1984

Intelligence Testing of American Indian Children: Sidesteps in Quest of Ethical Practice

Richard H. Dana
Portland State University

Citation Details
Dana, R. H. (1984). Intelligence testing of American Indian children: Sidesteps in quest of ethical practice. White Cloud Journal, 3(3), 35-43.

Intelligence Testing of American Indian Children: Sidesteps in Quest of Ethical Practice

Richard H. Dana

Previous literature reviews are updated. Recent findings impugn the WISC-R internal consistency and document item bias for American Indian children. A pattern of Spatial> Sequential> Conceptual> Acquired Knowledge holds across ages, tribes, and heterogeneous referral sources, except for acculturated children. Kaufman's three factors are replaced by Verbal and Performance factors for Papago and Navajo children. A culturally-learned basis for intellectual functioning among traditional children supports alternative assessment functions for traditional reservation lifestyle and for acculturation and entree into mainstream society. Performance measures, SOMPA, Piagetian and Luria-derived tasks may ultimately provide less biased intelligence estimates. Recent legislation outlines ethical practice although assessors are still making do with conventional measures that are ethically-questionable and discriminatory. Suggestions for increasing awareness of responsible practice include training in cultural contents and constant monitoring of research findings.

Intellectual assessment of American Indian children has been carried out utilizing tests that have not been standardized on these populations and are generally inappropriate for description of their intellectual functions or for prediction of their educational outcomes. Performance tests administered by the multidisciplinary Indian Educational Research Project in 1941 provided similar average scores and ranges of scores for American Indian and white children (Havighurst, 1958). Although these results have been replicated consistently across tribes for a variety of performance measures (e.g., Cundick, 1970; Telford, 1932; West & MacArthur, 1964), the continued use of verbal measures such as the Wechsler verbal subtests or the Stanford-Binet perpetuates the erroneous belief that these children have limited intellectual endowment and can anticipate only nominal educational achievement. This paper suggests a rationale for a lack of representation of the state-of-the-science research in current assessment practice, describes alternative measures, and offers a perspective on usage of intelligence tests with traditional and acculturated American Indian children.

Literature Reviews

Two recent reviews have described usage of the Wechsler Intelligence Scales with American Indian children (Hynd & Garcia, 1979; McShane, 1980).

Acknowledgement is made to Donna Thornton for assistance in compiling references for an earlier draft, to the School of Professional Psychology, University of Denver for facilitation, and to the University of Arkansas for off-campus leave.

Richard H. Dana is Professor of Psychology, University of Arkansas — Fayetteville. He is a Diplomate in Clinical Psychology, and Past President of the Society for Personality Assessment. His concern with services for Native Americans has appeared in *A Human Science Model for Personality Assessment with Projective Techniques* and *Human Services for Cultural Minorities* (Ed.) and a TAT chapter in Newmark's *Major Psychological Assessment Instruments*. Future research will focus on development of holistic health and intervention practices conducted by Native Americans.

McShane (1980) statistically examined the ranked subtest scores and Verbal-Performance scale differences from 12 studies including over 600 children. Although these children had significantly lower scores than white children, relatively consistent high scores were obtained on Mazes, Block Design, Object Assembly, and Picture Completion while Vocabulary was the single lowest subtest score. These findings were related to language deficiency, prevalent middle-ear disease (otitis media), items of inconsistent difficulty level and items which were biased, inappropriate, or anxiety-producing for this population. These children did not analyze experience in verbal terms leading to internalized rules for application in new situations. Their problem-solving was more determined by fluid than by crystallized intelligence.

On the basis of different literature, Hynd and Garcia (1979) concluded that performance subtests provide an estimate of intellectual potential while verbal subtests measure acculturation or current functioning within an English language academic setting. Their review suggested that verbal subtest scores are also affected by typical nonassertive behaviors which minimize rapport and contain little spontaneous verbal interaction or eye contact. Vocabulary differences and fewer nouns preclude adequate translations while test content that is irrelevant to reservation lifestyle consistently affects scores. Furthermore, the belief that self-disclosure to a stranger reduces self-control implies that examiner pressure for responsiveness may result in passivity or even withdrawal.

Conventional Standardized Intelligence Tests: Wechsler Intelligence Scale for Children-Revised (WISC-R)

Recent research with the WISC-R has clarified an empirical rationale for content or item bias, documented inadequate internal consistency, and modified interpretation based on factor structure and categorization of subtests. Flaugher (1978) argued for an empirical approach to content bias exemplified by Mishra (1982) who examined cultural bias on the 77 items from Information, Similarities, and Vocabulary subtests by comparing Anglo and Navajo children matched for grade level. Fifteen items, or 19 percent, were biased against these Navajo children. The difference between percentages of Anglo and Navajo children passing items were all extreme and the biased items were relatively evenly distributed across the three subtests. Since there were only 40 children in each group, a Navajo replication is mandatory prior to examining items for other tribes. Other methodologies, however, such a comparison of item correlation matrices may produce similar Anglo-American Indian factor structures (Martin, Pine & Weiss, 1979), at least in an aptitude test.

Mishra and Lord (1982) have also demonstrated remarkably low internal consistency for an identical number of Navajo children, although it is not clear whether the same data were used in both studies (Mishra, 1982). These children were from low-income, isolated, reservation families and thus may represent a minimally acculturated group. Split-half reliabilities for Information, Comprehension, and Arithmetic were .46, .55, and .59, respectively, while .63 was obtained for Verbal IQ. The range of Verbal coefficients was from .46 to .79 (Vocabulary) and from .59 (Picture Arrangement) to .75 (Block Design) for Performance subtests. In addition, correlations with the Wide Range Achievement subtests - Reading, Spelling, and Arithmetic - were non-significant.

Interpretation of the WISC-R has been based on factor analytic findings (Bannatyne, 1971, 1974; Kaufman, 1975). Three factors (Verbal Comprehension, Perceptual Organization, and Freedom from Distraction) were identified in the standardization data that were stable across age levels and factor analytic methods (Kaufman, 1975; Silverstein, 1977). The Verbal Comprehension factor includes Information (I), Similarities (S), Vocabulary (V), and Comprehension (C) subtests. Perceptual Organization includes Picture Completion (PC), Picture Arrangement (PA), Block Design (BD), Object Assembly (OA), and Mazes subtests. Freedom from Distraction includes Arithmetic (A), Digit Span (DS), and Coding (Co) subtests. Reschly (1978), however, was unable to replicate these three factors with a large Papago sample. He found only two factors, Verbal (V, I, C, S) and Performance (BD, OA, PA, PC). Zarske, Moore, and Peterson (1981) replicated these two-factor findings for Papago and extended them to include Navajo children.

Although it is tempting to use the two-factor clustering of subtests, McShane (1980) reports that random samples of American Indian children show mean Verbal-Performance differences of 10 points. Children from traditional families have a 25-point difference while children from acculturated families

have no Verbal-Performance difference whatsoever.

Several authors recommend use of Performance scores to measure learning potential for normal American Indian children (Hynd, Quackenbush, Kramer, Conner & Weed, 1979; Teeter, Moore & Peterson, 1982). In a large sample, WISC-R Navajo study of nonhandicapped, educationally disadvantaged, and learning disabled children, Verbal scores constituted an index of academic proficiency in English-speaking classrooms while Full Scale IQ scores led to bias which may be due to contamination of learning potential (Performance) with acculturation (Verbal) (Teeter et al., 1982). In addition, these learning disabled children have lower performance scores and the WISC-R may be unable to measure learning potential for this population.

Bannatyne (1971, 1974) developed new categories of mean scaled subtest scores as a substitute for the Verbal and Performance subtest clustering: Spatial (BD, PC, OA), conceptual (C, S, V), Sequential (DS, Co, A), and Acquired Knowledge (I, A, V). Learning disabled, white children across many studies and a variety of specific disabilities had a consistent pattern (Rugel, 1974): Spatial> Conceptual> Sequential. Zarske and Moore (1982) did not find this pattern among 192 rural, reservation, learning disabled Navajo children. They found instead a Spatial> Sequential> Conceptual pattern and believed that the very low Conceptual scores were due to English as a second language, culture, and the specific learning disability. McShane and Plas (1982a) extended these findings across tests (WISC, WISC-R, WPPSI), tribes (Ojibwa, Sioux), and a more heterogeneous sampling of school children (referrals for school difficulties, otitis media, and giftedness). Their findings established an American Indian pattern of Spatial> Sequential> Conceptual> Acquired Knowledge categories. This pattern was present only in children from traditional families.

Summary. Conventional instruments will continue to be used in school and employment settings. Nonetheless, the manner in which these instruments are used is now a matter of responsible professional practice and ethics. Certain suggestions for WISC-R usage are explicit in recent literature:

1. While Full Scale scores may continue to be obtained, they should not be reported as an IQ measure due to lack of internal consistency and potential item-bias on verbal subtests.

2. Verbal subtests may be interpreted as an aptitude test equivalent for estimation of performance in conventional, English language classrooms.

3. Verbal-Performance differences should be inspected for an estimate of degree of acculturation. Extreme differences may be hypothesized to reflect traditional culture, although independent confirming evidence is needed.

4. While performance subtest scores provide an estimate of IQ, or possibly learning potential, these scores will be artificially depressed for children with learning disabilities.

5. Kaufman's (1975) three factors should not be used for interpretation and the limitations mentioned above for Verbal-Performance differences apply.

6. Bannatyne (1971) categories should only be used in the Indian pattern for traditional children: Spatial> Sequential> Conceptual> Acquired Knowledge. Since this pattern occurs in diverse groups, it should not be used for diagnosis of learning disability without additional information.

Non-Wechsler Performance Tests

It is also possible to use performance tests other than the Wechsler Performance subtests. Havighurst and Hilkevitch (1944) reported data for Hopi, Navajo, Papago, Sioux, Zia and Zuni children from the Grace Arthur Performance Scale, short form, that includes Knox Cube, Seguin Form Board, Mare and Foal, Porteus Maze, and Kohs Block Design with pantomime directions. In addition, the Goodenough Draw-A-Man test may be used (Dennis, 1942; Havighurst, Gunther & Pratt, 1946; Levensky, 1970), especially in tribes where art work is culturally-relevant. However, the ability to draw a human figure, especially hands, is highly correlated with artistic skill. These performance tests all lack up-to-date norms and local norms would need to be developed for purposes of screening and placement in special classes. The advantage of being able to use pantomime directions, or an unstructured and potentially less formidable administration procedure for the figure drawing test, for example, may in some settings compensate for a history of less frequent usage and research.

Performance tests, however even when they constitute legitimate emic measures merely reiterate the truism that some measures are more fair and

relevant than others for a particular culture. It is necessary to have some strategy for selecting specific measures from among the thousands of available and potentially usable instruments.

System of Multicultural Pluralistic Assessment (SOMPA)

Reschly (1981) has used SOMPA (Mercer & Lewis, 1978) to provide corrections for culture that may be applied to WISC-R scores for American Indian children. The SOMPA contains an Estimated Learning Potential (ELP) measure (effects of using sociocultural background information) and an Adaptive Behavior for Children (ABIC) measure (non-cognitive performance). When these corrections were applied to Papago WISC-R records (N = 122), the numbers of students classified as mild mental retardation (using a -2 SD or IQ <70 criterion) were reduced from 22 to 4 with the ELP and to zero using both ELP and ABIC. The SOMPA provides separate norms for different groups of minority persons since standardized tests usually omit these persons altogether or use too few persons to be representative. While the SOMPA combination of medical, social, and cultural measures provides an IQ correction that may keep minority pupils out of the classes for the educable mentally retarded, adding bonus points to test scores based on minority status has been criticized (Oakland, 1979).

Test Bias and Culture-Specific Abilities

Intellectual assessment of Native Americans exhibits all of the problems inherent in testing persons from a different culture. Test item content is often biased and internal consistency reliabilities may be so low as to preclude effective use of standard intelligence tests. The conventional use of an equal number of verbal and performance subtests is contrary to primary cultural learning of performance abilities. The format for test administration is alien and potentially threatening, providing what Flaugher (1978) has called "atmosphere" bias. Finally, the basis for determination of IQ is found in norms derived from non-Indian populations.

These problems have been largely ignored in part because many psychologist assessors tacitly accept a deficit interpretation of American Indian intelligence. This belief asserts that poor persons, particularly those with minority status, have disorganized lifestyles leading to expression of deficit by intelligence test scores. Cole and Bruner (1971) have refuted this specious argument on the grounds of the anthropological doctrine of psychic unity or intellectual equality. Intellectual equality may also be inferred from the similarity of all languages in degree of development and from "situation-bound" or emic measures.

In spite of these problems that fuse ethnocentricism, racial bias, and insistence upon homogenization of all persons in accord with American middle-class behaviors, we do possess some knowledge that is applicable to intellectual assessment of American Indians. First, their traditional cultures foster the development of performance abilities and minimize opportunities for processing information in abstract verbal terms. High visual-spatial abilities have been reportedly consistently across tribes (e.g., Berry, 1971; Lombardi, 1970). Historically, adaptation to Native American lifestyles required superior perceptual-motor skills, emphasis upon concrete reality (practicality), and individuality of objects at the expense of generalizations. English conceptual abilities may be lower because traditional American Indian children in some tribes do not analyze experience in verbal terms (Schubert & Cropley, 1972). Instead these children learn primarily by imitation and lack the pervasive early experience with adult models who process information in abstract verbal terms.

Second, the extent to which verbal skills are developed among American Indians is a function of acculturation to middle-class, white society. An acculturated parent, and/or early experience with a white peer group, are necessary for development of an approach to learning that fosters the use of concepts and generalization across contexts. It is now mandatory to apply routine corrections for acculturation to intelligence test scores as has been done with the Traditional Experience Scale (McShane & Plas, 1982a) or statistically (Reschly, 1981) by using SOMPA measures to modify the interpretation of WISC-R scores. However, the application of such corrections for acculturation/traditionalism to standard tests does not mitigate the harm that these tests inflict upon persons who employ different learning styles, alternative modes of experience, and function on the basis of values not shared with the majority culture.

Theory-Derived Assessment

Assessment of intelligence implies more than a decision that a certain set of measurement operations is sufficient. Phenomenon naming is not causal naming. Historically, we have accepted empirical definitions such as Binet's age-related task proficiencies, or Wechsler's population norms for specific verbal and performance tasks. The emphasis on individual differences and task content was culturally determined. We wanted a rank-order of merit for assignment of persons to scarce educational resources and occupational opportunities. Intelligence was a prediction of goodness-of-fit with the established, middle-class Western society that enabled survival and facilitated mobility and access to material goods. The attempt to understand intelligence occurred within this framework of assumptions such that empirical scrutiny did yield communalities among tasks/tests that had been derived for prediction of future personal status rather than understanding of individual intellectual development and functioning.

Theory-derived assessments have been almost totally obscured by blind faith in empiricism and disinclination to consider alternatives that violate the implicit values and assumptions which define our identity as assessors. Nonetheless, it is feasible to approach intelligence by careful and detailed observation of individuals as Piaget has done in order to understand the basic nature of the phenomenon prior to interest in its range of variations (Furth, 1973). Psychologists in their haste to be acknowledged as scientists have often minimized or omitted adequate description of human phenomena.

Piagetian tasks offer promise for understanding the intellectual functioning of Native American children (Clarizio, 1982). A general intelligence factor is common to Piagetian and psychometric tests (Humphreys & Parsons, 1979). Piagetian tasks measure fluid intelligence and use theory-based attempts to define developmental benchmarks in logical reasoning. These tests constitute genuine ordinal scales that are relatively unstructured and are qualitatively scored. They presuppose no interest in individual differences or invidious comparisons with age and grade norms. When lower-class, bilingual American Indian (tribes unspecified) and white children were compared on a Piagetian scale (Goldschmid's Concept Assessment Kit; Goldschmid & Bentler, 1968), there were no differences across various conservation tasks in rates of concept acquisition (Gaudia, 1972). Such tasks are more likely to be culture-fair, child-relevant, and essentially descriptive of current developmental status. While there has been disinterest in predictive validity by researchers in this area, it is possible to demonstrate that a 25-minute Piagetian battery outperforms a conventional IQ test (Lorge-Thorndike Intelligence Test, Level 1, Form A) in estimation of a Stanford Achievement criterion (Kaufman & Kaufman, 1972).

It is also feasible to combine a biological approach with analysis of determining social conditions as Luria has done. Luria's contributions to verbal regulation of behavior during specific learning stages, especially as applicable cross-culturally, have been presented in detail as a developmental theory by Josef Schubert (1983). Schubert developed a reliable, verbal regulation of behavior (VRB) apparatus for presentation of problems at three levels of difficulty that include training, statement of the rule, and discovery of the rule. The VRB has been administered to about 800 white school children and has been used in two studies (Schubert & Cropley, 1972; Steinberg, 1974 cited in Schubert, 1983) with remote, rural (traditional) and urban-fringe (acculturated) reservation, Cree children. While both groups were approximately equivalent on motor tasks, the remote rural children were unable to state the rule that guided their behavior. Failure to formulate the rule for previously solved problems resulted in perseveration when the child faced more complex problems while formulation of the rule led to easy solution of the more complex problems.

Luria (1966) has also described successive and simultaneous modes of information processing located in specific parts of the brain that are represented by a broad range of memory and reasoning tasks identified by factor analysis (Das, 1973). The successive processing factor has high loadings for Serial and Free Recall, Cross-Modal Coding, and Raven's Progressive Matrices while the simultaneous factor has high loadings on Memory for Designs and Figure Coping Tests. A Speed factor has Stroop Word Reading Speed as the highest loading but also included higher loadings on WISC Verbal and Performance components than occurred on the other factors. Krywaniuk and Das (1976) used a remedial intervention program based on identification of these factors for Canadian Cree children that improved auditory and visual memory and reading (sequential strategies) as demonstrated by post-treatment factors and factor loadings. Initially

these children had well-developed simultaneous strategies which they used instead of more efficient sequential processing. This study indicates greater utility for non-WISC tests in assessment of learning problems for these children. McShane and Plas (1982b) could not replicate the Cree findings with Ojibwa children using the WISC-R subtests, perhaps because the WISC-R is an inappropriate measure of these cognitive processes and loads most highly on a speed factor.

When assessment devices are constructed on the basis of developmental theory, the evidence does not indicate cultural bias. In spite of this fact, most assessors have accepted the WISC-R subtests as representative of the entire range of intellectual abilities for all persons. The two landmark studies using instruments derived from theories that stress development of intellectual functions with Native American children are not recent (Gaudia, 1972; Schubert & Cropley, 1972). These studies have not been influential in affecting assessment practices.

Discrimination and Some Assessment Palliatives

Oakland (1982) suggests that discrimination occurs whenever pupils are not tested in their native or dominant language and tests are used that reflect white, middle-class abilities. The outcomes are overrepresentation and longevity in classes for the educable mentally retarded and underrepresentation in learning disabled classes. Parents are seldom informed or consulted on potentially dehumanizing decisions based on meager information. Although Oakland recommends a wide spectrum of ameliorations that include help for pupils in taking tests, pluralistic norms, culture-fair and culture-specific tests, he indicates that all of these procedures have problems.

A major change in school assessment has resulted from the current guidelines for nonbiased assessment (Public Law 94-12). The major provisions of this legislation are for a professional team to develop an annual, educationally-relevant Individual Educational Plan that provides diagnostic information usable for intervention and programming with reevaluation every three years. Participation by parents is encouraged with provision for written notices in their native language prior to review for special class assignment. Multidimensional assessment using nonbiased tests is required. It is mandatory to assess areas of educational need with tests validated for these specific purposes and administered by competent examiners using standard procedures. While laws do not work miracles, they do provide an additional incentive for schools to alter their current assessment practices and a viable recourse for individual professionals or parents.

Discussion

There are no easy solutions to the problems inherent in the assessment of American Indians. Cost-benefit ratios and human values are often discrepant. By espousing values of cultural homogenization and expressing prejudice toward the culturally-different, we have attempted to minimize or eradicate the traditional heritage of all minorities, including American Indians. Our failure is evidenced in the resurgence of racial identity and pride that has characterized all minorities in recent years. For American Indians, maintenance of language (Leap, 1981), and cultural-religious traditions, as seen at Sinte Gleska College, has been coupled with aggressive confrontation of the larger society exemplified by the American Indian Movement. Josephy (1982) has traced the history of these evidences for an expanding American Indian self-consciousness and the determination to remain traditionally Indian.

Intellectual assessment is mandatory for provision of adequate learning environments for all persons in this society. Idealistically American Indians need two specific learning environments, one designed to train them in skills necessary for entry and good functioning within the mainstream culture and the other designed for reservation living in traditional cultural lifestyles, including instruction in their native languages. There can be no restoration of Native American past, or return to a society in which communal values and harmony with nature assured continuity and quality of life. There must be preservation of the identity-forming ingredients of Native American culture and provision for adequate environments to sustain the entire range of acculturated and traditional lifestyles.

A reservation lifestyle that preserves belief systems constituting the bases for identity—language, medicine, ceremony and social practices—can provide a meaningful personal future for traditional American Indians. There is documentation for assessment that is designed to enhance traditional lifestyles. Traditionality among the

Lakota Sioux is characterized by a functioning Tiospaye, an extended family providing a way of life following a set of values with rules for social interaction, rituals for transition, identity acquisition, and healing. Using Likert-Type scales to measure community Tiospaye, Mohatt and Red Bird (undated) found that high Tiospaye did reduce or prevent serious stress in dispersed communities as defined by legal data and hospital records.

Clinical psychologist assessors should distinguish clearly between relevant measures of general intellectual functioning and readiness measures for specific educational objectives. We should use established measures only in ways that have been validated in research with American Indians. The development of local norms is desirable since most research has been done with Navajo and Papago tribes. Finally there is an ethical necessity to develop measures that are tribe-specific and setting-specific in order to be responsive to local conditions. In addition, there is need for more general theory-derived culture-fair instruments that are primarily descriptive in nature. Such instruments promote understanding of the intellectual processes and can serve as a basis for curricula that are designed for all students.

As clinical psychologist assessors we have been unwitting conspirators in the preservation of American Indian status quo by using our instruments for caricature, dehumanization, and discrimination. This middle-class bias among clinical psychologists has been described in other contexts (Beit-Hallahmi, 1974). We have thereby contributed to a denial or limitation of individual access to educational and vocational opportunity in the dominant culture.

The application of inappropriate instruments remains an ethical issue. Intellectual assessment that is fair and ethical requires instruments constructed on the basis of theory, or especially designed for American Indians and using these persons as the primary reference group with administration conditions, instructions, and tasks that are culturally appropriate. Furthermore, the training of assessors must include perspective-taking (Brady, Manni & Winikur, 1983), culturally-relevant professional experiences (Dana, in press-a, 1984; Oakland, 1983) and explicit monitoring of the extant literature as it applies to practice (Dana, 1981). Ultimately what is required is a change in attitude toward culturally different persons that legitimatizes their construction of reality and honors their human condition. Ethical assessment practice would follow since we already possess the necessary knowledge and instruments. However, assessment practice is part-and-parcel of a professional stance that is bound to middle-class culture and enmeshed with a larger social system that nourishes the fiction of ethnocentricism.

References

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

Bannatyne, A. (1974). Diagnosis: A note on recategorization of the WISC scale scores. *Journal of Learning Disabilities, 7*, 13-14.

Beit-Hallahmi, B. (1974). Salvation and its vicissitudes: Clinical psychology and political values. *American Psychologist, 29*, 124-129.

Berry, J. W. (1971). Psychological research in the North. *Anthropologica, 13*, 142-157.

Brady, P. M., Manni, J. L. & Winikur, D. W. (1983). A three-tiered model for the assessment of culturally and linguistically different children. *Psychology in the Schools, 20*, 52-58.

Clarizio, H. F. (1982). Piagetian assessment measures revisited: Issues and applications. *Psychology in the Schools, 19*, 421-430.

Cole, M., & Bruner, J. S. (1971). Cultural differences and inferences about psychological processes. *American Psychologist, 26*, 867-876.

Cundick, B. P. (1970). Measures of intelligence on Southwest Indian students. *Journal of School Psychology, 81*, 151-156.

Dana, R. H. (Ed.). (1981). *Human services for cultural minorities*. Baltimore: University Park Press.

Dana, R. H. (in press - a.). Clinical assessment. In G. S. Tryon (Ed.), *The professional practice of psychology*. Norwood, NJ: Ablex Publishing Corp.

Dana, R. H. (1984). Personality assessment: Practice and teaching for the next decade. *Journal of Personality Assessment, 48*, 46-57.

Das, J. P. (1973). Structure of cognitive abilities: Evidence for simultaneous and successive processing. *Journal of Educational Psychology, 65*, 103-108.

Dennis, W. (1942). The performance of Hopi Children on the Goodenough Draw-A-Man Test. *Journal of Comparative Psychology, 34,* 341-348.

Flaugher, R. L. (1978). The many definitions of test bias. *American Psychologist, 33,* 671-679.

Furth, H. G. (1973). Piaget, IQ and the nature-nurture controversy. *Human Development, 16,* 61-73.

Gaudia G. (1972). Race, social class, and age of achievement of conservation Piaget tasks. *Development Psychology, 6,* 158-167.

Goldschmid, M. & Bentler, P. (1968). *The conservation concept diagnostic kit.* San Diego, CA: Educational and Industrial Testing Service.

Havighurst, R. J. (1958). Education among American Indians: Individual and cultural aspects. *The Annals of the American Academy of Political and Social Science, 311,* 113.

Havighurst, R. J. & Hilkevitch, R. R. (1944). The intelligence of Indian children as measured by a performance scale. *Journal of Abnormal and Social Psychology, 39,* 419-433.

Havighurst, R. J., Gunther, M. K. & Pratt, I. E. (1946). Environment and the Draw-A-Man Test: The performance of Indian children. *Journal of Abnormal and Social Psychology, 41,* 50-65.

Humphreys, L., & Parsons, C. (1979). Piagetian tasks measure intelligence and intelligence tests assess cognitive development. A reanalysis. *Intelligence, 3,* 369-382.

Hynd, G. W. & Garcia, W. I. (1979). Intellectual assessment of the Native American student. *School Psychology Digest, 8 (4),* 446-454.

Hynd, G. W., Quackenbush, R., Kramer, R., Conner, R., & Weed, W. (1979). Clinical utility of the WISC-R and the French Pictorial Test of Intelligence with Native American primary grade children. *Perceptual and Motor Skills, 49,* 480-482.

Josephy, A., Jr. (1982). *Now that the buffalo's gone.* New York: Knopf.

Kaufman, A. (1975). Factor analysis of the WISC-R at eleven age levels between 6½ years and 16½ years. *Journal of Consulting and Clinical Psychology, 43,* 135-147.

Kaufman, A. S., & Kaufman, N. L. (1972). Tests built from Piaget's and Gessell's task as predictors of first grade achievement. *Child Development, 43,* 521-535.

Krywaniuk, L. W., & Das, J. P. (1976). Cognitive strategies in Native children: Analysis and intervention. *The Alberta Journal of Educational Research, 22*(4), 271-280.

Leap, W. L. (1981). American Indian language maintenance. *Annual Review of Anthropology, 10,* 209-236.

Levensky, K. (1970). *The performance of American Indian children on the Draw-A-Man Test.* Washington, DC: National Study of American Indian Education. (ERIC Document Reproduction Service No. ED 039 081)

Lombardi, T. P. (1970). *Psycholinguistic abilities of Sioux Indian children.* Unpublished doctoral dissertation, University of South Dakota.

Luria, A. R. (1961). *The role of speech in the regulation of normal and abnormal behavior.* London: Pergamon Press.

Luria, A. R. (1966). *Higher cortical functions in man.* New York: Harper & Row.

Martin, J. T., Pine, S. M., & Weiss, D. J. (1979). An item bias investigation of a standardized aptitude test. *Catalog of Selected Documents in Psychology, 9,* 47-48 (Ms. No. 1870).

McShane, D. A. (1980). A review of scores of American Indian children on the Wechsler Intelligence Scales. *White Cloud Journal, 1*(4), 3-10.

McShane, D. A., & Plas, J. M. (1982a). Wechsler Scale performance patterns of American Indian children. *Psychology in the Schools, 19*(1), 8-17.

McShane, D. A. & Plas, J. M. (1982b). WISC-R factor structures for Ojibwa Indian Children. *White Cloud Journal, 2*(4), 18-22.

Mishra, S. P. (1982). The WISC-R and evidence of item bias for Native-American Navajos. *Psychology in the Schools, 19,* 458-464.

Mishra, S. P. & Lord, J. (1982). Reliability and predictive validity of the WISC-R with Native-American Navajos. *Journal of School Psychology, 20*(2), 150-154.

Mohatt, G., & Red Bird, S. (undated). *Identity through traditional Lakota methods* (Rep. 5R01 MH23697-03). Rockville, MD: National Institute of Mental Health.

Oakland, T. D. (1979) Research on the ABIC and ELP: A revisit to an old topic. *The School Psychology Digest, 8,* 211-213.

Oakland, T. D. (1982). Nonbiased assessment in counseling: Issues and guidelines. *Measurement and Evaluation in Guidance, 15*(1), 107-116.

Oakland, T. D. (Ed.). (1983). *Nonbiased assessment.* Minneapolis: University of Minnesota School Psychology Inservice Training Network.

Reschly, D. J. (1978). WISC-R factor structures among Anglos, Blacks, Chicanos, and Native-American Navajos. *Journal of Consulting and Clinical Psychology, 46,* 417-422.

Reschly, D. J. (1981). Evaluation of the effects of SOMPA measures on classification of students as mildly mentally retarded. *American Journal of Mental Deficiency, 85*(1), 16-20.

Rugel, P. R. (1974). WISC subtest scores of disabled readers: A review with respect to Bannatyne's recategorization. *Journal of Learning Disabilities, 7,* 57-64.

Schubert, J. (1969). The VRB apparatus: An experimental procedure for the investigation of the development of verbal regulation of behavior. *Journal of Genetic Psychology, 114,* 237-252.

Schubert, J. (1970). The VRB apparatus: A cross-validation. *Psychological Reports, 27,* 571-574.

Schubert, J. (1983). The implications of Luria's theories for cross-cultural research on language and intelligence. In B. Bane (Ed.) *The sociogenesis of language.* (pp. 59-77). New York: Plenum.

Schubert, J., & Cropley, A. J. (1972). Verbal regulation of behavior and IQ in Canadian Indian and white children. *Developmental Psychology, 7,* 295-301.

Silverstein, A. (1977). Alternative factor analysis solution for the Wechsler Intelligence Scale for Children - Revised. *Educational and Psychological Measurement, 37,* 121-124.

Steinberg, R. H. (1974). *Psychometric and operative intelligence in an Indian school population.* Unpublished doctoral dissertation, University of Saskatchewan, Regina Campus. (Ottawa: National Library of Canada, 1975, Microfiche).

Telford, C. W. (1932). Test performance of mixed blood North Dakota Indians. *Journal of Comparative Psychology, 14,* 123-145.

Teeter, A., Moore, C. L., & Peterson, J. D. (1982). WISC-R Verbal and Performance abilities of Native American students referred for school learning problems. *Psychology in the Schools, 19*(1), 39-41.

West, L. W., & MacArthur, R. S. (1964). An evaluation of selected intelligence tests for two samples of Metis and Indian children. *Alberta Journal of Educational Research, 10*(1), 17-27.

Zarske, J. A., & Moore, C. L. (1982). Recategorized WISC-R scores of learning disabled Navajo Indian children. *Psychology in the Schools, 19*(2), 156-159.

Zarske, J. A., Moore, C. L., & Peterson, J. D. (1981). WISC-R factor structures for diagnosed learning disabled Navajo and Papago children. *Psychology in the Schools, 18*(4), 402-407.