DOCUMENT RESUME

ED 052 848 RC 005 332

AUTHOR	Mercer, Jane R.
TITLE	Current Retardation Procedures and the Psychological and Social Implications on the Mexican-American. A Position Paper.
SPONS AGENCY	Southwestern Cooperative Educational Lab., Albuquerque, N. Mex.
PUB DATE	Apr 70
NOTE	40p.
EDRS PRICE	EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS	Intelligence Tests, *Mental Retardation, *Mexican Americans, Special Education, *Student Evaluation, *Student Placement, Surveys, *Testing
IDENTIFIERS	California

ABSTRACT

In view of the disproportionately large numbers of Mexican American children in special education classes, this position paper presents past and current research findings examining 2 hypothetical causes of the disproportion: discriminatory referral procedures and discriminatory clinical procedures. Examining 1 medium-sized California school system, it was determined that the pupil personnel department's clinical testing procedures rather than referral processes produced the ethnic misrepresentation in special classes. Recommendations are that (1) more refined scales need to be developed for assessing the child's adaptive behavior outside the school and (2) pluralistic norms need to be used in interpreting the meaning of both the IQ and the adaptive behavior score for children of Mexican American heritage. Five tables and 16 references are included. (MJB)

CURRENT RETARDATION PROCEDURES AND THE PSYCHOLOGICAL AND SOCIAL IMPLICATIONS ON THE MEXICAN-AMERICAN

A Position Paper

Prepared By

Jane R. Mercer, Ph.D.

Associate Professor, Sociology
University of California, Riverside

For

Southwestern Cooperative Educational Laboratory
Albuquerque, New Mexico
April, 1970

The Problem

Disproportionately large numbers of children of Mexican-American heritage are labeled as mentally retarded by the public schools and placed in special education classes. This phenomenon appears to be true throughout the Southwestern states and in most communities with a sizable Mexican-American population. For example, in the State of California in 1966, there were 86,781 children enrolled in special classes and 23,103 had Spanish surnames, 26.6% of the total. The comparable percentage for 1967 was 26.3%. Since approximately 13% of the children attending public schools in California have Spanish surnames, there were twice as many children with Spanish surnames in special education classes as would be expected from their proportion in the total population of the public schools. The ratio varied from county to county. In 1966, some of the counties with the highest overrepresentation of Spanish surname children in the special classes were Orange (2.48), Ventura (2.3), San Bernardino (2.23), Santa Clara (2.21), Riverside (2.02), Los Angeles (1.92), Fresno (1.96), and Contra Costa (1.65).[1] Although other counties have lower ratios, Spanish surname children are consistently overrepresented.

When these statistics are presented as rates per thousand, the disproportions are more readily interpreted in terms of children. For example, 19.9 Spanish surname children in every thousand are in special education classes in Alameda County compared to 9.1 Anglo children. Comparable percentages for other counties with 12,000 or more Spanish surname children enrolled were 21.8 to 10.7 in Contra Costa County, 21.2 to 7.1 in Fresno County, 25.7 to 12.5 in Kern County, 45.3 to 13.0 in Los Angeles County, 32.1 to 10.9 in Orange County, 44.5 to 13.2 in Riverside County, 51.8 to

15.1 in San Bernardino County, 49.1 to 16.0 in San Diego County, 16.7 to 9.5 in San Francisco County, 25.5 to 8.8 in Santa Clara County, 27.9 to 15.7 in Tulare County, and 32.4 to 9.9 in Ventura County. The rates of placement for Spanish surname children are consistently two to three times higher than Anglo rates.

The purpose of this paper is to present findings from the author's current and past research which relate to the processes and procedures which appear to be producing these disproportions. Two explanatory hypotheses are explored: discriminatory referral procedures and discriminatory clinical procedures.

The Referral Process in the Public Schools

Disproportionately high rates of placement in classes for the mentally retarded of Spanish surname children may result from a referral process in the public schools which discriminates against Spanish surname children. Proportionately more Mexican-American children may be placed in special education classes because teachers and principals may refer them at a proportionately higher rate than they refer Anglo children. To explore this hypothesis, we examined the referral procedures in one medium sized California school system whose procedures are similar to those used in other school systems of its size in the state. Like other districts in the State of California, it, too, had disproportionately large numbers of Mexican-American children enrolled in classes for the mentally retarded. In the year of the referral study, 45.3% of the children placed in special classes for the mentally retarded were of Mexican-American heritage although only 11.0% of the population of the school district consisted

of children from Mexican-American homes.

At the time of the study, the school district had approximately 25,300 students enrolled in 22 elementary schools, 6 junior high schools, and 3 senior high schools. In addition, there was a special school for physically handicapped children. Although the schools varied greatly in the size of their student bodies and the size of their administrative staffs, all schools depended upon the staff of the Pupil Personnel Department located in the central administrative offices of the district for psychological evaluations of children. The two primary responsibilities of the department were to organize and evaluate findings from the extensive group testing program of the district and to administer the individual intelligence tests which are required by state law before a child may be placed in a program for the mentally retarded. This department, directed by a psychologist with the rank of Associate Superintendent, controlled the movement of children into the many special programs offered by the school district. Any child who was removed from the status of "normal" student in the regular classroom to "mentally retarded" student in special education classroom had to have his placement certified and legitimated by one of the three certified psychologists working for the Pupil Personnel Department. They, and only they, could certify a child as mentally retarded and eligible for special education.

In order to study the referral processes in operation, we investigated the characteristics of all 1,234 children who were referred to the Pupil Personnel Department for any reason during a single school year. We traced the steps through which a child must pass in being placed in a class for the mentally retarded to determine at what juncture in the

referral process the disproportions so evident in final placement statistics appeared. Potentially, any child who is referred can be tested by a psychologist and found eligible for placement in a class for the mentally retarded. In actual practice, however, the 66.2% of the children referred by a teacher or principal were more likely to be tested than children referred from other sources. In the year studied, 31.0% of the children were referred as candidates for classes for the gifted, 29.8% were referred exclusively for behavior and disciplinary problems, 19.5% were referred for unspecified academic difficulties with no mention of possible retardation, and 8.3% were referred for evaluation specifically because the teacher-principal team believed they should be placed in a class for the mentally retarded. There was a consistent decline in the rate of referral as grade level increased.

We found that Mexican-American children were underrepresented among the total referrals. Only 6.9% of the children sent to the Pupil Personnel Department for evaluation were Mexican-American.

Because school psychologists did not have enough time to administer individual intelligence tests to every child referred for evaluation, they had to decide which children to test. Any child who is not tested cannot be placed in a class for the retarded. When we examined the characteristics of the children psychologists elected to test, once again we found that Mexican-American children were underrepresented. Only 7.6% of the children administered an individual IQ test during the year of the referral study were of Mexican-American heritage.

Among those children tested, there were 134 who had an IQ of 79 or below and, hence, were eligible for placement in a class for the mentally retarded. Of these children, 32.7% were Mexican-American. Among the 81 children who were recommended for placement from this group, 40.9% were Mexican-American and, among the 71 who were actually placed, 45.3% were Mexican-American. Thus, there were four times more Mexican-American children among those who were eligible for classes for the mentally retarded because of their low IQs than we would expect from their percentage in the population tested. In addition, Mexican-American children with low IQs were somewhat more likely to be placed than the Anglo children with low IQs.

The discontinuities in the referral process are very distinct. Although teacher-principal teams referred Mexican-American children at a lower rate than their percentage in the population and proportionately fewer were given IQ tests by school psychologists, three times as many Mexican-American children appeared among those failing the IQ test as we would expect from their proportion in the population of the school district. Subsequently, this disproportion increased so that four times as many Mexican-American children were placed in special education classes as would be expected from their percentage in the district because proportionately more children with low IQs from Mexican-American backgrounds were recommended for placement and were ultimately placed. It is at the point in the referral process when the IQ test is administered that the sharp ethnic disparities first appeared. The referral process in this district was not discriminatory. Disproportions appeared only in the clinical process of IQ testing.

Evidence from another source supports this conclusion. There were 509 Mexican-American children in the regular classes of the same school district who were administered a Wechsler Intelligence Scale for Children as part of a study conducted for the Public Health Service.[2] We found that 15.3% of these children scored 79 or below and thus would have been eligible for a class for the mentally retarded, if the teacher had referred them for evaluation and they had been tested. Only 1.2% of the 500 Anglo children similarly tested were found to have IQs of 79 or below. This indicates that many Mexican-American children who are potentially eligible for classes for the mentally retarded on the basis of IQ are not referred by teachers, but very few Anglo children who are eligible remain in the regular class. Thus, the preponderance of Mexican-American children in classes for the educable mentally retarded does not appear to be the result of over-referral from teachers and principals but rather results from the diagnostic process itself.

The Clinical Process in the Public Schools

Our evidence points to the clinical process rather than the referral process as the primary factor in disproportionate placement of Mexican-American children in classes for the mentally retarded. The primary instrument used to diagnose mental retardation is the standardized individually administered intelligence test, usually the Stanford-Binet LM or the Wechsler Intelligence Scale for Children. Although the clinical definition for mental retardation adopted by the American Association for Mental Deficiency defines a mental retardate as one who is subnormal in both intelligence and adaptive behavior, there are no recognized scales for

measuring adaptive behavior and this dimension is not systematically evaluated as part of the clinical diagnostic process in the public school. In actual practice, a child's score on an intelligence test is the primary factor in his diagnosis as a mental retardate. If we are to understand the disproportionate placement of Mexican-American children in classes for the mentally retarded, we must focus specifically on the nature of intelligence tests.

The Standardized Intelligence Test

The IQ is ordinarily treated as a measure of an individual's intellectual capacity, his mental ability. Obviously, intellectual capacity cannot be measured directly because that would require assessment of the genetic component of performance, the genotype. An individual's genotype can only be expressed through behavior learned in a social and cultural setting, his phenotype. The IQ test, of necessity, measures what a person has learned, his phenotype. On the basis of this performance, psychologists make inferences about the nature of the person's genotype. The clinician assumes that it is possible to make valid inferences about the genotype from a properly normed and administered intelligence test. Psychologists who make diagnoses of children in the public schools are constantly making inferences about the characteristics of genotypes on the basis of the performance of the phenotype. The logic behind these inferences is relatively simple but the assumptions are rarely if ever met in actual practice.

These inferences are based on the following sets of assumptions. If two persons have had an equal opportunity to learn certain types of cognitive,

linguistic, and mathematical skills and to acquire certain types of information; if they were equally motivated to learn these skills and to acquire this information; if they are equally motivated to exert themselves in a test situation and equally familiar with the demands of the test situation; if they are equally free of emotional disturbances and anxieties which might interfere with their test performance; and if they are equally free of biological dysfunctions and organic difficulties which might interfere with their performance; then any difference between their performance on the test probably measures differences in their genetic intellectual endowment. Simply stated, if learning opportunities and all other factors are equal, those persons who learn the most and perform the best probably have greater innate mental capacity than those who learn and perform most poorly.

Of course, the major difficulty in applying logic based on these assumptions in actual test situations, especially situations which involve children from very different sociocultural backgrounds, is that "all other factors are never equal." Children from Mexican-American homes do not have the same opportunities to learn the types of cognitive, linguistic, and mathematical skills and to inquire the kinds of information covered in the Stanford-Binet LM or the WISC as Anglo children from middle and upper middle class families. Therefore, the assumptions on which the standardized intelligence tests are based are not met for most Mexican-American children and valid inferences about comparative innate mental ability cannot be made on the basis of performance on these tests. When we consider that Mexican-Ameri an children were not even included in the samples on which the norms of the Stanford-Binet and WISC were established,

the inappropriateness of a direct comparison of scores of Anglo and Mexican-American children is even more evident.

In general, there have been three different approaches to the issues raised in cross-cultural intelligence testing: attempts to develop "culture-free" tests; adapting present tests for cross-cultural applications; and developing multiple tests, each adapted for a particular sociocultural group.

Culture-Free Tests

Serious attempts have been made to develop tests which are "culture-free," such as the Davis-Eells Games, Raven Progressive Matrices, and the Cattell Culture-Free Test. Among the major tests used in international studies have been the Goodenough Draw-A-Man Test (Harris, 1963) and the Raven Progressive Matrices while the Peabody Picture Vocabulary Test has been used domestically to test children from various subcultures (Torrance, 1968).

Many writers have discussed the inherent errors and difficulties in developing a culture-free test (Biesheuvel, 1952, Verhaegen T., 1956). Workers who have attempted to develop such tests have made extensive use of apparatus tests. These are used mainly for low-level screening (Schwarz, 1963). Paper-and-Pencil tests present special difficulties in cross-cultural testing, although work done in Africa by Schwarz (1961) indicates that basic adaptations in the number and kinds of test operations and the test format may make such tests feasible.

There are unknown complexities in comparing scores, cross-culturally, on any of these tests. For example, Dennis reports that scores on the Goodenough are positively correlated with modernization in forty different cultures and subcultures which he studied (Dennis, 1966). Irvine (1966), working in African cultures, concluded that the Raven is not an unbiased test simply because it is based on a figural and low verbal content but is biased in different and undefined ways. Cross-cultural efforts have, at best, met with only limited success because an intelligence test can only measure phenotypic behavior, the product of learning in a specific social and cultural setting. Beyond simple reflex acts and basic organic processes, there are few if any human behaviors which are culture-free.

Modification of Existing Tests

One of the most common procedures in cross-cultural testing has been to attempt to modify either the language or administrative procedures when tests standardized in one society are used in making cross-cultural comparisons. Torrance (1968), in his program of international testing for creative development modified existing tests by translating them into the native language of the subjects. Darcy (1963) reviewed numerous studies of the effects of bilingualism on the measurement of intelligence in which Spanish versions of standard intelligence tests were used. The Cattell Culture-Free Test has both English and Spanish versions, as does the Stanford-Binet. Keston and Jiminez (1954) tested fifty Spanish-American children using both the English and Spanish versions of the Stanford-Binet and found the mean IQ on the English version 14 points higher than that on the Spanish version. They concluded that this resulted from the fact that the children's formal education was in English and that the development of their Spanish had essentially stopped when they entered the public schools.

They concluded that the Spanish version of the Stanford-Binet was not appropriate because the quality of Spanish spoken in the area studied varied significantly from that used in the test.

The WISC has been translated and adapted to Puerto Rican Spanish. Norms were established on the basis of a representative sample of Puerto Rican school children. Morán (1962) found that his sample of Puerto Rican children still had a mean IQ of 88 with a standard deviation of 22 even using the Spanish version. His study illustrates the difficulty of securing scores which can be interpreted clinically even when translated forms of standardized tests are used. In other cases alternate wording has been necessary to adjust for culturally different meanings. For example, Coyle (1965) found that the wording of one of the items in the Information Subtest of the WISC had multiple interpretations relating to sexual matters. He proposed a different phrasing which would prevent such misinterpretations in cross-cultural testing.

Thus, the translation approach has limited value. The content of the examination is culture bound. The content of information questions and the concepts used in recognizing conceptual similarities are related to a particular society. The materials used to test comprehension and perceptional skills all reflect the content of the culture in which the test was developed. Simply translating the content of a test designed for persons socialized in one culture into the language of another culture does not eradicate the cultural differences. Persons from backgrounds other than the culture in which the test was developed will always be penalized. It is difficult to interpret the meaning of IQs when this is the case.

Another common adaptation of existing tests is to weight the non-verbal portion of the test more heavily than the verbal portion in

cross-cultural testing. The rationale for this procedure is that performance tests are believed to be more "culture free" than verbal tests which have a high language component. Most investigators have found, for example, that the Performance IQ of Mexican-American children averages 5 to 7 points higher than their Verbal IQ on the WISC. In our research in Riverside, we found that the Performance IQ ranges from 7.6 points higher for Mexican-American children whose family backgrounds most closely approximate the cultural configuration of the Anglo community to 11.6 points higher for children from homes which are not highly acculturated to the Anglo society.

In a study of 235 bilingual Puerto Rican children in New York public schools, Darcy found an 8.3 point difference in favor of the non-language test using the Pintner General Ability Test. She concluded that the language and nonlanguage tests were not measuring the same function to a large enough extent to warrant substituting one for the other. She recommends that both verbal and non-verbal scores should be used in making any diagnosis (Darcy, 1952). Johnson (1953) used the Goodenough Draw-A-Man Test and the Otis Self-Administering Test of Mental Ability in evaluating 30 bilingual Spanish boys. Although he found the mean Goodenough scores were over 12 points higher than the Otis scores he concluded that there are many complex problems in making such comparisons and results are open to question. Studies of Welsh bilinguals have found a similar pattern of higher non-verbal scores among bilingual children. Jones (1952) has found that as "Welshness" of background decreased, mean IQ scores increased, a finding similar to that for Mexican-American children which will be reported later in this paper.

Most investigators report relatively high correlations between verbal and non-verbal scores but it is difficult to know the extent to

which one may be substituted for the other. Factor analytic studies indicate that they are factorially separate dimensions.

A third adaptation of the standardized testing situation which has been used in cross-cultural testing is to vary the speed and power components of the test. The rationale for this procedure is that persons from cultures which do not place a great value on the speed of performance do poorly on tests which are timed. Given adequate time, as in a power test situation, they would do significantly better. Some studies seem to confirm this hypothesis. For example, Knapp (1966) studied 100 Mexican male immigrants using the Cattell Culture-Free Test, Form 2A and 2B. He found that the group averaged higher under power conditions and improved more than a comparable group of Anglos when the speed limits were removed.

However, Schwarz (1963) contends that there are unpredictable complexities when the speed-power factor in test administration is varied. In African samples, persons with little prior knowledge of tests varied greatly in the strategies which they used under the two conditions. This was especially true if the test required both speed and accuracy, a situation especially confusing to the uninitiated. He found that there is "no easy way of encouraging more thoughtful responses by the examinees who use only a fraction of the available time, or of speeding up those for whom exceptionally generous limits are not enough." He concludes that adequate power tests are, in general, the most difficult to devise.

We must conclude that any modifications in standard procedures in order to accommodate for cross-cultural differences is fraught with unknown biases, whether the modification be translation into the language

of the subject, emphasis on non-verbal rather than verbal tests, or altering the procedures in test administration. Such modifications make it difficult to interpret the meaning of scores produced under these varying conditions and leave the evaluator without a normative framework within which he can evaluate the meaning of a particular performance.

A fourth approach which has been proposed but has not been pursued to any great extent, is to develop intelligence tests for each of the major cultural groups in American society. If this were done, each child could be tested using the test designed for persons from his linguistic and cultural background. While this approach is certainly worth further exploration, the creation and standardization of multiple intelligence tests would be expensive. It would be difficult to determine the cultural boundaries for each of the cultural groups for which such tests should be developed. The interpretation of the scores on such culture-specific tests would also pose problems for the educator. While a person's score on a culture-specific test would indicate his relative position in the distribution of scores for other persons of his sociocultural background, it would not indicate his position relative to the larger society or his probability of success in the public schools. If nothing else, present IQ tests are relatively good indicators of academic performance.

Schwarz (1963) provides a detailed discussion of the complexities of developing specific tests adapted to varied cultural settings in Africa. Each test must be backed by adequate environmental supports. The test writer must be cognizant of the extent and nature of the environmental support for each test before it is introduced. Test instructions, whether written, oral, or pictorial, must be adapted to the local vernacular.

For example, he found that when pictures were used which had a referent symbol, African children were often not able to recognize drawings of highly familiar objects for what they are. In general, he suggests that pictures should be avoided. Tests instructions may need to be pantomimed, but this is difficult for more advanced tasks. In summary, he concluded that the entire testing process requires a large number of cultural adaptations when culture-specific tests are devised. Even when such tests are developed, it is difficult to know how performance on them can be generalized beyond the culture for which they were designed. In this case, we have the same problem, in reverse, as we now have with the present tests standardized in American society.

We have been exploring a fifth alternative, pluralistic diagnosis. A pluralistic diagnosis is based on the American Association of Mental Deficiency definition of mental retardation and involves the use of both an IQ test and a measure of adaptive behavior based on pluralistic norms. A person is diagnosed as mentally retarded only when he scores in the lowest 3.0% of his own sociocultural group on both IQ and adaptive behavior.

Pluralistic Evaluation of Intelligence

The logic of the pluralistic evaluation of intelligence is based on a three-step process:

1. Identifying those sociocultural characteristics which account for the greatest amount of the variance in the IQs of Mexican-American children;

2. Developing a sociocultural index using those characteristics most highly correlated with IQ for placing children into categories according to the sociocultural grouping most closely approximating their background characteristics; and

3. Interpreting the meaning of the IQ against two normative standards: the standardized norms of the test as published in the test manuals and the pluralistic norms based on the distribution of scores for persons from comparable sociocultural backgrounds.

Each aspect of this procedure will be discussed in turn.

Identifying Sociocultural Characteristics Correlated With IQ for Mexican-American Children. During the school year 1966-67, 598 Mexican-American, and 576 Anglo elementary school children in the Riverside Unified School District were tested using the Wechsler Intelligence Scale for Children. The team of psychometrists who tested these children were given supplementary training, in addition to the training they had received in regular university

courses in testing, to assure that they followed standardized administration and scoring procedures. They were closely supervised throughout the testing. The Mexican-American children in the sample included the total Mexican-American population of the three segregated minority elementary schools which then existed in the district. The Anglo children were a random selection of elementary school children in 11 predominatly Anglo schools. The parents of all of these children were interviewed, providing extensive information on the family background characteristics of all the children who were studied.

Seventeen of the characteristics which were found to differentiate most highly between Mexican-American families and the modal sociocultural configuration of the community were used as independent variables in a multiple regression analysis with used Full Scale WISC IQ as the dependent variable. In other words, we attempted to determine the amount of the variance in WISC Full Scale IQ which could be accounted for by the socio-cultural characteristics of the child's family. Ten of the variables studied focused on characteristics of the head of household and were dichotomized as follows: born in Mexico vs born in the United States; reared in Mexico vs reared in the United States; reared in a rural vs reared in an urban area; citizen of the United States vs non-citizen; Catholic vs non-Catholic; white-collar occupation vs blue-collar occupation; Duncan Socioeconomic Index Score for occupation 0-29 vs 30 or higher;[3] formal education 8 years or less vs 9 or more; married and living with spouse vs divorced, separated or deceased; male head of household vs female head of household. Six family variables focused on characteristics

of the household: whether the household had moved two or fewer times in the past 10 years vs moved three or more times; owning or buying a house vs renting; speaking English all the time or most of the time in the family vs seldom or never; five or fewer family members vs six or more; family structure consisting of parents and their children vs all other types of family structures; 1.4 persons per room or less vs 1.5 persons per room or more. The 17th variable used in the regression was the number of years of schooling the mother expected the child to complete. These responses were dichotomized, a high school education or less vs more than a high school education.

Insert Table 1 About Here

Table 1 shows the correlation coefficients for the five variables which predicted the largest amount of the variance in Full Scale IQ and for all 17 variables. All 17 variables were correlated .39 with Full Scale IQ, .42 with Verbal IQ, and .30 with Performance IQ. Together they account for approximately 15% of the variance in Full Scale IQ, 18% of the variance in Verbal IQ, and 9% of the variance in the Performance IQ.

We noted earlier that some clinicians use Performance IQ as the primary score in assessing the measured intelligence of the Mexican-American children, because they believe that this score is less influenced by sociocultural factors than the Verbal IQ. Our findings support this belief, but they also show that this expedient does not eliminate cultural factors entirely. The correlation between sociocultural factors and Performance IQ is statistically significant ($p < .01$) and even 9% of the variance cannot be ignored.

Table 1

Regression of Sociocultural Variables on WISC Full Scale, Verbal, and Performance IQs for 598 Mexican-American Children (6-14 years)[a]

Sociocultural Variable	Full Scale R	Verbal R	Performance R
Overcrowding	.24	.26	.18
School expectancy	.32	.33	.24
Education of head	.34	.37	-
Frequency English spoken in home	.36	.39	.25
Own home vs renting	.37	.40	.27
Head reared, rural vs urban	-	-	.27
All 17 variables	.39	.42	.30
Variance accounted for	15.2%	17.6%	9.0%

[a] All correlations are significant beyond the .01 level.

Overcrowding is the primary indicator, having a linear correlation with Full Scale IQ of .24. The next variable in order of its contribution to the multiple correlation coefficient is the number of years of schooling the mother expects the child to complete. It is correlated .23 with the Full Scale IQ and adds significantly to the multiple correlation coefficient, raising it from .24 to .32. Education of the head of household, use of English in the home, whether the family owns or is buying its home, and whether the head of household was reared in an urban or rural area appear in that order.

In brief, Mexican-American elementary school children with higher IQs tend to come from less crowded homes and have mothers who expect them to have some education beyond high school. They have fathers who were reared in an urban environment (over 10,000 population) and have a ninth grade education or more. They live in a family which speaks English all or most of the time and is buying its home. Thus, the more the family is like the modal sociocultural configuration of the community, the higher the IQ of the Mexican-American child on the WISC. Clearly, sociocultural factors can not be ignored in interpreting the meaning of a standardized intelligence test when evaluating a child from a Mexican-American home.

Developing a Sociocultural Index for Classifying Children by Family Background

The findings from the multiple regression were used to group each Mexican-American elementary school child who was given the WISC into one of five groups according to the extent to which their family background conformed to the modal configuration for the Anglo community. Each child was given one point for each of his family background characteristics which

were like the dominant society on the five primary sociocultural variables predicting Full Scale IQ. For example, a child received one point for each of the following: if he came from a family with 1.4 or fewer persons per room; if his mother expected him to go to school beyond high school; if the head of household had a ninth grade or higher education; if his family speaks English all or most of the time at home; and if his family owns or is buying its home. If his family was similar to the modal configuration on all five characteristics, he received a score of 5. If his background was similar to the dominant configuration on four characteristics, he received a score of 4, and so forth.

There were 25 Mexican-American children who came from backgrounds that included all five characteristics, only 4.2% of the total group. When Anglo children were scored, using the same variables and scoring procedures, 68.7% scored 5 points. The large difference between the percentage of Mexican-American and Anglo children who came from family backgrounds correlated with high performance on the WISC emphasizes that relatively few of the Mexican-American children in Riverside come from sociocultural backgrounds which correlate with high performance on the WISC, a test developed and standardized on the general population of the United States. However, the mean IQ for the group of 25 children whose families were most like the dominant cultural configuration, was slightly above the norms for the test--mean Full Scale IQ was 104.4; mean Verbal IQ was 100.6; and mean Performance IQ was 108.2. Those Mexican-American children with cultural backgrounds most similar to those of Anglo children did as well on the WISC as the Anglo children on whom the tests were normed.

The 174 children with scores of 4 had a mean Full Scale IQ of 95.5,

a mean Verbal IQ of 91.8, and a mean Performance IQ of 100.5. All scores were approximately nine points lower than those for the group most like the dominant society, a half standard deviation difference.

Insert Table 2 About Here

The 126 Mexican-American children with three sociocultural characteristics similar to the modal configuration for the community had a mean Full Scale IQ of 89.0; the 146 with two such characteristics had a mean Full Scale IQ of 88.1; the 127 with zero or one such characteristic had a mean Full Scale IQ of 84.5. The group least like the modal sociocultural configuration for the community is one standard deviation below the Full Scale norm for the general population. The drop of 20 points in Verbal IQ is even more precipitous, 1.3 standard deviations. In other words, about half the 127 children coming from backgrounds least like the dominant Anglo society would qualify as retardates if only an IQ test were used for diagnosis. In fact, several of these children had been labeled as mentally retarded and were in special education classes at the time of our study.

Interpreting the Meaning of the IQ Against Two Normative Standards

A pluralistic diagnosis evaluates each Mexican-American child in terms of the sociocultural group to which he belongs. For example, the child who has only one sociocultural characteristic that is like the dominant configuration for the community would fall into the group we have called 0-1 in Table 2. For him an IQ of plus or minus one standard deviation, 11.3 points, from 84.5 is within the normal range for his sociocultural group using standard statistical definitions. An IQ of 73.2 to 95.3 is "normal"

Table 2

Sociocultural Groupings and IQ for Mexican-American Children

Most like modal configuration----------Least like modal configuration

WISC IQ	5 (N=25) (4.2%)	4 (N=174) (29.1%)	3 (N=126) (21.1%)	2 (N=146) (24.4%)	0-1 (N=127) (21.2%)	All Groups (N=598)
Full Scale						
X	104.4	95.5	89.0	88.1	84.5	90.4
S	10.4	12.1	11.8	11.6	11.3	12.7
Verbal						
X	100.6	91.8	85.0	83.9	80.3	86.3
S	11.3	12.7	10.8	10.9	11.1	12.5
Performance						
X	108.2	100.5	95.5	95.0	91.9	96.6
S	12.1	12.2	13.6	13.0	12.9	13.5

for a child from his sociocultural background. An IQ between 61.9 and 73.2 would be "low normal" but not clinically retarded, if we regard only the lowest 3% of the population as intellectually subnormal. At the other end of the scale, an IQ between 95.3 and 106.6 would indicate "high normal" ability. An IQ above 106.6 would indicate very superior performance for a child from such a non-Anglo cultural background. Only an IQ below 61.9, the lowest 3% for his group, would be diagnosed as clinical retardation.

A culturally aware pluralistic diagnosis would thus evaluate the intelligence of each person only in relation to others who have come from similar sociocultural backgrounds and who have had approximately the same opportunity to acquire the knowledge and skills needed to answer questions on an IQ test designed for an Anglo-American society. If the person scores more than one standard deviation above the mean for his group, then he probably has "high normal" ability, even if his actual IQ is 100--average by the standard norms of the test. Conversely, a Mexican-American child who manages to achieve a score of 75 on an IQ test when he comes from an overcrowded, Spanish-speaking home in which the father has less than an eighth grade education and was reared in a rural area, and his mother does not expect him to go beyond high school, is well within the normal range for persons, like himself, who have had little exposure to cultural materials needed to pass an Anglo IQ test. His educational program can be planned on the assumption that he is a person with normal learning ability who needs special help with English as a second language and emphasis on socialization to the ways of the dominant society. His educational opportunities are greatly reduced if his education is planned on the

assumption that he is mentally deficient and is not capable of acquiring any facility in intellectual matters. Table 3 presents the range of IQs which would be regarded as subnormal, low normal, normal, high normal, and supranormal for Mexican-American children in the five sociocultural groupings.

Insert Table 3 About Here

Table 3

Pluralistic Norms for Mexican-American Children

on WISC IQ Test

Sociocultural Groupings	IQ	Subnormal -2 or more	Low Normal -1 to -2	Normal -1 to +1	High Normal +1 to +2	Supranormal 2 +
0-1	Full-Scale IQ	-61.8	61.9-73.1	73.2-95.7	95.8-107.0	107.1+
	Verbal IQ	-58.0	58.1-69.1	69.2-91.3	91.4-102.4	102.5+
	Performance IQ	-66.0	66.1-78.9	79.0-104.7	104.8-117.6	117.7+
2	Full-Scale IQ	-64.8	64.9-76.4	76.5-99.6	99.7-111.2	111.3+
	Verbal IQ	-62.0	62.1-72.9	73.0-94.7	94.8-105.6	105.7+
	Performance IQ	-68.9	69.0-81.9	82.0-107.9	108.0-120.9	121.0+
3	Full-Scale IQ	-65.3	65.4-77.1	77.2-100.7	100.8-112.5	112.6+
	Verbal IQ	-63.3	63.4-74.1	74.2-95.7	95.8-106.5	106.6+
	Performance IQ	-68.2	68.3-81.8	81.9-109.0	109.1-122.6	122.7+
			Standard Norms			
4	Full-Scale IQ	-69.9	70.0-84.9	85.0-114.9	115-129.9	130+
	Verbal IQ	"	"	"	"	"
	Performance IQ	"	"	"	"	"
5	Full-Scale IQ	"	"	"	"	"
	Verbal IQ	"	"	"	"	"
	Performance IQ	"	"	"	"	"

Pluralistic Evaluation of Adaptive Behavior

Although adaptive behavior is the most ancient basis for judging individual competence, it has never been measured systematically. Empirical referents for a "socially adequate person" are largely unspecified and clinical judgments depend primarily on unstandardized, intuitive evaluations. Lack of a standardized measure of adaptive behavior is one of the primary criticisms of the AAMD classification system (Sheerenberger, 1965). This lack means that the dimension has not been extensively used in clinical diagnosis.

As part of an epidemiological study of mental retardation conducted as a field survey in the city of Riverside, a series of scales were developed in an attempt to screen the population for adaptive behavior. The American Association of Mental Deficiency two dimensional definition of mental retardation was employed, i.e., subnormality on both intelligence and adaptive behavior.

Although Vineland Social Maturity Scale and Gesell Developmental Scales are widely used in clinical practice to rate infants and young children, neither of these scales could be incorporated directly into field screening procedures because they are designed for use in a clinical setting. However, the content of those scales was a valuable source of material for developing many of the standard questions finally used in the adaptive behavior scales of the field epidemiology. For older children and adults, none of the existing measures were sufficiently comprehensive for our purposes. Therefore, we proceeded to develop additional items covering other facets of adaptive behavior and pretested

the procedure. Essentially, the adaptive behavior acale for older children and adults assesses the number and complexity of social roles which the individual plays. The larger the number and the more complex the social roles, the higher the adaptive behavior.

There are 25 scales graded by the age level of the person to be screened. The scales are 7-9 months, 10-12 months, 13-15 months, 15-18 months, 19-21 months, 22-24 months, 25-29 months, 30-35 months, 36-47 months, 48-59 months, 60-71 months, 6 years, 7 years, 8 years, 9 years, 10 years, 11 years, 12 years, 13 years, 14 years, 15 years, 16-19 years, 20-29 years, 30-39 years and 40-49 years.

The adaptive behavior questions appropriate for the age of the person being screened were asked in the interview of a parent or an adult to whom the individual was related. An individual's score on the scales consists of the number of failing responses.

The total population of the community was divided into three age groups. The lowest 3% of each age group was the general cutoff for subnormality on the adaptive behavior scales. The population of the community was then divided into Anglos and Mexican-American, and each ethnic group into the same three age groups. Table 4 presents the norms for the entire community without respect to ethnic group and also presents the pluralistic norms on the Adaptive Behavior Scales for the Anglo and Mexican-American populations separately.

Insert Table 4 About Here

Table 4

Pluralistic Norms for the Adaptive Behavior Scales

Ethnic Group	Age	Subnormal (Lowest 3%)	Low Normal (Below 16% - Above 3%)
Anglo	7 months-5 years	7+ failures	5-6 failures
	6-15 years	9+ failures	6-8 failures
	16-49 years	8+ failures	5-7 failures
Mexican-American	7 months-5 years	8+ failures	5-7 failures
	6-15 years	11+ failures	6-10 failures
	16-49 years	14+ failures	9-13 failures

General Norms for the Adaptive Behavior Scales

All Ethnic Groups	7 months-5 years	7+ failures	5-6 failures
	6-15 years	9+ failures	6-8 failures
	16-49 years	8+ failures	5-7 failures

When the community-wide norms were used, we found that adults who failed the Adaptive Behavior Scales are more likely to be from lower socioeconomic levels and from ethnic minority groups; however, there were no socioeconomic or ethnic differences for children. In a multiple regression analysis, we found that socioeconomic status plus physical disability plus ethnic group account for only 4.1% of the variance in adaptive behavior.

On the other hand, both children and adults from lower socioeconomic status and from ethnic minority groups have significantly lower IQs than persons from higher statuses and persons from the Anglo majority. Ethnic group plus socioeconomic level account for 30.3% of the variance in IQ. Therefore, it was concluded that adaptive behavior is not so highly related to sociocultural background in the total population as is IQ.

Evaluation Using Community-Wide Norms

When IQ and adaptive behavior were combined to identify those in the lowest 3% on both measures, using community-wide rather than pluralistic norms, we found that the clinically retarded by these definitions are significantly more likely to come from sociocultural environments which are least like the modal configuration for the community: they are more likely to have parents with significantly less education; they are more likely to come from ethnic minority groups; they come from significantly larger families; they come from homes in which English is less likely to be spoken all the time; they come from lower socioeconomic levels; and so forth.

When adaptive behavior was evaluated as well as IQ, 40% to 60% of the Mexican-Americans who failed the IQ test passed adaptive behavior. Clearly, the IQ test and the psychometric situation do not adequately measure the ability of persons from non-Anglo backgrounds in coping intelligently with the complexities of living but focus on a narrow band of verbal and cognitive skills.

The "quasi-retarded," i.e., those who failed IQ but passed adaptive behavior, were children who were performing their school roles acceptably and were not having any difficulties in their family or community roles. Most of the adults in this category had graduated from high school, over half had white-collar jobs, and most were performing their family and community roles in a fashion similar to other adults in the community. We concluded that the present one-dimensional diagnosis used by clinicians in which only IQ is evaluated may be relatively effective for Anglos. All the Anglos who failed had IQ below 70 also failed adaptive behavior at the 3% level. However, a one-dimensional diagnosis is not equitable for persons from non-Anglo backgrounds.

What happens to prevalence rates for clinical retardation if pluralistic, culturally-aware screening procedures are systematically applied? We reanalyzed the field survey data using pluralistic procedures and secured the rates which are presented in Table 5.

Insert Table 5 About Here

Nothing happens to the Anglo rates. They are 4.4 per 1,000 regardless of whether adaptive behavior is evaluated or not. Since the standard norms

Table 5

Rates per 1,000 of Clinical Retardation by Ethnic Group Using Successively More Pluralistic Evaluative Procedures and the Traditional 3% Criterion

Screening Procedure Used at Traditional Criterion (lowest 3%)	Mexican-American	Anglo	Total Population Rate
One-Dimensional Definition-IQ Only Using Standard Test Norms	149.9	4.4	21.4
Two-Dimensional Definition-IQ and Adaptive Behavior Using Standard Norms	60.0	4.4	9.7
Two-Dimensional Definition-IQ Using Standard Norms--Adaptive Behavior Using Pluralistic Norms	30.4	4.4	6.8
Two-Dimensional Definition-IQ Using Pluralistic Norms--Adaptive Behavior Using Pluralistic Norms	15.3	4.4	5.4

are Anglo norms, there is no change in norms for either the adaptive behavior scales or the intelligence test when pluralistic procedures are used. Anglos are those for whom the diagnostic procedures were designed and there is nothing in our epidemiological findings which would suggest the need for significant modification in screening procedures for Anglos.

Mexican-American rates, however, are greatly reduced when pluralistic norms are applied. When adaptive behavior is evaluated as well as IQ, rates drop from 149.9 per 1,000 to 60.0 per 1,000. When pluralistic adaptive behavior norms are applied but standard IQ norms are maintained, the rate drops to 30.4. When pluralistic norms are used to interpret both adaptive behavior and IQ, the rate is 15.3. This is still slightly higher than the Anglo rates but the difference between 4.1 per 1,000 (.41%) and 15.3 per 1,000 (1.53%) is relatively small compared to the large differentials in rates which are found when the entire population is treated as a group. When screening procedures are modified to interpret adaptive behavior scores and IQs in relation to the sociocultural background of the person being evaluated, gross differences between rates of clinical retardation in the three ethnic groups disappear. Approximately the same proportion of persons in each population are screened as clinically retarded when the lowest 3% are regarded as subnormal--between .5% and 1.5%.

A Pluralistic Clinical Evaluation

A pluralistic evaluation defines a person as clinically retarded only if he scores in the lowest 3% of his own sociocultural group on both IQ and adaptive behavior. In this case, he is regarded as mentally

retarded.

Specifically, in a pluralistic IQ evaluation, the WISC is administered and scored in the usual fashion. The Full Scale, Verbal, and Performance IQs are calculated using procedures published in the manuals.

Questions are asked in an interview with the parents to determine the extent to which the sociocultural background of the child conforms to the modal sociocultural configuration of the community. Each child is then assigned to the sociocultural group to which he belongs on the basis of family characteristics. The IQ of the child is interpreted according to the norms for his sociocultural group as shown in Table 2.

Discrepancies between the child's placement on the two distributions are interpreted diagnostically as indicative of the type of special education program that he will need to achieve in the public school. If he scores in the lowest 3% on both adaptive behavior and intelligence for his sociocultural group, then he would be diagnosed as mentally retarded and would be placed in a class for the mentally retarded. However, if he scores within the normal range for his sociocultural group on either adaptive behavior or IQ, he would not be diagnosed as mentally retarded even though his score might be below the norm for the larger community. Instead, the child would be regarded as coming from a different socio-cultural background which has not adequately prepared him for achievement in the dominant society and he would be placed in programs specifically designed to provide for his educational needs. He would be treated as a child of normal intellectual potential who needs special programs to prepare him for participation in the dominant society.

Conclusion

In the series of studies briefly summarized in this position paper, we have shown that the disproportionate number of children of Mexican-American heritage in special education classes for the mentally retarded in the State of California probably does not result from discriminatory practices in teacher referral. In the school district studied, children of all socio-economic levels and ethnic groups were referred to the Pupil Personnel Department and were tested at approximately the rate expected for their proportion in the population of the school district. Ethnic biases were introduced when a standardized intelligence test was administered. At this point, a larger percentage of children from Mexican-American backgrounds scored below 80 than children from Anglo background. Thus, we concluded that it is the clinical procedures rather than the referral process which is producing the overrepresentation of Mexican-American children in classes for the mentally retarded.

The usual procedure in the public school is to make the diagnosis for mental retardation and place children in a class for the mentally retarded primarily on the basis of an intelligence test, even though the official definition for mental retardation advocated by the American Association for Mental Deficiency is that a person being subnormal in both intelligence and adaptive behavior before being diagnosed as retarded.

In the field epidemiology of mental retardation in which both IQ and adaptive behavior were systematically measured, 60% of the Mexican-American population who had IQs below 70 passed the adaptive behavior measure at the 3% level. This indicates that the inclusion of some measure of adaptive

behavior in a clinical evaluation is likely to identify persons who are coping intelligently with problems posed by their social environment even though these persons may not do well on a standard tests of intelligence which are based on materials selected from the dominant Anglo American culture and normed on an Anglo American population. In a two-dimensional definition of retardation, persons who pass the adaptive behavior skills would not be diagnosed as mentally retarded.

If, in addition to the two-dimensional definition, pluralistic norms are used for interpreting the meaning of both the IQ and adaptive behavior scores, then the rate of clinically defined mental retardation among the Mexican-American population is reduced so it is approximately the same as the rate for the Anglo population. By distinguishing children who are low performers primarily because of cultural differences from those children who are probably mentally retarded, a pluralistic approach would make it possible to plan an educational program better adapted to the needs of these children.[4]

On the basis of these findings, there appear to be two avenues which could usefully be explored in relation to clinical practices and procedures in screening for mental retardation in the public schools. First, more refined scales need to be developed for assessing the child's adaptive behavior outside the school. Second, pluralistic norms need to be used in interpreting the meaning of both the IQ and the adaptive behavior score for children of Mexican-American heritage.

Footnotes

[1] California State Department of Education. Racial and ethnic survey of California public schools. Part I: Distribution of Pupils, Fall 1966. Sacramento, 1967.

[2] Data in this report has been collected under the auspices of the following grants: Public Health Service Research Grant No. MH-08667, Socio-Behavioral Study Center in Mental Retardation, Pacific State Hospital, Pomona, California; Public Health Service Grant No. PH43-67-756; McAteer Grant No. M8-14A and M9-14 from the California State Department of Education, Office of Compensatory Education.

[3] Reiss, Albert J. Occupations and social status, The Free Press of Glencoe, Inc. 1961.

[4] We are currently re-evaluating all children in special education classes in the Riverside and Alvord School Districts using pluralistic procedures. Reports of our findings will be forthcoming.

References

Biesheuvel, S. The Study of African Ability. African studies, Vol. XI (1952), pp. 105-117.

Coyle, F. A. Another Alternate Wording on The WISC. Psychological reports, Vol. 16, No. 3, Part 2 (1965), p. 1276.

Darcy, N. T. Performance of Bilingual Puerto Rican Children on Verbal and on Non-Language Tests of Intelligence. Journal of educational research, Vol. 45 (1952), pp. 499-506.

Darcy, Natalie P. Bilingualism and The Measurement of Intelligence: Review of a Decade of Research. Journal of genetic psychology, Vol. 103 (1963), pp. 259-282.

Dennis, Wayne. Goodenough Scores, Art Experience, and Modernization. Journal of social psychology, Vol. 68 (April 1966), pp. 211-228.

Harris, Dale B. Children's drawings as measures of intellectual maturity New York: Harcourt, Brace, and World, 1963, p. 367.

Irvine, S. H. Toward A Rationale For Testing Attainments and Abilities in Africa. British journal of educational psychology, Vol. 36 (February 1966), pp. 24-32.

Johnson, G. B., Jr. Bilingualism as Measured by a Reaction-Time Technique and The Relationship Between a Language and a Non-Language Intelligence Quotient. Journal of genetic psychology, Vol. 82 (1953), pp. 3-9.

Jones, W. R. The Language Handicap of Welsh-Speaking Children: A Study of Their Performance In an English Verbal Intelligence Test in Relation to Their Non-Verbal Ability and Their Reading Ability in English. British journal of educational psychology, Vol. 22, 1952, pp. 114-123.

Keston, J. J., Jiminez, C. K. A Study of the Performance on English and Spanish Editions of the Stanford-Binet Intelligence Test by Spanish-American Children. Journal of genetic psychology, Vol. 85 (1954), pp. 263-269.

Knapp, R. R. Effects of Time Limits on the Intelligence Test Performance of Mexican and American Subjects. Journal of educational psychology, Vol. 51, 1960, pp. 14-20.

Morán, Roberto E. Observations and Recommendations on the Puerto Rican Version of the Wechsler Intelligence Scale for Children. Pedagogia, Rio Piedras, Vol. 10 (1962), pp. 89-98.

Schwarz, Paul A. Adapting Tests to the Cultural Setting, Educational and psychological measurement, Vol. XXIII, No. 4, 1963, pp. 673-686.

Schwarz, Paul A. Aptitude Tests for Use in the Developing Nations. Pittsburgh: American Institute for Research, 1961.

Torrance, E. Paul. Testing the Educational and Psychological Development of Students from Other Cultures and Subcultures. Review of educational research, Vol. XXXVIII, (February 1968), pp. 71-76.

VerNaegen, P. Utilite Actuelle des Tests pour l'Etude Psychologique des Autochtones Congolais, Review de psychologie appliqueem Vol. VI (1956), pp. 139-151.