the Deaf (TSD). TSD, a state school for students with diagnosed hearing impairments, serves as both a residential and day facility for students in the East Tennessee area.

Forty-six protocols from the UNIT, the TONI-II, the PIAT-R and thirty-five protocols from the SAT were analyzed. These protocols were obtained from all available subjects at the Tennessee School for the Deaf who meet the minimum criteria of "Hearing Impairment," as identified by the state mandated evaluation and for whom signed parental permission slips were obtained. The protocols were analyzed to determine the relative predictive and correlational validity of the UNIT and the TONI-II on the achievement levels of students with hearing impairments. Descriptive information regarding the demographics and the auditory functioning of the student population at TSD and their parents include a review of the attendance records, the school lunch program data, audiology reports, and hearing status of parents, i.e., parents with hearing impairments and those with normal hearing. Formal parental permission forms were sent to the students' parents, with an enclosed envelope for return mail. Written consent from the administration of TSD was also obtained. Because participation in the study was voluntary, no payment nor any other form of renumeration was given to participants. Furthermore, no test results were reported to participants, their legal guardians, or to the administration at TSD.

Instruments

UNIT

The UNIT is comprised of six subtests which are structured in a 2 by 2 scheme. Several scores are used to calculate a Full Scale score (FSIQ), a Memory Quotient (MQ), a Reasoning Quotient (RQ), a Symbolic Quotient (SQ), and a Nonsymbolic Quotient (NSQ). UNIT subtests are described below:

Memory Subtests

1. Symbolic Memory: the examinee recalls and recreates sequences of visually presented universal symbols (e.g., green boy, black woman). This subtest is amenable to verbal mediation. Reliability data from a sample of 104 children yielded a Spearman-Brown of .89.
2. Spatial memory: the examinee must remember and recreate the placement of black and/or green chips on a 3x3 or 4x4 cell grid. This subtest is not amenable to verbal mediation via symbols. Reliability data from a sample of 104 children

yielded a Spearman-Brown of .87.

3. Object Memory: The examinee is shown a visual array of common objects (e.g., shoe, telephone, tree) for five seconds, after which the examinee identifies the pictured objects from a larger array of pictured objects. This subtest is amenable to symbolic mediation. Reliability data from a sample of 104 children yielded a Spearman-Brown of .83.

Reasoning Subtests

4. Cube Design: the examinee completes a three dimensional block design task using as many as nine green and white blocks. This subtest is not amenable to symbolic mediation. Reliability data from a sample of 104 children yielded a Spearman-Brown of .92.

5. Analogic Reasoning: the examinee completes a matrix analogies task using common objects (e.g., hand/glove, foot/________ ?) and novel geometric figures (e.g., dice). This subtest is amenable to symbolic mediation. Reliability data from a

sample of 104 children yielded a Spearman-Brown of .85.

6. Mazes: the examinee completes a maze task by tracing a path through each from the center starting point to an exit. This subtest is not amenable to symbolic mediation. Reliability data from a sample of 104 children yielded a Spearman-Brown of .60.

Completion of all six subtests requires approximately 45 minutes. Administration is nonverbal; gestures, demonstration, and pantomime may be used to convey task demands.

TONI-II

The TONI-II consists of two equivalent forms containing 55 items each. All of the items require abstract/figural problem solving. The subject looks at the stimulus items and responds by pointing or some other form of meaningful motor activity. The test is norm referenced and yields a deviation quotient score. Reliability data provided by the TONI-II manual reports a Cronbach Alpha of .81.

One or more of the following rules are used in each TONI-II item. (1) Simple matching in which all figures share the same number of critical attributes. (2) Analogies in which the relationship among the figures in one of the rows or columns is the same relationship found among the figures in the other rows and columns. (3) Classification in which the figure in the stimulus is a member of one of the sets of figures in the response alternatives. (4) Intersections in which a new figure is formed by joining parts of figures in the rows and columns. And (5) Progressions in which the same change continues between or among figures.

PIAT-R

The PIAT-R is an individually administered achievement test providing wide-range assessment in six content areas: General Information, Reading Comprehension, Mathematics, Spelling, and Written Expression.

Subtests

1. General Information: measures the subject's

general encyclopedic knowledge. Reliability data provided by the PIAT-R manual reports a Spearman-Brown reliability coefficient of .94.

2. Reading Recognition: measures the subject's ability to recognize the sounds associated with the printed letters and the ability to read words aloud. Reliability data provided by the PIAT-R manual reports a Spearman-Brown reliability coefficient of .96.

3. Reading Comprehension: measures the subject's understanding of what is read. The subject reads a sentence silently, and then chooses the one picture out of four which best illustrate the sentence. Reliability data provided by the PIAT-R manual reports a Spearman-Brown reliability coefficient of .92.

4. Mathematics: measures the subject's knowledge and application of mathematical concepts and facts, ranging from recognizing numbers to solving geometry and trigonometry problems. Reliability data provided by the PIAT-R manual reports a

Spearman-Brown reliability coefficient of .94.

5. Spelling: measures the subject's ability to recognize letters from their names or sounds and the subject's ability to recognize standard spellings by choosing the correct spelling of a word spoken by the examiner. Reliability data provided by the PIAT-R manual reports a Spearman-Brown reliability coefficient of .94.
6. Written Expression: measures the subjects's written language skills at two levels. Level I tests pre-writing skills - copying and writing letters, words, and sentences from dictation. In Level II the subject writes a story in response to a picture prompt.

Two composites summarize the subject's achievement: **Total Reading**, a combination of Reading Recognition and Reading Comprehension, is an overall measure of reading ability. **Total Test** describes the subject's overall level of achievement on the first five subtests. The derived scores provide for interpreting the five subtests and the two composites are grade and age equivalents, percentile ranks, normal curve equivalents, and stanines. Norms are provided from kindergarten through Grade 12 (Ages 5-0 to 18-11). Items within each subtest which did not lend themselves to a nonverbal administration or which required listening skills were omitted from the testing. Therefore, only the Reading Comprehension and the Mathematics subtests were administered as part of this research study.

SAT-8

The 8th Edition Stanford Achievement Test is an achievement test which measures the abilities and skills of students in a number of different subtest or subject areas: e.g., reading, language, spelling, mathematics, science, and social science. The SAT-8 is available at eight difficulty levels which measure content that is considered appropriate for HEARING students in specific grades in school. The eight difficulty levels and their corresponding grade levels for use with HEARING students are:

Primary 1	1.5 to 2.5
Primary 2	2.5 to 3.5
Primary 3	3.5 to 4.5
Intermediate 1	4.5 to 5.5

Intermediate 2	5.5 to 6.5
Intermediate 3	6.5 to 7.5
Advanced 1	7.5 to 8.5
Advanced 2	8.5 to 9.9

Assignment of students with hearing impairments to one of these eight levels is determined by a screening test. Although the test items and questions used to test students with hearing impairments are identical to those used to test hearing children, the screening test was developed by the Center for Assessment and Demographic Studies (CADS) at Gallaudet University to be administered prior to the actual test to assist in the assignment of students to the appropriate difficulty level. Since English language development is more gradual for students with hearing impairments the test is often used throughout high school and even into postsecondary years.

The SAT-8 is a group administered achievement test. TSD personnel administered the test to students in grades 1 through 12 in the Spring of 1996. Raw scores from the reading and mathematics subtests were used in this research study. Kuder-Richardson-20 subscale coefficients for all ages and most tests are in excess of

.85 or .90, with the exception of Listening, Environment and Language Mechanics subtests which tended to be in the upper .7s.

Procedure

Participants responded to test items which required performance on several different tasks, such as, construction of designs using blocks, identifying patterns in figure drawings, working puzzles and mazes. The tests were administered by advanced graduate students trained to use the test using gestures. Administration required no verbal responses from the participants. The nature of the test administration was explained to the subjects prior to the test session. The measures were administered in counterbalanced order to prevent order effects. The tests took an average of one and one half hours to complete.

Data Analysis

The following statistical procedures were performed.

I. The relationship between the constructs of reasoning and memory operationalized by the UNIT, and figural problem solving operationalized by the TONI-II, and achievement were determine according to the following:

A. To determine predictive and concurrent relationships multiple regression analyses were calculated for UNIT Memory Quotient, Reasoning Quotient, TONI-II Total Score, using SAT Total Score as the criterion.

B. To determine predictive and concurrent relationships multiple regression analyses were calculated for UNIT Memory Quotient, Reasoning Quotient, TONI-II Total Score using the PIAT-R and SAT Reading Scores as the criteria.

C. To determine predictive and concurrent relationships multiple regression analyses were calculated for UNIT Reasoning Quotient, Memory Quotient, TONI-II Total Score, using the PIAT-R and the SAT Math Scores as the criteria.

II. The relationship between the Symbolic and Nonsymbolic Quotients of the UNIT with Achievement were determined through the following:

A. To determine predictive and concurrent relationships multiple regression analyses were calculated for UNIT Symbolic Quotient and UNIT Nonsymbolic Quotient using the SAT Total Score as the criterion.

B. To determine predictive and concurrent relationships multiple regression analyses were calculated for UNIT Symbolic Quotient and UNIT Nonsymbolic Quotient using the PIAT-R and the SAT Reading Scores as the criteria.

C. To determine predictive and concurrent relationships multiple regression analyses were calculated for UNIT Symbolic Quotient and UNIT Nonsymbolic Quotient using the PIAT-R and

SAT Math Scores as the criteria.

III. The Experiential Deficit Hypothesis was tested by the following analyses:

A. The mean Symbolic Quotient of hearing impaired children of hearing impaired parents was compared to the mean of hearing impaired children of hearing parents; a one-tailed t-test was calculated because the mean of hearing impaired children of hearing impaired parents was expected to be higher than the mean of the children of hearing parents.

B. The mean Non-Symbolic Quotient of hearing impaired children of hearing impaired parents was compared to the mean of hearing impaired children from hearing parents; a two-tailed t-test was calculated because the mean of hearing impaired children of hearing impaired parents was expected not to differ from the mean of the children of hearing parents.

C. The mean Total Score of the TONI-II for hearing impaired children of hearing impaired parents was compared to the mean of hearing impaired children from hearing parents; a one-tailed t-test was calculated because the mean of hearing impaired children of hearing impaired parents was expected to be higher than the mean of the children of hearing parents.

CHAPTER III

RESULTS

Results provide evidence to address three research questions. The first question focuses on the relationship between the theoretical constructs of reasoning and memory operationalized by the UNIT, and figural problem solving operationalized by the TONI-II, and achievement. The second focuses on the relationship between the Symbolic and the Non-Symbolic Quotients of the UNIT and achievement. The third question is directed toward exploring the relationship between experience, defined as having parents who share the same language structure as their children versus those who do not, and children's cognitive ability. Each of these questions are addressed with relevant descriptive statistics provided.

Preliminary Analyses

Preliminary statistical analysis included investigation of the assumption of normality in the distribution of scores. Relevant descriptive statistics including means and standard deviations for each variable are included in Table 1. The Coefficients of Skewness for each variable are included in Table 2.

Tests for skewness showed normality for UNIT Memory, Reasoning, Nonsymbolic Global Scores, TONI-II Total raw scores and area achievement scores. SAT Total Scores and Symbolic Global Scores were asymmetrical. For UNIT and SAT scores found to exhibit nonnormality in distribution, reciprocal transformations were conducted. According to Glass and Hopkins (1984) conversion of original scores to another metric often results in less skewness and a more normal distribution. Using reciprocal transformations normality in distribution was achieved for all skewed

Table 1

Raw Score Means and Standard Deviations for the PIAT-R, the SAT, UNIT Global Scores, and the TONI-II

	Mean	SD	N
	PIAT-R Subtests		
Reading Comp	39.91	19.04	46
Mathematics	30.91	18.89	46
	SAT Subtests		
Reading Comp	22.91	9.85	35
Math Application	18.29	6.84	35
Total	70.09	21.68	35
	UNIT Global Scores		
Reasoning	81.26	28.43	46
Memory	55.15	18.09	46
Symbolic	57.17	16.78	46
Nonsymbolic	79.24	30.09	46
Full Scale	136.41	43.95	46
	TONI-II		
Total Raw Score	17.02	7.75	46

Table 2

Normality Diagnostics for raw scores of the Universal Nonverbal Intelligence Test(UNIT), the Peabody Individual Achievement Test (PIAT-R), the Stanford Achievement Test (SAT), and the Test of Nonverbal Intelligence (TONI-II)

Area	Skewness
UNIT	
Full Scale	-0.52
Reasoning	-0.02
Memory	-0.88
Symbolic	-1.04
Nonsymbolic	0.07
PIAT-R	
Mathematics	-0.09
SAT	
Reading Comprehension	-0.33
Reading Comprehension	0.88
Math Application	0.12

(table continues)

Area	Skewness
Total Score	-1.09
TONI-II	
Total Raw Score	0.99

data. Following reciprocal transformations, additional tests for normality indicated normal distributions for all previously skewed data. Multiple regression analyses were conducted utilizing both converted and nonconverted data. Comparison of multiple regression results yielded no substantial differences in the findings. For example, using reciprocal transformation data, multiple regression analysis of the Symbolic Quotient of the Unit with the PIAT-R Math achievement scores used as the dependent variable resulted in a Multiple R of .73, R Square of .53 and an Adjusted R of .52. Similarly, multiple regression analysis, using the original nonconverted data with the aforementioned variables, resulted in a Multiple R of .76, R Square of .58 and an Adjusted R of .57. Because the results are very similar, they are reported in original, nonconverted scores. (See Table 3).

Table 3

Pearson r Correlations of Raw Scores Among All Variables

	UNIT FS	MEM	REA	SYM	NONS	TONI
PIAT R	.80	.76	.75	.76	.75	.61
PIAT M	.74	.71	.69	.70	.69	.56
SAT R	.27	.41	.03	.42	.15	.06
SAT M	.56	.64	.14	.61	.44	.14
SAT TOTAL	.61	.70	.42	.71	.43	.21

Note. UNIT FS = Universal Nonverbal Intelligence Test - Full Scale; MEM = UNIT Memory Quotient; REA = UNIT Reasoning Quotient; SYM = UNIT Symbolic Quotient; NONS = UNIT Nonsymbolic Quotient; TONI-II = Test of Nonverbal Intelligence; PIAT R = Peabody Individual Achievement Test Reading Comprehension; PIAT M = Peabody Individual Achievement Test Math Application; SAT R = Stanford Achievement Test Reading Comprehension; SAT M = Staford Achievement Test Math Application; SAT Total = Stanford Achievement Test Total Score.

Relationships among UNIT and TONI-II Total Scores and achievement

To investigate the relationship between the theoretical constructs of reasoning, memory, and figural problem solving, and achievement UNIT Memory and Reasoning Global Scores and the TONI-II total raw score were used in a multiple regression. In terms of total achievement the UNIT Memory quotient was the single best predictor of achievement accounting for 48% of the variance. (See Table 4). Neither the UNIT Reasoning Quotient nor the TONI-II Total Raw Score entered significantly in the prediction.

Two multiple regressions were conducted to investigate the relationship between the constructs of reasoning, memory, and figural problem solving and reading achievement as measured by the PIAT-R Reading Comprehension and the Stanford Achievement Test-Reading Comprehension. The Unit Memory and Reasoning Quotients and the TONI-II Total Raw Scores were used in multiple regression analysis to predict Reading Achievement as

Table 4

Raw Scores of the UNIT Reasoning and Memory Global Scores and the TONI-II as Predictors of Total Achievement as measured by the SAT.

Area	Best Pred.	Stand. Beta	Mult. R	Adj. R Square	F	p
SAT						
Total	Mem	.70	.70	.48	32.42	.01

Note. Mem = UNIT Memory Quotient; Rea = UNIT Reasoning Quotient; Toni = TONI-II Total Raw Score.

*F statistic allows for test of the significance of change in R squares as a result of each independent variable being entered into the equation.

measured by the SAT and the Piat-R Reading Comprehension score. The Memory Quotient was found to be the best predictor of reading as measured by the SAT reading comprehension subtest, yielding a Multiple R of .41, R Square of .14, and an Adjusted R Square of .14. Neither the Reasoning Global Score nor the TONI-II Total Raw Scores added significantly to the equation.

When the PIAT-R reading score was used as the criterion, both Memory and Reasoning Quotients enter significantly into the prediction. The Memory Quotient yielded a Multiple R of .76, an R Square of .58, and an Adjusted R Square of .57. There was a significant increase in the dependent variable when the Reasoning Quotient was added to the model, accounting for a .07 change in the prediction. This model accounted for 64% of the variance. The TONI-II was not significant in the prediction. (See Table 5)

To determine the ability of the Reasoning Quotient of the UNIT, the Memory Quotient of the UNIT, and the Total Raw Scores of the TONI-II, to predict achievement

Table 5

Raw Scores of the UNIT Reasoning and Memory Global Scores and the TONI-II as Predictors of Reading Achievement as measured by the SAT and the PIAT-R.

Area	Pred	Stand. Beta	Mult. R	Adj. R Square	F	p
PIAT						
Read	Mem	.46	.76	.57	61.37	.01
	Rea	.41			7.83	.01
SAT						
Read	Mem	.41	.41	.14	6.72	.01

Note. Mem = UNIT Memory Quotient; Rea = UNIT Reasoning Quotient; Toni = TONI-II Total Raw Scores.

*F statistic allows for test of the significance of change in R squares as a result of each independent variable being entered into the equation.

as measured by the PIAT-R Mathematics subtest and the Stanford Achievement Test math application subtest, multiple regressions were conducted. The Memory Quotient was the best predictor of math achievement when the SAT math score was used as the criterion yielding, a multiple R of .64, R Square of .40, and an Adjusted R of .39. When the Memory and Reasoning Quotients were added or when the Memory and TONI-II Total Raw Scores were added to the equation no significant increases were noted in the prediction. When the PIAT-R math score was used as the criterion the Memory Global scores entered significantly in the equation yielding a Multiple R of .75, R Square of .51, and an Adjusted R Square of .50. The Reasoning Global scores were also important in the prediction adding a significant .05 to the prediction. This model accounted for 53% of the variance. The TONI-II Total Raw Scores did not contribute to the equation. (See Table 6).

Table 6

Raw Scores UNIT Reasoning and Memory Global Scores and the TONI-II as Predictors of Math Achievement as measured by the SAT and the PIAT-R.

Area	Best Pred.	Stand. Beta	Mult. R	Adj. R Square	F	p
PIAT-R						
Math	Mem	.45	.75	.50	45.74	.01
	Rea	.33			4.33	.04
SAT						
Math	Mem	.41	.64	.39	22.25	.01

Note. Mem = UNIT Memory Quotient; Rea = UNIT Reasoning Quotient; Toni = TONI-II Total Raw Scores.

*F statistic allows for test of the significance of change in R squares as a result of each independent variable being entered into the equation.

Relationship Among UNIT Symbolic and Nonsymbolic Quotients and Achievement

The Symbolic Quotient, as operationalized through the symbolic subtests of the UNIT was expected to predict achievement, as measured by the SAT (Total Score), significantly better than the Non-Symbolic Quotient, as operationalized by the non-symbolic subtests of the Unit. The Unit Symbolic and Nonsymbolic Quotients were used in multiple regression analyses. The Symbolic Quotient was found to be the better predictor of total achievement, yielding a multiple R of .71, R Square of .50, and an Adjusted R Square .49. The Unit Non-Symbolic Quotient did not contribute. (See Table 7).

The Symbolic Quotient of the UNIT was expected to predict achievement significantly better than the Non-Symbolic Quotient of the UNIT. The UNIT Symbolic and Nonsymbolic Global Scores were used in multiple regression analyses to predict reading achievement. Both Symbolic and Non-Symbolic Quotients were significant in the prediction of PIAT-R reading achievement, with the

Table 7

Raw Scores for the UNIT Symbolic and Nonsymbolic Global Scores of the Universal Nonverbal Intelligence Test (UNIT) as Predictors of Total Achievement as measured by the SAT.

Area	Best Pred.	Stand. Beta	Mult. R	Adj. R Square	F	p
SAT						
Tot	Sym	.71	.71	.49	16.43	.00

Note. Sym - Symbolic Quotient of the UNIT; Non - Nonsymbolic Quotient of the UNIT.

*F statistic allows for the test of the significance of change in R squares as a result of each independent variable being entered into the equation. Symbolic Quotient first entering into the equation.

The Symbolic Quotient yielded a Multiple r of .76, R Square of .58 and an Adjusted R Square of .57. When the Non-Symbolic Quotient was added to the equation a significant .06 increase in the variance was noted. This model accounted for 63 % of the total variance. When the SAT reading comprehension score was used as the criterion variable, the Symbolic Global score was the better predictor yielding a Multiple R of .42, an R Squared of .17, and an Adjusted R of .15. The Non-Symbolic scores did not contribute. (See Table 8).

The Symbolic Quotient of the UNIT was expected to predict achievement, as measured by the PIAT-R and the SAT math scores significantly better than the Non-Symbolic Quotient of the UNIT. When the PIAT-R mathematics score was used as the criterion variable in multiple regression analyses, the Symbolic Quotient contributed significantly to the equation yielding a Multiple R of .70, R Square of .50, and an Adjusted R Square of .48. There was a significant .06 increase in the variance when the Non-Symbolic Quotient was added to the equation. This model accounted for 53% of the

Table 8

The Symbolic and Nonsymbolic Global Scores of the Universal Nonverbal Intelligence Test (UNIT) as Predictors of Reading Achievement as measured by the PIAT-R and SAT.

Area	Best Pred.	Stand. Beta	Mult. R	Adj. R Square	F	p
PIAT-R						
Read	Sym	.46	.76	.57	61.37	.01
	Non	.39			7.83	.01
SAT						
Read	Sym	.42	.42	.15	6.90	.01

Note. Sym - Symbolic Quotient of the UNIT; Non - Nonsymbolic Quotient of the UNIT.

*F statistic allows for the test of the significance of change in R squares as a result of each independent variable being entered into the equation.

variance. However, when the SAT mathematics score is used as the criterion variable, the Symbolic Quotient was the better predictor of math achievement yielding a multiple R of .61, an R Square of .37 and an Adjusted R of .35. The Non-Symbolic Quotient did not contribute significantly to the equation. Thus, partial support for the hypothesis was obtained. (See Table 9).

Test of the Experiential Deficits Hypothesis

To investigate the relationship between experience, defined as having parents who share the same language structure as their children versus those who do not, and children's cognitive ability, directional hypotheses were tested. The mean Symbolic Quotient of the UNIT for hearing impaired children of hearing impaired parents was predicted to be significantly better than the mean for hearing impaired children of hearing parents. The mean Symbolic Quotient of the UNIT for hearing impaired children of hearing impaired parents(M = 69.50, sd = 7.09) was not significantly different from the mean

Table 9

The Symbolic and Nonsymbolic Global Scores of the Universal Nonverbal Intelligence Test (UNIT) as Predictors of Math Achievement as measured by the PIAT-R and SAT.

Area	Best Pred.	Stand. Beta	Mult. R	Adj. R Square	F	p
PIAT-R						
Math	Sym	.40	.70	.48	42.24	.01
	Non	.38			5.85	.02
SAT						
Math	Sym	.53	.61	.35	19.24	.01

Note. Sym - Symbolic Quotient of the UNIT; Non - Nonsymbolic Quotient of the UNIT.

*F statistic allows for the test of the significance of change in R squares as a result of each independent variable being entered into the equation. Symbolic Quotient for hearing impaired children of hearing parents ($M = 56.00$, $sd = 16.97$). The obtained t of 1.54, was in the direction hypothesized, yet not statistically significant ($p > .05$).

The mean Non-Symbolic Quotient of the UNIT for hearing impaired children of hearing impaired parents was predicted not to differ from the mean for hearing impaired children of hearing parents. The mean Non-Symbolic Quotient of the UNIT for hearing impaired children of hearing impaired parents ($M = 56.00$, $sd = 16.97$) was not significantly different from the mean Symbolic Quotient for hearing impaired children of hearing parents ($M = 69.00$, $sd = 7.09$). Results from a two-tailed t yielded a p value of .06.

The mean Total Score of the TONI-II, as operationalized as figural problem solving, of hearing impaired children from hearing impaired parents ($M= 16.52$, $sd = 7.82$) was expected to be significantly higher than the mean of hearing impaired children from hearing parents ($M = 22.25$, $sd = 4.27$). Use of the Total Score of the TONI-II in a one-tailed test shows no significant difference ($p>.05$). However, it is noteworthy that means were in the hypothesized direction.

CHAPTER 4

DISCUSSION

One purpose of this study was to investigate the relative predictive and concurrent utility of the cognitive constructs purported to be assessed by two nonverbal tests of intelligence (i.e., the Universal Nonverbal Intelligence Test (UNIT) and the Test of Nonverbal Intelligence - II (TONI-II) with achievement for a population of hearing impaired students. Specifically, the UNIT's constructs of memory, reasoning, symbolic and nonsymbolic functioning and the TONI-II's construct of figural problem solving were used to determine the practical applications of such constructs in the prediction of and correlation with achievement.

Additionally, those interested in determining the usefulness of intellectual measures for hearing impaired populations may consider this population as two distinct groups, separating hearing impaired students with deaf parents from those with hearing parents. Since these two groups have been shown to differ significantly on both academic and intelligence measures, studying these groups as separate populations may lend support for the experiential deficits hypothesis.

The following discussion will first compare current research results to previous research conducted with the UNIT and TONI-II. Then, the practical implications for use of the UNIT and the TONI-II in assessment of hearing impaired individuals will be addressed. Research regarding the use of achievement as a criterion variable for intelligence measures with a hearing impaired population will be discussed. Finally, limitations of this study and suggestions for future research will be presented.

UNIT and TONI-II Constructs and Achievement

Multiple regression analyses were used to identify subsets of cognitive constructs that best predict total achievement (i.e., SAT Total Achievement). The Memory Global Score was found to be the best predictor of overall achievement. The entry of either the Reasoning Global Score or the TONI-II Deviation Quotient resulted in no significant increases in the Multiple R as compared to the Memory Score individually. When the SAT reading math scores are used as the criterion variables similar results were obtained.

However, when the PIAT-R reading and math scores are used as the criterion variables, both the memory and reasoning constructs were found to be strongly related to achievement. These results are consistent with a previous predictive validity study by Williams (1994). In her predictive validity study with hearing students, Williams suggests both Memory and Reasoning quotients of the UNIT are valuable indices of general intelligence construct and that the use of achievement as the criterion variable likely reflects a broad spectrum requiring both rote learning and reasoning abilities. That is, both retrieval skills and problem solving skills, are predictive of achievement.

Jensen hypothesized that measures of the construct of general intelligence is defined by both Level I Associative Skills, (i.e., rote learning and short term memory) and Level II Cognitive Skills (i.e., problem solving and reasoning). He also postulated that Level I and Level II abilities typically predict achievement differentially: that is, Level I tasks are better predictors of factual learning, and, Level II abilities are better predictors of academic skills requiring reasoning. Applying Jensen's hypotheses to the present study, results would appear to indicate that SAT scores for a hearing impaired population may require a narrower scope of skills, rote memory and factual learning, and PIAT-R, a broader range encompassing both Level I and Level II skills.

As cited in Chapter 1 Braden surveyed available published and unpublished works, and identified only one predictive study reporting correlations for a population of hearing impaired students (Hiroshoren, Hurley & Kavale, 1979). In this research the correlations between WISC-R PIQ and Stanford Achievement Test grade equivalents ranged from .09 to .35, and were considerably less than the expected moderate range (rs of .50 to .70) of criterion-related validity coefficients found with other populations. In a study of the correlations between SAT Subtests and Nonverbal IQs only limited correlations were revealed and these were well below any practical significance. As a result of his extensive meta-analysis, Braden suggests that achievement may be an inappropriate predictive criterion for IQ scores because deafness results in a drastic reduction in incidental learning of language and culturally-specific information. Additionally, he reasons that the low correlation between PIQ and academic achievement may result from a lack in overlap among the psychological constructs tapped by the tests.

Jensen (1985) and Braden (1984) argue that "deprived" deaf children do not conform to "Spearman's

Supposition" that the magnitudes of the mean differences between certain groups of individuals is directly related to the loadings of those measures on the hypothetical general intelligence factor "g." Jensen asserts that "Although there are significant ability differences between normally hearing and congenitally deaf children, the differences are not positively related to "g"; if anything, they are negatively related to the tests' g loadings. The language deprivation caused by deafness evidently takes its toll mostly on common factors other than "g", or on tests' specificities (i.e., nonerror variance that is not shared with any other tests in the battery) (Jensen, 1985, pp. 204-206). Jensen further asserts that many verbal tests are merely a reflection of a person's ability to recall stored information. On the other hand, nonverbal tests require more complex manipulations of elements to produce a correct answer. The difference in "g" loading for verbal and nonverbal tests indicate that nonverbal tasks rely more heavily on reasoning and inference than verbal tasks. "The verbal-nonverbal distinction is much more important than the "g"-loadedness of items for groups in which English is a second language" (Jensen, p. 161).

In view of the aforementioned research with hearing impaired individuals it is encouraging to find moderate to high correlations between the UNIT memory and reasoning constructs and achievement. Previous predictive research with nonverbal measures of intelligence and achievement have generally resulted in coefficients from .10 to .20 lower than those for verbal intelligence measures (Jensen, 1980). Coefficients obtained between UNIT Global Scores (i.e., Full Scale, Memory Quotient, Reasoning Quotient) with PIAT-R and SAT achievement ranged from .03 to .80, with the UNIT Full Scale and Memory Quotients yielding moderate to high correlations with achievement. One possible explanation for the high correlations relate to the age span of subjects (aged 6 to 17 years) and the use of raw scores in the data analysis. Both these factors may have contributed to the large correlations found between the independent and dependent variables.

Data depicting the relationship between the UNIT

Symbolic and Nonsymbolic Quotients and the prediction of achievement using SAT as the criterion measure shows the Symbolic Quotient to be the better predictor. Williams, 1994, in her study with a hearing population found similar results. However, when the PIAT-R is used as the criterion, both the Symbolic and Nonsymbolic Quotients enter significantly in the equation.

According to Bracken and McCallum (1992) the Symbolic Quotient is hypothesized as representing memory and reasoning skills amenable to symbolic mediation. The Nonsymbolic Quotient is purported to also tap reasoning and memory skills but to rely less heavily on symbol-laden content. The differential predictive strength of these constructs is likely influenced by the use of selected criterion variables, i.e., various indices of achievement.

There are different correlations between measures of intelligence and the two measures of achievement (i.e., PIAT-R and SAT). That is, all IQ measures correlate better with the PIAT-R than the SAT. What accounts for this difference? Several possible explanations exist.

First, the different task demands of each test may have confounded the analysis. The SAT reading comprehension test requires subjects to read a selection and answer questions about the reading. The PIAT-R reading comprehension test presents the student with a sentence which is followed by an array of four pictures. The student is required to read the sentence, then select the one picture which most clearly represents the meaning of the sentence.

A second plausible explanation is that the method of administration may have impacted the results. The SAT is a group administered test and the PIAT-R is an individually adminstered test. Students may have been more persistent in their efforts in a one-on-one situation.

A third possibility is the time differnce between administration of the two tests. The PIAT-R was administered at the same time as the ability measures and the SAT was an end of the year achievement test.

An equally plausible explanation for the differences obtained between independent and criterion variables may lie in the correlation between the criterion variables themselves. Correlation between PIAT-R and SAT reading and mathematics raw scores for the the present population resulted in the following: PIAT-R and SAT reading subtests resulted in a Pearson Correlation of .56 ($p < .05$ - 2-tail); PIAT-R and SAT math subtests resulted in a Pearson Correlation of .68 ($p < .05$ - 2-tail). Therefore, much of the variance not shared by the two measures accounted for the differences. [NOTE: Correlation studies have not been conducted between the SAT-8 and PIAT-R for hearing students, (personal communication with the research department of Harcourct-Brace publisher of the SAT). However, a 1977 concurrent validity study conducted between earlier versions of both tests, i.e., PIAT and STANFORD yield correlations of .50 for Reading Comprehension and .63 for Math Application for a group of economically deprived children.]

TONI-II authors give theoretical justification for selection of abstract/figural problem solving as the core for their test by referring to Sternberg's hypothesis that problem solving is a general construct of intelligent behavior rather than a subcomponent or splinter skill (Resnick & Glaser, 1976, and Gardner, 1983). The construct of abstract figural problem solving was found to be moderately related to achievement when the PIAT-R was used as the criterion. However, when the SAT scores were used as the criteria for achievement, no significant relationship was found. These findings do not support the predictive validity study conducted by Brown, Sherbinou, and Johnsen in 1982. In their research on 16 deaf students ranging in age from 16.1 to 18.8 they report predictive relationships between the TONI and the SAT reading comprehension and math concepts of .70. However, since the current investigation included a larger number of students, ages 5 through 18, the difference in correlations may be a reflection of hearing impaired students' more gradual English language development.

In the present investigation the predictive and concurrent utility of the memory construct operationalized by the UNIT memory subtests consistently revealed moderate to high correlations, from .41 to .70, across achievement measures. Additionally, the Symbolic Quotient, revealed similar findings with correlations ranging from .42 to .76 across criterion variables. Thus, the UNIT's memory and symbolic constructs demonstrated predictive and concurrent validity equivalent to or better than other nonverbal assessment measures currently available.

Exploration of the Experiential Deficits Hypothesis

Comparison of the mean differences of Symbolic, Nonsymbolic Quotients of the UNIT and Deviation Quotients of the TONI-II for deaf children of deaf parents and deaf children of hearing parents did not yield expected results. The hypotheses that deaf children of deaf parents would perform significantly better than deaf children of hearing parents on symbolically mediated tasks, as found on the UNIT, and figural problem solving items, as found on the TONI-II were not supported. However, all mean difference analyses were in the predicted direction, suggesting mild support of Levine's

Experiential Deficit Hypothesis (i.e., that persons with hearing impairments possess normal intelligence, but that differences in academic achievement or performance on intelligence tests is a result of experiential deficits rather than sensory loss per se.) Levine's notion of experiential deficits, i.e., that deaf children of hearing parents do not have ready access to the language around them and consequently suffer from a lack of much information which language conveys may be one explanation for current results. However, considering the small number of participants in the category of deaf children of deaf parents, any conclusions regarding Levine's Hypothesis is extremely tentative.

Limitations and Implications for Future Research

There are several limitations associated with the internal and external validity of this investigation. Demographic characteristics of the sample limit generalizability. The sample was comprised of students attending a special school for the hearing impaired.

Students with hearing impairments attending regular public school programs were not included. Additionally, only those students with signed parental release forms were included in the study. Only four students with deaf parents were included in this study, therefore, further research using larger populations of hearing impaired students is necessary for a better test of the Experiential Deficit Hypothesis. Generalizations of these results should be limited to students sharing similar demographic characteristics.

Historically, the development of intellectual measures for persons with hearing impairments has lacked strong theoretical underpinnings making interpretation of their results tentative at best. Nonverbal measures have generally addressed only a narrow scope of abilities and the efficacy of their predictive validity is questionable. The UNIT provides a theory-driven, multi-dimensional method of intelligence assessment. The method of administration, pantomimed instructions, lends itself well to persons for whom traditional verbal measures are inappropriate, i.e., hearing impaired individuals. Predictive validity of the UNIT, as evidenced by this study, suggests moderate to high correlations between UNIT Global Scores and achievement for a group of hearing impaired students. Since few currently available nonverbal intelligence measures have conducted correlational and predictive validity studies with this population, the UNIT fulfills a void in the ability testing of hearing impaired students. Extensive testing of students with hearing impairments to provide norms for the newly developed UNIT may be warranted.

REFERENCES

References

Anderson, R.J. & Sisco, F.H. (1977). Standardization of the WISC-R for deaf children. (Office of Demographic Studies Publication Series T, No. 1). Washington, DC: Gallaudet College.

Altshuler, K., Deming, W., Wollenweider, J., Rainer, J. & Tender, R. (1976). Impulsivity and profound deafness: a cross cultural inquiry. American annals of the Deaf, 121, 331-339.

Ballantyne, J., Martin, M. C., & Martin, Antony. (1993). Deafness. London: Whurr Publishers.

Binet, A. & Simon, T. (1961). The development of intelligence in children. In T. Shapley (Ed.). Classics in psychology (pp. 872-919). New York: Philosophical Library. (Reprinted from L'Annee Psychological, 1905, 11, 163-191.)

Bodner-Johnson, B. (1986). The family perspective. In D. M. Luterman (Ed.) Deafness in Perspective. San Diego, CA: College-Hill Press.

Boyd, J. & Shapiro, A. (1986). A comparison of the Leiter-International Performance Scale to WPPSI performance with preschool deaf and hearing impaired children, Journal of Rehabilitation of the Deaf, 20, 1, 23-26.

Bracken, B., & McCallum, R.S., (in press). Universal Nonverbal Intelligence Test. Chicago: Riverside Publishing.

Braden, J.P. (1985). The structure of nonverbal intelligence in deaf and hearing subjects. American Annals of the Deaf, 131, 496-501.

Braden, J.P. (1989). The Criterion-Related Validity of the WISC-R Performance Scale and Other Nonberbal IQ Test for Deaf Children, American Annals of the Deaf, 134, 5, 329-332.

Braden, J.P. (1989). Fact or Artifact? An empirical Test of Spearman's Hypothesis, Intelligence, 13, 149-155.

Brill, R.G. (1974). The superior IQs of deaf children of deaf parents. In P.J. Fine (Ed.) Deafness in infancy and early childhood. New York: Medcom.

Brown, L., Sherbenou, R.J., & Johnsen, S.K. (1990). Test of Nonverbal Intelligence - Second Edition. Austin, TX: Pro-Ed.

Brunich, P.(1981). Relationship between intellectual functioning and communicative competence in deaf children. Journal of Communication Disorders, 14, 429-434.

Burgemeister, B.B., Blum, L.A., & Lorge, I. (1972). Columbia Mental Maturity Scale - Third Edition. New York: Harcourt, Brace, Jovanovich.

Cattell, J. M. (1890). Mental tests and measurements. Mind, 15, 373-380.

Cole, D.A. (1987). Methodological contribution to clinical research: Utility of confirmatory factor analysis in test validation research, Journal of Consulting and Clinical Psychology, 55, 4, 584-594.

CTBS Macmillian/McGraw Hill (1989). Comprehensive Test of Basic Skills, Fourth Edition Technical Bulletin, 1. Monterrey, CA: CTBS.

Das, J. P. (1973). Cultural deprivation and cognitive competance. In N. R. Ellis (Ed.), International review of research in mental retardation (Vol. 6, pp. 1-53). New York: Academic Press.

Dillon, R. (1979). Improving validity by testing for competence: Refinement of a paradigm and it's application to the hearing impaired. Educational and Psychological Measurement, 39, 363-371.

Dillon, R. (1980). Cognitive style and elaboration of logical abilities in hearing impaired children. Educational and Psychological Measurement, 39, 363-371.

Dweck, C.S. & Elliot, E.S. (1983). Achievement motivation. In E.M. Herrington (Ed.), Socialization, personality, and social development (pp. 643-691). New York: Wiley.

Elliott, H., Glass, L., & Evans, J. W. (1987). Mental Health Assessment of Deaf Clients: A Practical Manual. Boston: Little, Brown and Company.

Erikson, E. H. (1968). Identity, youth and crisis. New York: Norton.

Freeman, R. D., Carbin, C. F., & Boese, R. J. (1981). Can't your child hear? A guide for those who care about deaf children. Baltimore, MD: University Park Press.

Gallaudet Research Institute (1987). Factors predictive of literacy in deaf adolescents with deaf parents/Factors predictive of literacy in deaf adolescents in total communication programs (Report submitted to the National Institute of Neurological and Communicative Disorders and Stroke: Project No. NIH-NINCDS-83-19). Washington, DC: Gallaudet University.

Gardner, H. (1983). Frames of mind: the theory of multiple intelligences. New York: Basic Books.

Gardner, E., Rudman, H., Karlsen, B., & Merwin, J. (1990). Stanford Achievement Test, Eighth Edition. The Psychological Corporation.

Glass, G.V., & Hopkins, K.D. (1984). Statistical methods in education and psychology (2nd ed.). Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

Guilford, J.P. (1967). The nature of human intelligence. New York: McGraw-Hill.

Harris, R. (1978). The relationship of impulse control to parent hearing status, manual communication, and academic achievement in deaf children. American Annals of the Deaf, 123, 52-67.

Harvey, M. A. (1989). Psyhotherapy with deaf and hard-of-hearing persons: a systemic model. New Jersey:

Lawren Erlbaum Associates.

Hiskey, M.S. (1966). Hiskey-Nebraska Test of Learning Aptitude. Lincoln, Nebraska: Union College Press.

Hiskey, M.S. (1979). A Study of the Intelligence of Deaf and Hearing Children, American Annals of the Deaf, 124, pp. 134-149.

Hiskey-Nebraska Test with deaf children (1978). American Annals of the Deaf, 122, 392-394.

Hoffenmeister, R.J. (1988). Cognitive assessment in deaf preschoolers. In T.D. Wachs & R. Sheehan (Eds.), Assessment of young developmentally disabled children (pp. 109-126). New York: Plenum Press.

Hoffmeister, R. J. (1985). Families with deaf parents: A functional perspective. In K. Thurman (Ed.). Handicapped families: Functional perspectives. New York: Academic Press.

Hirshoren, A., Hurley, O.L., & Hunt, J.T. (1977). The WISC-R and the Hiskey-Nebreska Test with deaf children. American Annals of the Deaf, 122, 392-394.

Horn, J. L. (1988). Thinking about human abilities. In J. R. Nesseroade & R. B. Cattell (Eds.). Handbook of multivariate experimental psychology, second edition (pp. 645-685).

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

Hubert, T., & Koller, J. (1984). A test of the learning potential hypothesis with hearing and deaf students. Journal of Educational Research, 78, 22-27.

Ishan, R.P. & Kamin, L.J. (1993). blackness, Deafness, IQ, and g. Intelligence, 17, 37-46.

Itard, J. (1801). De l'education d'un homme sauvage. [Translated by Edmund Fawce, Peter Ayrton, and Joan White], New York, Monthly Review Press, 1972.

Jensen, A.R. (1980). Bias in mental testing. New York: Macmillan.

Jensen, A.R. (1985). The nature of black-white differences on various psychometric tests: Spearman's hypothesis. The Behavioral and Brain Sciences, 8, 193-219.

Jensen, A.R. (1993). Spearman's Hypothesis Tested with Chronometric Information-Processing Tasks. Intelligence, 17, 47-77.

Kaufman, A.S. & Kaufman, N. L. (1983). K-ABC. Kaufman Assessment Battery for Children. Circle Pines, MN: American Guidance Service.

Keane, K. (1983). Application of mediated learning theory to a deaf population: A study in cognitive modifiability. Unpublished doctoral dissertation, Columbia University, New York.

Keane, K. (1987). Assessing deaf children. In C.S. Lidz (Ed.) Dynamic assessment (pp.360-378). New York: Guilford.

Keeny, B. P. (1983). Aesthetics of change. New York: Guilford.

Kelly, R.R. & Tomlinson-Keasey, C. (1977). Hemispheric laterality of deaf children for processing words and pictures visually presented to the hemifields. American Annals of the Deaf, 122, 525-533.

Leiter, R.G. (1979). Leiter International Performance - Instruction Manual. Chicago: Stoelting.

Levine, E.S. (1974). Psychological tests and practices with the deaf: A survey of the state of the art. Volta Review, 76, 298-319.

Levine, E. S. (1977). The preparation of psychological service providers to the deaf," Journal of Rehabilitation of the Deaf, 4.

Levine, E. (1981). The ecology of early deafness. New York: Academic Press.

Luckner, J. (1992). Problem Solving: A comparison of Hearing-Impaired and Hearing Individuals. American Annals of the Deaf, 25, 4, 21-27.

Lui, Herman Chan-En. Non-verbal Intelligence Test for Use in China. Teachers College, Columbia University Contribution to Education, No. 126, New York 1922.

Luria, A (1961). The role of speech in the regulation of normal and abnormal behaviors. New York: Liveright.

MacKane, K. A comparison of the intelligence of deaf and hearing children; a study of the reaction of comparable groups of deaf and hearing children to three performance scales and a non-language test. Teachers

College, Columbia University, no.585, New York, AMS Press, 1972 (Reprint of the 1933 ed., issued in series)

McCallum, R.S. & Bracken, B.A. (in press). The Nonverbal Intelligence Test. In Flanagan, Genshaft, & Harrison (Eds.), Beyond Traditional Intellectual Assessment: Contemporary and Emerging Theories, Tests and Issues.

Markwardt, F.C. (1989). Peabody Individual Achievement Test - Revised. Minnesota: American Guidance Service.

Mayberry, R., & Woodlinger-Cohen, R. (1987). After the revolution: Educational practice and the deaf child's communication skills. In E. D. Mindel & M. Vernon (Eds.), They Grow in Silence, (2nd ed.). Boston, MA: Little Brown.

Marschark, M. (1993). Psychological Development of Deaf Children. New York: Oxford Univrsity Press.

Meadow, K.P (1975). The development of deaf children. In E.M. Hetherington (Ed.), Review of child development research (Vol. 5, pp. 441-508). Chicago: University of Chicago Press.

Meadow, K. & Dyssegard, B. (1983). Social-emotional adjustment of deaf students. Teachers' ratings of deaf children: an American-Danish comparison. International Journal of Rehavilitation Research, 6(3), 345-348.

Moores, D. (1985). Reactions from a researcher's point of view. In D. Martin (Ed.). Cognition, education and deafness (pp. 224-228). Washington, DC: Gallaudet College Press.

Myklebust, H. (1953). Towards a new understanding of the deaf child. American Annals of the Deaf, 98, 345-357.

Myklebust, H. (1960). The psychology of deafness. New York: Grune & Stratton.

Naglieri, J.A. & Jensen, A.R. (1987). Comparison of black-white differences on the WISC-R and the K-ABC: Spearman's hypothesis. Intelligence, 11, 21-43.

Naglieri, J.A. & Prewett, P.N. (1990). Nonverbal intelligence measures: A selected review of instruments and their use. In R.W. Kamphaus & C.R. Reynolds (Eds.) Handbook of psychological and educational assessment of children (pp. 348-370). New York: Guilford Press.

Naglieri, J.A. & Welch, J.A. (1991). Use of Raven's and Naglieri's Nonverbal Matrix Texts. Journal of Rehabilitation of the Deaf, 24, 3 & 4, 98-103.

Naglieri, J.A. & Prewett, P.N. (1989). Nonverbal Intelligence: A selected review of instruments and their use. In R. W. Kamphaus & C. R. Reynolds (Eds.), Handbook of psychological and educational assessment of Children. (pp. 348-370). New York: Guilford Press.

Pascual-Leone, J. (1970). A mathematical model for the transition rule in Piaget's developmental stages. Acta Psychologica, 32, 301-345.

Pitner, R., and Patterson, D.G.: "The Binet Scale and the Deaf Child," Journal of Educational Psychology, Vol. VI (1915), 202.

Raven, J.C., (1965) Progressive Matrices. London, H. K. Lewis.

Raven, J.C., Court, J.H., & Raven, J. (1986). Manual for Progressive Matrices and Vocabulary Scales (1986 Edition, with U.S. norms). London: Lewis.

Raven, J.C. (1994). Raven Progressive Matrices. England: J.H. Court, H.K. Lewis & Co. Ltd.

Rayson, B. (1987). Deaf parents of hearing children. In E. D. Mindel & M. Vernon (Eds.) They grow in Silence (2nd ed.). Boston, MA: Little Brown.

Ray, S. (1979). An adaptation of the `Wechsler Intelligence Scales (Performance) for Children-Revised' for the deaf. Unpublished dissertation.

Reed, M.T. & McCallum, R.S. (1995). Construct validity of the Universal Nonverbal Intelligence Test (UNIT). Psychology in the schools, 12, 277-290.

Robinson, L. D. (1978). Sound minds in a soundless world. U. S. Department of Health, Education and Welfare, Public Health Service, Alcohol, Drug Abuse, and Mental Health Administation, National Institute of Mental Health.

Rodda, M. & Groves, C. (1987). Language, cognition and deafness. Hillsdale, NJ: Erlbaum.

Sattler, J.M. (1992). Assessment of children (3rd ed.). San Diego: Publisher, Inc.

Schlesinger, H. S. & Meadow, K. (1972). Sound and sign: Childhood deafness and mental health. Berkeley, CA: University of California Press.

Seguin, E. (1907). Idiocy: and its treatment by the physiological method. New York: Press of Brandow Printing.

Shockley, W. (1987). Jensen's data on Spearman's hypothesis. No artifact. Behavioral and Brain Sciences, 10, 512.

Shouksmith, G. (1970). Intelligence, creativity, and cognitive style. New York: Wiley.

Smith, M.K. & McManis, D.L. (1977). Concurrent validity of the Peabody Achievement Test and the Wide Range Achievement Test. Psychol Rep, 41, 1279-1284.

Spearman, C. E. (1927). The abilities of man. New York: Macmillan.

Spearman, C. E. (1923). The nature of intelligence and the principles of cognition. London: Macmillan.

Stern, W. The Psychological Methods of Testing Intelligence. Translated by G.M. Whipple. Warwick and York, Baltimore (1914).

Sternberg, R. J. (1985). Beyond IQ: A triarchic theory of human intelligence. New york: Cambridge University Press.

Stein, L. K., Mindel, E. D. & Jabaley, T. (1981). Deafness and Mental Health. New York: Grune & Stratton.

Thorndike, E.L. "A Standard Group Examination Of Intelligence Independent of Language." Journal of Applied Psychology, Vol.III, No.I (March, 1919) pp. 13-32.

Thurstone, L. L. (1924). The nature of intelligence. New York: New York: Harcourt, Brace.

Todman, J., & Cowdy, N. (1993). Processing of Visual-Action Codes by Deaf and Hearing Children: Coding Orientation or M-Capacity? Intelligence 17, 3, 237-249.

Tzuriel, David, & Caspi, Naomi. Cognitive Modifiability and Cognitive Performance of Deaf and Hearing Preschool Children. The Journal of Special Education, Vol. 26/No. 3 1992 pp. 235-252.

Vernon, M. (1968). Fifty years of research on the intelligence of deaf and hard of hearing children: A review of literature and discussion of implications. Journal of Rehabilitation of the Deaf, I, 1-12.

Vernon, M. (1976). Psychological evaluation of hearing-impaired children. In L.L. Lloyd (ed.).

Communication assessment and intervention strategies. Baltimore: University Park Press.

Watts, W.J. (1979). The influence of language on the development of quantitative, spatial, and social thinking in deaf children, American Annals of the Deaf, 124, 46-56.

Wechsler, D. (1974). Wechsler Intelligence Scale for Children-Revised. New York: The Psychological Corporation.

White, K. (1982). Defining and prioritizing the personal and social competancies needed by hearing impaired students. Volta Review, 84, 266-274.

Williams, P.C. & McCallum, R.S. (in press). Predictive Validity of the Universal Nonverbal Intelligence Test with Achievement. Assessment in Rehabilitation and Exceptionality.

Woodcock, R.W. & Johnson, M.B. (1989). Woodcock-Johnson Psychoeducational Battery-Revised. Chicago: Riverside.

Whorf, B.J. (1956). Language, thought and reality. New York: John Wiley & Sons.

Zigler, E. & Butterfield, E.C. (1968). Motivational aspects of changes in IQ test performance of culturally deprived nursery school children, Child Development, 39, 1-14.

VITA

Connie Theriot England was born in St. Martinville, Louisiana. She attended Our Lady of Mercy Elementary School and Mercy High School. She graduated with a Bachelor of Arts Degree from the University of Southwestern Louisiana in 1976, with Majors in Speech Pathology/Audiology and Secondary English Education. She was a Speech Therapist and Deaf Educator for two years, then moved with her family to Knoxville, Tennessee, to attend the University of Tennessee. She earned a Master of Science Degree in Special Education in 1979. She taught high school to students with hearing impairments for three years. In 1982, she returned to Louisiana and was employed as a supervisor of Special Education in the Jefferson Parish School District, New Orleans, Louisiana. She attended the University of New Orleans, earning a Master's + 30 hours and receiving certification as a Supervisor of Instruction. After working in New Orleans for ten years, she and her husband once again returned to

Knoxville, to open an antique shop and to allow Connie the opportunity to work on a Doctor of Philosophy Degree in Education with a major in School Psychology.