12-2001

A Comparison of Intellectual Test Performance Across Native American and European American Children

Austin P. Keith

A COMPARISON OF INTELLECTUAL TEST PERFORMANCE ACROSS NATIVE AMERICAN AND EUROPEAN AMERICAN CHILDREN

by
Austin P. Keith
Black Hills State University, 2000

A Thesis

Submitted to the Graduate Faculty

of the

University of North Dakota

in partial fulfillment of the requirements

for the degree of

Master of Arts

Grand Forks, North Dakota
December
2001

This thesis, submitted by Austin P. Keith in partial fulfillment of the requirement for the degree of Master of Arts from the University of North Dakota, has been read by the Faculty Advisory Committee under whom the work has been done and is hereby approved.

(Chairperson)

This thesis meets the standards for appearance, conforms to the style and format requirements of the Graduate School of the University of North Dakota, and is hereby approved.

Dean of the Graduate School

November 29, 2001

Date

PERMISSION

Title A Comparison of Intellectual Test Performance Across Native American and European American Children

Department Psychology

Degree Master of Arts

In presenting this thesis in partial fulfillment of the requirements for a graduate degree from the University of North Dakota, I agree that the library of this University shall make it freely available for inspection. I further agree that permission for extensive copying for scholarly purposes may be granted by the professor who supervised my thesis work or, in his absence, by the chairperson of the department or the dean of the Graduate School. It is understood that any copying or publication or other use of this thesis or part thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of North Dakota in any scholarly use which may be made of any material in my thesis.

Signature Austin Keith

Date 11-29-01

TABLE OF CONTENTS

LIST OF TABLES

ACKNOWLEDGMENTS

I express gratitude to the members of my Thesis committee; Dr. Doug McDonald, Dr. Alan King and Dr. John Tyler for their support and commitment to completing this project. Also, thanks to Dr. Thomas Petros and Dr. Andrea Zevenbergen who, though not on my committee, nevertheless gave generously of their time. Most of all, thanks to my wife Colette and my daughters Meredith and Lucy who continue to provide the motivation for me to succeed.

ABSTRACT

Differential scoring across Native American and European American groups on the Wechsler Intelligence Scales for Children - Third Edition (WISC-III) was examined while controlling for Full Scale IQ and age. Comparisons were made across the groups on Verbal IQ and Performance IQ, individual subtests, individual subtest items and two composite index scores. Results indicated no significant scoring differences between the two groups on Performance IQ or, on the Perceptual Organization Index. Significant scoring differences were also not evident on the Coding, Picture Arrangement, Arithmetic, Block Design, Vocabulary, Object Assembly and Comprehension subtests. Significant scoring differences were found across the groups on Verbal IQ and the Verbal Comprehension Index. Significant differences were also found between the groups on the Picture Completion, Information and Similarities subtests as well as on three verbal subtest items and several performance subtest items. Previous research finding a pattern of higher performance relative to verbal scores for Native American children was supported; while other research suggesting cultural bias against Native American children on individual subtest items was not.

INTRODUCTION

Intelligence assessment is a widespread practice in the United States. In measuring intellectual abilities, scientists and practitioners rely heavily upon the results of standardized psychometric instruments. Most widely used intelligence tests were developed and standardized primarily on white, middle-class, adults and children (Dana, 1993; Sue & Sue, 1990). These instruments continue to play an important role in assessing not only white, middle-class, adults and children but also individuals from diverse ethnic and cultural backgrounds. Such widespread application has resulted in a decades' long dispute between practitioners and scientists regarding the tests' potential bias against minorities.

Although a considerable amount of attention has been directed toward examining the nature and extent of bias in standardized intelligence tests when assessing Native American children, the attention has prompted only a little research (e.g. Devers, Bradley-Johnson, and Johnson, 1994; Mishra, 1982). Claims of cultural bias on psychometric instruments when assessing minority children are based upon the fact that the experiential background of minority children differs from that of middle-class European Americans upon which the tests are typically normed (Mishra, 1982). It is useful to begin a discussion of the research on test bias by defining some prominent differential characteristics between Native American and European American cultures.

European American Worldview

The values, behavior and beliefs of individuals within a given culture are typically incorporated under the rubric "worldview" (Katz, 1985). A core construct of the worldview of the dominant culture in the United States according to Waterman, (1981) is individualism. Besides an expressive component of individualism, there is a utilitarian component in which "human life becomes an effort to maximize self-interest in the form of power" (Waterman, cited in Dana, 1993, p.14).

Other characteristics generally accepted as descriptive of Western, European American culture include acquisitiveness, inquisitiveness, and competitiveness. Such characteristics, founded in the Western, Lockean, scientific tradition (Becvar & Becvar, 2000) presuppose behaviors and beliefs that support the preeminence of objective versus subjective interpretations of reality and a value-free acquisition of knowledge. As Dana (1993) points out, "Scientific method dictates structure of thinking that is quantitative, dualistic, objective, rational, and linear, with regard to cause-and-effect relationships" (p.14). Thus, the dominant culture, in its continued effort to understand the nature of reality, eschews the subjective and champions the objective with regard to the interpretation of phenomena. An essential premise in this epistemology is an assumed duality concerning the mind/body conundrum (Dana, 1993). Evidence of this dualism can be found in the practice of Western medicine where successful treatments for injuries are typically considered cures (Locust, 1988) and physical health merely the absence of disease (Dana, 1993). Sampson (1985) states further that the American worldview is egocentric with individuals believing that personal control gained through money, power,

and recognition will lead to an orderly world and a satisfactory self-definition. These assumptions, underlying the worldview of the dominant culture, may differ and in some respects, be contrary to assumptions underlying the worldview of some Native Americans (Coleman & Barker, 1991).

Native American Worldview

Definition of Native American

McDonald, Morton, and Stewart (1993) define a Native American as an individual who, through blood quantum, descendency or, tribal ceremonial adoption belongs to any federal, state, or locally-recognized tribe and attempts to live within the customs of that tribe. The degree to which an individual identifies with and participates in a given culture and its attendant values and customs is termed his or her level of acculturation (Oetting & Beauvais, 1990). Individuals may identify with one or more cultures simultaneously and be acculturated to varying degrees in each (Oetting & Beauvais, 1990). Thus, great variation exists in the degree to which Native Americans believe in and practice their traditional tribal values, language and customs (Tyler, Cohen, & Clark, 1982). This variation notwithstanding, definitive contrasts do exist between the worldview of the Native American and that of the dominant culture.

Mind/ body dualism

Many indigenous cultures do not hold strongly a view of mind/body dualism (Chapleski, Lamphere, Kaszynski, Lichtenberg, & Dwyer, 1997). For most Native Americans, there is not only an integration of mind, body and spirit, but spirit is considered the very essence of one's being (Dana, 1993). Thus, religion or spirituality and

health are intertwined (Coggins, 1990) and healing may not be easily separated from culture or religion (Locust, 1988). For Native Americans, an illness or injury might be viewed from a spiritual rather than a physical perspective and treatment might involve addressing the disharmony in one's mind, body and spirit perceived to have caused the pathology (Locust, 1988).

Interrelatedness

The phrases "Akwe:kon" of the Mohawk and "Mitakuye Oyasin," of the Lakota are translated as "All of us" and "All my relatives," respectively (Simonelli, 1994). For the Lakota, the words reflect their deepest sentiments regarding the interdependence and interconnectedness of all things. This belief in the value of interrelationships is also reflected in family compositions that include many extended family members with whom relationships are considered as close as members of their nuclear families. Relationships are further established through ceremonies such as the Lakota "hunka" or "making of relatives" ceremony where individuals are adopted into a family and thereafter considered as close as those who are biologically related.

Few individuals socialized into such a cultural worldview could comprehend isolation, in the existential sense, which holds that an unbridgeable abyss exists between the person and the rest of the world. As French (1981) asserts, the identities and self-concepts of many traditional Native Americans are bound to their tribal groups and to nature as a whole. Individualism, an integral part of the dominant culture's worldview, may be a difficult notion to comprehend for many Native Americans whose identity may best be described as ensembled, extended and sociocentric (Sampson, 1985).

The circle

The assumption of linear relationship and causality accepted by Western science may be foreign to some Native Americans whose culture embraces the concept of circularity in all things. A Lakota holy man named Black Elk explained the concept in this way "You have noticed that everything an Indian does is in a circle, and that is because the Power of the World always works in circles, and everything always works in circles, and everything tries to be round" (Neihardt, 1932, p.198). The circle is manifest in the construction of traditional Native American dwellings and ceremonial lodges. The circular nature of existence is further evidenced for indigenous cultures in the occurrence of natural phenomena such as the recurring seasons and in the movement and shape of celestial objects.

For many Native Americans, the notion of circularity also implies an underlying sense of community and universal reciprocity. Two traditional tribal ceremonies, the potlatch of the Kwakiutl (Ballantine & Ballantine, 1993) and the wopila of the Lakota, reflect these ideas. Both ceremonies acknowledge milestones in the lives of individuals and validate their accomplishments (Ballantine & Ballantine, 1993) while simultaneously strengthening relationships among all tribal members. Community ties are strengthened through feasting and the giving of gifts by the family of the individual being celebrated . Circularity, as represented in the examples given above reflect assumptions that are central to a basic understanding about the nature of reality for many Native Americans.

Epistemology

Traditionally, the education of Native Americans included sacred learning that was inseparable from secular knowledge or formal education (Beck, 1996). Patience, listening and not asking why were emphasized by tribal elders who believed that adherence to these practices would ensure that knowledge would remain aligned with experience and wisdom with divinity (Beck, 1996).

Larry Merculieff, an Aleut community leader, gives and example from his own personal experience of learning from a tribal elder and in so doing, illustrates a traditional method of passing knowledge from one generation to the next:

> My generation was the last that had a fully intact traditional upbringing, with the extended family and the entire community involved. For example, in order for me to get to know my grandfather, I was with him every hour of the day for two years. He would go to work at 5 o'clock in the morning and I'd go with him. I'd go to church, I'd go hunting. One day my grandfather and I were walking on the beach. I commented that the sun was just beautiful, and he said to me in Aleut, 'Be quiet. When you talk you lose the essence of the sun.' I think that single sentence captures how our people have survived and thrived in the Bering Sea for almost 9,000 years. I also had an acha, which is a spiritual relationship between a younger person and an older person. He taught me most of what I know about the environment and Aleut ways, but in the 13 years that he took me under his wing he might have said 200 words to me. (Shute, 1996, p. 23).

A Keres man named Larry Bird explains that "you don't ask questions when you grow up. You watch and listen and wait, and the answer will come to you . . . not like learning in school" (Tedlock, 1995, p. xxxi).

Beck (1996) further illustrates how the behaviors of Native Americans may in fact be antithetical to those of the dominant culture. A woman from Taos Pueblo had this to say about the proper way of learning: "All through this time I never asked of them (grandmother and grandfather) or anyone, 'why?' It would have meant that I was learning nothing -- that I was stupid. And in Western Society if you don't ask why they think you are stupid" (p. 49).

Carole Anne Heart, a Lakota and President of the National Native American Education Association, shares an early experience of attending boarding school on the Yankton Sioux reservation in South Dakota. Her story may help illustrate the potential conflict between a "traditional" upbringing and the expectations and practices of the dominant culture's educational system:

> I went to a Catholic school on the Yankton reservation until eighth grade. I got straight A's. In class, when nobody knew the answer to a question, the teacher would say, 'Carole Anne, you tell them the answer.' I hated that, because being held up to be smarter or better than others is a no-no in Indian tradition, so I hardly had any friends (Katz, 1995, p. 287-8).

These examples illustrate how the epistemology of Native Americans emphasizes contemplation over verbalization and relationship over competitiveness. Differences may exist, therefore, between European American and Native American views on culturally

appropriate ways of acquiring and expressing knowledge. These differences may have the potential to introduce bias or otherwise complicate psychoeducational assessment in a variety of ways. For example, Native Americans, in their desire to not call attention to themselves and not be viewed as superior, may provide insufficient responses to questions and exhibit hesitancy in asking critical questions. More research is needed, therefore, on intelligence test performance of Native Americans to help uncover some of these potentially influential factors.

The need for this research is made even clearer when one considers that projected socioeconomic and legal trends may lead to increasing numbers of psychologists working with Native American populations (Hynd & Garcia, 1979). Educators and practitioners need accurate assessment data with which to decide such things as appropriate grade assignments and identifying students in need of remedial instruction.

Literature Review

A large body of research has examined various factors of intellectual testing that could lead one to conclude that a particular test is biased toward a particular cultural group. In seeking to understand further test bias in Native American populations a number of studies have investigated the nature of how bias is manifest by using various measures and approaches.

Predictive Validity

One method of investigating the existence of bias in a test of intellectual ability is to determine if scores on the instrument predict a criterion such as achievement test scores or grades in school equally well for various samples. For example, in conducting

cross-cultural comparisons of predictive validity of an intelligence test one would examine the slopes of the regression lines used to predict the criterion scores for two groups differentiated according to ethnicity or culture. Similarity between the slopes of the regression lines for each group would indicate that the degree of predictive validity of the instrument is similar across these two groups. To the degree that the regression lines are different one could conclude that the predictive validity of the instrument is different for each group.

In this vein, Weiss and Prifitera (1995) examined the predictive validity of the WISC-III in predicting Wechsler Individual Achievement Test (WIAT; The Psychological Corporation, 1992) standard scores for the Reading, Writing, Mathematics and Language domains. The authors compared the slopes and intercepts of the regression lines using IQ to predict achievement scores in European Americans, African Americans, and Hispanics. The results showed a minimal degree of differential prediction and the authors concluded that there was no bias in using WISC-III scores to predict WIAT achievement test scores.

Hale, Raymond, and Gajar (1982) examined regression equations relating Wechsler Intelligence Scale Children-Revised (WISC-R; Wechsler, 1974) Verbal IQ (VIQ) scores to scores on the Wide Range Achievement Test-Revised (WRAT-R; Jastak & Wilkinson, 1984) Reading subtest. Subjects were 144 children aged 7-10 years from a southwestern Virginia city school system. The children were randomly selected from a larger population of children identified for or placed in special education classes. Subjects were categorized into three classes of mental disability: emotionally disturbed (28%),

learning disabled (45%), and educable mentally retarded (27%). The Hollingshead Index of Social Position (Hollingshead, 1957) was used to classify subjects further into either middle or low socioeconomic status (SES) groups. Ethnic representation within the dichotomous SES classifications included Anglo (74%, 53%), Black (22%, 42%) and other (4%, 5%) in the middle and low SES categories, respectively. Additionally, students were selected on the availability of their WISC-R and WRAT-R scores. The WISC-R VIQ was used as the predictor variable in the regression analysis while the WRAT-R Reading subtest scores were used as the criterion. No significant difference was found between the regression lines of the low and middle SES groups, thus supporting the validity of the WISC-R in predicting reading achievement in children regardless of their socioeconomic status. Limitations of this study include a lack of subjects from the highest socioeconomic levels and the use of a single measure of academic achievement as a criterion measure. A further limitation is the failure to evaluate the Performance IQ and Full Scale IQ scores for bias.

Mishra and Lord (1982) endeavored to study if assessment instruments may differentially predict a common criterion across ethnicity by examining the reliability and predictive validity of the WISC-R in a sample of Navajo children. Subjects in the study were 40 Navajo students from the fourth and fifth grades attending a Navajo reservation elementary school. Students were randomly selected from the entire fourth and fifth grade population stratified on ethnicity, grade and sex. All students came from low socioeconomic backgrounds. All subjects were individually administered the ten regular subtests of the WISC-R and the Reading section of the Wide Range Achievement Test

(WRAT; Jastak, Bijou, & Jastak, 1976) while the Arithmetic and Spelling sections of the WRAT were administered as a group test. Predictive validity coefficients were derived using the WRAT scores as a criterion measure. WRAT scores were correlated with the Verbal, Performance and Full Scale IQs as well as the individual subtest scores of the WISC-R. Reliability coefficients for individual WISC-R subtests, except Coding, were derived using the split-half technique. Of the three WISC-R scales, the highest reliability coefficient for this sample was found for the Performance Scale (.86). The results also showed little correlation between WISC-R subtests and IQ scale scores and the WRAT scores. This study suggests the need for further investigation into the utility of the WISC-R as a measure of achievement potential in ethnic populations.

Item Bias

A second way of determining the existence of test bias is to examine individual test items. Mishra (1982) asserts that good predictive validity for a test does not preclude the existence of item bias. In fact Mishra (1982) asserts that the most serious allegations of test bias are directed at the level of individual test items. A study of item bias on tests might reveal the relative appropriateness of using specific test items for measuring a construct, such as intelligence, in culturally different samples. Mishra (1982) examined whether or not item bias on the WISC-R existed by studying individual test items, more specifically, the verbal items of the test. His study examined the WISC-R protocols of 40 European American and 40 Navajo fourth and fifth grade students attending elementary school in Tucson, Arizona, and on the Navajo reservation. All subjects were from low socioeconomic backgrounds and were randomly selected from school attendance rosters.

Each subject was individually administered all ten regular subtests of the WISC-R. Two groups differentiated by ethnicity were further subdivided by Full Scale IQ (FSIQ) scores into high and low ability categories. Items from the Information, Similarities and Vocabulary subtests were examined for ethnic bias while controlling for the effects of ability. Results indicated that on 5 of the 30 items on the Information subtest, significantly fewer Native American students passed the item than European Americans. Also, a similar pattern was found on 4 of the 17 items on the Similarities subtest and 6 of the 30 items on the Vocabulary subtest. The author concluded that these specific items were biased against Navajo subjects. Mishra suggested that the results be generalized with caution because only children of lower SES were tested. He also suggested that the results are limited by the fact that the Native American students had little opportunity for interracial exchange.

Verbal/Performance differences

A third method used by researchers to look for potential bias on IQ tests between Native Americans and European Americans is to examine differences between the Verbal and Performance scores of these two groups.

Teeter, Moore, and Petersen (1982) examined differences in WISC-R scores in three groups of Native American children identified by their level of learning deficits. These groups were identified as non-handicapped (NH), educationally disadvantaged (ED) (i.e., children with low academic achievement but without signs of other handicapping conditions), and learning disabled (LD). Subjects were 452 teacher-referred Navajo students from 11 different school systems on the Navajo Indian reservation who

were experiencing school-related problems and who had met differential criteria placing them in one of the three previously identified diagnostic categories. Students ranged in age from 6 to 16, had differing levels of interracial contact, and identified Navajo as their primary language. The results indicated that for all three groups, Verbal IQ was below the average range while Performance IQ was within the average range. The authors suggest that the results support the idea that Performance IQ is the most unbiased estimate of intellectual potential of Native American children especially when English is not their primary language.

McCullough, Walker, and Diessner (1985) examined whether or not consistent discrepancies would be found between the Verbal and Performance scale scores of the WISC-R and WAIS (Wechsler, 1955) in a sample of 75 Native American students. All students spoke English as their primary language and attended a tribally-operated reservation school in the Columbia River Basin area of the Pacific Northwest. The majority (88%) of the students were from the Columbia River Basin area while (12%) were from Great Plains tribes. The WISC-R was administered to 42 of the students (i.e., aged 12 to 16 years) while the remainder, aged 16-19 years, were administered the WAIS. Gender distribution was equally divided among the subjects taking the WISC-R while 27% of subjects taking the WAIS were female. Three of the students were considered learning disabled but were mainstreamed the majority of the day. The authors found Verbal scale scores on the WAIS and the WISC-R to be significantly below the normative mean while Performance scale scores were at or above the normative mean. The authors state that Verbal-Performance differences have been found across several Native

American tribes and that predictive validity of the Wechsler tests may vary across tribes as well. They advise caution, therefore, when assessing Native Americans with the WISC-R and WAIS.

Subtest scoring patterns

A fourth way in which studies have examined IQ differences between Native Americans and European Americans has been to compare these groups on subtest scoring patterns. This is done by recategorizing subtest scores into sets of factored abilities and then investigating the existence of differential scoring patterns on these categories between minority and majority culture children.

One such recategorization scheme for subtest comparisons is the Bannatyne scales (McShane & Plas, 1982). When utilizing these scales one can differentiate the WISC-R subtests according to Spatial, Conceptual, Sequential and Acquired Knowledge categories. With the Bannatyne recategorization scheme a Spatial score is derived by computing a mean scaled score from the Block Design, Picture Completion and Object Assembly subtest scores, a Conceptual score from the Comprehension, Similarities and Vocabulary subtest scores, a Sequential score from the Digit Span, Coding and Arithmetic subtest scores and an Acquired Knowledge score from the Information, Arithmetic and Vocabulary subtest scores (Zarske & Moore, 1982). Studies with children with learning disabilities have suggested that an identifiable scoring pattern is evident when WISC-R subtest scores are grouped into Bannatyne's categories. The existence of a Spatial > Conceptual > Sequential pattern in learning disabled children has been identified in a number of studies using heterogeneous samples (Zarske & Moore, 1982).

McShane and Plas (1982) using the Bannatyne categories, sought to identify Wechsler Scale performance patterns within a sample of Native American students. The authors hypothesized that Spatial scores would be significantly higher than Sequential scores and that the Sequential scores would in turn, be significantly higher than Conceptual and Acquired Knowledge scores for less acculturated Native American children. Subjects were 142 Native American children, evenly distributed on sex and ranging from 4.5 to 16 years of age. More than 2/3 of the sample was Ojibwa while the remainder was primarily Sioux with 8% being from other tribal groups. The sample included children referred for educational difficulties (n = 105), learning problems possibly related to otitis media (n = 20), and giftedness screening (n = 17). Data from 78 previously administered WISC, 32 WISC-R and, 12 Wechsler Preschool and Primary Scale of Inteiligence (WPPSI; Wechsler, 1967) protocols were used in their analysis. The WPPSI is a series of eleven tests for children age four to six and one half years. The WPPSI provides Verbal, Performance and Full Scale IQs. Comparisons were made between a traditional group and an acculturated group by assigning the WISC and WISC-R protocols to either group based upon differences between Verbal and Performance IQ. Protocols were assigned to the traditional group if the Verbal IQ was lower by more than 9 points than the Performance IQ and to the acculturated group if this difference was less than 9 points. This difference was deemed by the authors to have descriptive utility since random samples of Native American children consistently showed mean differences of 10 points or greater (e.g., McShane, 1980) while Anglo acculturated children showed differences of nine points or less. The authors validated this methodology by comparing

group assignment with parental ethnicity and level of parental and child reservation contact. It should be noted, however, that this method of differentiation is atypical and a potential weakness of this study. A more common method is to group subjects prior to examining scoring differences. The results of the Bannatyne recategorized WISC and WISC-R scales revealed significance in the expected pattern of Spatial > Sequential > Conceptual and Acquired Knowledge skills for the total group and traditional group but not for the acculturated group. Analysis of the twelve WPPSI protocols revealed Spatial abilities to be highest with no significant differences between scores of the remaining three factors. McShane and Plas (1982) concluded that the findings validated the claims of a typical ordered pattern of performance on the Wechsler Scales for less acculturated Native American children.

Zarske and Moore (1982) examined whether a group of Navajo children with learning disabilities would demonstrate the same pattern on the Bannatyne Scales (i.e., Spatial > Conceptual > Sequential > Acquired Knowledge) as that found with other learning disabled samples (e.g., Bannatyne, 1974; Rugel, 1974; Smith, Coleman, Dokecki & Davis, 1977). Their subjects were 192 Navajo children with learning disabilities aged 6 years through 16 years, 11 months speaking English as a second language [illegible] rural reservation schools. Results revealed that, although significant differences were found between scores derived using the Bannatyne scales, the pattern exhibited in previous studies using heterogeneous samples was not present. Instead, the scores for this sample were characterized by a Spatial > Sequential > Conceptual pattern, replicating that found in Gutkin's (1979) sample of Mexican-American children. The authors concluded

that absence of the Bannatyne pattern does not contraindicate learning disabilities in Navajo Indian children. These findings also lend credibility to suggestions by Smith, Coleman, Dokecki, and Davis (1977) that the Bannatyne pattern is usually not found in groups of learning disabled children with Full Scale IQs of 75 or less (i.e. the sample used in this study had a mean Full Scale IQ of 72.68). The authors cautioned, however, that the low IQ scores of this sample may be indicative of language differences rather than low intellectual ability. The authors further posited that, when considered with Gutkin's findings, the results of this study indicate a unique pattern for ethnic minority children where Conceptual scores are usually very low compared to Spatial scores that are, in turn, typically higher than Sequencing scores. The authors recommended further studies using matched groups of Navajo children with and without learning disabilities to help determine if such findings are a result of learning disabilities or ethnic differences.

Factor Structure

Beyond the Verbal, Performance and Full Scale IQ scores, four factor-based index scores can be calculated from the subtests of the WISC-III. These are the Verbal Comprehension Index (VCI), a Perceptual Organization Index (POI), a Freedom from Distractability Index (FDI) and a Processing Speed Index (PSI) (Wechsler, 1991). Subtests making up the four indices are as follows: VCI: Information, Similarities, Vocabulary and Comprehension; POI: Picture Completion, Picture Arrangement, Block Design and Object Assembly; FDI: Arithmetic and Digit Span; PSI: Coding and Symbol Search.

A fifth way to examine differential scoring among cultural groups is to determine if a common factor structure exists across groups. For example, an intelligence test such as the WISC-III could be administered to a sample of European American and a sample of Native American children and a factor analysis conducted separately for each cultural group.

If cultural affiliation is not an influence on test performance then the WISC-III factor coefficients, error variances, factor variances and covariances would be similar for both groups. If the factor solutions were not the same for each group then potential bias would be indicated. No study on differential factor structure between Native American and European Americans was found in the literature to address this.

Problems with existing studies

Extant literature (e.g., McCullough, Walker & Diessner, 1985; Hale, Raymond & Gajar, 1982; Teeter, Moore & Petersen, 1982) shows that Native Americans score higher on the Performance than Verbal component of the Wechsler scales. When comparisons are made between the IQ test scores of European Americans and Native Americans, the differences are always largest on the Verbal scales.

Concluding that this scoring differential is characteristic of Native American cultural groups raises several concerns. For example, many studies that found differential scoring between these two groups failed to account for the fact that for many Native American children, English may have been their second language. Similar studies comparing scores between European immigrants and European Americans show similar results again suggesting limited English proficiency as a potential factor. An additional

problem with the existing studies in the literature mentioned above is that many were conducted with clinic-referred Native American children. An extensive body of literature suggests that large Verbal-Performance IQ differences, of 15 points or more, are associated with clinical problems (Bloom, Topinka, Goulet, Reese, & Podruch, 1986). This raises the question of potential problems in interpretation when examining the results of these studies. For example, are the patterns of scores on these IQ tests a result of subjects being Native American or are they the result of the subjects having behavioral or psychological problems that led to the referral?

Beiser and Gotowiec (2000) raise the issue of potential confounds to the conclusions reached in studies of intelligence test bias against minorities. Among others, these potential confounds include: prenatal health, hearing problems, parental education, and socioeconomic status. They investigated the relationship between culture, language and environment, and IQ scores in a sample of Native American children relative to a sample of non-Native children. The 691 Native subjects were from tribes indigenous to four geographically circumscribed, culture areas.

The 234 Non-Native subjects composing the comparison sample either attended the same schools as the Native students or were from an area possessing similar ecological characteristics as that of the Native sample. All subjects attended English-language schools on or immediately adjacent to the respective American reservation or Canadian reserve. IQ scores were derived from the WISC-R administered to all subjects. Data on biological and psychosocial factors were obtained through interviews with parents and teachers. Teachers completed a 3-item inventory rating the children's English

language skills while parents were administered an 86-item standardized Biodemographic Interview (Beiser, 1989). The psychometric interview schedule probed the following eight domains of the subjects' macro and micro environments: 1) prenatal health care as assessed by the frequency of medical contact during pregnancy, 2) mother's prenatal health measured as the degree of hypertension and excessive weight gain during pregnancy, 3) hearing problems, 4) interviewee education, 5) household amenities, 6) satisfaction with school, 7) cultural separation attitudes, and 8) life events. Comparisons between Native and non-Native IQ scores validated findings from previous studies. Native students in this study scored lower overall, and scored considerably lower on the Verbal than Performance subscales than the comparison sample. Fewer household amenities and more negative [illegible] life events were recorded for Native subjects. Non-Native parents endorsed more frequently, items indicating hearing problems in their children potentially related to otitis media and they were less likely than Native parents to endorse items that indicated a preference for cultural separation. After controlling for differences in English language fluency, IQ Performance score differences between the two samples were no longer significant. Proficiency in the use of the English language thus provided perhaps the most parsimonious explanation for intersite differences as well as differences between Natives and non-Natives in Performance scores.

Present Study

The general hypothesis of the present study was that Native American children would exhibit a pattern of performance on the WISC-III different from that of European American children. More specifically, I hypothesized that, I would find a significant

pattern in the Native American group whereby Performance IQ scores would be higher than Verbal IQ scores while no significant difference would be found between Performance and Verbal IQ scores for the European American sample.

With regard to the index scores I hypothesized that Native American children would have a higher Perceptual Organization Index Score (POI) than a Verbal Comprehension index Score (VCI) and that no significant difference would be found between these two scores for the European American children. Additionally, I hypothesized that the European American children would have a significantly higher mean VCI score than the Native American children.

Finally, although I do not have a specific hypothesis regarding individual subtest items, an exploratory analysis was conducted on each item to determine if differential response patterns existed across the groups.

METHODOLOGY

Subjects

The data used in this study were derived from Wechsler Intelligence Scale for Children - Third Edition (WISC-III) protocols which were administered previously to Native American and European American children. The assessment instruments were administered to European American children at the University of North Dakota and to Native American children at Four Winds Community School a tribally controlled school located at Fort Totten, North Dakota. Children in the Native American group were members of federally recognized tribes.

Materials

Consent to use data of the European American subjects was obtained via a consent form signed at the time of the test administration. Consent to use data of the American Indian subjects was obtained from the principal, superintendent and school board of Four Winds Community School as well as the Spirit Lake Sioux Tribal Council. Beyond age and race of subjects no other demographic information was required in this study. Thus, consent forms and demographic sheets were dispensed with and the archival WISC-III protocol was the only instrument used for this study.

Wechsler Intelligence Scale for Children - Third Edition

The test examined in this study was the WISC-III, one of the most widely used tests of intelligence among psychologists (Hutton, Dubes, & Muir, 1992; Stinnett, Havey, & Oehler-Stinnett, 1994). It is administered individually and is a measure of intellectual ability in children aged six through 16 years and 11 months. The WISC-III measures both Verbal and Performance capabilities and divides 13 subtests into the Verbal and Performance Scales. The six subtests making up the Verbal Scale are Information, Similarities, Arithmetic, Vocabulary, Comprehension, and Digit Span and the seven making up the Performance Scale are Picture Completion, Coding, Picture Arrangement, Block Design, Object Assembly, Symbol Search, and Mazes. These subtest scores combine to produce a composite Verbal IQ, a composite Performance IQ and a Full Scale IQ. It is useful to divide the thirteen subtests into Verbal and Performance categories and provide a definition of each.

Verbal Subtests:

The Information subtest assesses the child's knowledge of events, objects, places and people and measures the range of factual information the child has absorbed from the environment and has stored in long-term memory.

The Similarities subtest consists of questions that assesses the child's ability to articulate similarities between two objects, concepts or other phenomena that are paired and presented serially to them. The subtest calls upon the child's verbal concept formation and long-term memory.

The Arithmetic subtest measures a child's numerical reasoning ability, concentration, attention, short-term memory and long-term memory as they are applied to solving a series of arithmetic problems orally presented to them.

The Vocabulary subtest assesses word knowledge, verbal skills, language development and long-term memory as a child utilizes them to recall and articulate definitions to words orally presented to them.

The Comprehension subtest assesses social judgement, common sense and knowledge of conventional standards of behavior as articulated by the child in seeking to resolve a series of orally presented problems.

The Digit Span subtest assesses short-term memory, rote memory, attention and concentration as they are utilized by the child to immediately recall a series of numbers presented orally to them.

Performance Subtests:

The Picture Completion subtest assesses the child's ability to discriminate between essential and non-essential details by asking the child to identify an essential detail missing from each of a series of pictures presented to them.

The Coding subtest assesses the child's speed and accuracy in information processing as the child is asked to match a series of symbols with corresponding shapes or numbers.

The Picture Arrangement subtest assesses the child's attention to detail, alertness, planning ability and ability to comprehend and evaluate social situations as they are applied to correctly sequencing a series of pictures presented to them.

The Block Design subtest assesses the child's spatial visualization, non-verbal reasoning ability and visual-motor coordination as the child uses colored cubes to replicate designs presented to them.

The Object Assembly subtest assesses the child's visual-motor coordination, organizational ability and sense of spatial relations as the child assembles a series of puzzles by connecting the various pieces.

The Symbol Search subtest assesses the child's attentional skills, hand-eye coordination, speed and accuracy and short-term memory as the child scans two groups of symbols and must indicate if a target symbol from one group is present in the other group.

The Mazes subtest assesses the child's spatial visualization ability and visual-motor coordination as the child attempts to identify a correct pathway from the center of a maze to the exit.

In addition to the subtest scale scores the WISC-III provides 4 factor scores. One of the factors, Verbal Comprehension (VCI) is a composite score resulting from a combination of the Information, Similarities, Vocabulary, and Comprehension subtests. A second factor, Perceptual Organization (POI) is a composite score resulting from a combination of the Picture Completion, Picture Arrangement, Block Design and Object Assembly subtests. A third factor, Freedom from Distractibility (FDI) is a composite score resulting from a combination of the Arithmetic and Digit Span subtests. The fourth factor, Processing Speed (PSI) is a composite score resulting from a combination of the Coding and Symbol Search subtests. Two of the scores that contribute to the FDI and the PSI are derived from supplemental subtests (Digit Span and Symbol Search) that are not

consistently administered because their scores are not required to compute Verbal, Performance and Full Scale IQ scores. Thus, scores derived from these two subtests are not always available and computation of the FDI and PSI for the two groups was not possible for this study and only the VCI and POI factor scores were used in a comparative analysis.

Procedure

The data used in this study was archival and was retrieved from WISC-III protocols previously administered at UND or Four Winds Community School for other studies or for clinical purposes by licensed psychologists, special education professionals or graduate students under supervision. After securing approval for this study from the Four Winds Community School, the Spirit Lake Tribal Council and the Institutional Review Board (IRB), the archival data was collected at Four Winds Community School and the UND Psychology Department and placed on coded data sheets. The sample for this study consisted of data from 56 protocols equally divided into European American and Native American groups (See table 1). All subjects were between the ages of six and 16 and were matched as closely as possible according to FSIQ, age and gender. The average age of the two groups did not significantly differ, $t(54) = 1.68$, $p = .098$. Participation in this study was anonymous. Only the subjects' WISC-III data with the exclusion of the subjects' names were recorded on a coded data sheet. These data sheets were secured in a locked file cabinet in the principal investigators office to further ensure the anonymity of the subjects.

Table 1. Information Regarding Sample

Characteristic	Native American	European American
Age		
Mean	10.57	9.62
Standard Deviation	(2.44)	(1.72)
Gender		
Female	9	12
Male	19	16
Total	28	28

RESULTS

Verbal IQ, Performance IQ and Full Scale IQ scores

T-Tests were conducted to examine the Native American and European American groups on the Wechsler Intelligence Scale for Children-III Verbal IQ, Performance IQ, and Full Scale IQ scores. No significant difference was found between the two groups on Full Scale IQ scores, $t(54) = .379$, $p = .706$, or Performance IQ scores, $t(54) = 1.838$, $p = .072$ however, the European American group had a significantly higher mean Verbal IQ score than did the Native American group, $t(54) = -2.230$, $p = .030$. Additionally, there was a significant discrepancy within the Native American group between the Verbal and Performance IQ Scales, $t(54) = -5.206$, $p = .000$, while no such difference was identified in the European American group (see Table 2).

Table 2. Means and Standard Deviations (within parentheses) of WISC-III Verbal IQ, Performance IQ, and Full Scale IQ for Native American and European American groups.

WISC-III	Native American n = 28	European American n = 28
Verbal IQ	81.68 (12.07)	88.57 (11.03)
Performance IQ	87.46 (10.56)	91.54 (13.39)
Full Scale IQ	87.93 (9.57)	88.96 (10.84)

Verbal Comprehension Index and Perceptual Organization Index scores

T-Tests were also conducted to compare the Native American and European American Groups on the two index scores of Verbal Comprehension and Perceptual Organization. There was a significant discrepancy between these two index scores for the Native American group with the group having a higher mean Perceptual Organization Index Score, $t(54) = -5.140$, $p = .000$ than a Verbal Comprehension Index Score. No significant difference between these scores was evident within the European American group. Comparing the two index scores across groups revealed that the European American group had a significantly higher mean Verbal Comprehension Index Score, $t(54) = -2.731$, $p = .009$ than did the Native American group. The means and standard deviations (in parentheses) for the two WISC-III Index scores compared in this study are listed in Table 3.

Table 3. Means and Standard Deviations (within parentheses) of Verbal Comprehension Index Scores and Perceptual Organization Index Scores by race for WISC-III.

WISC-III Index	Native American $n = 28$	European American $n = 28$
Verbal Comprehension Index	81.71 (12.70)	90.32 (10.80)
Perceptual Organization Index	97.29 (9.78)	92.54 (13.64)

Table 4. Means and Standard Deviations for Comparison Groups on WISC-III Verbal and Performance Subtest Scaled scores (Picture Completion, Information, Coding, Similarities, Picture Arrangement, Arithmetic, Block Design, Vocabulary, Object Assembly, and Comprehension)

WISC-III	Native American	European American
Picture Completion	10.14 (2.57)	8.36 (2.61)
Information	6.39 (3.10)	8.43 (2.55)
Coding	9.93 (2.89)	8.32 (4.17)
Similarities	6.17 (2.69)	8.50 (2.58)
Picture Arrangement	8.18 (2.81)	7.57 (2.98)
Arithmetic	7.29 (2.52)	7.11 (2.78)
Block Design	9.71 (2.46)	8.79 (3.42)
Vocabulary	6.18 (2.37)	7.29 (2.35)
Object Assembly	9.82 (2.26)	10.00 (3.54)
Comprehension	7.14 (3.18)	8.57 (3.22)

Individual Subtest Scaled Scores

Scoring was pass/fail for every item of the WISC-III with 0 designating failure on an item and 1 designating a success regardless of whether the subject scored greater than 1 on any particular item. T-Tests were conducted on each of the ten WISC-III subtest scaled scores (Picture Completion, Information, Coding, Similarities, Picture Arrangement, Arithmetic, Block Design, Vocabulary, Object Assembly, and Comprehension) for comparison across the Native American and European American groups. Significant differences were found across the groups on three of the WISC-III Subtest Scaled scores. These were Information, $t(54) = -2.67$, $p = .010$, Similarities, $t(54) = -2.53$, $p = .014$ and Picture Completion, $t(54) = 2.57$, $p = .013$ with European Americans having higher mean Subtest Scaled scores on the Information and Similarities subtests. No other significant differences were observed across the two groups on the other Subtest Scaled Scores of the WISC-III (See Table 4).

Item analysis

A series of chi square analyses were conducted on items from nine (Picture Completion, Information, Similarities, Picture Arrangement, Arithmetic, Block Design, Vocabulary, Object Assembly and Comprehension) of ten WISC-III standard subtests for comparisons across the Native American and European American groups. The Coding subtest was the only standard subtest excluded from the item analysis since all items of the Coding subtest are similar in content and could not be differentiated in terms of any meaningful examination of item bias.

In terms of differential scoring on individual items of the Verbal subtests, the European American group scored significantly higher on item 4 of the Similarities subscale while the Native American group scored significantly higher on items 16 and 12 of the Information and Comprehension subtests, respectively (See tables 5 & 6).

Table 5. WISC-III Verbal IQ Subtest Item found Significant for European American group.

WISC-III Subtest	Item	Question
Similarities	4	In what way are PIANO and GUITAR alike?

Table 6. WISC-III Verbal IQ Subtest Items found Significant for Native American group.

WISC-III Subtest	Item	Question
Information	16	In what direction does the sun set?
Comprehension	12	Tell me some advantages of getting the news from a newspaper rather than a television news program.

Examination of individual items from the Performance subtests revealed that, the Native American group scored significantly higher on the WISC-III Picture Completion subscale items 14, 18, 20, 22 and 23; Picture Arrangement subscale items 5 and 9; and Block Design subscale items 9, 11 and 12. All items for these subtests along with their associated chi square values are illustrated in tables 7 through 15.

Table 7. WISC-III Picture Completion Item comparisons by group using chi square analysis.

WISC-III Subtest	Item Number	Native American		European American		X^2	p
		Pass	Fail	Pass	Fail		
Picture Completion	1	28	0	28	0	.00	1.00
	2	28	0	28	0	.00	1.00
	3	28	0	28	0	.00	1.00
	4	26	2	28	0	2.07	.149
	5	28	0	26	2	2.07	.149
	6	27	1	27	1	.00	1.00
	7	28	0	26	2	2.07	.149
	8	27	1	26	2	.35	.552
	9	22	6	18	10	1.40	.236
	10	19	9	22	6	.82	.365
	11	21	7	20	8	.09	.762
	12	21	7	19	9	.35	.554
	13	22	6	23	5	.11	.736
	14	23	5	16	12	4.14	*.041
	15	25	3	20	8	2.83	.092
	16	19	9	16	12	.69	.407
	17	20	8	15	13	1.90	.167
	18	18	10	8	20	7.18	*.007
	19	18	10	11	17	3.50	.061
	20	19	9	10	18	5.79	*.016
	21	11	17	6	22	2.11	.146
	22	12	16	3	25	7.38	*.006
	23	15	13	2	26	14.27	*.000
	24	7	21	2	26	3.31	.068
	25	5	23	2	26	1.47	.225
	26	7	21	2	26	3.31	.068
	27	4	24	2	26	.75	.387
	28	1	27	0	28	1.02	.313
	29	3	25	0	28	3.17	.075
	30	0	28	0	28	.00	1.00

Note. * Denotes significance at the $p < .05$ level.

Table 8. WISC-III Information Item Comparisons by Group using chi square analysis.

WISC-III Subtest	Item Number	Native American		European American		X^2	p
		Pass	Fail	Pass	Fail		
Information	1	28	0	28	0	.00	1.00
	2	28	0	28	0	.00	1.00
	3	27	1	28	0	1.02	.313
	4	26	2	27	1	.35	.552
	5	25	3	26	2	.22	.639
	6	23	5	26	2	1.47	.225
	7	25	3	25	3	.00	1.00
	8	22	6	21	7	.10	.751
	9	17	11	19	9	.31	.577
	10	17	11	18	10	.08	.782
	11	20	8	17	11	.72	.397
	12	8	20	14	14	2.70	.100
	13	6	22	9	19	.82	.365
	14	9	19	10	18	.08	.777
	15	3	25	2	26	.22	.639
	16	5	23	0	28	5.49	*.019
	17	4	24	6	22	.49	.485
	18	5	23	3	25	.58	.445
	19	2	26	1	27	.35	.552
	20	3	25	2	26	.22	.639
	21	1	27	2	26	.35	.552
	22	2	26	1	28	.35	.552
	23	3	25	3	25	.00	1.00
	24	3	25	1	27	1.08	.299
	25	1	27	1	27	.00	1.00
	26	1	27	0	28	1.02	.313
	27	2	26	0	28	2.07	.149
	28	0	28	0	28	.00	1.00
	29	1	28	0	28	.00	.313
	30	1	27	0	28	1.02	.313

Note. * Denotes significance at the $p < .05$ level.

Table 9. WISC-III Similarities Item Comparisons by Group Using Chi Square Analysis.

WISC-III Subtest	Item Number	Native American		European American		X^2	p
		Pass	Fail	Pass	Fail		
Similarities	1	20	8	24	4	1.70	.192
	2	27	1	25	3	1.08	.299
	3	25	3	24	4	.16	.686
	4	23	5	28	0	5.49	*.019
	5	25	3	28	0	3.17	.075
	6	28	0	27	1	1.02	.313
	7	25	3	22	6	1.19	.275
	8	20	8	22	6	.38	.537
	9	21	7	21	7	.00	1.00
	10	9	19	13	15	1.20	.273
	11	14	14	11	17	.65	.420
	12	5	23	5	23	.00	.00
	13	5	23	4	24	.13	.716
	14	5	23	3	25	.58	.445
	15	3	25	4	24	.16	.686
	16	4	24	6	22	.49	.485
	17	2	26	3	25	.22	.639
	18	1	27	1	27	.00	1.00
	19	1	27	1	27	.00	1.00

Note. * Denotes significance at the $p < .05$ level.

Table 10. WISC-III Picture Arrangement Item Comparisons by Group Using Chi Square Analysis.

WISC-III Subtest	Item Number	Native American		European American		X^2	p
		Pass	Fail	Pass	Fail		
Picture Arrangement	1	28	0	28	1	.00	1.00
	2	28	0	27	4	1.02	.313
	3	24	4	24	4	.00	1.00
	4	23	5	23	5	.00	1.00
	5	25	3	19	9	3.82	*.050
	6	18	10	14	14	1.17	.280
	7	18	10	15	13	.66	.415
	8	15	13	14	14	.07	.789
	9	16	12	7	21	5.98	*.014
	10	12	16	10	18	.30	.584
	11	8	20	3	25	2.83	.092
	12	11	17	8	20	.72	.397
	13	7	21	5	23	.42	.514
	14	1	27	1	27	.00	1.00

Note. * Denotes significance at the $p < .05$ level.

Table 11. WISC-III Arithmetic Item Comparisons by Group Using Chi Square Analysis.

WISC-III Subtest	Item Number	Native American		European American		X^2	p
		Pass	Fail	Pass	Fail		
Arithmetic	1	28	0	28	0	.00	1.00
	2	27	1	28	0	1.02	.313
	3	28	0	28	0	.00	1.00
	4	28	0	27	1	1.02	.313
	5	27	1	28	0	1.02	.313
	6	27	1	24	4	1.98	.159
	7	26	2	26	2	.00	1.00
	8	26	2	25	3	.00	1.00
	9	26	2	24	4	.16	.686
	10	27	1	26	2	.35	.552
	11	25	3	21	7	1.95	.162
	12	24	4	22	6	.49	.485
	13	22	6	18	10	1.40	.236
	14	20	8	16	12	1.24	.264
	15	11	17	7	21	1.31	.252
	16	8	20	7	21	.09	.762
	17	4	24	4	24	.00	1.00
	18	2	26	2	26	.00	1.00
	19	3	25	0	28	3.17	.075
	20	1	27	1	27	.00	1.00
	21	1	27	0	28	1.02	.313
	22	1	27	1	27	.00	1.00
	23	0	28	0	28	.00	1.00
	24	0	28	0	28	.00	1.00

Note. * Denotes significance at the $p < .05$ level.

Table 12. WISC-III Block Design Item Comparisons by Group Using Chi Square Analysis.

WISC-III Subtest	Item Number	Native American		European American		X^2	p
		Pass	Fail	Pass	Fail		
Block Design	1	28	0	28	0	.00	1.00
	2	28	0	28	0	.00	1.00
	3	28	0	28	0	.00	1.00
	4	27	1	23	5	2.99	.084
	5	26	2	21	7	3.31	.068
	6	23	5	20	8	.90	.342
	7	21	7	21	7	.00	1.00
	8	13	15	8	20	1.20	.273
	9	17	11	9	19	4.59	*.032
	10	12	16	6	22	2.95	.086
	11	6	22	1	27	4.08	*.043
	12	6	22	0	28	6.72	*.009

Note. * Denotes significance at the $p < .05$ level.

Table 13. WISC-III Vocabulary Item Comparisons by Group Using Chi Square Analysis.

WISC-III Subtest	Item Number	Native American		European American		X^2	p
		Pass	Fail	Pass	Fail		
Vocabulary	1	27	1	27	1	.00	1.00
	2	27	1	28	0	1.02	.313
	3	26	2	28	0	2.07	.149
	4	26	2	28	0	2.07	.149
	5	25	3	27	1	1.08	.299
	6	27	1	27	1	.00	1.00
	7	23	5	24	4	.13	.716
	8	24	4	23	5	.13	.716
	9	24	4	23	5	.13	.716
	10	20	8	19	9	.08	.771
	11	16	12	14	14	.29	.592
	12	11	17	12	16	.07	.785
	13	10	18	7	21	.76	.383
	14	3	25	6	22	1.19	.275
	15	4	24	6	22	.49	.485
	16	4	24	3	25	.16	.686
	17	4	24	3	25	.16	.686
	18	4	24	3	25	.16	.686
	19	1	27	1	27	.00	1.00
	20	1	27	0	28	1.02	.313
	21	1	27	4	24	1.98	.159
	22	0	28	0	28	.00	1.00
	23	2	26	1	27	.35	.552
	24	3	25	2	26	.22	.639
	25	0	28	0	28	.00	1.00
	26	0	28	0	28	.00	1.00
	27	1	27	0	28	1.02	.313
	28	0	28	0	28	.00	1.00
	29	0	28	0	28	.00	1.00
	30	0	28	0	28	.00	1.00

Note. * Denotes significance at the $p < .05$ level.

Table 14. WISC-III Object Assembly Item Comparisons by Group Using Chi Square Analysis.

WISC-III Subtest	Item Number	Native American		European American		X^2	p
		Pass	Fail	Pass	Fail		
Object Assembly	1	28	0	28	0	.00	1.00
	2	28	0	27	1	1.02	.313
	3	28	0	28	0	.00	1.00
	4	24	4	19	9	2.50	.113
	5	26	2	25	3	.22	.639

Note. * Denotes significance at the $p < .05$ level.

Table 15. WISC-III Comprehension Subtest Item Comparisons by Group Using Chi Square Analysis.

WISC-III Subtest	Item Number	Native American		European American		X^2	p
		Pass	Fail	Pass	Fail		
Comprehension	1	27	1	26	2	0.35	.552
	2	27	1	28	0	1.02	.313
	3	26	2	28	0	2.07	.149
	4	26	2	28	0	2.07	.149
	5	21	7	22	6	.10	.751
	6	24	4	24	4	.00	1.00
	7	19	9	23	5	1.52	.217
	8	23	5	21	7	.42	.514
	9	18	10	19	9	.08	.777
	10	15	13	14	14	.07	.789
	11	12	16	16	12	1.14	.285
	12	13	15	5	23	5.24	*.022
	13	4	24	3	25	.16	.686
	14	4	24	4	24	.00	1.00
	15	8	20	4	24	1.70	.192
	16	8	20	3	25	2.83	.092
	17	3	25	1	27	1.08	.299
	18	3	25	3	25	.00	1.00

Note. * Denotes significance at the $p < .05$ level.

DISCUSSION

My hypotheses for this study suggested potential differential scoring by Native American children relative to European American children on the Wechsler Intelligence Scale for Children - Third Edition (WISC-III). I predicted a scoring pattern in the Native American group whereby mean Performance IQ scores would be higher than mean Verbal IQ scores thus, supporting other studies (Teeter, Moore, and Petersen, 1982; McCullough, Walker, and Diessner, 1985). Such findings were indeed supported in the present study in that Performance IQ scores were significantly higher than Verbal IQ scores for Native American children. Similar results, though not significant, were found for the European American group. Additional within group comparisons were conducted for both groups between the Verbal Comprehension Index (VCI) scores and the Perceptual Organization Index (POI) scores yielding similar results in that the VCI scores for both groups were higher than their respective POI scores with significance being found only for the Native American group. The subtest scores that contribute to the VCI score are all the Verbal IQ scaled scores except Arithmetic while the POI score is comprised of all the Performance IQ scaled scores except Coding. An examination of subtest scaled scores revealed the European American group scored significantly higher on the Verbal subtests Information and Similarities while the Native American group scored significantly higher on the Performance subtest Picture Completion.

With regard to significantly higher scoring by the European American group on item 4 from the verbal subtest, it is possible that the content of the question is related to SES in that individuals from differing socio-economic backgrounds may not have comparable exposure to musical instruments. Given that chronic levels of poverty and unemployment have been historically and are presently attributed to Native Americans, particularly on reservations, it is possible that, if measured, the Native American group would score lower than their European American counterparts on a measure of SES. If this were in fact true then the differential scoring on this item would possibly be related to lower SES. Culture may offer an alternative explanation in that Native Americans have historically used the drum and to a lesser extent the flute to create music. Stringed instruments may be unfamiliar to them. Perhaps the most parsimonious explanation is that this item is from a subtest that overall, significantly favored the European American group and that a finding of significance on any one item is to be expected.

The Native American group scored significantly higher than the European American group on items 16 and 12 from the Verbal subtests Information and Comprehension respectively These findings appear counter-intuitive since the European American group had higher mean scores than their Native American counterparts on both of these subtests and significantly so on the Information subtest.

The Information subtest assesses the child's knowledge of events, objects, places and people and measures the range of factual information the child has absorbed from the environment and has stored in long-term memory. One explanation for the finding of significance on item 16 of the Information subtest might be contained in an aspect of

Native American culture that is particularly ubiquitous. This is the concept of the four directions. In recent years there has been a resurgence of many tribal religious ceremonies. In practicing these ceremonies explicit instructions must be adhered to particularly regarding the direction participants and ceremonial structures are to face. For example, the entrance of the Lakota Sun Dance arbor is toward the East and the doorway of an inipi (sweatlodge) is always pointed toward the West. Siouan cultures use cloth of the four sacred colors representing the four cardinal directions in many of these ceremonies. Most Native Americans readily identify these colors as red, yellow, black and white. When the four quadrants of a circle are filled with one each of these colors most Native Americans will understand what the colors represent and recognize the design as a medicine wheel. Item 16 from the Information subtest asks the question; "In what direction does the sun set?" The significant differences in the frequency of correct responses to this question across the Native American and European American groups may perhaps be correlated with culture in the ways described above. As Lewis (1995) states:"Native Americans have long had an immediate relationship with their physical environments . . . Land and place remain the substance of Native American identity and rulership, reflected in their life, sacred places and rituals" (p. 423). Another explanation for the apparent anomalous scoring results on these two Verbal subtest items is a difference in the mean ages between the two groups that approaches significance. The Native American group has a mean age that is 11.43 months higher than the European American group. Beiser (1998) states that IQ tests are less a measure of inherent ability than they are of intellectual skills resulting from experience. The scaled scores of the

WISC-III are indeed adjusted according to age. It follows then, that age and its corollaries, experience and environmental exposure, become increasingly important variables as individuals respond to progressively difficult items particularly on the Verbal subtests. It would also be likely that differential interaction effects would be present across different items and ages. In controlling Full Scale IQ, the significant between group differences in scoring on the higher numbered items 16 and 12 from the Information and Comprehension subtests may represent such interaction effects. This potential age by race confound may be a factor related to the item analysis portion of this study only since, as stated earlier, scale and Index scores are age adjusted. Conversely, differences in scoring on lower numbered items (below 5) might typically be explained as being due to differences in inherent ability, however, since item 4 is the only lower numbered Verbal subtest item on which significant differential scoring was found and Full Scale IQ is controlled for, it is more likely that the significance found on this item is due to SES, culture or some other variable rather than age or inherent ability.

A limitation of this study related to the issue of Native American identity is that although subjects from the Native American group were members of federally recognized Northern Plains tribes, over 500 federally recognized tribes exist in the United States, many of whom have their own language and customs therefore, results from this study may not be able to be generalized to other Native Americans. Additionally, the archival nature of the data precluded determination of the individuals' first language as well as their level of cultural and ethno-racial identification. Not controlling for psychiatric diagnoses was another potentially important limitation in this study. Many of the

available protocols for this study were completed by subjects who were administered the WISC-III as a result of being referred for suspected educational difficulties. Small sample size is another limitation of this study. The archival nature of the data presented additional limitations in that demographic information such as socioeconomic status (SES) was not available. Low SES has been shown to be a reliable predictor of low scores on IQ tests both within and across ethnic and racial groups (Suzuki & Valencia, 1997). Further, Willcutt and Pennington (2000) found that children with a reading disability were more likely than those without a reading disability to come from families with lower SES. Native Americans are more likely than any other race in the United States to live in poverty (Beals, Piasecki, Nelson, Jones, Keane, Dauphinais, Red Shirt, Sack and Manson, 1997; Rolo, 1999; U.S. Department of Health and Human Services, 1994) thus, controlling for SES might have helped to elucidate the findings in this study. Another limitation potentially related to SES is that Native Americans have the shortest life expectancy and highest mortality rates of any racial or ethnic group (Rolo, 1999; U.S. Department of Health and Human Services, 1994; Beals et al., 1997).

Upon closer examination these statistics reveal that since 1979, the leading cause of death for Native Americans has been motor vehicle crashes with suicide and homicide alternating as the second and third leading causes (Wallace, Calhoun, Powell, O'Neil, James, 1996). Thus, a measure to determine recent exposure to trauma or stressful life event experiences might have contributed to the findings in this study since these have been linked to various psycho-social problems including school difficulties (Pynoos, Frederick, Nader, Arroyo, Steinberg, Eth, Nunez, & Fairbanks, 1987).

In conclusion, despite the limitations, as outlined above, this study was useful in that it provided support for earlier findings on differential scoring on the WISC-III Verbal and Performance IQ scores. This study may also contribute to discussions regarding the question of item bias on the WISC-III (Mishra, 1982). Since the results of this study did not appear to support claims of cultural bias on individual WISC-III items perhaps it will invoke efforts to further examine such claims in an attempt to identify true sources of potential bias. The WISC-III is used in conjunction with other protocols and interviews in order to fully assess the various dimensions of an individual's functioning. Though it is frequently only part of an assessment battery the WISC-III is an important component and its resulting scores contribute to decisions relating to such things as placement, remedial education and diagnoses of learning disabilities. Findings of scoring discrepancies on the WISC-III across Native American and European American samples have led to such questions regarding the appropriate use of such instruments when assessing Native American children and may cast doubt on the validity of decisions such as those outlined above. Beiser (1998) asserts that IQ scores, particularly those of non-majority children, would provide greater utility if they were utilized as an index of school readiness rather than as a measure of academic potential. Findings from this study may help advance such a debate as well as contribute to the area of American Indian mental health by helping to identify factors, beyond ethnicity, race or culture, that may influence intellectual assessment outcomes for Native Americans. Identifying such factors may enable practitioners in mental health, education and other professions to provide interventions that will prove of greater utility in eliminating the underlying causes of scoring

discrepancies on the WISC-III and ultimately lead to greater academic success for Native American youth.

REFERENCES

Achenbach, T. M. (1991). Integrative guide for the 1991 CBCL/4-18, YSR, and TRF profiles. Burlington, VT: University of Vermont, Department of Psychiatry.

Ballantine, B., & Ballantine, I. (Eds.) (1993). The Native Americans: an illustrated history. Menomonee Falls, WI: Inland Press

Bannatyne, A. (1968). Diagnosing learning disabilities and writing remedial prescriptions. Journal of Learning Disabilities, 1, 242-9.

Bannatyne, A. (1971). Language, reading and learning disabilities. Springfield, IL: Charles C. Thomas.

Bannatyne, A. (1974). A note on recategorization of the WISC scaled scores. Journal of Learning Disabilities, 7, 272-3.

Beals, J., Piasecki, J., Nelson, S., Jones, M., Keane, E., Dauphinais, P., Red Shirt, R., Sack, W.H.. Manson, S.M. (1997). Psychiatric disorder among American Indian adolescents: prevalence in Northern Plains youth. Journal of the American Academy of Child and Adolescent Psychiatry, 36 (9), 1252-1258.

Beck, P. G., Walters, A. L., & Francisco, N. (1996). The sacred: Ways of knowledge, sources of life. Tsaile, AZ: Navajo Community College Press.

Becvar, D. S. & Becvar, R. J. (2000). Family therapy: A systemic integration.(4th Ed.). Needham Heights, Massachusetts: Allyn & Bacon.

Beiser, M., & Gotowiec, A. (2000). Accounting for Native/Non-Native differences in IQ scores. Psychology in the Schools, 37 (3), 237-52.

Beiser, M., Sack, W., Manson, S. M., Redshirt, R., & Dion, R. (1998). Mental health and the academic performance of first nations and majority-culture children. American Journal of Orthopsychiatry, 68(3), 455-467.

Beiser, M. (1989). Flower of Two Soils: Final Report for NIMH Grant No. MH 36678-04. Rockville, MD: National Institute of Mental Health.

Bloom, A. S., Topinka, C. W., Goulet, M., Reese, A., Podruch, P. E. (1986). Implications of large WISC/WISC-R Verbal-Performance IQ discrepancies. Journal of Clinincal Psychology, 42, 353-6

Chapleski, E. E., Lamphere, J. K., Kaszynski, R., Lichtenberg, P. A. & Dwyer, J. W. (1997). Structure of a depression measure among Native American elders: Confirmatory factor analysis of the Center for Epidemiological Studies Depression (CES-D) Scale. Research on Aging, 19 (4), 462-86.

Coggins, K. (1990), Alternative pathways to healing: The recovery medicine wheel. Health Communications. Deerfield Beach, FL.

Coleman, V. D., & Barker, S. A. (1991). Barriers to the career development of multicultural populations. Educational and Vocational Guidance, 52, 25-9.

Dana, R.H. (1993). Multicultural assessment perspectives for professional psychology. Massachusetts: Allyn &Bacon.

Devers, R., Bradley-Johnson, S., & Johnson, C.M. (1994). The effect of token reinforcement on WISC-R performance for fifth-through ninth-grade Native Americans. The Psychological Record, 44 (3), 441-9.

Flaugher, R. L. (1978). The many definitions of test bias. American Psychologist, 33, 671-9.

French, L. (1981). Counseling Native Americans. International Journal of Offender Therapy and Comparative Criminology, 25, 139-55.

Gutkin, T. B. (1979). Bannatyne patterns of European American and Mexican-American learning disabled children. Psychology in the Schools, 16, 178-83.

Hale, R. L., Raymond, M. R., & Gajar, A. H. (1982). Evaluating socioeconomic status bias in the WISC-R. Journal of School Psychology, 20, (2), 145-9

Hollingshead, A. B. (1957). Two-factor index of social position. Unpublished manuscript.

Hutton, J. B., Dubes, R., & Muir, S. (1992). Assessment practices of school psychologists: Ten years later. School Psychology Review, 21, 271-84.

Hynd, G. W., & Garcia, W. I. (1979). Intellectual assessment of the Native American student. School Psychology Review, 8 (4), 446-54.

Jastak, J., Bijou, S., & Jastak, S. (1976) The Wide Range Achievement Test. Wilmington: Guidance Associates of Delaware.

Jastak, S., & Wilkinson, G. S. (1984). Wide Range Achievement Test-Revised. Wilmington, DE: Jastak Associates.

Katz, J. H. (1985). The sociopolitical nature of counseling. The Counseling Psychologist, 13, 615-24.

Katz, J. (Ed.) (1995). Messengers of the wind. New York: Ballantine Books.

Lewis, D. R. (1995). Native Americans and the environment: a survey of twentieth-century issues. The American Indian Quarterly, 19, (3), 423-28.

Locust, C. (1988). Wounding the spirit: Discrimination and traditional Native American belief systems. Harvard Educational Review, 58, 315-30.

McCullough, C. S., Walker, J., & Diessner, R. (1985). The use of Wechsler Scales in the assessment of Native Americans of the Columbia River Basin. Psychology in the Schools, 22, (1), 23-8.

McDonald, J. D., Morton, R., & Stewart, C. (1993). Clinical concerns with Native American patients. Innovations in Clinical Practice: A Source Book, 12, 437-54.

McShane, D.(1980). A review of scores of Native American children on the wechsler intelligence scales. White Cloud Journal, 1, (4), 3-10.

McShane, D. M., & Plas, J. M. (1982). Wechsler scale performance patterns of Native American children. Psychology in the Schools, 19, 8-17.

Mishra, S. P. (1982). The WISC-R and evidence of item bias for Native-American Navajos. Psychology in the Schools, 19, 458-64.

Mishra, S. P., & Lord, J. (1982). Reliability and predictive validity of the WISC-R with Native-American Navajos. Journal of School Psychology, 20, (2), 150-4.

Neihardt, J. G. (1932). Black Elk Speaks. New York: W. Morrow and Company

Oetting, E. R. & Beauvais, F. (1990). Orthogonal cultural identification theory: The cultural identification of minority adolescents. The International Journal of the Addictions, 25 (5A & 6A), 655-685.

Pynoos, R. S., Frederick, C., Nader, K., Arroyo, W., Steinberg, A., Eth, S., Nunez, F., & Fairbanks, L. (1987). Life threat and posttraumatic stress in school-age children. Archives of General Psychiatry, 44 1057-1063.

Rolo, M. A. (1999). Indian country's hidden healthcare CRISIS. Civil Rights Journal, 4, (1), 44.

Rugel, R. P. (1974). WISC subtest scores of disabled readers: A review with respect to Bannatyne's recategorization. Journal of Learning Disabilities, 7, 48-55.

Sampson, E. E. (1985). The decentralization of identity: Toward a revised concept of personal and social order. American Psychologist, 40, 1203-11.

Shute, N. (1996). Larry Merculieff: Profile of an Aleut community leader. Winds of Change 11, (4) 23.

Simonelli, R. (1994). Finding balance: Looking beyond the scientific method. Winds of Change, 9, (4), 106-12.

Smith, M. D., Coleman, J. M., Dokecki, P. R., & Davis, E. E. (1977). Intellectual characteristics of learning disabled children. Journal of Learning Disabilities, 10, 437-43.

Stinnett, T. A., Havey, J. M., & Oehler-Stinnett, J. (1994). Current test usage by practicing school psychologists: A national survey. Journal of Psychoeducational Assessment, 12, 331-50.

Suzuki, R. J., Valencia, R. R. (1997). Race-ethnicity and measured intelligence: Educational implications. American Psychologist, 52, 1103-14.

Sue, D. W. & Sue, D. (1990). Counseling the culturally different (2nd ed.). New York: Wiley & Sons.

Tedlock, B. & Tedlock, D. (1995). Teaching from the American earth, New York: Liverright Press.

Teeter, A., Moore, C. L., & Petersen, J. D. (1982). WISC-R verbal and performance abilities of Native American students referred for school learning problems. Psychology in the Schools, 19, (1), 39-44.

The Psychological Corporation. (1992). Manual for the Wechsler Individual Achievement Test. San Antonio, TX.

Tyler, J. D., Cohen, K. N. & Clark, J. S. (1982). Providing community consultation in a reservation setting. Journal of Rural Community Psychology, 3 (1), 49-58.

U.S. Department of Health and Human Services. (1994). Trends in Indian Health. Public Health Service, Indian Health Service. Washington, DC: US Government Printing Office.

Wallace, L. J. D., Calhoun, A. D., Powell, K. E., O'Neil, J., & James, S. P. (1996). Homocide and suicide among Native Americans 1979-1992. Atlanta: U. S. Department of Health and Human Services Center for Disease Control; Violence surveillance summary series, 2.

Waterman, A. S. (1981). Individualism and interdependence. American Psychologist, 36, 96-112.

Wechsler, D. (1955). Manual for the Wechsler Adult Intelligence Scale. New York: Psychological Corporation.

Wechsler, D. (1967). Manual for the Wechsler Preschool and Primary Scale of Intelligence. San Antonio: The Psychological Corporation.

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

Wechsler, D. (1991). Wechsler Intelligence Scale for Children-Third Edition: Manual. New York: Psychological Corporation.

Weiss, L. G., & Prifitera, A. (1995). An evaluation of differential prediction of WIAT achievement scores from WISC-III FSIQ across ethnic and gender groups. Journal of School Psychology, 33, (4), 297-304.

Willcutt, E. G., Pennington, B. F. (2000). Comorbidity of reading disability and attention-deficit/hyperactivity disorder: differences by gender and subtype. Journal of Learning Disabilities, 33, (2), 179.

Zarske, A. & Moore, C. L. (1982). Recategorized WISC-R scores of learning disabled Navajo Indian children. Psychology in the Schools, 19, 156-59.