INTELLECTUAL ASSESSMENT IN A MULTI-CULTURAL POPULATION

PAMELA ROSE PENNINGTON

INTELLECTUAL ASSESSMENT IN A MULTI-CULTURAL POPULATION

A Research Paper

Presented to the

Graduate and Research Council of

Austin Peay State University

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

---

by

Pamela Rose Pennington

May 1988

To the Graduate and Research Council:

I am submitting herewith a Research Paper, written by Pamela Rose Pennington entitled "Intellectual Assessment in a Multi-cultural Population". I have examined the final copy of this paper for form and content, and I recommend that it be accepted in partial fulfillment of the requirements for the degree Master of Arts, with a major in Clinical Psychology.

John D. Martin
Major Professor

Accepted for the Graduate and Research Council:

William H. Ellis
Dean of the Graduate School

ACKNOWLEDGEMENTS

The writer wishes to express her appreciation to Dr. John Martin for his assistance in finally making this research paper a reality; to Dr. Susan Kupisch and Dr. Jean Lewis for their helpful suggestions; and to her friend Thelma Budias for her invaluable typing expertise.

The writer also wishes to express enormous gratitude to her family for their constant encouragement, and to her husband, Jeff, for his patience and support throughout the completion of this writer's graduate study.

TABLE OF CONTENTS

CHAPTER 1

Introduction

For centuries, philosophers and researchers have studied the human mind in an attempt to understand what we call "intelligence." As a result of these studies, a multitude of different theories and definitions of intelligence emerged. According to David Wechsler, one of the most well known pioneers in the study of intelligence, "intelligence, as a hypothetical construct, is the aggregate or global capacity of an individual to act purposefully, to think rationally, and to deal effectively with his environment" (cited in Matarazzo, 1972, pg. 79).

E.L. Thorndike proposed subdividing intelligence into three main types: (1) abstract or verbal intelligence, involving facility in the use of symbols; (2) practical intelligence, involving facility in manipulating objects; and (3) social intelligence, involving facility in dealing with human beings (cited in Matarazzo, 1972). Charles Spearman thought that intellectual abilities could be expressed as functions of two factors: a general or intellectual factor (g) common to every ability, and a specific factor (s), specific to any particular ability and in every case different from that of all others (cited in Matarazzo,

1972). Raymond Cattell is well known for his theory of fluid and crystalized intelligence. According to this theory, fluid ability is usually described as a general ability to perceive relationships, to analyze, and to reason, especially in abstract or novel situations. Crystallized ability involves skills that may be enhanced by one's environment and educational background, such as reading, numerical operations, and verbal ability (cited in Matarazzo, 1972).

Jerome Sattler (1988) cited Phillip Vernon who believes that three meanings are usually associated with intelligence: one meaning is that intelligence refers to innate capacity or genetic equipment. This meaning reflects the genotypic form of intelligence and cannot be measured directly. A second meaning of intelligence refers specifically to behaviors involving learning, thinking, and problem solving. It is a function of the interaction of genes with the prenatal and postnatal environment, the phenotypic form. Sattler described a third meaning for intelligence which refers to results obtained on intelligence tests that sample verbal, nonverbal, or mechanical abilities.

Once the basic constructs of intelligence were formulated, problems in operationalizing it emerged.

Researchers needed ways of measuring the range of human intelligence. Thus, the intelligence test was born in an attempt to meet that need.

With this idea in mind, Binet and Simon provided a definition of intelligence as they assessed it in 1908:

> A subject has the intellectual development of the highest age at which he passes all the tests, with the allowance of one failure in the tests for age...(providing that you) give him the benefit of an advance of one year every time he passes at least five of the tests beyond (this) level, and the benefit of an advance of two if he has passed at least ten above (this) level (cited in Matarazzo, 1972, pg. 67).

Wechsler (1981) provided a modernized view of Binet and Simons definition:

> IQs thus obtained have the same basic meaning regardless of the subject's age. An IQ of 120 obtained by a 60-year-old and by a 20-year-old reflects the same relative standing among people of the subjects age groups. But in one sense, identical IQs do not have the same meaning at different ages. This is because average test scores change with age, typically rising to a

> peak in the years of young adulthood, then falling off somewhat later on. Thus, a higher level of test performance - that is, more ability in an absolute sense - is needed to obtain a given IQ at age 25 than at age 65, or conversely, a lower level of test performance is needed to obtain a given IQ at age 65 than at age 25 (pg. 9).

According to Ausubel and Sullivan (cited in Sattler, 1988), three factors play an important role in the stability and change of IQ scores: measurement factors, genetic factors, and environmental factors. Measurement factors involve things that can affect the test results, such as situational factors, administration and scoring errors or test taking experience. Genetic factors are the developmental changes that may occur in an individual. Environmental factors include outside stimuli that might affect test performance, such as physical and emotional factors, or changes in cognitive stimulation and motivation.

Kaufman (1979), stated that specialization of the hemispheres must also be considered in order to fully understand human intelligence. The left brain is primarily language oriented and is responsible for

abilities such as spoken and written language, number skills, scientific skills and reasoning. The right hemisphere is responsible for processing of visual-spatial information and includes less tangible abilities such as music and art awareness, insight and imagination. Kaufman thought that in order to obtain a global respresentation of intelligence, measures of intellectual assessment should include items that utilize the abilities of each hemisphere separately and in unison.

For example, the Wechsler Adult Intelligence Scale-Revised contains Verbal and Performance Scales which employ tasks aimed at measuring different areas of ability. However, as Kaufman (1979) pointed out, while the Verbal Scale basically utilizes left-brain processing, the Performance Scale requires much interhemispheric integration. Coding, for example, uses the analytic and sequencing abilities associated with the left hemisphere. Picture arrangement requires temporal sequencing skills which are also associated with the left hemisphere. Most intelligence tests require good verbal comprehension in order to understand the instructions of the nonverbal tests, which automatically suggests utilization of both

hemispheres.

Some research has suggested that many tests may be culturally biased because of the emphasis placed on the respective hemispheres. Ornstein (cited in Kaufman, 1979) thought that some cultures may place more emphasis on different skills, and consequently emphasize different hemispheres. He suggested that Eastern and Western societies are cognitively different in this respect. For example, the emphasis placed on language and logical thinking in Western societies suggests a well developed left hemisphere. Eastern societies appear to be more intuitive, focusing on cultures, mysticism, and religion, which are right hemispheric traits.

Bogen (cited in Springer and Deutsch, 1981), compared the intellectual performance of 1220 individuals from different cultural backgrounds. This population included Hopi Indians, urban blacks, and rural and urban whites. The two tests selected for use in the study were the Street Gestalt Completion Test and the Similarities subtest of the Wechsler Adult Intelligence Scale. These tests were chosen because of their emphasis on the abilities of the right and left hemispheres, respectively. Results of the study

revealed that the Hopi Indians and urban blacks tended to rely more heavily on their right hemispheres in thinking than the other groups tested.

Since cross-cultural differences in processing verbal and nonverbal information are so varied, it is often difficult to find appropriate measures of intellectual assessment. Many different tests have been developed which place emphasis on nonverbal abilities. This allows the examiner to have a choice in selecting the method that is best suited for the individual being assessed.

However, caution must be exercised in interpretation of tests which measure only a partial component of intelligence. According to Hopkins and Bracht (1975), "all too frequently, the long-term stability of IQ scores is treated as if the findings for one type of intelligence test generalize to all intelligence tests" (pg. 470). From the results of their ten year study, they concluded that nonverbal IQ scores are less stable over time than verbal IQ scores, and thus cautioned use of nonverbal tests for long term prediction or placement.

Gordon and Terrell (cited in Suinn, 1985) elaborated further on the potential misuse of test

scores:

> Critics of testing have objected not so much to the tests themselves as to the ways in which tests have been used. In keeping with the current social and political context for testing...their concerns have more to do with the theoretical and technical aspects of testing (pg. 680).

When used correctly, the advantages of using nonverbal intelligence tests in a multi-cultural population appear to outweigh any disadvantages. First, they can serve as a culture fair measure to compare with other standard measures of assessment. Also, most are less time consuming to administer and score. Many require minimal equipment and training to administer. Differentiating between the many tests available can be difficult, however. The present research paper will attempt to assist in simplifying this decision.

The Problem

With all of the methods of intellectual assessment now available, there are still some institutions struggling to find an appropriate means of assessment for special populations for whom verbal intelligence

tests are inappropriate. An example is the military population referred for assessment in a United States Military Hospital.

Military hospitals frequently see patients of a variety of nationalities and cultures, many who speak little or no English. Like many mental health professionals, the mental health clinics of these hospitals often use intelligence tests to assist in the diagnosis of a variety of mental and emotional disorders. While the emergence of nonverbal intelligence tests have been helpful in providing feedback for these special populations, some problems still exist in choosing an appropriate test.

A review of the current literature reveals a surprising lack of recent research on nonverbal intelligence testing for adults. The majority of available studies focus on children. This research paper will review the available data on nonverbal intelligence testing as it relates to the specific problems of testing a multi-cultural population of adults in a United States Military Hospital. The Test of Nonverbal Intelligence (TONI) will be evaluated as a possible alternative to other nonverbal tests which do not fully meet the special needs of this population.

CHAPTER 2

Review of the Literature

Numerous attempts have been made at developing a culture-fair or culture-free intelligence test. However, it was finally pointed out by Frijada and Jahoda (1966) that a truly culture-fair test is unattainable, since all tests are "anchored" in an originating culture or culture area and are unfair to people who reside in another culture or who belong to an ethnic minority group. The term "culturally reduced" has been used by Sattler (1988) to refer to tests that are "less dependent on exposure to specific language symbols" (pg. 579).

Arthur Jensen, one of the leading authorities of test bias, believed that culturally reduced tests are different from traditional intelligence tests in several important ways. In his book Bias In Mental Testing, Jensen (1980) suggested ways of reducing the culture loading of tests. In order to produce a culturally reduced test, Jensen stated that performance tests should replace paper-and-pencil tests. Pantomimed instructions should be used in place of written or oral instructions. Practice items should be employed to ensure the subject fully comprehends instructions. Items should be composed of abstract

content instead of culturally loaded pictures or passages. Untimed items should replace timed items. Items should be novel so that test items will not be biased by the content of previously learned information.

The test-taking attiudes or response sets of the culture being tested also must be considered. According to Lonner (1985), "differences among test scores that may have resulted from less than optimal opportunities to present oneself fully (for example, cultural differences in the use of time, cooperativeness, familiarity with tests, or in following instructions) could invalidate test results" (pg. 606). Lonner further explained that response sets can also vary cross culturally. For example, individuals from some cultures may tend to agree with nearly everything, perhaps out of politeness. Some may give only socially desirable answers, while others tend to be very careless in their responses.

In choosing an appropriate method of intellectual assessment for a multi-cultural population, the central issue is not whether a test is biased because it yields differences among populations. The controversy centers on the decisions that may be made when differences are

found (Cronbach, 1984). Sundberg and Gonzales (1981) state that "a test is unfair when it is irrelevant or misleading for the decisions it is intended to facilitate" (pg. 481).

Several tests have been developed in an attempt to meet the special assessment needs of a variety of cultures. The Culture Fair Intelligence Test (CFIT) was designed to be a culturally reduced test that would measure fairly the general intelligence level of individuals having different cultures, national languages, or countries of origin. It employs Cattell's concept of fluid intelligence. The CFIT is a paper-and-pencil test with strict time limits. The instructions for taking the test are verbal; however, the test may be obtained in 23 foreign language editions. The test may be used with individuals ranging from age four to adult. Problems with the CFIT are reported in a critique by Koch (1984). He stated that a modernization of the CFIT is in order, since it has been more than 20 years since it has been revised. Koch further elaborated that the format of the test booklets is quite old fashioned, the figural drawings used for the test are badly dated, and the elaborate oral instructions need to be simplified. The actual

printing quality of many of the items results in pictures that are difficult to see clearly and are sometimes misleading. Also, the norming of the CFIT has never really been adequate. The standardization samples employed in the various revisions over the years have always been relatively small by usual test publishing standards, the samples used for norming have been convenience samples, and the nature of the samples has never been clearly specified.

Another test developed specifically for multi-cultural assessment is the System of Multicultural Pluralistic Assessment (SOMPA). The SOMPA was designed to incorporate medical, social and pluralistic information in the assessment of the cognitive, perceptual-motor and adaptive behavior of black, white, and Hispanic-American children between the ages of 5-0 and 11-11 years. It employs the use of several different procedures, such as the Adaptive Behavior Inventory for Children and the Wechsler-Intelligence Scale For Children-Revised. While the SOMPA is for children only, it gives yet another example of a culturally reduced method of assessment (cited in Sattler, 1988).

The Goodenough-Harris Drawing Test requires

verbal instructions, but can be modified to fit many cultures. In this test, the subject is asked to draw a picture of a man, of a woman, and of self. The various features of the drawings can be quantified to establish a mental-age score. However, the scoring of various aspects of clothing may make the test highly culturally loaded outside of Western societies. Also, the test has been found to work best with children below the age of 12 or 13 (cited in Jensen, 1980).

Since one of the major problems in testing a multi-cultural population is the language barrier, a test in which the examiner pantomimes the instructions appears advantageous and appropriate. Tests such as the Progressive Matrices, the Leiter International Performance Scale, and the Test of Nonverbal Intelligence all include pantomimed instructions.

The Progressive Matrices was developed in 1938 by J.C. Raven. It was designed to measure a person's ability to form perceptual relations and to reason by analogy independent of language and formal schooling, and may be used with persons ranging in age from six years to adult (cited in Llabre, 1984). It is composed of three instuments: The Standard Progressive Matrices, The Coloured Progressive Matrices and

The Advanced Progressive Matrices. All three tests are measures of Spearman's g. The nonverbal nature of the Progressive Matrices makes it useful for testing persons from different linguistic backgrounds as well as those with communication disorders or limited language proficiency. It can serve as a screening device of intellectual ability.

The Leiter International Performance Scale is a 54-item nonverbal test of intelligence appropriate for use with individuals ages two through adult (cited in Matey, 1984). According to Matey, Russell Leiter was originally interested in examining the differences in native intelligence between children of different races. His intention was to develop a test of intellectual ability that would eliminate the language function and thus enable a fair comparison of children from different racial backgrounds. Unfortunately, both the Progressive Matrices and the Leiter scale contain outdated norms (Sattler, 1988).

Since the Test of Nonverbal Intelligence (TONI) is the most recently developed nonverbal intelligence test available at this writing, it merits further investigation to evaluate its potential for multi-cultural assessment with an adult population.

The TONI was developed in 1982 by Linda Brown, Rita J. Sherbenou, and Susan J. Dollar. It was designed to be a nonbiased test of intellectual ability for use with handicapped or minority populations who may require language-free testing formats. There are two equivalent forms of the TONI, each containing 50 items. The administration of the TONI contains no reading, writing, or verbalization: instructions are pantomimed by the examiner and the subject responds by pointing. Because of this format, the TONI can be used for intellectual assessment of subjects suspected of having reading, writing, speaking or listening problems. These include people who are bilingual or nonEnglish speaking, speech or language handicapped, deaf, learning disabled, mentally retarded, or victims of stroke or other brain injury. The TONI may be used with subjects ranging in age from 5-0 through 85-11 years. It requires approximately 20-30 minutes to administer and may be given individually or in small groups up to five subjects. The test yields percentile ranks and a TONI quotient with a mean of 100 and a standard deviation of 15.

The developers of the TONI used the guidelines established by Jensen (1980) on preparing a culturally

reduced test. Unlike the other nonverbal tests previously mentioned, the TONI meets all of Jensen's criteria.

The TONI focuses primarily on the problem solving aspect of intelligence, employing a content of an abstract/figural nature. This problem solving format helps to decrease the cultural loading found in many of the traditional intelligence tests and is similar to that of the Progressive Matrices and Leiter International Performance Scale.

According to the test manual, the TONI was standardized on a large, nationally representative population of 1,929 subjects from 28 states. It is highly reliable with normal subjects and with populations of retarded, learning disabled, and deaf students. Internal consistency and alternate forms reliability coefficients are in the .80s and .90s at most ages. Concurrent validity of the TONI was established by correlating performance on the TONI with performance on other measures of intelligence and achievement, including the Wechsler Intelligence Scale for Children-Revised, Raven's Progressive Matrices, the Leiter International Performance Scale, the Otis-Lennon Mental Ability Test, the Iowa Tests of Basic Skills,

the Stanford Achievement Test and the SRA Achievement Series.

The TONI manual does not give the details of these studies, but does report the results. The Progressive Matrices and Leiter are the only adult tests that were compared with the TONI. The studies revealed high correlations on the Progressive Matrices and Leiter with a deaf population. As of this writing, no correlational studies of the TONI using a normal adult population could be found. The few studies available on the TONI also commented on the lack of research examining concurrent validity.

In a study by Haddad (1986), TONI scores were correlated with scores from the WISC-R and Wide Range Achievement Test (WRAT) by learning disabled children. This study yielded relatively low correlation coefficients between the TONI and standard scores of the WRAT. The correlation between the TONI and the WISC-R was found to be nonsignificant for the Full scale IQ and Verbal IQ, however a significant difference was found between the mean score of the TONI and Performance IQ of the WISC-R. Haddad concluded from these results that the TONI may be measuring an aspect of nonverbal intelligence that is different from

that measured by the Performance Scale of the WISC-R, except for visual organizataion skills. Haddad also suggested that the differences found between the TONI and the WISC-R could indicate that the 12-year-old norms of the WISC-R may not be relevent for today's youth.

Another study conducted by Bond (1982) compared the TONI with the WISC-R and the Slosson Intelligence Test (SIT). Contrary to the previous study cited, the results of this study found high correlations between the TONI and the Performance scores of the WISC-R and moderate correlations with the Verbal scores. It also revealed high correlations between the TONI and the SIT. Bond concluded from this study that the TONI appears more strongly related to a language free or composite measure of intelligence than to verbal ability.

While these two particular studies tend to focus on assessment of children using the TONI, it is interesting to see the discrepancies in the results obtained. Not only has a review of the literature revealed a deficiency in research on the TONI using an adult population, but it also supports the need for expanded studies of concurrent validity on the TONI.

CHAPTER 3

Summary and Conclusions

A review of the literature reveals a deficiency in current research available on nonverbal intelligence testing for adults. This applies not only to the TONI, but to other tests as well. Many popular measures of nonverbal intelligence such as the Leiter International Performance Scale and Progressive Matrices may become undesirable due to outdated norms.

The TONI has many favorable qualities which support its usefulness in a multi-cultural population. The author's efforts at producing a culturally reduced test are to be commended, since the TONI appears to meet all the necessary criteria cited previously by Jensen.

One of the major obstacles in providing intellectual assessment for a miliary hospital setting has been the language differences among the individuals referred. While a nonverbal intelligence test such as the TONI does not provide assessment of global intellectual ability, it would be helpful in providing feedback on one aspect of intelligence, which would be problem solving ability. This information would provide a measure for comparison with individuals of similar or different languages and cultures.

The small amount of research available on the TONI focuses primarily on children. The majority of these studies revealed significant correlations between the TONI and other traditional measures of intelligence. It is desirable to determine how highly the TONI correlates with other valid measures of nonverbal intelligence using a sample of adults.

In conclusion, the TONI appears to have a great deal of potential for use as an appropriate instrument of intellectual assessment in a multi-cultural population. Further research needs to be conducted in order to support its usefulness with an adult population. Correlational studies with adult intelligence tests such as the Wechsler Adult Intelligence Scale-Revised (WAIS-R) (Wechsler, 1981) would be helpful. It is the opinion of this researcher that a high correlation between the TONI and the Performance scale of the WAIS-R would be found if such a study were conducted. Research of this nature would provide a basis for comparison between the TONI and some of the more traditional, well researched tests of intellectual ability.

REFERENCES

Bond, C.L. & Kennon, R.W. (1982). TONI: A Language Free Substitute for the SIT or WISC-R? (Report No. CS-008-270). Department of Curriculum and Instruction, Memphis State University. (ERIC Document Reproduction Service No. ED 264-549).

Brown, L., Sherbenou, R.J., Dollar, S.J. (1982). Manual of the Test of Nonverbal Intelligence, Austin, Texas: PRO-ED.

Cronbach, L.J. (1984). Essentials of psychological testing. New York: Harper and Row.

Frijada, N. & Johoda, G. (1966). On the scope and methods of cross-cultural research. International Journal of Psychology, 1, 109-127.

Haddad, F.A. (1986). Concurrent validity of the Test of Nonverbal Intelligence with learning disabled children. Psychology in the Schools, 23, 361-364.

Hopkins, K.D. & Bracht, G.H. (1975). Ten year stability of verbal and nonverbal IQ scores. American Educational Research Journal, 12, (4), 469-477.

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

Kaufman, A.S. (1979). Intelligent testing with the WISC-R. New York: Wiley.

Koch, W.R. (1984). Culture Fair Intelligence Test. Test Critiques, 1, 233-238.

Llabre, M. (1984). Standard Progressive Matrices. Test Critiques, 1, 595-602.

Lonner, W.J. (1985). Issues in testing and assessment in cross-cultural counseling. The Counseling Psychologist, 13, (4), 599-614.

Matey, C. (1984). Leiter International Performance Scale. Test Critiques, 1, 411-419.

Matarazzo, J.D. (1972). Wechsler's measurement and appraisal of adult intelligence (5th ed.). Baltimore: Williams and Wilkins.

Sattler, J.M. (1988). Assessment of children (3rd ed.). San Diego: Author.

Springer, S.P. & Deutsch, G. (1981). Left brain, right brain. New York: W.H. Freeman and Company.

Sundberg, N.D. & Gonzales, L.R. (1981). Cross-cultural and cross-ethnic assessment: overview and issues. Advances in Psychological Assessment, 5, San Francisco: Josey-Bass.

Wechsler, D. (1981). Manual for the Wechsler Adult Intelligence Scale-Revised. New York: The Psychological Corporation.