DOCUMENT RESUME

ED 312 811 EC 221 269

AUTHOR Rueda, Robert; And Others
TITLE Performance of Hispanic Educable Mentally Retarded, Learning Disabled, and Nonclassified Students on the WISC-RM, SOMPA, and S-KABC: Short-Term Study One. Final Report.
INSTITUTION Southwest Regional Laboratory for Educational Research and Development, Los Alamitos, Calif.
SPONS AGENCY Special Education Programs (ED/OSERS), Washington, DC.
PUB DATE 85
CONTRACT 300-83-0273
NOTE 184p.; Paper presented at the Annual Convention of the Council for Exceptional Children (67th, San Francisco, CA, April 3-7, 1989). Product of Handicapped-Minority Research Institute.
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143) -- Guides - Non-Classroom Use (055)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Classification; *Cognitive Measurement; Educational Diagnosis; Elementary Education; Evaluation Methods; *Hispanic Americans; *Learning Disabilities; *Limited English Speaking; *Mild Mental Retardation; Performance; *Student Evaluation; Testing
IDENTIFIERS California; Kaufman Assessment Battery for Children (Spanish); System of Multicultural Pluralistic Assessment; Wechsler Intelligence Scale for Child Rev Mexicano

ABSTRACT

The study examined performance of limited-English proficient Hispanic students on a battery of psychometric instruments designed to appropriately assess linguistic minority students. Subjects consisted of three groups: 44 nonhandicapped, 45 learning-disabled, and 39 mildly mentally retarded elementary-level students. Instruments included the Wechsler Intelligence Scale for Children Revised, Mexicano; Kaufman Assessment Battery for Children--Spanish Edition; and Physical Dexterity Tasks and Bender Visual Motor Gestalt Test of the System of Multicultural Pluralistic Assessment. Results indicated that the psychometric properties of the instruments were within acceptable ranges for use with students such as those in the sample. In addition, in general, the patterns of scores on the instruments were in the expected directions, given the diagnostic classifications assigned to the students in the school setting. A second part of the analysis utilized California state eligibility criteria to calculate the numbers of students in the school-assigned diagnostic groups who met the state requirements for these designations. Results indicated large discrepancies between the school-assigned categories and the categories suggested through the use of the study instruments in conjunction with applicable state guidelines. Appendices include a manual for the determination of a severe discrepancy between intellectual ability and achievement as defined by California regulations. (Author/JDD)

FINAL REPORT—SHORT-TERM STUDY ONE

Performance of Hispanic Educable Mentally Retarded, Learning Disabled, and Nonclassified Students on the WISC-RM, SOMPA, and S-KABC

Robert Rueda, Richard Figueroa,
Pauline Mercado, and Desdemona Cardoza

Submitted to the U.S. Department of Education
Office of Special Education Programs

Contract No. 300-83-0273

Handicapped-Minority Research Institute
at the
SWRL Educational Research & Development
4665 Lampson Avenue
Los Alamitos, CA 90720
(213) 598-0481

This document was produced by the Handicapped Minority Research at the Southwest Regional Laboratory. This project has been funded through federal funds from the U.S. Department of Education under contract number 300-83-0273. The contents of this publication do not necessarily reflect the views or policies of the U.S. Department of Education, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

FINAL REPORT—SHORT-TERM STUDY ONE

Performance of Hispanic Educable Mentally Retarded, Learning Disabled, and Nonclassified Students on the WISC-RM, SOMPA, and S-KABC

Robert Rueda, Richard Figueroa,
Pauline Mercado, and Desdemona Cardoza

Submitted to the U.S. Department of Education
Office of Special Education Programs

Contract No. 300-83-0273

Handicapped-Minority Research Institute
at the
SWRL Educational Research & Development
4665 Lampson Avenue
Los Alamitos, CA 90720
(213) 598-0481

This document was produced by the Handicapped Minority Research at the Southwest Regional Laboratory. This project has been funded through federal funds from the U.S. Department of Education under contract number 300-83-0273. The contents of this publication do not necessarily reflect the views or policies of the U.S. Department of Education, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

HANDICAPPED-MINORITY RESEARCH INSTITUTE

Senior Staff

Victor E. Rodriguez, Ph.D.
Director

Robert S. Rueda, Ph.D.
Research Coordinator

Desdemona Cardoza, Ph.D.
Linda Carpenter, M.S.
Vicki Jax, M.A.

Pauline Mercado, Ph.D.
Lois A. Weinberg, Ph.D.
Richard Mendoza, Ph.D.

Trainees

Atala Castro
Alberto Chueca y Mora
Margaret Glazer
Marquita Grenot-Scheyer
Kathy Hayes-Latimer
Su Hsu
Michelle Huguet
Georgia Malcolm
Sharon Menke
Inez Salman
Sister Phyllis Supancheck

Research Assistants

Sandy Clegg
Margarita Collado
Karen Hecht
Anthony C. Hernandez
Bridget Kent
Kim Payne
Susan Purviance
Dolores Segura
Suzanne Shipley
Gloria Sosa
Dora Ruiz Twigg

Consultants

Richard Figueroa, Pn.D.
University of California,
Davis

Jane Mercer, Ph.D.
University of California,
Riverside

University Coordinators

Robert Berdan, Ph.D.
California State University,
Long Beach

Kathy Harris, Ph.D.
California State University,
Los Angeles

Barbara Hecht, Ph.D.
University of California,
Los Angeles

Reynaldo Macias, Ph.D.
University of Southern
California

Virginia Kennedy, Ph.D.
University of California,
Los Angeles

Advisory Panel

Dr. Maria Estela Brisk
Boston University

Dr. Elva Duran
University of Texas, El Paso

Mr. Amador Mojica
Special Education Teacher
Hartford, CT

Dr. Eligio Padilla
University of New Mexico

Ms. Martha Chavez
Denver Association for
Retarded Citizens

Ms. Amalia Guerrero
Community Advocate for
Developmentally Disabled

Maria Vasquez-Chairez
State Dept. of Education
Sacramento, CA

Support Staff

Lydia Jewell
Lydia Villarreal

Abstract

The purpose of this study was to examine performance of a sample of limited English proficient Hispanic students on a battery of psychometric instruments designed to appropriately assess linguistic minority students. The instruments included the WISC-RM, tne Spanish KABC, and the SOMPA assessment system measures. The psychometric qualities of the intsturments, such as the validity, reliability and standard error of measurement, were examined for the various subgroups in the study.

The students in the sample consisted of three groups based on the diagnostic categories assigned by the schools which the students were attending. These three groups included a nonhandicapped group (N = 44), a group of learning disabled students (N = 45), and a group of mildly mentally retarded (ER) students (N = 39).

The results indicated that the psychometric properties of the instruments, within the limits of the coefficients calculated in this study, were within acceptable ranges for use with students such as those in the present sample. In addition, in general, the patterns of scores on the instruments were in the expected directions given the diagnostic classifications assigned to the students in the school setting.

A second part of the analysis utilized the State of California eligibility criteria (for both learning disability and mild mental retardation) to calculate the numbers of students in the school assigned diagnostic groups who met the requirements for these designations. The results of this procedure indicated large discrepancies between the school assigned categories and the categories

suggested through the use of the instruments of the study in conjunction with applicable state guidelines. These results were discussed with respect to the issues of non-biased assessment and the testing of linguistic minority students.

Table of Contents

Table of Contents (continued)

FINAL REPORT--SHORT TERM STUDY 1

Performance of Hispanic Educable Mentally Retarded, Learning Disabled, and Nonclassified Students on the WISC-RM, SOMPA, and S-KABC

I. Introduction

Misdiagnosis and Hispanic Students

Students whose language background is not English have always had a difficult time in U.S. public schools. It is a particular phenomenon of this century, however, that their linguistic problems have been perceived either as a handicapping condition or one that can be addressed in classes for the handicapped.

In 1902, the New Haven schools segregated "incorregible boys, defective children, and children who speak no English" (Connecticut Special Education Association, 1936, p. 23).

In 1916, H. A. Miller (1916) reported the following about the Cleveland public schools:

> At the present time such cases are often handled in a most unsatisfactory manner. The non-English speaking child cannot keep up with his companions in the regular grades. For this reason he is sent to a special class, but if there is not a class available, the pupil is all too frequently assigned to the backward class. This is not because the backward class is the right place for him, but rather because it furnished an easy means of disposing of a pupil, who, through no fault of his own, is an unsatisfactory member of a regular grade (p. 74).

In 1933, Reynold's (1933) report to the United States Department of the Interior on "The Education of Spanish-speaking children in five southwestern states" noted that in Los Angeles:

A number of Mexican children are in development rooms which handle children who are for the most part below 65 IQ. In the words of a member of the research division staff, "The proportion of Mexican pupils in development rooms is probably somewhat higher than is their relative number in the general [pupil] population" (p. 51).

Interestingly, at about this same time, the role of psychometric tests in the overrepresentaton/misplacement of Hispanic students in special classes came to be scrutinized. Researchers began questioning the validity English language tests on linguistically different populations (Sanchez, 1932; 1934). In other instances, school officials went beyond questioning the use of tests:

> Results of testing programs administered to Spanish-speaking pupils have impressed a certain number of school administrators as so unreliable that they have lost faith in the use of tests for the members of this group.
>
> The following citation shows the attitude of one investigator:
>
> "No attempt was made to give tests to the pupils in the Lexington School, who are all foreign, as it has been quite conclusively shown on various occasions that valid results cannot be obtained with the ordinary tests. A specially designed test must be used for testing the intelligence of non-English-speaking children and, no such test was available, it was deemed unwise to attempt to use the standardized tests, prepared and standardized for American children" (Reynolds, 1933, p. 46).

By the 1950's, the issue of testing "bilinguals" had accumulated a substantial amount of literature (Darcy, 1963; 1952) and had come to

acknowledge the adverse impact of a language background other-than-English on psychometric test scores. Also, research on the measurement of intelligence of "bilinguals" (Darcy, 1963) had arrived at several conclusions relative to the complexity of "bilingualism" and its impact on IQ scores.

It was during the 1960's, however, that the most comprehensive study of the impact of IQ test scores on Mexican American pupils was conducted in Riverside, California (Mercer, 1973). In a classic epidemeological study on mental retardation, Mercer discovered: that Mexican American pupils were excessively overrepresented in classes for the mentally retarded, that their overrepresentation could be directly traced to language-background characteristics and low IQ scores, and that this form of "mental retardation" was unique to the school and existed nowhere else in the pupils' environment.

In 1969, the issue found its way to the U.S. Ninth Circuit Court in San Francisco. A small group of Mexican American pupils sued the state of California alleging that they had been erroneously diagnosed as educable mentally retarded (ER) because they had been tested on an IQ test that was both linguistically and culturally invalid. Diana vs. California State Board of Education (1969) was settled out of court when the state acknowledged that its ER classes were significantly overrepresented with Mexican American pupils. The out-of-court settlement is unique both in terms of testing and the law. It provides: for testing in English and in the primary language, for developing a test of Mexican American intelligence, and for monitoring the statewide ER representation of Mexican American pupils.

Regrettably, the Diana case has not succeeded in resolving the problems of misdiagnosing LEP pupils. Data exists showing that the nearly one hundred year-old problem persists in California (Anderson, Gallego, Twomey, Williamson, & Williamson, 1980; Figueroa, 1985) and throughout the United States (Heller, Holtzman, & Messick, 1982). The National Academy of Sciences' study, in fact, underscores several national trends associated with the problem of Hispanic misdiagnosis.

> "In summary, the apparently similar ER placement rates for Hispanic and nonminority students disguise variation in practices among school districts. There are a number of districts in which Hispanic students are assigned to ER programs in large proportions. They are distinguished from other districts by having small enrollments that are often--but not always--largely Hispanic; furthermore, they have small black enrollments, small or nonexistent bilingual programs, and high percentages of Hispanic students in SLD classes as well. Among large districts with the greatest pool of resources, low ER disproportion and low SLD disproportion occur where many Hispanic students participate in bilingual programs" (Heller, Holtzman, & Messick, 1982, p. 374).

Yet, inspite of the intractable nature of this issue of misdiagnosis, progress has been made in understanding the complexity and the dimensions of the problem. The next three sections explore these. The one concluding comment that can be made here is that up until recently (when the Spanish-language tests were developed and published), distinguishing between LEP, ER/LD Hispanic students and LEP, underachieving Hispanic students has been either an extremely easy (and invalid) or an exceedingly difficult (and unreliable) process.

Bias

During the last three decades, few topics in American psychology have produced as much controversy and literature as the subject of test bias (Cronbach, 1975; Flaugher, 1978; Hunter, & Schmidt, 1974). Clearly, given the phenomena of minority overrepresentation in some special education classes (Tucker, 1980; Brady, Manni, & Winikur, 1983), two "explanations" stand out. Either the tests are valid and accurate and therefore minority students are more prone to handicapping conditions; or, the tests are biased.

The various positions, studies, and definitions of test bias, for all their complexity, can be subsumed under three categories: political, psychometric, and statistical.

Political discussions of test bias (Kamin, 1974; Garcia, 1977) invariably begin with the assumption that all groups are equal with respect to talent. Differences in test scores, particularly intelligence test scores, are the result of differences in opportunity to learn what a test assumes should have been learned in a given culture and by a given age. When large groups of children from a legally protected entity, such as racial minorities, are overrepresented or underrepresented in educational (or employment) programs because of decisions based on tests, political considerations (i.e., legal and ethical) come into play. Political arguments about test bias are exceedingly difficult to sort out and address because they involve, personal opinion (Sandoval & Miille, 1980), ethnocentrism (Jirsa, 1983), fairness (Novick & Peterson, 1976), due process, segregation (Hobsen vs. Hansen) civil rights, etc. Court cases such as Larry P. vs. Riles and PASE vs. Hannon typify the difficulty inherent

in addressing test bias from this perspective. Both cases looked at virtually identical data, but came up with diametrically opposed conclusions.

Less complex but not less controversial are psychometric considerations of bias. Here the literature abounds (Jensen, 1982; Reynolds, 1982; Cleary, Humphreys, Kendrick, & Wesman, 1975). The most cogent discussion on test bias and minority pupils was produced by a special committee of the American Psychological Association (Cleary, Humphreys, Kendrick, & Wesman, 1975). Though several dimensions of test validity and test bias were considered, the position paper by Cleary et. al. (1975) essentially opts for a predictive criterion of validity and bias. A test is valid (or its inverse, biased) if it predicts in the same manner and with the same accuracy (as determined by regression procedures) for minority as for majority groups. The problem with this approach for Hispanic pupils from diverse language backgrounds (which according to various sources (Keller & Van Hooft, 1982) may include up to 5,000,000 of the U.S. Hispanic student population) is that bias due to language and culture may affect both the predictor and the criterion in any given analysis of bias based on predictive validity (c.f., Linn & Werts, 1971, on the methodological problems associated with this). In other words, an English test of intelligence with LEP students may yield low IQ scores and these may be highly predictive of low academic achievement for an English (biased) instructional program. Prediction in this instance is virtually tautological and may hide the fact that bias is likely to exist in the IQ test scores of the LEP pupils. As Goldman and Hewitt (1975) noted about investigations of bias between SAT and college GPA's of Mexican

American college students:

> "If the criterion itself is biased, all bets are off. We mention this to stress the tentative nature of all investigations of test bias. Unless the criterion is itself free of bias, such investigations are of questionnable value" (p. 196).

Prediction, however, is not the only psychometric approach to the study of bias. Indices of bias can also be ascertained through measures of a test's internal consistency. Typically, reliability coefficients (Hurt, & Mishra, 1970; Dean, 1977) item analyses, (Sandoval, 1979; Sandoval, Woo, & Zimmerman, 1983) and factor structures (Reynolds, 1982; Alston, Say, & Thompson, 1978) are used to study whether a test's internal consistency holds across various ethnic groups. By and large, the literature indicates that in terms of internal consistency, tests show little evidence of bias with ethnic groups including Hispanics (Laosa, 1982; Reschly, & Sabes, 1979). However, Figueroa (1983) has suggested that item analyses are insensitive to both cultural and linguistic differences in Hispanic children. Using the SOMPA sample of 700 Hispanic students, he divided this group according to the language-background rating given to each of the mothers of the pupils. The items in the verbal subtests of the WISC-R showed the same difficulty levels for the three language-background subgroups. As Figueroa explains it, the expectation of cultural/linguistic bias in English test items (Sandoval, 1979) assumes that English language acquisition is different for non-English speakers than for English speakers. This assumption is false and seriously calls into question the use of internal - consistency methods of evaluating verbal test bias with Hispanic

pupils. Low verbal test scores of Hispanic pupils may have high reliabilities, expected item difficulty levels, high Cronbach Alpha's, or comparable factor structures and still be biased. The low scores are depressed by differences in English-language exposure and the indices of internal consistency are high because the acquisition of English is sequentially and structurally similar regardless of primary language background. Difficult/easy words, concepts, or ideas are difficult/easy for all and are in all likelihood acquired in the same cumulative, sequential manner. This means that reliabilities will be similar, that item difficulties and, in all likelihood, item intercorrelations, will be similar. But the scores will be biased if they are taken as a gauge of ability rather than ability-and-linguistic-exposure (Hickey, 1972; Darcy, 1963).

The final area in the study of bias is construct validity. Does a test measure the same construct across ethnic groups (Tryon, 1979) Psychometrically, this critical and most difficult area for establishing validity has been studied by determining if tests that purport to measure the same thing are closely related to each other. Congruence essentially supports construct. For Hispanic pupils, however, the same arguments questioning high reliabilities as indices of internal evidence of non-bias apply, particularly in the case of verbal tests. The fact that Binet and WISC-R scores of Hispanic pupils may be highly correlated does not mean they are both measuring the construct of intelligence independent of bias contaminated because of differences in language exposure. In fact, a singularly ignored study of construct validity with Mexican American pupils (Jensen, 1974) concluded that:

"The fact that the Mexican group is very similar to the white in rank order of p values and p decrements on both the PPVT and the Raven, yet has lower scores on the PPVT than on the Raven, suggests that some factor is operating to depress the PPVT performance more or less uniformly for all items and that this factor does not depress Raven performance, at least to the same degree. It seems plausible to suggest that this factor is verbal and may be associated with bilingualism in the Mexican group" (Jensen, 1974, p. 239-240).

Referring to the same analysis in another publication, Jensen (1976) further states:

"Thus, there is some evidence that a vocabulary test in English may be a biased test of intelligence for Mexican-Americans" (p. 342).

Other studies of construct validity concur with Jensen's (1976) conclusion (e.g., Gutkin, 1979).

Notwithstanding the criticism made so far about psychometric studies of bias with Hispanic children, one other major caveat needs to be highlighted. For LEP and perhaps even bilingual Hispanic students, many of the studies of test bias may be flawed. Not only do these studies fail to adequately control for English language proficiency, (Dean, 1979; Killian, 1971; Kirk, 1972; Henderson, & Rankin, 1973; Gutkin, & Reynolds, 1980; Dean, 1977a; Dean, 1977b), they also fail to include LEP pupils in their samples. Obviously, to do so would be ludicrous, if not cruel. Yet, the use of English language tests in school-based assessments continues to reflect actual practice with LEP pupils and some researchers even suggest that tests are adequate, as

they are, because "bias" in the psychometric sense has not been unequivocally demonstrated (Sattler, 1982). [The fact is, however, that some studies indicate otherwise (Gandara, Keogh, & Yoshioka-Maxwell, 1980; Jensen, 1974; Gutkin, 1979).]

Studies on bias that rely on descriptive statistics and mean-score differences make up the last category of this area of research. It is closely related to the political category of studying this phenomenon. Ethnic differences in average group scores on tests are prima facie evidence of bias (Zirkel, 1972; Dean, 1979; Goldstein, & Myers, 1979). Psychometrists, by and large, argue against this manner of studying bias (Jensen, 1982; Sattler, 1982). For Hispanic pupils, however, this way of examining bias has been pursued by the court (Diana vs. California State Board of Education) and has been discussed in several texts (DeBlassie, 1980; Erickson, & Omark, 1983; Cummins, 1984; Lambert, & Peal, 1962). The reason for this is that there is a unique pattern of "bias" in Hispanic pupils' test scores: English, verbally-loaded scores are usually significantly depressed. For example, on the WISC-R, the SOMPA sample of 700 Hispanic pupils (Mercer, 1979) scored 87.7 in Verbal IQ and 97.9 on nonverbal IQ whereas Black and Anglo children scored nearly the same on their respective verbal and non-verbal IQ scores (89 and 102.9 for Black and Anglos respectively). Further, when Figueroa (1983) examined the subtest profiles of the WISC-R of the three language subgroups of the Hispanic SOMPA sample (mother spoke only English, Spanish-English, or only Spanish), he found that the verbal (and to a lesser extent the non-verbal subtest scores) varied directly according to the home language background. Table 1 depicts these data.

Table 1

WISC-R Subtest Means for Anglo, Hispanic, and Hispanic Language Subsamples

	Anglo	Hispanic			
		Total	English	English/Spanish	Spanish
Verbal	10.39	7.98	9.08	7.80	6.70
Non-verbal	10.71	9.67	10.06	9.66	9.18

Also, there is some evidence that for Hispanic pupils there may be cultural bias in the way that individually given tests are administered (Dryman, Fernandez, Hertzig, & Thomas, 1971; Mishra, 1980; Diaz-Guerrero, Holtzman, & Swartz, 1975).

The ironic aspect of bias in testing Hispanic children is that the "language factor" has been acknowledged and known for so long.

> "For purposes of comparing individuals or groups, it is apparent that tests in the vernacular must be used with only individuals having equal opportunities to acquire the vernacular of the test. This requirement precludes the use of such tests in making comparative studies of individuals brought up in homes in which the vernacular of the test is not used, or in which two vernaculars are used. The last condition is frequently violated here in studies of children born in this country whose parents speak another tongue. It is important, as effects of bilingualism are not entirely known" (Brigham, 1930, p. 165).

The present study is unique in examining the issues from a dual language perspective: comparing the diagnostic result of testing done in English (in school) and testing done in the primary language with linguistically appropriate instruments administered by trained bilingual staff. Though this model has been suggested in the literature (Mowder, 1979; Mowder, 1980), it has usually been implemented using simple-minded, straight translations of English-language tests (Henderson, Rankin, & Valencia, 1981; Jimenez, & Keston, 1954).

Public Law 94-142

In 1975, "The Education for All Handicapped Children Act" established a unique precedent relative to minority assessment and special education placement:

Section 612(5)(C)

> "procedures to assure that testing and evaluation materials utilized for the purposes of evaluation and placement of handicapped children will be selected and administered so as not to be racially or culturally discriminatory. Such materials or procedures shall be provided and administered in the child's native language or mode of communication, unless it is clearly not feasible to do so . . . ."

Regrettably, the definition of nondiscriminatory "testing and evaluation materials" was never specified. Though two other mandates relative to assessment are critically germaine to this provision (the use of many tests rather than a single IQ score, and the need to undertake modifications of the regular program prior to testing), the notion of "nondiscriminatory assessment" owes its origins to the

problems of overrepresentation investigated in the Riverside Study (Mercer, 1973) and litigated before the Ninth Circuit Court (Diana & Larry P.). In that sense, nondiscriminatory assessment calls for assessment procedures that lead to ethnically balanced special education placements. Others, of course, argue that a psychometric definition of nondiscriminatory assessment is the only empirically justifiable one (Reschley, 1982; Reynolds, 1982; Sattler, 1982) regardless of ethnic disparities in special education classes. Given the National Academy of Sciences report (Heller, Holtzman, & Messick, 1982), it appears that it is precisely this meaning of "nondiscriminatory assessment" that professions such as school psychology have accepted and followed, and accordingly there continue to be ethnic disparities in diagnosed mental handicaps.

There have been only two major responses to 94-142's intended meaning of nonbiased assessment: The System of Multicultural Pluralistic Assessment (SOMPA) (Mercer, 1979) and the Kaufman Assessment Battery for Children (K-ABC) (Kaufman & Kaufman, 1983). SOMPA, developed and normed on a Black (700), Anglo (700), and Hispanic (700) sample attempts nondiscriminatory assessment by: assessing with many instruments (see Table 2), using three models of assessment, providing Spanish versions of some tests (Physical Dexterity Tasks, Bender, Health History Inventories, Adaptive Behavior Inventory for Children, and Sociocultural Scales), and adjusting IQ scores in such a way that norms are estimated within sociocultural configurations thereby producing an Estimated Learning Potential (ELP) score for WISC-R Verbal IQ, Performance IQ, and Full Scale IQ. ELP's tend to produce ethnic rates in special education classes commensurate with the

ethnic representation of student populations (Talley, 1979). These adjusted IQ's have also produced some of the most vitriolic criticisms from the testing profession (Jirsa, 1983; Goodman, 1979; Sattler, 1982).

Table 2

SOMPA Measures

Medical Model	Social System Model	Pluralistic Model
Tests	Tests	Tests
Physical Dexterity Tasks	Adaptive Behavior Inventory for Children (ABIC)	Estimated Learning Potential (ELP)
		Verbal
Bender Visual Motor Gestalt Test	Family	Performance
	Community	Full Scale
	Peer Relations	
	Nonacademic School Roles	
Weight by Height	Earner/Consumer	Sociocultural Scales
	Self-Maintenance	
		Family Size
Visual Acuity	ABIC Average Scaled Score	Family Structure
		Socioeconomic Status
		Urban Acculturation
Auditory Acuity	School Functioning Level (SFL)	
Health History Inventories	Verbal	
	Performance	
	Full Scale	

The K-ABC was nationally normed on a stratified sample of Anglo, Black, Hispanic, and "other" children. There are sixteen subtests in the K-ABC that are divided into three areas of assessment: Sequential Processing Scale, Simultaneous Processing Scale, and Achievement Scale. Table 3 presents these.

Table 3

K-ABC Scales and Subtest

Sequential Processing Scale
Hand Movements
Number Recall
Word Order
Simultaneous Processing Scale
Magic Window
Face Recognition
Gestalt Closure
Triangles
Matrix Analogies
Spatial Memory
Photo Series
Achievement Scale
Expressive Vocabulary
Faces and Places
Arithmetic
Riddles
Reading/Decoding
Reading/Understanding

The K-ABC's main contributions to nondiscriminatory assessment are in the provision of separate norms for Black children based on their sociocultural norms and the distinction between achievement tests (which in many ways take the place of verbal sections of IQ tests but

acknowledge their "cultural loading" by labeling them Achievement subtests) and cognitive tests (the Nonverbal Scales). Spanish instructions are provided in the latter.

Notwithstanding the contributions of these two batteries of tests to nondiscriminatory assessment, they have limited relevance to LEP or bilingual children. Neither attempts to control for the confounding variable in Hispanic assessment (i.e., English language proficiency) and neither really provides norms for LEP or bilingual pupils. As Cleary et al. (1975) noted about the latter:

> "The intellectual repertoire of a bilingual child.... can only be sampled by testing in both languages, on the basis that the repertoires in the separate languages will rarely overlap completely" (p. 22).

Fortunately a substantial degree of feasibility for testing bilingually became possible in the early 1980's. The Mexican National Office of Special Education (Dirección General de Educación Especial) under the direction of Dr. Margarita Gómez-Palacios, translated, adapted and normed many of the SOMPA tests and K-ABC subtests on a large representative sample of Mexico City public school pupils. Of these, the following will be used in this study: the WISC-R Mexicano (WISC-RM), and K-ABC Spanish edition (S-KABC) subtests (15). Other tests will include the Spanish versions of the California SOMPA's Adaptive Behavior Inventory for Children, the Physical Dexterity Tasks, and the Bender.

Behind the idea that nondiscriminatory assessment should eliminate ethnic disparities in special education classes (ER and LD), there is a critical outcome that the present study will attempt to effect. This

is that assessment should lead to accurate diagnosis of handicapping conditions as opposed to misdiagnoses of temporary conditions brought on by linguistic or cultural differences. The attempt to justify ethnic disparities in special classes because only present (rather than chronic) conditions can be assessed (Reschley, 1982) or because there "exists" a "cultural-familial" retardation ("retardation due to psychosocial disadvantage" (Balla, & Zigler, 1982, p. 3) versus a "real" retardation does not adequately confront the need to accurately differentiate between linguistic/cultural difficulties faced by Hispanic children in U.S. schools and chronic, life-long disabilities. One of the initial steps in the determination of what constitutes a handicapping condition is to examine within the context of linguistically appropriate, nondiscriminatory assessment, the placement decisions already made about the pupils.

Diagnosis

Individual States determine the criteria by which public school pupils are diagnosed as Educable Mentally Retarded and Learning Disabled. These criteria invariably meet the broad guidelines for determining these handicapping conditions included in PL 94-142 Regulations. They also ostensibly meet the "nondiscriminatory/linguistically appropriate" stipulations in federal law.

In California, Educable Mental Retardation no longer exists as a category of a handicapping condition. This is due to the state's Master Plan for Special Education which aims at delabeling. The "Eligibility Criteria for Individuals with Exceptional Needs, Title 5, California Administrative Code, Sections 3030-3031" does include,

however, a regulation establishing criteria for determining mental retardation. Generally, districts vary in their policies as to the designation of ER, MR labels, and programs. By and large Mentally Retarded students are schooled in special day classes. Educable Mentally Retarded students can be found in special day classes and in resource rooms. The "Eligibility Criteria" for Mentally Retarded is:

> 3030(h) A pupil has significantly below average general intellectual functioning existing concurrently with deficits in adaptive behavior and manifested during the developmental period, which adversely affect a pupil's educational performance.

Though this criteria is quite vague, the "significantly below average general functioning" provision is relatively well established as meaning an IQ of 70 or below. "Deficits in adaptive behavior," however, is technically undefined and there is either wide variation in complying with this provision or substantial noncompliance with it (Anderson, Gallegos, Twomey, Williamson, & Williamson, 1980).

The application of these criteria to Hispanic, LEP pupils needs to consider several additional points.

First, California Title 5 Sections 3022, 3023 (a) and 3023 (b) apply:

> 3022. ASSESSMENT PLAN
>
> In addition to the assessment plan requirements of Education Code Section 56321, the proposed written assessment plan shall include a description of any recent assessments conducted, including any available independent assessments and any assessment information the parent requests to be considered, and information indicating the pupil's primary language and the pupil's language

proficiency in the primary language as determined by Education Code Section 52164.1.

Authority cited: Education Code Section 56100(a), (i), (j); 20 USC 1414(c)(2)(B)

Reference: Education Code Sections 56321, 53629; 34 CFR 300.500-502-*300.530-541

3023. ASSESSMENT

(a) In addition to provisions of Education Code Section 56320, assessments shall be administered by qualified personnel who are competent in both the oral and written skills of the individual's primary language or mode of communication and have a knowledge and understanding of the cultural and ethnic background of the pupil. If it clearly is not feasible to do so, and interpreter must be used, and the assessment report shall document this condition and note that the validity of the assessment may have been affected.

(b) The normal process of second-language acquisition, as well as manifestations of dialect and sociolinguistic variance shall not be diagnosed as a handicapping condition.

Authority cited: Education Code Section 56100(a), (i), (j); 20 USC 1414(c)(2)(B)

Reference: Education Code Sections 56001, 56320, 56324, 56327; 34 CFR 300.530, 300.532, 300.543

Second, the Diana out-of-court settlement permitting the use of nonverbal test scores to determine intellectual functioning also applies. Admittedly, this provision would seem to be unnecessary in the present study since a linguistically appropriate IQ test will be

used. However, since the LEP sample in this study has been in the United States for some time, use of Performance IQ's as a supplement to full scale IQ's, may provide a check on primary language loss.

Third, with regards to adaptive behavior, the criterion most directly and empirically linked with significantly lowering the overrepresentation of Mexican American pupils in ER classes was found by Mercer (1973) to be a $\leq$ 2 SD cutoff on both IQ and adaptive behavior. These criteria have been subsequently endorsed by the AAMD (Grossman, 1977) though not without controversy (Balla, Hodapp, & Zigler, 1984).

Learning Disabilities or Learning Handicaps (LH) are exceedingly problematic to assess because aside from the well-documented intelligence/achievement discrepancy demonstrated by LH pupils, there is no professional agreement nationally or locally as to how to operationalize, augment, or substitute this broad guideline. California, however, has decided on an elaborate statistical procedure to diagnose Learning Handicaps. Title 5 Regulations specify that:

> 3030(j) A pupil has a disorder in one or more of the basic psychological processes involved in understanding or in using language, spoken or written, which may manifest itself in an impaired ability to listen, think, speak, read, write, spell, or do mathematical calculations, and as a severe discrepancy between intellectual ability and achievement in one or more of the academic areas specified in Sections 563337(a) of the Education Code. For the purpose of Section 3030(j):
>
> (1) Basic psychological processes include attention, visual processing, auditory processing, sensory-motor skills,

cognitive abilities including association, conceptualization and expression.

(2) Intellectual ability includes both acquired learning and learning potential and shall be determined by a systematic assessment of intellectual functioning.

(3) The level of achievement includes the pupil's level of competence in materials and subject matter explicitly taught in school and shall be measured by standardized achievement tests.

(4) When standardized tests are considered to be valid for a specific pupil, a severe discrepancy shall be determined by: first, converting into common standard scores the achievement test score and the ability test score to be compared; second, computing the difference between these common standard scores; and third, comparing this computed difference to the product of 1.5 multiplied by the standard deviation of the distribution of computed differences of students taking these achievement and ability tests. A computed difference which equals or exceeds the product of 1.5 multiplied by the standard deviation of this distribution of computed differences indicates a severe discrepancy when such discrepancy is corroborated by other assessment data which may include other tests, scales, instruments, observations and work samples, as appropriate.

(5) When standardized tests are considered to be invalid for a specific pupil the discrepancy shall be measured by alternative means as specified on the assessment plan.

(6) The discrepancy shall not be primarily the result of limited school experience or poor school attendance.

A specific Manual for the Determination of a Severe Discrepancy as Defined by Title 5, CAC, Section 3030(j) (Appendix A) is used to operationalize this process.

For LEP Hispanic pupils, the determination of a Learning Handicap has been technically impossible to carry out because state directives (Appendix A) stipulate that:

> The discrepancy must not be due to factors of environment, cultural differences or economic disadvantage. Also, the discrepancy must not be the result of visual, hearing or motor handicaps, mental retardation, limited school experience or poor attendance.

and that:

> "When standardized tests are considered to be invalid for a specific pupil, the discrepancy shall be measured by alternative means as specified on the assessment plan."
>
> Section 3030(j)(5), Title 5, California Administrative Code

If it is determined that the use of standardized tests is an invalid assessment tool, the assessment personnel are required to use professional judgment, based on such data as the results of informal or criterion-referenced assessments, analysis of pupil work samples, classroom performance and observations to determine the evidence of a severe discrepancy. The need for professional

judgment will apply to the areas of written expression and listening comprehension since there are few, if any, standardized tests which measure these skill areas. Special attention should be given in the assessment of pupils whose primary language is other than English, or whose cultural background might mitigate against the use of a certain standardized test.

This study will provide the first true application of California's LH criteria for LEP, Hispanic pupils since for the first time it will be possible to follow the "General Computational Procedures" in Appendix A to diagnose LH with linguistically appropriate tests. It remains to be seen, of course, if in this study the discrepancy criteria functions differently than those followed by school district and if the criteria itself can be validated by using other indices of LH found in bilingual populations (Cummins, 1984) or established through information-processing experiments (Brown, Campione, & Ferrara, 1982). In addition, the study will permit an examination of ER classification procedures based on a ≤ 2 standard duration criterion on the adaptive behavior and intellectual functioning measures used in the study.

Summary and Research Questions

As the review has indicated, the assessment of limited-English proficient pupils has come to hold a special place in 94-142's provision on nondiscriminatory assessment (Figueroa, 1980; Cummins, 1984) as well as in the empirical literature's acknowledgment that a language background other-than English poses unique problems for the test (Jensen, 1974), the tester (Figueroa, Merino, & Sandoval, 1984) and the pupil being assessed (Mercer, 1973; Heller, Holtzman, &

Massick, 1982). Until recently, there were no norm-referenced tests available in Spanish to assess LEP pupils whose school problems may have been caused by a handicapping condition (ER, LD). Recently, however, several such instruments have become available, predominantly normed in Mexico City. As useful as they may be in helping to evaluate the cognitive, social, and physical aptitudes of LEP pupils, their technical properties with Hispanic, LEP pupils in the United States remain unknown. This study is the first empirical look at these instruments with U.S. Hispanic students in three school-diagnosed clinical groups: ER, LD, and regular classes.

The purpose of this study was to examine the characteristics of student performance and to empirically determine whether diagnoses suggested by the normed, Spanish versions of the WISC-RM, the K-ABC tests, and the SOMPA lead to outcomes that are different from those arrived at in school-based assessments. Mexican American, Spanish-proficient pupils, ages 5-5 to 12-5, placed in ER (N = 39), LD (N = 45), and nonhandicapped (N = 44) classes were re-assessed using the Spanish WISC-RM, K-ABC tests, and the SOMPA battery. It was expected that these linguistically appropriate instruments would indicate the presence or absence of discrepancies in diagnosis (that have been historically pervasive with Hispanic LEP students) in the present sample.

The following were the specific hypotheses as well as the research questions addressed in this study:

1. The diagnostic judgements generated with the Spanish-language instruments in this study will not correspond with the diagnostic categories assigned in the schools through the use of traditional assessment procedures.

a) Do the diagnostic judgments derived from the Spanish-language instruments used in this study correspond with the diagnostic judgments assigned by the schools?

2. The instruments used in this study will be able to reliably and validly distinguish ER, LD, and nonhandicapped students based upon cognitive profiles and adaptive functioning levels.

a) Are the reliability coefficients, standard errors of measurement (SEM), discriminant and concurrent validities of the Spanish-language tests comparable to those of the English-language versions?

II. Method

Presented in this section is a description of the methods employed in this study. This information is divided into three subsections describing the subjects who participated in the study, the assessment instruments employed, and the procedures used to collect the data.

Subjects

A total of 128 students between the ages of 5-5 to 12-5 were part of this study. All subjects were Hispanic (primarily Mexican American) limited-English-proficient students from low socioeconomic backgrounds. Subjects for the present study were classified by the public schools as learning disabled (45), educationally retarded (39) and nonhandicapped students (44) and were attending elementary schools in a large metropolitan school district located in California.

The students were volunteers whose parents responded to a written request from the project director to participate in a study designed to examine non-biased assessment of limited-English-proficient Hispanic students. Previous to this, district and elementary school records were searched to identify a pool of potential subjects. Students were selected according to age (5-5 to 12-5), language classification (limited-English-proficient) educational status (learning disabled, educationally retarded and nonhandicaped) and low socioeconomic status. Once a random selection of students had been drawn from the pool, a letter containing a description of the research project was sent to the parents which requested the approval of parents and students alike. This letter informed parents that they would be contacted by a parent interviewer. The parent interviewer described the scope of the study, the extent of the student's and parent's participation and the

safeguards for confidentiality and privacy. During this contact, written informed consent was obtained from those parents who agreed to participate in the study. Copies of the letter sent to the parents and the consent form may be found in Appendices B and C respectively.

Learning Handicapped - Learning Disabled Students

In terms of educational status, 45 of the 119 subjects were identified as pupils with specific learning disabilities. These students were receiving special education services in accordance with their primary handicapping conditions (i.e., learning disabilities). These services were provided through the Resource Specialist Program.

The Resource Specialist Program was available at all elementary and secondary schools within the school district. This program is considered the appropriate educational setting for pupils with specific learning disabilities when:

(1) The pupil requires instruction with specialized techniques or strategies for a limited period of time; such a pupil is enrolled in a regular class setting for the majority of the day;

(2) The frequency, intensity and duration of needed remediation require no more than two hours per day;

(3) Remediation goals as determined by the Individual Educational Plan Team are expected to be achieved in the duration of time specified.

Eligibility criteria. The specific criteria used by the school district to identify students with learning disabilities is stated below:

(1) Public school students who fit the description of learning disabilities as stated in Title 34, Code of Federal Regulations, 300a.5(b)(9) (see Appendix D).

(2) Public school students who satisfy the eligibility criteria for pupils who have a specific learning disability as set forth in California Administrative Code (see Appendix D).

(3) Public school students who were currently enrolled in the Resource Specialist Program but who spend the majority of the school day in a regular class setting.

Learning Handicapped - Educationally Retarded Students

Thirty nine of the subjects were identified as educationally retarded. (This diagnostic classification, known as "ER," is essentially the equivalent of the EMR designation most commonly used. However, in order to be consistent with school district labels, the ER designation is used throughout the report.) These students were receiving special education services in accordance with their primary condition (i.e., educationally retarded). These services were provided through the Special Day Class.

The Special Day Class was offered at various school sites within the school district. This program is considered the appropriate educational setting for pupils identified as educationally retarded when:

(1) The pupil required specialized instructional techniques for the majority of the school day;

(2) The frequency, intensity, and duration of needed remediation require more than two hours per day;

(3) Remediation goals and objectives are set significantly be ow regular classroom goals and therefore require more intensive instruction.

Eligibility criteria. The specific criteria used by the school district to identify students as educationally retarded is described below:

(1) Public school students who fit the description of educationally retarded as set forth in Title 34, Code of Federal Regulations, 300a.5(b)(4).

(2) Public school students who meet the eligibility criteria as stated in California Administrative Code.

(3) Public school students who were currently enrolled in the Special Day Class.

Nonhandicapped Students

A total of 44 of the subjects were identified as pupils without any handicapping condition. These students were enrolled in the regular elementary school program and were not receiving any special education services.

Eligibility criteria. The specific criteria used to identify students as nonhandicapped is as follows:

(1) Public school students who were enrolled in a regular school program and who were not receiving any special education services.

(2) Public school students whose level of achievement was commensurate with their intellectual abilities. Furthermore, these intellectual abilities were assumed to be within the average range.

Additional subject variables. Aside from educational classification subjects were selected based upon age level, language classification and socioeconomic status. The criteria for each of these subject variables is described below:

(1) Age level. Students who ranged in age between 5 years 5 months and 12 years 5 months were included in this study. All of the subjects were enrolled in the elementary school program (excluding pre-school).

(2) Language classification. Subjects included in this study were identified as Limited-English-Proficient (LEP) at the school district level. LEP classification refers to a student who is considered a non-English-speaker or limited-English speaker. Spanish was the primary language of all of the subjects. This type of language classification is persuant to the requirements of California Education Code, Title 5 of the California Administrative Code; the Bilingual Education Improvement and Reform Act of 1980 (Assembly Bill 507); and, administrative directives from the California State Department of Education regarding the education of students of limited-English-proficiency as of January, 1984.

(3) Socioeconomic status. Subjects included in this study were predominately from low socioeconomic (SES) households. The determination of SES was based upon two sources: parent interviews and school district compliance reports. Parents were administered the Sociocultural Scale which yielded occupational information as well as family income. School district compliance/records provided the rank order of

elementary schools for 1984-1985 Chapter 1 and Economic Impact Aid programs. Ranking was based upon the percentage of pupils at each elementary school who were recipients of Aid to Families and Dependent Children (AFDC) and Free Lunch program.

III. Assessment Instruments

The selection of assessment instruments was based upon several considerations. First, it was necessary that the instruments be able to measure cognitive abilities, achievement, sensory-motor abilities and adaptive behavior. It was also considered desirable that the instruments provide an estimate of the subject's learning potential. Secondly, the instruments were required to provide non-biased assessment which would not penalize subjects for their unfamiliarity with the English language and culture. At the same time these instruments had to be comprehensible to students of varying grade levels, different educational status, and of either sex. Third, since information would be obtained both directly from the student subject and indirectly from observers, it was necessary that instruments be flexible in their administration. A battery of assessment instruments satisfied these requirements and were subsequently employed in this study. Since the subjects in the study were Spanish speakers, the Spanish versions of these instruments were used. These instruments were the Wechsler Intelligence Scale for Children-Revised, Mexicano (Padilla, Palacios, & Roll, 1984; Wechsler, 1949), the Kaufman Assessment Battery for Children-Spanish Edition (Kaufman, Kaufman & Padilla, 1981), the System of Multicultural and Pluralistic Assessment (Mercer, & Lewis, 1978) and the Contact with Mexico Questionnaire. Each of these instruments is described below.

Wechsler Intelligence Scale for Children-Revised, Mexicano

The WISC-RM is an adaptation of the Wechsler Intelligence Scale for Children-Revised, intended for use with children from six years of age to sixteen. Both the WISC-R and the WISC-RM are

based on the assumption that intelligence is global; both are also designed to assess intelligence at different age levels. Also, like the WISC-R, the WISC-RM test yields a Full Scale I.Q. Verbal I.Q., and a Performance I.Q. The standard WISC-R battery as well as the WISC-RM includes the following subtests:

Verbal	Performance
Information	Picture Completion
Comprehension	Picture Agreement
Arithmetic	Block Design
Similarities	Object Assembly
Vocabulary	Coding or Mazes

The WISC-R, Mexicano instrument is based upon a sample of 1,100 pupils in the primary and secondary schools governed by the Secretaría de Educación Pública (SEP) of Mexico. These pupils, male and female, were residents of Mexico City. Three substantive changes were made in the WISC-RM as compared to the original WISC-R. First, experimental items were included in the following subtests: Information Comprehension, and Vocabulary. These items contained references to the Latin culture. Secondly, certain items were excluded from original subtests, as in the Picture Completion subtest. Last of all, within the Verbal scale, items were reordered so that the level of difficulty was appropriate to the normative sample.

Kaufman Assessment Battery for Children-Spanish Edition

The Spanish Kaufman Assessment Battery for Children (S-KABC) is a translated and adapted version of the original K-ABC. New normative data have been obtained for this version. The Spanish

K-ABC, like the original English K-ABC, is a test of mental processes and achievement. Intellectual functioning is described as problem-solving ability, that is, the ability to be flexible and adaptable when faced with unfamiliar problems. Intellectual functioning is divided into two modes of processing information: sequential processing and simultaneous processing. Sequential processing focuses on serial or temporal order, while simultaneous processing uses a gestalt approach or spatial integration. In contrast to these problem-solving skills, academic achievement is viewed as a set of acquired skills. These three dimenions--sequential processing, simultaneous processing, and achievement--are measured through 15 subtests on the Spanish K-ABC:

1. La Ventana Mágica
2. Retención de Caras
3. Movimiento de las Manos
4. Cierre Gestalt
5. Retención de Números
6. Triángulos
7. Retención de Palabras
8. Analogías de Matríz
9. Retención de Lugares
10. Serie de Fotografía
12. Vocabulario Visual
13. Aritmética
14. Adivinanzas
15A. Decodificación de Lectura
15B. Comprensión de Lectura

From these 15 subtests, four global scales are computed: sequential processing, simultaneous processing, mental processing composite, and achievement.

Aside from the translation of the test, there were other substantive changes made by the publisher in the Spanish version. First, the achievement subtest Faces and Places has been dropped. Second, although the English edition of the K-ABC is designed for use with 2 1/2-year-old children, the Spanish edition should be not given to any children below age 3. Third, the total Reading score is a combination of subtests Decodificacion de Lectura and Comprension de Lectura. All of these changes yield fifteen subtests. The four global scales remain intact.

Sociocultural Scales, System of Multicultural Pluralistic Assessment

The Sociocultural Scales are a part of the System of Multicultural Pluralistic Assessment (SOMPA). These scales provide information relative to the social and cultural characteristics of the child's family, and are used in interpreting the child's performance in the WISC-R. (In the present study, the WISC-RM scores were used in place of WISC-R scores.) The child's performance in the WISC-R is compared to those of other children with similar sociocultural characteristics. The Sociocultural Scales consist of four scales: Family Size, Family Structure, Socioeconomic Status, and Urban Acculturation. These are administered during the Parent Interview.

Adaptive Behavior Inventory of Children, System of Multicultural Pluralistic Assessment

The ABIC measures the child's adaptive behavior within selected social systems such as family, the peer group, and the community. The ABIC includes six scales, relating to the six spheres of activity: Family, Community, Peer Relations, Nonacademic School Roles, Earner/Consumer, and Self-Maintenance.

Health History Inventories, System of Multicultural Pluralistic Assessment

The Health History Inventories provides information in the health history of a child. This information may serve as a preliminary screening of high-risk children. The Health History Inventories are identified as follows: Prenatal/Postnatal Inventory, Trauma Inventory, Disease and Illness Inventory, Vision Inventory, and Hearing Inventory.

Physical Dexterity Tasks, System of Multicultural Pluralistic Assessment

The Physical Dexterity Tasks are a series of 29 tasks that measure the intactness and capability of the subject's motor and sensory pathways. The tasks are grouped into six scales: Ambulation, Equilibrium, Placement, Fine Motor Sequencing, Finger-Tongue Dexterity, and Involuntary Movement.

Bender Visual Motor Gestalt Test, System of Multicultural Pluralistic Assessment

The Bender-Gestalt Test is a measure of the subject's visual-perceptual ability. The test consists of nine cards each with a geometric figure. The subject is required to copy the

figure as presented. The Koppitz (Koppitz, 1975) scoring system is employed to evaluate the quality of the drawings.

Contact with Mexico Questionnaire

The Contact with Mexico Questionnaire assesses the degree of exposure and contact the child's family has with Mexico or other Latin American Countries. This contact may be due to economic; social or familial factors. (Although this was administered, data was not used in the final analyses.)

In summary, the Wechsler Intelligence Scale for Children-Revised, Mexicano, the Kaufman Assessment Battery for Children-Spanish edition, the System of Multicultural Pluralistic Assessment, and the Contact with Mexico Questionnaire constituted the assessment battery for the present study. These instruments yielded quantitative information in the following domains: cognitive, achievement, sensory-motor and adaptive behavior.

Data Collection

In order to complete the study, information from both students and parents as well as school records was obtained. The following provides a description of the sequence of the procedures that were completed.

Parent Interviews

The parent interview constituted the first phase of data collection in this study. A letter, written in English and Spanish, was sent to the parents/guardians of a randomly selected sample of students who met all selective criteria. This letter described the nature of the study and provided the name of the interviewer who was assigned to meet with the parent.

Interviewers were specially trained, bilingual, research associates who had been selected for their exceptional interpersonal skills. The interviewer contacted the parents at their home. During the home visit the interviewer related the purpose of the study and the scope of the testing that the child would complete. After this explanation, the parent was asked to sign a consent form for his/her child to participate in the study and for release of specified school records. If the parent agreed to sign the consent forms, a parent interview immediately followed. Without such consent the home visit was terminated and another pupil from the sample was contacted.

The parent interview took approximately one to one and a half hours to complete. The interviewers administered the Sociocultural Scales, the Health History Inventories, Adaptive Behavior Inventory for Children and the Contact with Mexico Questionnaire. Parents were given the choice of responding in English or Spanish to any of the inventories. All of the interviews were conducted in Spanish or a combination of English and Spanish.

Assessment of Subjects

Once each parent gave informed, written consent for his/her child to participate in the study, the assessment phase began. The subjects were individually administered the following battery of instruments:

(1) Wechsler Intelligence Scale for Children-Revised, Mexicano

(2) Kaufman Assessment Battery for Children-Spanish edition

(3) Physical Dexterity Tasks

(4) Bender Visual Motor Gestalt Test

These instruments were administered in Spanish by bilingual licensed psychologists, school psychologists and supervised school

psychologists. The test administrators had previously undergone in-service training in the use of these instruments, as part of the study.

The battery of instruments were administered in two sessions, according to the following schedule. The sessions were counter-balanced within each subgroup (i.e., learning disabled, educationally retarded and nonhandicapped) to control for order effects. Time between sessions ranged from one day to fourteen days.

Session One

Activity	Approximate Time
1. rapport building	10 min.
2. administration of WISC-RM	90 min.
3. administration of Bender Gestalt	10 min.

Session Two

1. rapport building	10 min.
2. administration of S-KABC	90 min.
3. administration of physical dexterity tasks	10 min.

Students were generally tested at school sites either during or after school.

School Records

In addition to parent interviews and student assessments, school based data was collected for each subject. A coding sheet was utilized to gather data in the areas of: student information, school information, family characteristics, educational disruptions, teacher evaluation and academic achievement, teacher evaluation, bilingual language proficiency and language instruction, instruction, psychological assessment findings, and IEP data goals and objectives.

Trained coders using the coding sheet examined student cumulative records at the individual school site and at the district level. This information was obtained in order to answer questions relative to the validity of instruments used in student assessment, however, these analyses, currently underway will be reported at a later point.

IV. Results

As was indicated earlier, the two central questions addressed in this study were focused on 1) characteristics of students performance, as well as psychometric properties of the instruments, especially with respect to the ability to differentiate the groups of students in the study, and 2) the correspondence of the diagnostic judgements assigned by the schools with those suggested by the combination of instruments used in the study. Each of these will be addressed in turn in the following section.

Data Analysis

Comparison of student performance on each of the measures was carried out through examination of trends and patterns in the mean subscale as well as of total composite mean scores (where appropriate) for each of the measures. Analyses of variance on each of the measures was used to indicate the significance of difference between groups. In addition to the calculation of mean scores, standard deviations and standard errors of the mean were used as estimates of variance for each of the diagnostic groups.

In addition to the calculation of summary descriptive information on each of the measures by subgroup, reliability estimates for each of the measures were calculated. The estimates used in the analyses were the split half reliability coefficient using the Spearman-Brown correction for unequal test lengths, and Cronbach's alpha, essentially measures of internal consistency and internal stability. In addition, the standard error of measurement for each of the measures was calculated, which is an estimate of the extent to which a test can be said to provide a "true" score. Finally, the subscale items on the

tests were intercorrelated to provide an estimate of concurrent validity.

The final part of the analysis examined the numbers of students in each of the subgroups in the study who met the eligibility criteria of the State of California for learning disabled (LD) and mildly mentally retarded (ER). The analysis focused on the percentage of students whose classification by the school matched or did not match the classification which would have been given according to the criteria outlined by the state of California using the assessment instruments which were part of the present study.

A. Summary of Student Performance and Properties of the Measures

Character.stics of student performance: Intellectual measures. The primary interest in this part of the investigation was to examine the assessed performance of the students according to the diagnostic category assigned by the public schools. The first step in this process was to calculate summary scores for each of the diagnostic subgroups in the sample. The mean summary scores for the intellectual assessment measures are presented in Table IV(1). In general terms, the performance scores are higher than the verbal scores in all cases, and the WISC-RM scores tend to be lower than the corresponding adjusted SOMPA scores. The highest score of the S-KABC in nearly all cases was the on the simultaneous processing subtest. In addition, the regular classroom students tended to score higher on all measures than the learning disabled students, who in turn scored higher than the ER students.

Table IV(1)

Summary Statistics on Intellectual Assessment Measures by Diagnostic Category

Measure*	$\bar{X}$	s.d.	std. error of mean	
WISC-RM verbal	92.21	15.6	2.38	Nonhandicapped
WISC-RM performance	112.42	11.11	1.69	
WISC-RM full scale	103	12.47	1.90	
SOMPA ELP verbal	112.79	16.19	2.47	
SOMPA ELP performance	119.49	13.16	2.01	
SOMPA ELP full scale	118.78	14.36	2.19	
S-KABC sequential	105.79	17.10	2.64	
S-KABC simultaneous	110.83	12.54	1.94	
S-KABC mental processing comp.	110.48	13.8	2.13	
S-KABC nonverbal	111.05	13.53	2.09	
WISC-RM verbal	78.39	13.66	2.22	Learning Disabled
WISC-RM performance	107.84	12.47	2.02	
WISC-RM full scale	93.89	13.16	2.14	
SOMPA ELP verbal	96.95	15.51	2.52	
SOMPA ELP performance	114.34	15.09	2.45	
SOMPA ELP full scale	108.11	16.47	2.67	

Table IV(1) continued

Measure*	$\bar{X}$	s.d.	std error of mean
S-KABC sequential	89.6	17.63	2.79
S-KABC simultaneous	105.08	16.37	2.59
S-KABC mental processing comp.	98.75	17	2.69
S-KABC nonverbal	104.08	17.84	2.82
WISC-RM verbal	56.46	16.13	2.73
WISC-RM performance	72.17	20.45	3.46
WISC-RM full scale	60.49	19.89	3.36
SOMPA ELP verbal	71.69	15.88	2.68
SOMPA ELP performance	74.68	20.57	3.48
SOMPA ELP full scale	70.89	19.72	3.33
S-KABC sequential	63.58	21.58	3.6
S-KABC simultaneous	69.47	25.84	4.31
S-KABC mental processing comp.	61.86	26.57	4.43
S-KABC nonverbal	65.28	25.49	4.25

Educable Retarded

*In each case, the mean is 100, and the standard deviation is 15.

In order to provide a more detailed picture of student performance, the subscale scores for the both the WISC-RM and the S-KABC Mental Processing Subtests are presented in Tables IV(2) and IV(3). As the tables demonstrate, the performance scales on the WISC-RM tend to have higher means than the verbal scales.

Table IV(2)

Summary Statistics on the WISC-RM Subtest Scaled Scores by Diagnostic Category

	Diagnostic Group					
	Nonhandicapped		LD		ER	
Subscales	$\bar{X}$	s.d.	$\bar{X}$	s.d.	$\bar{X}$	s.d.
Información (Information)	5.6	3.0	4.68	9.58	2.67	2.37
Semejanzas (Similarities)	11.63	3.26	12.57	11.95	6.17	2.95
Aritmética (Arithmetic)	8.14	(2.86)	5.94	(2.55)	1.69	(2.20)
Vocabulario (Vocabulary)	8.81	2.91	6.32	2.08	3.91	2.34
Comprensión (Comprehension)	9.74	3.0	8.19	3.77	4.11	3.07
Retención de Digitos (Digit Span)	9.34	3.04	7.64	2.92	3.45	3.52
Figuras Incompletas (Picture Completion)	10.38	2.26	9.69	2.74	5.94	3.91
Ordenación de Dibujos (Picture Arrangement)	12.14	2.16	11.35	2.51	6.59	2.96
Diseños con Cubos (Block Design)	10.98	2.87	10.27	2.61	5.43	3.22
Composición de Objetos (Object Assembly)	12.33	2.67	11.59	2.95	6.54	3.53
Claves (Coding)	11.65	2.86	10.43	2.62	6.23	4.07
Laberintos (Mazes)	11.6	3.1	11.31	2.67	4.11	3.45

*$\bar{X}$ = 10, s.d. = 3.

Table IV(3)

Summary Statistics on the S-KABC Mental Processing Subtest Scaled Scores by Diagnostic Group

	Nonhandicapped
Subtest	Scaled Score (s.d.)
Ventana Mágica (Magic Window)	—
Retención de Caras (Face Recognition)	--
Movimiento con la Mano (Hand Movements)	11.55(3.06)
Cierre Gestalt (Gestalt Closure)	12.05(2.67)
Retención de Números (Number Recall)	11.24(3.23)
Triángulos (Triangles)	11.79(3.52)
Retención de Palabras (Word Order)	12.05(3.33)
Analogías de Matríz (Matrix Analogies)	12.29(2.23)
Retención de Lugares (Spatial Memory)	11.57(2.52)
Serie de Fotografías (Photo Series)	12.72(3.05)

Table IV(3) continued

Subtest	Learning Disabled Scaled Score (s.d.)
Ventana Mágica (Magic Window)	--
Retención de Caras (Face Recognition)	--
Movimiento con la Mano (Hand Movements)	10.48(3.49)
Cierre Gestalt (Gestalt Closure)	11.85(2.89)
Retención de Números (Number Recall)	8.53(3.31)
Triángulos (Triangles)	10.73(2.59)
Retención de Palabras (Word Order)	8.83(3.0)
Analogías de Matríz (Matrix Analogies)	11.23(3.06)
Retención de Lugares (Spatial Memory)	10.67(3.24)
Serie de Fotografías (Photo Series)	11.55(3.61)

Table IV(3) continued

Subtest	ER Scaled Score (s.d.)
Ventana Mágica (Magic Window)	--
Retención de Caras (Face Recognition)	--
Movimiento con la Mano (Hand Movements)	5.61(3.52)
Cierre Gestalt (Gestalt Closure)	7.89(4.4)
Retención de Números (Number Recall)	4.67(4.17))
Triángulos (Triangles)	5.39(4.64)
Retención de Palabras (Word Order)	5.44(3.5)
Analogías de Matríz (Matrix Analogies)	8.69(3.77)
Retención de Lugares (Spatial Memory)	4.63(3.66)
Serie de Fotografías (Photo Series)	4.61(3.86)

*$\bar{X}$ = 10, s.d. = 3 for all scores.

In addition to the mean scores for the WISC-RM and the S-KABC the intercorrelation matrices for the subtests of each measure are presented in Tables IV(4) and IV(5).

Table III(4)

Intercorrelation Matrix of WISC-RM Subscales

	Infor.	Simil.	Arith.	Vocab.	Comp.	Digit Span	Pict. Comp.	Pict. Arrang.	Blk. Design	Obj. Assoc.	Coding
Similarities (Semejanzas)	.18										
Arithmetic (Aritmética)	.17	.40***									
Vocabulary (Vocabulario)	.31	.36***	.63***								
Comprehension (Comprensión)	.24**	.32***	.66***	.80***							
Digit Span (Retención de Digitos)	.23*	.22*	.74***	.65***	.68***						
Picture Comprehension (Figuras Incompletas)	.24**	.40***	.54***	.53***	.63***	.57***					
Picture Arrangement (Ordenación de Objetos)	.16	.35***	.67***	.66***	.64***	.61***	.69***				
Block Design (Diseños con Cub s)	.19*	.24**	.68***	.55***	.58***	.61***	.65***	.67***			
Object Assembly (Composición de Objetos)	.12	.26**	.61***	.54***	.63***	.66***	.69***	.68***	.73***		
Coding (Claves)	.19*	.38***	.66***	.54***	.64***	.62***	.69***	.72***	.62***	.71***	
Mazes (Laberintos)	.19*	.32***	.64***	.46***	.57***	.63***	.66***	.67***	.68***	.72***	.64***

* = < .05

Table IV(5)

Intercorrelation Matrix of S-KABC Scaled Scores for Entire Sample

	Hand Movement	Gestalt Closure	Number Recall	Triangles	Word Order	Matrix Analogies	Spatial Order
Hand Movement (Movimiento de las manos)	1.0						
Gestalt Closure (Cierre Gestalt)	.52***	1.0					
Number Recall (Retención de Números)	.64***	.53***	1.0				
Triangles (Triángulos)	.62***	.67***	.62***	1.0			
Word Order (Retención de Palabras)	.67***	.44***	.82***	.62***	1.0		
Matrix Analogies (Analogías de Matríz)	.62***	.43***	.57***	.60***	.54***	1.0	
Spatial Order (Retención de Lugares)	.71***	.65***	.72***	.78***	.66***	.64***	1.0

*** = $p < .001$

One of the questions of interest with respect to this study was the ability of the measures to differentiate the groups of students in the different diagnostic categories. Therefore, a series of one way ANOVA's were carried out using the total score on each of the intellectual assessment measures as the dependent variable. In addition, Bonferroni multiple contrast procedures were used to evaluate the differences between the groups. Table IV(6) presents the group comparisons based upon this analysis.

Table IV(6)

One Way Analyses of Variance on Intellectual Assessment Measures

Measure	F-value	p	Group Comparisons
WISC-RM Verbal IQ	53.92	< .001	NH ≠ LD ≠ ER
WISC-RM Perf. IQ	80.63	< .001	NH & LD ≠ ER
WISC-RM Full Scale IQ	80.07	< .001	NH ≠ LD ≠ ER
SOMPA ELP Verbal	65.00	< .001	NH ≠ LD ≠ ER
SOMPA ELP Performance	83.45	< .001	NH & LD ≠ ER
SOMPA ELP Full Scale	83.61	< .001	NH ≠ LD ≠ ER
S-KABC Sequential Processing	49.53	< .001	NH ≠ LD ≠ ER
S-KABC Simultaneous Processing	54.07	< .001	NH & LD ≠ ER
S-KABC Mental Process. Composite	64.16	< .001	NH & LD ≠ ER
S-KABC Nonverbal	61.89	< .001	NH & LD ≠ ER

NH = Nonhandicapped, LD = Learning Disabled, ER = Educable Mentally Retarded.

As the table indicates, the ER students were differed significantly from the other two groups on all measures. However,the LD group differed from the nonhandicapped group only on the WISC-RM Verbal IQ, the WISC-RM Full Scale IQ, the ELP Verbal Score, the Full Scale ELP, and the S-KABC Sequential Processing Score. On the remaining measures, the LD and the nonhandicapped groups did not differ.

As part of the examination of the psychometric properties of the measures, reliability indices were calculated for each of the measures. The two types of reliability indices included in this analysis were the split half reliability using the Spearman-Brown correction, and Cronbach's alpha. The split half method, essentially a method for determining internal consistency, is based upon a comparison of the odd and even items in a test, thus permitting the estimation of a reliability coefficient even though the measure was administered at one point in time. Cronbach's alpha, on the other hand, although also a measure of internal consistency, is based upon the individual items in a measure. In essence, this coefficient allows a determination of the extent to which all the items in a measure assess the same thing. The two reliability coefficients for the subtests in each of the intellectual assessment measures is presented in Table IV(7). In addition, the table contains the standard error of measurement (SEM) of each of the subtests, in essence an "error band", or an index of the accuracy with which a test can be said to provide a "true" score.

Table IV(7)

Reliability Coefficients for the Intellectual Functioning Measures for Entire Sample

Measure		Split Half	Cronbach's Alpha	SEM
WISC-RM	Information (Información)	.90	.89	.90
	Similarities (Semejanzas)	.90	.85	1.82
	Arithmetic (Aritmética)	.92	.88	1.42
	Vocabulary (Vocabulario)	.88	.92	2.73
	Comprehension (Comprensión)	.85	.85	1.90
	Digit Span (Retención de Digitos)	--	--	--
WISC-RM	Picture Completion (Figuras Incompletas)	.90	.90	1.72
	Picture Arrengement (Ordenación de Dibujos)	.89	.86	1.88
	Block Design (Diseños con Cubos)	.86	.84	4.38
	Object Assembly (Composición de Objetos)	.82	.86	5.00
	Coding (Claves)	--	--	--
	Mazes (Laberintos)	.95	.93	3.31
S-KABC	Magic Window (Ventana Mágica)	--	--	--
	Face Recognition (Retención de Caras)	--	--	--

Table IV(7) continued

Measure	Split Half	Cronbach's Alpha	SEM
Hand Movement (Movimientos de la mano)	.86	.83	1.58
Gestalt Closure (Cierre Gestalt)	.86	.84	1.61
Number Recall (Retención de Números)	.89	.80	.96
Triángles (Triangulos)	.95	.90	1.13
Word Order (Retención de Palabras)	.89	.83	1.26
Matrix Analogies (Analogías de Matríz)	.82	.83	1.76
Spatial Memory (Retención de Lugares)	.91	.90	1.56
Photo Series (Serie de Fotografías)	.94	.91	1.52

In addition to the measures of reliability, an intercorrelation of the intellectual assessment measures was performed as an index of concurrent validity. The intercorrelation matrix is presented in Table IV(8). As the table demonstrates, the correlations between the measures tend to be high, at about .90 and higher. The highest correlation was between the ELP and the WISC-RM IQ, and the lowest was between the ELP and the S-KABC Mental Processing Composite Score.

Table IV(8)

Intercorrelation Matrix of the Intellectual Functioning Measures

	WISC-RM Full Scale IQ	SOMPA Full Scale ELP
SOMPA Full Scale ELP	.98***	
S-KABC Mental Processing Composite	.91***	.89***

***p < .001

Although factor analytic procedures were considered as a means of estimating the number and composition of domains which the subscales on the measures were tapping, it was not possible to conduct this type of analysis within the constraints of the present study. Although the number of subjects for the entire sample was sufficient for factor analytic procedures, it was assumed that the three diagnostic groups were different and therefore it did not appear appropriate to combine the three groups for the purposes of analysis. On the other hand, the individual sample sizes for each of the three diagnostic groups were not deemed sufficiently large for this type of procedure.

Characteristics of Student Performance: Achievement. In addition to the measures of intellectual functioning, the battery administered to the subjects in the present study included a measure of academic achievement. This measure is part of the S-KABC, and purportedly measures a dimension distinct from that tapped by the intellectual

functioning parts of the test. This Achievement measure is comprised of four subtests, including expressive vocabulary, arithmetic, riddles, and reading. The mean standard scores, standard deviations, and standard errors of the mean for each of the subscales on the measure are presented in Table IV(9).

In general, the order of the groups on the mean scores is as expected. That is, the nonhandicapped students tend to score the highest, the learning disabled next, and the ER students last. Interestingly, the learning disabled students tend to score noticeably lower on the reading subtest than on the other subtests. A one way ANOVA on the total achievement score indicated that there were significant differences between groups, $F(2, 114) = 64.24$, $p < .001$. A Bonferroni multiple contrast procedure indicated that the three groups were all different from each other.

In order to investigate the relationship between the achievement measure and the measures of intellectual functioning, correlation coefficients were calculated between the Total Achievement Score and the scores on the three intellectual measures. The correlation coefficients are presented in Table IV(10). As the table indicates, the correlation coefficients are significant and relatively high for all measures.

Reliability coefficients were also calculated for the subscales of the Achievement measure. These coefficients are reported in Table IV(11), along with the standard error of the mean for each of the subscales.

Table IV(9)

Summary Statistics on the Subscales of the S-KABC Achievement Measure by Diagnostic Group

Measure	Nonhandicapped			Learning Disabled			ER		
	$\bar{X}$	s.d.	std. error of X	$\bar{X}$	s.d.	std. error of X	$\bar{X}$	s.d.	std. error of X
Expressive Vocabulary	99.26	13.57	2.09	90.95	17.58	2.81	71.92	25.48	4.25
Arithmetic	101.05	18.61	2.87	82.69	14.65	2.35	63.81	18.06	3.01
Riddles	99.26	12.12	1.87	89.56	12.83	2.05	74.44	19.21	3.20
Reading	95.38	16.98	2.62	66.33	20.27	3.25	57.92	14.98	2.50
Total Achievement	95.92	13.56	2.09	74.28	17.07	2.73	51.03	21.43	3.57

$\bar{X}$ = 100, s.d. = 15

Table IV(10)

Intercorrelations of Achievement Measure and Measures of Intellectual Functioning for Entire Sample

Measure	Correlation with S-KABC Total Achievement Score	p
WISC-RM Verbal	.89	< .001
WISC-RM Performance	.75	< .001
WISC-RM Full Scale	.86	< .001
ELP Verbal	.83	< .001
ELP Performance	.75	< .001
ELP Full Scale	.83	< .001
S-KABC Sequential Process.	.83	< .001
S-KABC Simultaneous Process.	.81	< .001
S-KABC Mental Process. Composite	.87	< .001
S-KABC Nonverbal	.82	< .001

Table IV(11)

Reliability Coefficients for Subscales of the S-KABC Achievement Measure for Entire Sample

Subscale	Split Half	Alpha	S.E.M.
Expressive Vocabulary (Vocabulario Visual)	.99	.87	.55
Arithmetic (Aritmética)	.96	.93	1.45
Riddles (Adivinanzas)	.95	.90	1.45
Decoding (Decodificación)	.99	.98	.88
Comprehension (Comprensión)	.96	.95	1.39

*The Standard error of measurement was not calculated for these two subtests because in the total scores these achievements were combined.

Characteristics of Student Performance: Adaptive Behavior. The measure of adaptive behavior employed in the present study was the Adaptive Behavior Invetory for children (ABIC), from the SOMPA assessment battery. This test consisted of six subscales which all demonstrated a considerable amount of internal consistency with alpha coefficients all above .90. The alpha coefficients for the six ABIC subscales were as follows: Family $\alpha = .95$, Community $\alpha = .93$, Peer $\alpha = .90$, School $\alpha = .93$, Consumer $\alpha = .92$, and Self Maintenance $\alpha = .94$. The summary statistics for this measure are provided in Table IV(12). As the table demonstrates, the trend for the scores on all the subscales indicates that the nonhandicapped subsample tends to score

Table IV(12)

Summary Statistics for the ABIC by Diagnostic Category

Measure	Nonhandicapped			Learning Disabled			ER		
	$\bar{X}$	s.d.	std. error of X	$\bar{X}$	s.d.	std. error of X	$\bar{X}$	s.d.	std. error of X
Family	47.05	10.[illegible]	[illegible].60	39.6	15.25	2.27	29.49	13.84	2.22
Community	35.14	8.61	1 30	31.56	12.33	1.84	23.85	10.66	1.71
Peer Relations	46.77	12.45	1.88	42.31	13.03	1.94	30.67	14.12	2.26
Nonacademic School Roles	44.80	11.68	1.76	36.44	13.13	1.96	26.15	11.34	1.82
Earner/Consumer	43.05	10.44	1.57	37.44	12.23	1.82	28.03	13.39	2.14
Self-Maintenance	44.66	12.85	1.94	40.58	11.71	1.75	28.38	13.0	2.08
ABIC Avg. Scaled Score	43.63	8.35	1.26	38.22	10.66	1.59	28.31	10.98	1.76

$\bar{X}$ = 50, s.d. = 15

highest, followed by the learning disabled group, followed by the ER group. A one way ANOVA on the ABIC Average Scaled Score indicated that there were significant differences by group ($F(2, 125) = 24.59$, $p < .001$). A Bonferroni multiple contrast procedure indicated that the three groups were all different from each other.

Characteristics of Student Performance: SOMPA Medical Model Measures. One of the features of the SOMPA battery is that it assumes that there are three separate models which correspond to different ways of examining student behavior and/or characteristics. One of these models is the Medical Model under which it is assumed that low scores on this type of measure indicate within-child deficits which may exist in spite of the fact that these deficits are not publicly identified. The SOMPA battery includes several of these measures, including the Physical Dexterity Tasks, the Bender Gestalt, and the Health History Inventories. The summary statistics for these measures are provided in Table IV(13).

Comparison of the scaled scores on the Physical Dexterity Tasks indicates that the nonhandicapped students and the learning disabled students tended to score at about the same level on the various subscales, and both groups tended to score higher than the ER group. A one-way ANOVA on the Physical Dexterity Average Scaled Score indicated significant differences between groups ($F(2, 108) = 36.73$, $p < .001$). A Bonferroni multiple contrast procedure indicated that the nonhandicapped and learning groups did not differ from each other, but both groups differed from the ER group. In addition, a one-way ANOVA on the Bender Gestalt test indicated that there were significant differences between the groups ($F(2, 108) = 36.73$, $p < .001$). A

Table IV(13)

Summary Statistics for the SOMPA Medical Model Measures by Diagnostic Group

Measure	Nonhandicapped			Learning Disabled			ER		
	$\bar{X}$	s.d.	std. error of mean	$\bar{X}$	s.d.	std. error of mean	$\bar{X}$	s.d.	std. error of mean
Physical Dexterity Tasks									
Ambulation*									
Equilibrium	37.64	21.39	3.43	31.61	19.16	3.19	14.98	12.57	2.09
Placement	61.59	10.94	1.75	60.67	13.97	2.33	29.69	21.97	3.66
Fine Motor Sequencing	32.53	19.68	3.15	31.11	20.21	3.37	18.17	15.49	2.58
Finger Tongue Dexterity	47.0	19.77	3.17	40.81	22.63	3.77	25.94	15.34	2.56
Involuntary Movement	52.41	18.38	2.94	37.42	18.32	3.05	15.83	13.10	2.18
Physical Dexterity Average Scaled Score	37.87	16.49	2.64	38.44	17/61	2.93	17.53	15.44	2.57
Bender Gestalt*	44.92	13.37	2.0	40.28	12.52	2.09	20.21	13.52	2.25
Health History Inventories	47.38	12.89	2.06	45.08	16.14	2.69	17.22	15.48	2.58
Prenatal/Postnatal**	1.86	2.73	.41	2.2	1.75	.26	4.87	5.9	.94
Trauma	3.2	6.6	.10	2.56	4.63	.69	5.95	9.62	1.54
Disease & Illness	1.84	2.93	.44	1.64	1.57	.23	4.18	4.30	.69
Vision	.48	1.07	.16	.56	1.45	.22	.95	1.49	.24
Learning	.05	.21	.03	.09	.29	.04	.13	.52	.08

*Scaled scores reported.

**Raw Scores reported.

Bonferroni multiple contrast procedure indicated that the nonhandicapped and learning disabled groups did not differ from each other, but both groups did differ from the ER group.

In order to compare the group differences on the subscales of the Health History Inventories, a series of one-way ANOVA's were done using the appropriate subscale score in each case. The results of these analyses are contained in Table IV(14).

Table IV(14)

Results of Analyses of Variance with the Subscales of the Health History Inventories

Subscale	F value	df	p	Group Differences
Prenatal/Postratal	8.19	2,125	<.001	NH & LD ≠ ER
Trauma	2.62	2,125	NS	No difference
Disease & Illness	8.57	2,125	<.001	NH & LD ≠ ER
Vision	1.44	2,125	NS	No difference
Hearing	.56	2,125	NS	No difference

As the table demonstrates, the only measures on which the groups differed were the prenatal/postnatal subscale and the disease and illness subscale. In each case, the ER group differed from the other groups, but the other groups did not differ from each other.

Characteristics of Student Performance: SOMPA Sociocultural Scales. The final set of measures on which the students in the study were compared consisted of the Sociocultural Scales of the SOMPA battery. These subscales are used in the SOMPA assessment system to

determine the "sociocultural space" which best characterizes a given student and to assist in the calculation of the estimated learning potential scores. However, these measures provide important information as an independent source of data about a given child's background. The four subscales of the Sociocultural Scales are family size, family structure, socioeconomic status, and urban acculturation. A summary of descriptive statistics is contained in Table IV(15).

Table IV(15)

Descriptive Statistics for the SOMPA Sociocultural Scales by Diagnostic Group

	Nonhandicapped			Learning Disabled			ER		
Measure	$\bar{X}$	s.d.	std. error of mean	$\bar{X}$	s.d.	std. error of mean	$\bar{X}$	s.d.	std. error of mean
Family Size	9. _	4.58	.68	9.6	3.8	.57	8.03	4.36	.70
Family Structure	15.02	5.10	.76	15.38	4.58	.68	13.18	6.24	.10
Socioeconomic Status	4.73	2.46	.37	3.9	2.65	.40	3.72	2.61	.42
Urban Acculturation	14.31	7.30	1.09	13.2	8.87	1.32	17.92	14.13	2.26

A series of one-way ANOVA's one each of the subscale measures indicated that there were no significant differences between the groups on any of the measures, indicating that the groups were equivalent on these important background indicators.

The final part of the results section is focused on a comparison of the diagnostic categories assigned by the public schools with those arrived at through the use of the measures in the present study.

B. Correspondence of School-Assigned Diagnostic Categories with Diagnostic Categories derived from Measures used in the Present Study

The final part of the data analysis was concerned with an examination of the diagnostic categories suggested through the use of the WISC-RM, the SOMPA, and the S-KABC from those assigned by the schools. This analysis was carried out as a two-step process. First, State of California guidelines for the determination of learning handicaps were used to calculate discrepancy scores (see Appendix A) and determine the numbers of students in the sample who would be categorized as learning disabled. The achievement measure used in each discrepancy calculation was the S-KABC Achievement Subscale. Using this procedure, discrepancy scores were calculated separately for each of th measures, and a determination was made using the State criteria what percentage of the subgroups in the study met the established criteria. The results of these calculations are prese in Table IV(16).

The second part of this analysis focused on determining the numbers of students in each group who would be considered ER using the measures of the present study. Students' intellectual functioning scores and adaptive behavior scores were compared in order to determine the numbers of students who fell below two standard deviations on both measures simultaneously. The adaptive behavior measure used in each calculation was the ABIC from the SOMPA battery. Using this criteria, a determination was made regarding the percentage of each diagnostic group that would be considered ER using each measure. The results of these determinations are presented in Table IV(17).

Table IV(16)

Numbers and Percent of Students from Original Diagnostic Categories who Meet the Discrepancy Criteria for Learning Disability

	Nonhandicapped		Learning Disabled		ER	
Measure	# Not Eligible (%)	# Eligible (%)	# Not Eligible (%)	# Eligible (%)	# Not Eligible (%)	# Eligible (%)
WISC-RM Verbal	42	6	39	8	34	10
	(87.5)	(12.5)	(82.98)	(17.02)	(77.27)	(22.73)
WISC-RM Perf.	23	25	14	33	23	21
	(47.9)	(52.08)	(29.79)	(70.21)	(52.27)	(47.73)
WISC-RM Full Scale	31	17	23	24	27	17
	(64.58)	(35.42)	(48.94)	(51.06)	(61.36)	(38.64)
ELP Verbal	18	30	16	31	16	28
	(37.5)	(62.5)	(34.04)	(65.96)	(36.38)	(63.64)
ELP Perf.	16	32	11	36	16	28
	(33.33)	(66.67)	(23.40)	(76.60)	(36.36)	(63.64)
ELP Full Scale	13	35	11	36	16	28
	(27.08)	(72.92)	(23.40)	(76.60)	(36.36)	(63.64)
S-KABC Sequential	31	17	24	23	29	15
	(64.58)	(35.42)	(51.06)	(48.94)	(65.91)	(34.09)
S-KABC Simultaneous	29	19	11	36	20	24
	(60.42)	(39.58)	(23.40)	(76.60)	(45.46)	(54.55)
S-KABC MPC	25	23	17	30	30	14
	(52.08)	(47.92)	(36.17)	(63.83)	(68.18)	(31.82)
S-KABC Nonverbal	29	19	13	34	29	15
	(60.42)	(39.58)	(27.66)	(72.34)	(65.91)	(34.09)

Table IV(17)

Numbers and Percent of Students from Original Diagnostic Categories who Meet the Criteria for ER

Measures	Nonhandicapped # Not Eligible (%)	Nonhandicapped # Eligible (%)	Learning Disabled # Not Eligible (%)	Learning Disabled # Eligible (%)	ER # Not Eligible (%)	ER # Eligible (%)
WISC-RM Verbal/ABIC	45 (93.75)	3 (6.25)	44 (93.62)	3 (6.38)	28 (63.64)	16 (36.36)
WISC-RM Perf./ABIC	45 (93.75)	3 (6.25)	45 (95.75)	2 (4.26)	30 (68.18)	14 (31.82)
WISC-RM Full Scale/ABIC	45 (93.75)	3 (6.25)	45 (95.75)	2 (4.26)	28 (63.64)	16 (36.36)
ELP Verbal/ABIC	45 (93.75)	3 (6.25)	45 (95.75)	2 (4.26)	30 (68.18)	14 (31.82)
ELP Perf./ABIC	45 (93.75)	3 (6.25)	45 (95.75)	2 (4.26)	31 (70.46)	13 (29.55)
ELP Full Scale/ABIC	45 (93.75)	3 (6.25)	45 (95.75)	2 (4.26)	29 (65.91)	15 (34.09)
S-KABC Sequential/ABIC	44 (91.67)	4 (8.33)	44 (93.62)	3 (6.38)	31 (70.46)	13 (29.55)
S-KABC Simultaneous/ABIC	44 (91.67)	4 (8.33)	45 (95.75)	2 (4.26)	31 (70.46)	13 (29.55)
S-KABC MPC/ABIC	44 (91.67)	4 (8.33)	45 (95.75)	2 (4.26)	30 (68.18)	14 (31.82)
S-KABC Nonverbal	44 (91.67)	4 (8.33)	45 (95.75)	2 (4.26)	30 (68.18)	14 (31.82)

As Table IV(16) shows, there are some discrepancies between the original diagnostic categories and the numbers of students who need the discrepancy criteron for learning disability. In general, the SOMPA measures tend to result in the largest numbers of students eligible, indicating a greater discrepancy between these intellectual functioning measures and Achievement. On the other hand, the WISC-RM measures tend to result in the fewest numbers of students eligible, indicating a smaller discrepancy between the measures used. Interestingly, there are sizable numbers of students in the learning disabled category who do not meet the discrepancy criteria based upon the present measures. The percentages of students in this group who were not eligible ranged from 23% using the S-KABC Simultaneous Processing Score, the ELP Performance Score, or the ELP Full Scale Score, to 83% using the WISC-RM Verbal Score. Obviously, the discrepancy is not the only factor that is used to determine that a given child is learning disabled, as reference to the State guidelines demonstrate. For example, there is provision made for "professional judgment," and for the measurement of the discrepancy "... by alternative means as specified on the assessment." Nevertheless, the table does indicate the numbers of students meeting the discrepancy criteria when the instruments from the present study are used as the basis for the calculation.

Table IV(17), indicating the numbers of students who would appear to be eligible for ER status, is somewhat more clearcut. For the nonhandicapped and learning disabled groups, the numbers of students who would be considered ER are small and fairly constant across instruments. On the other hand, there are large numbers of ER

students who would not be considered ER using the measures in the present study, and this did not vary greatly according to the measure used.

V. Discussion

As was indicated earlier, there were two major concerns in the present investigation. One major set of issues revolved around student performance on what might be considered linguistically appropriate psychometric measures for this population of students. A central concern was the ability of the measures to differentiate between the diagnostic subgroups used in the study, as well as the technical adequacy of each individual measure. A second set of issues concerned the diagnostic classifications suggested as a result of student performance on each of the measures. More specifically, the key issue here was the comparison of the school-assigned diagnostic labels with the labels derived from the measures used in the present study. Each of these sets of issues will be addressed in turn.

Student Performance and Technical Adequacy of the Measures

The intellectual functioning measures. Because of the importance of intellectual functioning estimates in the assignment of a variety of diagnostic labels, and the central place that these types of measures hold in educational decision making, the data on student performance was of key interest in this study. Examination of the scores on the WISC-RM and on the SOMPA indicate that there are large differences between the verbal and performance scores. This large discrepancy was characteristic of all three diagnostic groups, and suggests that students in the sample found the verbal portions of the measures problematic is spite of the fact that the assessments were conducted in Spanish. Comparison of the WISC-RM Full Scale IQ, the SOMPA ELP, and the S-KABC indicated that the students in all three groups scored lowest on the WISC-RM and highest on the SOMPA, with scores on the

S-KABC in between. The fairly large discrepancy between the WISC-RM and the SOMPA scores can be used as an estimate of the effect of sociocultural variables on performance, since the SOMPA scores are systematically adjusted to account for the effect of these variables.

When the WISC-RM was examined by means of mean subscale scores and the intercorrelations of those scores, an interesting pattern emerged. The information and similarities subscales tended to be lower and not to correlate well with the other subscale measures in either the performance or verbal portions of the test. In contrast, the intercorrelations of the S-KABC resulted in moderate to high coefficients. It appears that overall the two tests exhibit some internal consistency as indicated by the relationships of the subscale scores, with the exception of the Information and Similarities subtests of the WISC-RM.

The ability of the cognitive measures to discriminate between the three diagnostic groups was primarily investigated through the use of analysis of variance procedures. For example, according to the definitions of learning disability, the intellectual functioning levels should have been in the normal range, while the definition of ER stipulates that the functioning levels should have been significantly below average. Theoretically, this should have resulted in a pattern on the measures in which the nonhandicapped students and the learning disabled students did not differ form each other, but both groups should have been different from the ER students. Examination of the analyses for the Full Scale scores on the WISC-RM IQ and the SOMPA ELP indicated that these measures significantly differentiated among the three groups. That is, the learning disabled and nonhandicapped groups

were found to be significantly different. However, when the analyses with Verbal and Performance scores are examined, the nonhandicapped and the learning disabled students did not score differently from each other on the Performance IQ and Performance ELP, but did on the Verbal IQ and the Verbal ELP. Overall, when separate consideration is given to the performance scores for the learning disabled students, a functioning level comparable to the nonhandicapped group emerges. In essence, this suggests that the increased reliance on the Performance scores by the diagnostic team, even when the test is administered in Spanish, may be important in the evaluation of learning disabilities. In contrast to the pattern found with the learning disabled students, the verbal and performance scores for the ER group tend to be categorically low, suggesting a characteristic "flat' profile.

The S-KABC WAS found to differentiate between all three groups only on the Sequential Processing Score. On the remaining scores which this measure yields, including the Mental Processing Composite, the nonhandicapped and learning groups did not differ from each other, but both groups did significantly differ from the ER students. Overall, it appears that this intellectual functioning measure differentiates between the school-designated diagnostic categories in the expected fashion, with the exception of the Sequential Processing Scale

From a diagnostic perspective, it appears that the Verbal and Full Scale IQ scores on the WISC-RM may tend to yield depressed intellectual functioning estimates for learning disabled students, as compared to the Performance IQ's. Although the SOMPA adjustments had the effect of inflating the Verbal and Full Scale scores of the students in the sample, it did so for all the students and therefore this same pattern of results was found with the SOMPA ELP's.

In general, the WISC-RM and the S-KABC tend to yield fairly high reliability coefficients for all of the subscales. As will be recalled, the two reliability estimates used in the present study were essentially estimates of internal stability and stability or internal consistency. It should be noted that there are other methods of estimating reliability which could not be calculated with the present data, such as test-retest reliability. However, the reliability estimates from the present sample indicate that the two measures demonstrate acceptable reliability levels with this population. Further, the standard error of measurement, which provides an error estimate for the true range of a score, is fairly low for all of the subscales of the tests. Finally, the high correlations of the WISC-RM IQ, the SOMPA Estimated Learning Potential, and the Mental Processing Composite Score of the S-KABC indicates that the measures have a very high degree of association, and is one indication of concurrent validity.

Comparisons based on academic achievement. In addition to the intellectual functioning measures, one of the measures that was centrally important to the present study was the academic achievement measure. The measure used in this study, as will be recalled, was the Achievement Scales of the S-KABC. According to the established definitions, the rank order of scores on the test should have resulted in a pattern such that the nonhandicapped students would be first, followed by the learning disabled and then the ER students. This is exactly the pattern that was found in the data. That is, the analyses of variance indicated that the three groups were all significantly different from each other. Interestingly, the achievement measure was

fairly highly correlated with the intellectual functioning measures, in the range of .75 to .89, but not to the degree that the intellectual assessment measures were intercorrelated.

As was found with the intellectual assessment measures, both types of estimates of reliability calculated for this measure yielded fairly high reliability coefficients, indicating that it exhibits an acceptable degree of consistency for use with this population.

The role of adaptive behavior. Since the diagnostic category of mild mental retardation has occupied such a central place in much of the literature on non-biased assessment, the inclusion of an adaptive behavior measure was a key part of this investigation. The measure used, as will be recalled, was the ABIC from the SOMPA battery. Again, when reference is made to the definitions of the categories of learning disability and mild mental retardation, it would be expected that the rank order of performance on the ABIC would be expected that the rank order of performance on the ABIC would result in a pattern of the nonhandicapped students highest, followed by the learning disabled at or slightly below this level, followed finally by the ER students. In fact, examination of the scores on the subscales of the ABIC demonstrates just such a pattern. When the Average Scaled Score, a composite measure of overall performance is used as a dependent measure, all three groups differed significantly from each other. In general, the ABIC appeared to be an effective and sensitive measure for the purpose of differentiating adaptive behavior levels in the sample from a statistical standpoint.

Although academic skills are often the primary focus in the determination of learning disabilities, examination of the ABIC scores

among the groups in the study suggests that there may be a behavioral component associated with the learning disabled students. That is, slightly depressed adaptive behavior levels may be an important secondary indicator for students such as those in the sample, although the generality of this pattern with language minority students remains to be seen.

Medical and neurological differences. One of the often discussed but rarely measured aspects of student performance is that component associated with neurological intactness, and prior and/or early medical history. Fortunately, the SOMPA battery includes measures of both domains. It might have been expected that the measures would have differentiated the ER group from the other groups, given the association between early medical and health problems and later academic difficulties. In addition, it may have been expected that the learning disabled group would exhibit differences with respect to the nonhandicapped group, given the role of early trauma and neurological disturbances that some have focused on as being related to the development of later learning disabilities. Yet, comparison of the Physical Dexterity Score, the Bender Gestalt Score, and the Health History Inventories indicated that the learning disabled group was not different from the nonhandicapped. This suggests that factors other than neurological intactness and early medical/health problems likely play a more important role in any learning problems exhibited by the learning disabled group. In contrast, the role of these factors appears to be more substantial for the ER group, as they differed on the Physical Dexterity Score, the Bender Gestalt, and two of five subscales of the Health History Inventories (prenatal/postnatal problems, and disease and illness).

From a research design perspective, it was important that the groups in the study did not differ on important background characteristics that may have confounded differences on the other measures. Fortunately, the Sociocultural Scales of the SOMPA battery offered a way of empirically examining differences between the three groups on the domains of family size, family structure, SES and urban acculturation. As the results indicated earlier, there were no significant differences among the groups on any of the measures. These results add confidence that differences between the groups on the remaining measures can be considered as true differences and not extraneous variance due to prior unequal status on relevant background factors.

In sum, the measures used in the present study, which can be considered linguistically appropriate assessment tools for the population under consideration, was able to differentiate, in a statistical sense, the groups comprising the sample. In addition, given the admittedly limited measures of technical soundness calculated for each of the measures, the measures appear to be sound in a psychometric sense. Nevertheless, more comprehensive validity and factor analytic studies are needed to assure the appropriateness of these measures for students similar to those participating in this study.

Examination of School-Assigned Diagnostic Classifications

One of the central purposes of this study was to compare the diagnoses suggested by the heretofore unavailable linguistically appropriate psychometric measures with the prior categories assigned by the public schools in which the students were labeled. As described

earlier, this was a two step process, first focusing on the category of learning disability, and then on the diagnostic category of mild mental retardation or ER.

Eligibility for the category of LD. As stated before, the calculations for the required discrepancy between intellectual functioning and academic achievement for a learning disability were based upon State of California guidelines, and the academic achievement measure used in these calculations was the Achievement Scale of the S-KABC. These calculations resulted in the numbers of students in each of the diagnostic category who met the required discrepancy. In general, the greatest numbers of students in all of the categories met the required discrepancy with the SOMPA scores, and the fewest met the required discrepancy with the WISC-RM scores, especially when the Verbal IQ was used. It appears that the tendency of the SOMPA to provide somewhat higher estimates of intellectual functioning than the WISC RM or the S-KABC resulted in greater discrepancies between intellectual functioning and achievement, thereby resulting in greater numbers of students eligible for the learning disability designation. As an example, about 73% of the nonhandicapped students met the required discrepancy using the Full Scale ELP, about 76% of the LD students met the required discrepancy using the Performance or Full Scale ELP, and about 64% of the ER students met the required discrepancy using the Verbal, Performance, or Full Scale ELP's. In contrast, when the WISC-RM Verbal IQ was used in the calculations, about 13% of the nonhandicapped, 17% of the LD, and about 23% of the ER students met the required discrepancy. Again, this appeared to be primarily due to the fact that students exhibited depressed scores on the Verbal Subscales of the WISC-RM.

The fact that such a large number of students in the nonhandicapped group met the required discrepancy for learning disability is somewhat troubling, suggesting that many of these students are at risk for a variety of negative academic outcomes including referral for special education. On the other hand, it also suggests that alternatives other than special education referral are being used to address the lowered academic achievement of these students. The fact that there were significant numbers of nonhandicapped students who met the required discrepancy, as well as the fact that there appeared to be large numbers of LD students who did not meet the required discrepancy, cannot be taken as a direct measure of inaccuracy in school diagnostic procedures. Reference to the state law, for example, indicates that professional judgement, as well as consideration of informal testing, criterion referenced tests, work samples, and classroom performance and observation can all be used to determine the existence of a learning disability. From the present data, it was not possible to determine the extent to which these or other factors may have played a part in previous school-based decisions to refer or not refer low achieving students in the study. Nevertheless, the present data suggest that use of the measures indicated in this study may lead to different diagnostic decisions than those used by other measures or alternative assessment (i.e., informal) procedures. Further, even within the set of measures used in this study, there was a great deal of variance about the suggested diagnostic category. This would appear to indicate the importance of clinical judgment and consideration of multiple sources of information in arriving at appropriate diagnostic decisions for individual students.

Eligibility for the category of ER. Examination of the estimations of the numbers of students eligible for the ER classification provided equally interesting data. For example, a small percentage of nonhandicapped and LD students met the required cutoffs on both the ABIC and the individual measures of intellectual functioning, in the range of 4 to 8%. In contrast to the variation found between the instruments for the LD calculations, this percentage tended to be constant for these two groups across all the measures. Again, this suggests that some of the nonhandicapped students are at risk, and that current problems are being dealt with in ways that do not involve the special education system. It is possible that the major alternative intervention in these cases is provided through the bilingual education system. It also suggests that there are some LD students who exhibit significant deficits in adaptive behavior, and as mentioned previously, these may be secondary indicators characteristic of learning disabled students.

One of the major findings of the present study was that approximately a third of the students classified as ER would not be assigned this classification if the instruments employed in the present study were employed. This finding was fairly constant across all of the intellectual functioning measures for this group. Although the presence of lowered levels of academic achievement for this group cannot be denied, the battery of linguistically appropriate instruments used in this study suggest that ER placements might not be the appropriate educational response. The fact that the mean score on the ABIC for this group was greater (about eight points higher) than the ≤ 2 s.d. cutoff stipulated for a designation of mild mental retardation

suggests that many of the students in this group exhibit depressed but not severely deficient adaptive behavior skills.

One of the primary implications from this study is that consideration should be given to the use of linguistically appropriate assessment tools when consideri g educational decisions for students such as those in this sample. Although such instruments have been unavailable up to this point, failure to account for the linguistic backgrounds of language minority students in the assessment and diagnostic process invites the possibility of erroneous and inappropriate classification and intervention.

References

Brady, P. M., Manni, J. L., & Winikur, D. W. (1983). Implications of ethnic disproportion in programs for the educable mentally retarded. The Journal of Special Education, 17, 295-302.

Brigham, C. C. (1930). Intelligence tests of imigrant groups. Psychological Review, 37, 158-165.

Campione, J. C., Brown, A. L., & Ferrara, R. A. (1982). Mental retardation and intelligence. In R. Sternberg (Ed.), The handbook on human intelligence. Cambridge, MA: Cambridge University Press.

Cleary, A. T., Humphreys, L. G., Kendrick, S. A. & Wesman, A. (1975). Educational uses of tests with disadvantaged students. American Psychologist, 30, 15-41.

Connecticut Special Education Association. (1936). Development and progress of special classes for mentally deficient children in Connecticut. New Haven, CT: Columbia Printing Company.

Cronbach, L. J. (1975). Five decades of public controversy over mental testing. American Psychologist, 30, 1-14.

Cummis, J. (1984). Bilingualism and special education: Issues in assessment and pedagogy. San Diego, CA: College Hill Press.

Darcy, N. T. (1952). The performance of bilingual Puerto Rican children on verbal and non-language tests of intelligence. Journal of Educational Research, 45, 499-506.

Darcy, N. T. (1963). Bilingualism and the measurement of intelligence: Review of a decade of research. The Journal of Genetic Psychology, 103, 259-282.

Dean, R. S. (1977a). Internal consistency of the PIAT with Mexican American children. Psychology in the Schools, 14, 167-168.

Dean, R. S. (1977b). Reliability of the WISC-R with Mexican American Children. Journal of School Psychology, 15, 267-268.

Dean, R. S. (1979). Predictive validity of the WISC-R with Mexican American children. Journal of School Psychology, 17, 55-58.

Dean, R. S. (1979). Distinguishing patterns for Mexican American children on the WISC-R. Journal of Clinical Psychology, 35, 790-794.

DeBlassie, R. R. (1980). Testing Mexican American Youth. Hingham, MA: Teaching Resources Corporation.

Diana vs. California State Board of Education. United States District Court, Northern California District, C-70 37 RFP 1969.

Figueroa, R. A. (1980). Intersection of special education and bilingual education. In J. E. Alatis (Ed.), Current issues in bilingual education. Washington, DC: Georgetown University Press.

Figueroa, R. A. (1983). Test bias and Hispanic children. The Journal of Special Education, 17, 431-440.

Figueroa, R. A. (1985). Diana Revisted. Los Angeles: National Dissemination and Assessment Center, Los Angeles State University.

Figueroa, R. A., Sandoval, J., & Merino, B. School psychology and limited-English-proficient children: New competencies. Journal of School Psychology, 22, 131-144.

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

Gandara, P., Keogh, B. K. & Yoshioka-Maxwell, B. (1980). Predicting academic performance of Anglo and Mexican American kindergarten children. Psychology in the Schools, 17, 174-177.

Garcia, J. (1977). Intelligence testing: Quatients, quotas, and quackery. In J. L. Martinez (Ed.), Chicano Psychology. New York: Academic Press.

Goldman, R. D., & Hewitt, B. N. (1975). An investigation of test bias for Mexican American college students. Journal of Educational Measurement, 12, 187-196.

Goodman, J. F. (1979). Is tissue the issue? A critique of SOMPA's models and tests. School Psychology Digest, 8, 47-62.

Grossman, H. J. (1977) Manual on terminology and classification in mental retardation. Washington, DC: American Association on Mental Deficiency.

Gutkin, T. B. (1979). Bannatyne patterns of Caucasian and Mexican American learning disabled children. Psychology in the Schools, 16, 178-183.

Gutkin, T. B., & Reynolds, C. R. (1980). Factorial similarity of the WISC-R for Anglos and Chicanos referred for psychological services. Journal of School Psychology, 18, 34-39.

Heller, K. A., Holtzman, W. H., Messick, S. (1982). Placing children in special education: A strategy for equity. Washington, DC: National Academy Press.

Henderson, R. W., & Rankin, R. J. (1973). WPPSI reliability and predictive validity with disadvantaged Mexican American children. Journal of School Psychology, 11, 16-20.

Henderson, R. W., & Valencia, R. R. (In press). Nondiscriminatory school psychological servides: Beyond nonbiased assessment. In J. R. Bergan (Ed.), School psychology and contemporary society. Columbus, OH: Merrill.

Hickey, T. (1972). Bilingualism and the measurement of intelligence and verbal learning ability. Exceptional children, 39, 27-28.

Hobsen vs. Hansen. United States District Court, District of Columbia, 269F, Supp. 401 (D.O.C. 1967).

Holtzman, W. H., Diaz-Guerrero, R., & Swartz, J. D. (1975). Personality development in two cultures. Austin, TX: University of Texas Press.

Hurt, M., & Mishra, S. (1970). Reliability and validity of the Metropolitan Achievement Test for Mexican American children. Educational and Psychological Measurement, 30, 989-992.

Jensen, A. R. (1974). How biased are culturally-loaded tests? Genetic Psychology Monographs, 90, 185-244.

Jensen, A. R. (1976, December). Test bias and construct validity. Phi Delta Kappan, 340-346.

Jensen, A. R. (1982). Bias in mental testing. New York: The Free Press. Considerations, philosophical rationale, and implications for practice. Journal of School Psychology, 21, 13-21.

Jirsa, J. E. (1983). The SOMPA: A brief examination of technical

Kamin, L. (1974). The science and politics of I.Q. Potomac, MD: Erlbaum.

Kaufman, A. S., & Kaufman, N. L. (1982). Kaufman Assessment Battery for Children. Circle Pines, MN: American Guidance Service.

Keller, G. D., & Van Hooft, K. S. (1982). A chronology of bilingualism and bilingual education in the United States. In J. A. Fishman & G. D. Keller (Eds.), Bilingual education for Hispanic students in the United States. New York: Teachers College, Columbia University.

Keston, M. J., & Jimenez, C. (1954). A study of the performance on English and Spanish editions of the Stanford-Binet intelligence test by Spanish-American children. The Journal of Genetic Psychology, 85, 263-269.

Killian, L. R. (1971). WISC, Illinois Test Psycholinguistic Abilities, and Bender Visual-Motor Gestalt Test performance of Spanish-American kindergarten and first-grade school children. Journal of Consulting and Clinical Psychology, 37, 38-43.

Kirk, S. A. (1972). Ethnic differences in psycholinguistic abilities. Exceptional children, 39, 112-118.

Koppitz, E. M. (1975). The Bender Gestalt Test for young children, Vol. II. New York: Drew and Straton.

Laosa, L. M. (1982). Psychometric characteristics of Chicano and non-Hispanic white children's performance on the Preschool Inventory. Journal of Applied Development Psychology, 3, 217-245.

Larry P. vs. Wilson Riles, No. C-71-2270 RFP, U.S. District Court for the Northern District of Caliornia, October 16, 1979.

Linn, R. L., & Werts, C. E. (1971). Consideration for studies of test bias. Journal of Educational Measurement, 8, 1-4.

Mercer, J. R. (1973). Labelling the mentally retarded. Berkeley, CA: University of California Press.

Mercer, J. R. (1979). The System of Multicultural Pluralistic Assessment: Technical manual. New York: Psychological Corporation.

Miller, H. A. (1916). The school and the immigrant. Cleveland, OH: The Survey Committee of the Cleveland Foundation.

Mishra, S. P. (1980). The influence of examiners' ethnic attributes of intelligence test scores. Psychology in the schools, 17, 117-122.

Mowder, B. A. (1979). Assessing the bilingual handicapped student. Psychology in the schools, 16, 43-50.

Mowder, B. A. (1980). A strategy for the assessment of bilingual handicapped children. Psychology in the Schools, 17, 7-11.

Myers, B., & Goldstein, D. (1979). Cognitive development in bilingual and monolingual lower-class children. Psychology in the Schools, 16, 137-142.

Omark, D. R., & Erickson, J. G. (1983). The bilingual exceptional child. San Diego, CA: College-Hill Press.

P.A.S.E. vs. Hannon, No. 74C3586, U.S. District Court for the Northern District of Illinois, Eastern Division, 1980.

Peal, E., & Lambert, W. E. (1962). The relation of bilingualism to intelligence. Psychological Monographs: General and Applied, 76, 27

Peterson, N., & Novick, M. (1976). An evaluation of some models for culture-fair selection. Journal of Educational Measurement, 13, 3-29.

Reschly, D. J. (1982). Assessing mild mental retardation: The influence of adaptive behavior, sociocultural status, and prospectus for nonbiased assessment. In C. R. Reynolds & T. B. Gutkin (Eds.), The handbook of School Psychology. New York: Wiley.

Reschly, D. J., & Sabers, D. L. (1979). Analysis of test bias in four groups with regression definition. Journal of Educational Measurement, 16, 1-9.

Reynolds, A. (1933). The education of Spanish-speaking children in five southwestern states. Washington, DC: U.S. Department of the Interior (Bulletin No. 11, 1933).

Reynolds, C. R. (1982). The problem of bias in psychological assessment. In C. R. Reynolds & T. B. Gutkin (Eds.), The handbook of school psychology. New York: Wiley.

Reynolds, C. R., & Gutkin, T. B. (1982). The handbook of school psychology. New York: Wiley.

Sanchez, G. I. (1932). Group differences and Spanish-speaking children: A critical review. Journal of Applied Psychology, 16, 549-558.

Sanchez, G. I. (1934). Bilingualism and mental measures. The journal of Applied Psychology, 18, 765-772.

Sandoval, J. (1979). The WISC-R and internal evidence of test bias with minority groups. Journal of Consulting and Clinical Psychology, 47, 919-927.

Sandoval, J., & Miille, M. P. W. (1980). Accuracy of judgments of WISC-R item difficulty for minority groups. Journal of Consulting and Clinical Psychology, 48, 249-253.

Sandoval, J., Zimmerman, I. L., & Woo-Sam, J. M. (1983). Cultural differences on WISC-R verbal items. Journal of School Psychology, 21, 49-55.

Sattler, J. (1982). Assessment of children's intelligence and special abilities. Boston, MA: Allyn and Bacon.

Schmidt, F. L., & Hunter, J. E. (1974). Racial and ethnic bias in psychological tests: Divergent implications of two definitions of test bias. American Psychologist, 29, 1-8.

Talley, R. C. (1979). Evaluating the effects of implementing the System of Multicultural Pluralistic Assessment: A qualitative perspective. School Psychology Digest, 8, 71-78.

Thomas, A., Hertzig, M. E., Dryman, I., & Fernandez, P. (1971). Examiner effect in IQ testing of Puerto Rican working-class children. American Journal of Orthopsychiatry, 41, 809-821.

Thompson, B., Alston, H. L., & Say, M. W. (1978). Correspondence across three ethnic groups of constructs measured by the Iowa Test of Basic Skills. Psychology in the Schools, 15, 347-349.

Tryon, W. W. (1979). The test-trait fallacy. American Psychologist, 34, 402-405.

Tucker, J. A. (1980). Ethnic proportions in classes for the learning disabled: Issues in non-biased assessment. The Journal of Special Education, 14, 93-105.

Twomey, S. C., Gallegos, C., Anderson, L., Williamson, B., & Williamson, J. (1980). A study of the effectiveness of various nondiscriminatory and linguistically and culturally appropriate assessment criteria for placement of minority students in special education programs. Sacramento, CA: California State Department of Education, Office of Special Education.

Valencia, R. R. (1979). Comparison of intellectual performance of Chicano and Anglo third-grade boys on the Raven's Coloured Progressive Matrices. Psychology in the Schools, 16, 448-453.

Valencia, R. R., Henderson, R. W., & Rankin, R. J. (1981). Relationship of family constellation and schooling to intellectual performance of Mexican American children. Journal of Educational Psychology, 73, 524-532.

Zigler, E., & Balla, D. (1982). Mental Retardation: The developmental difference controversy. Hillsdale, NJ: Erlbaum.

Zigler, E., Balla, D., & Hodapp, R. (1984). On the definition and classification of mental retardation. American Journal of Mental Deficiency, 89, 215-230.

Zirkel, P. A. (1972, June). Spanish-speaking students and standardized tests. The Urban Review, 32-40.

APPENDIX A

A Manual for

The Determination of a Severe Discrepancy as Defined by Title 5, CAC, Section 3030(j)

Prepared for: California State Department of Education
Office of Special Education

by: California State Department of Education
Office of Program Evaluation and Research

CALIFORNIA STATE DEPARTMENT OF EDUCATION
Bill Honig—Superintendent of Public Instruction
Sacramento, 1983

THE DETERMINATION OF A SEVERE DISCREPANCY

The determination of whether a pupil has a severe discrepancy between intellectual ability and achievement requires a comparison of the difference between the pupil's ability and achievement scores against a specified criterion value.* There are two basic steps in determining whether a pupil manifests a severe discrepancy: 1) calculate the difference between the pupil's ability and achievement scores, and 2) compare that difference to a specified criterion value to see if the pupil's difference meets or exceeds the criterion value.

Included in this document are three different procedures which can be used to determine whether a pupil has a severe discrepancy between intellectual ability and achievement. The choice of procedure depends on the particular intellectual ability/achievement test combination you are using. The three procedures are: 1) reading specially prepared charts, 2) using a general computational procedure, and 3) using an estimation procedure. The following chart indicates which procedure you should use based on what information is available to you:

Information Available To You	Procedure Of Choice
1. Pupil scores on tests specified on page 2.	Charts
2. Pupil scores on tests <u>not</u> given above; test means and standards deviations; correlation between tests.	General computational procedure.

* The criterion value, as specified in CAC Section 3030 (j)(4), is defined as 1.5 times the standard deviation of the difference distribution.

Information Available To You	Procedure Of Choice
3. Pupil percentile rank scores on tests not given above; correlation between tests.	General computational procedure.
4. Pupil scores, no correlation between tests.	Estimation procedure.

Following is a brief description of the three procedures and page references for the more detailed explanation of each procedure in this document:

1. Charts of intellectual/achievement test combinations which can be used in determining if a pupil's scores indicate a severe discrepancy. Charts are presented for these commonly used test combinations:

Ability Test		Achievement Test
Wechsler Intelligence Scale for Children Revised (WISC-R)	&	Wide Range Achievement Test (WRAT)
Wechsler Intelligence Scale for Children Revised (WISC-R)	&	Peabody Individual Achievement Test (PIAT)
Wechsler Intelligence Scale for Children Revised (WISC-R)	&	Woodcock-Johnson Psycho-Educational Battery
Stanford-Binet Intelligence Scale (L-M)	&	Wide Range Achievement Test (WRAT)

These tests were selected based on a statewide survey of test use and on the availability of correlational data. Details are available in a technical supplement from the Office of Program Evaluation and Research (OPER), State Department of Education.

Beginning on page 5 are 26 charts which represent combinations of the above tests and subtest combinations. Charts could not be developed for combinations (e.g., Stanford-Binet with PIAT) where correlational data were unavailable.

To use the charts, follow the procedures outlined on page 4.

2. A general computational procedure is provided (pages 31 to 39) for use with other combinations of intellectual ability and achievement tests. To use this procedure, the means and standard deviations of both tests and the correlation between the tests must be available. A worksheet is provided (page 39) to help you use the computational procedure.

3. An estimation procedure for use when correlations for test pairs are unavailable is provided on pages 40-41.

The charts may be used to determine if a pupil's test scores indicate a severe discrepancy (i.e., at or greater than 1.5 standard deviations of the difference distribution) between measures of intellectual ability and achievement.

To use the charts, you must have the pupil's standard scores* on both tests. Note that the charts can be used only for the combination of tests (or subtests) named.

To Read A Chart

Follow this two step procedure to determine whether a pupil's test scores indicate a severe discrepancy:**

1. Convert the pupil's test scores on both tests to standard scores using a scale with mean of 100 and a standard deviation of 15. For the Stanford-Binet, the standard deviation is 16. No additional conversion is necessary once the deviation IQ is obtained. This conversion can be done from tables in the test manuals. (Test manual page references appear on each chart.)
2. Locate the pupil's standard score for each test on the appropriate chart. The point on the chart where the two score values meet determines if there is a severe discrepancy.

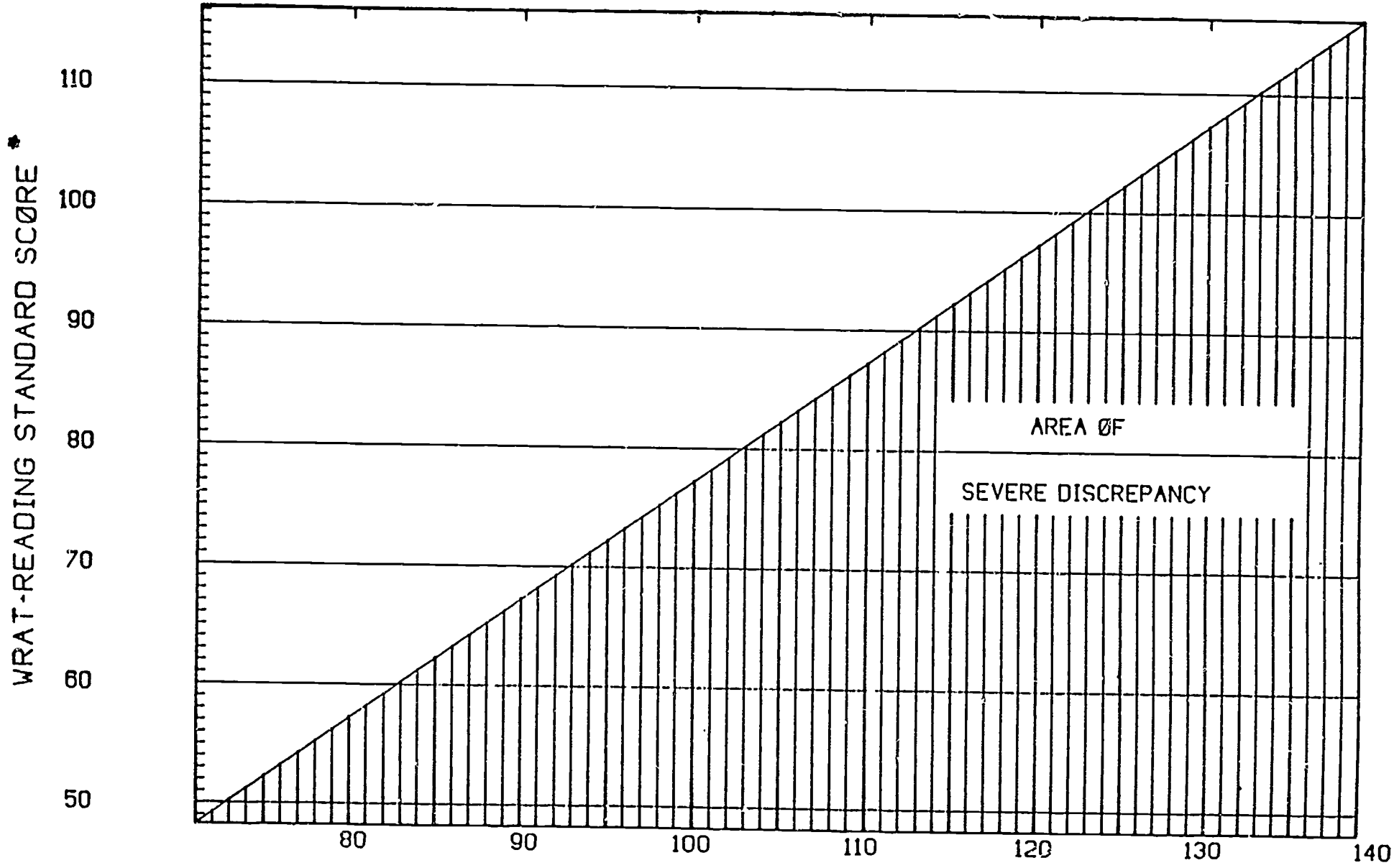
If the point lies on or below the diagonal line, the pupil's score difference does indicate a severe discrepancy.

If the point lies above the diagonal line, the pupil's score difference does not indicate a severe discrepancy.

* The standard score scale has a mean of 100 and a standard deviation of 15.

** Note that there is some variation in standard score starting points for the charts (e.g., in chart 1 the lowest IQ score is 71, whereas in Chart 14 the lowest IQ score is 90). This variation is due to different norming procedures established by the publishers. If the standard scores for a pupil do not appear in the chart, you may need to use a different test combination or to use another of the procedures described.

SEVERE DISCREPANCY CRITERIØN CHART 1:
WISC-R VERBAL AND WRAT-READING



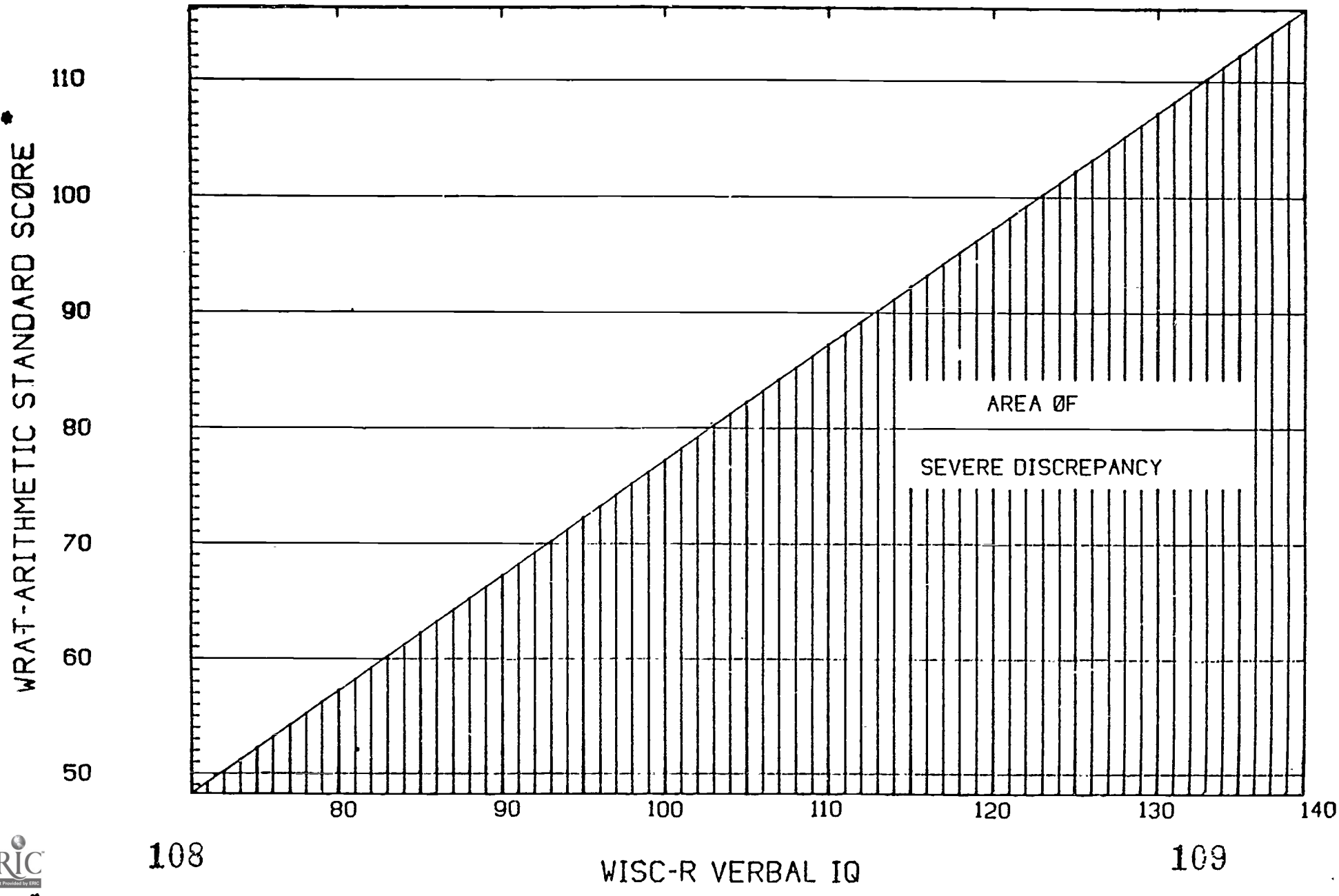
106

* WIDE RANGE ACHIEVEMENT TEST MANUAL, 1978 EDITION, PAGES 16-38

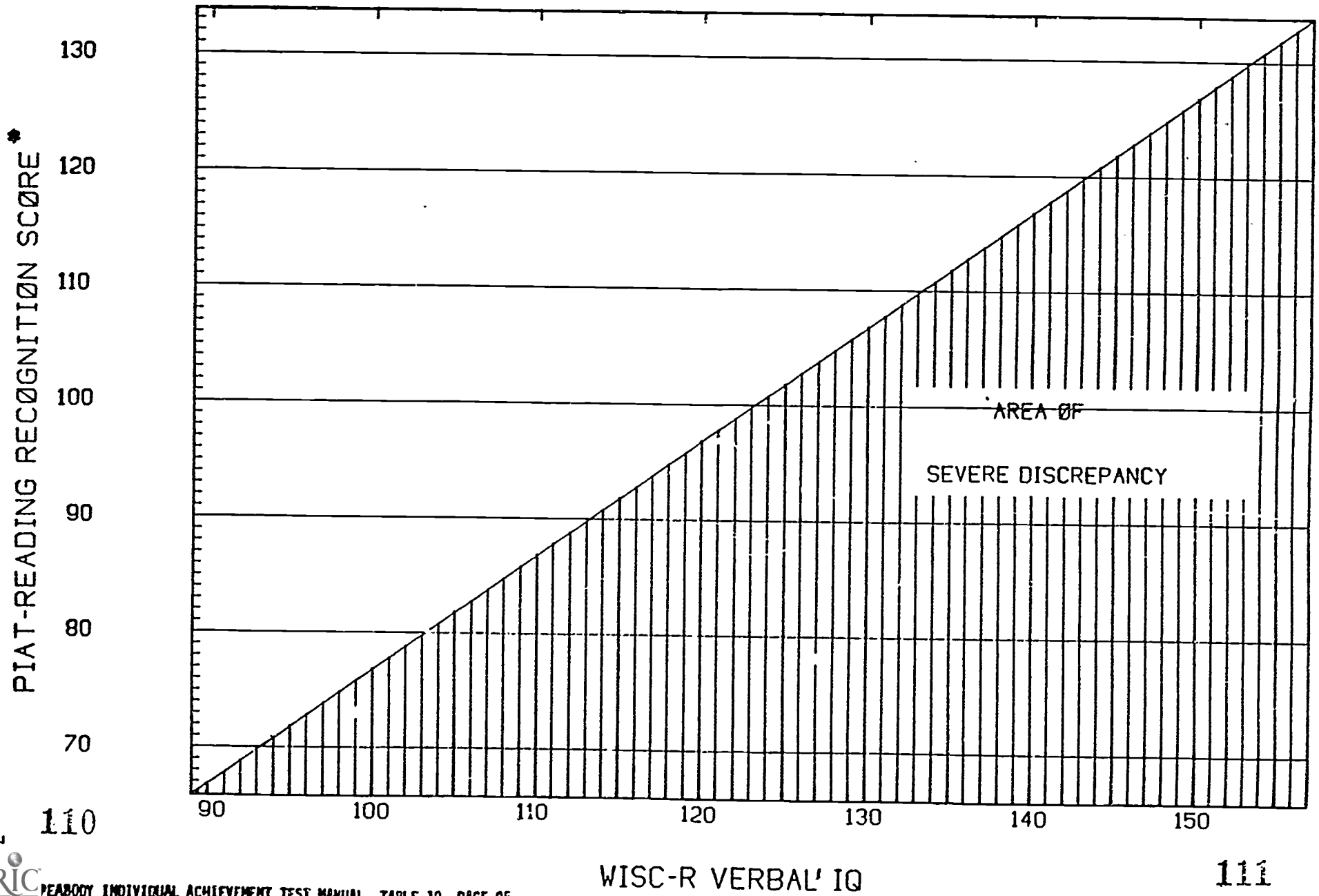
107

SEVERE DISCREPANCY CRITERIØN CHART 2:

WISC-R VERBAL AND WRAT-ARITHMETIC



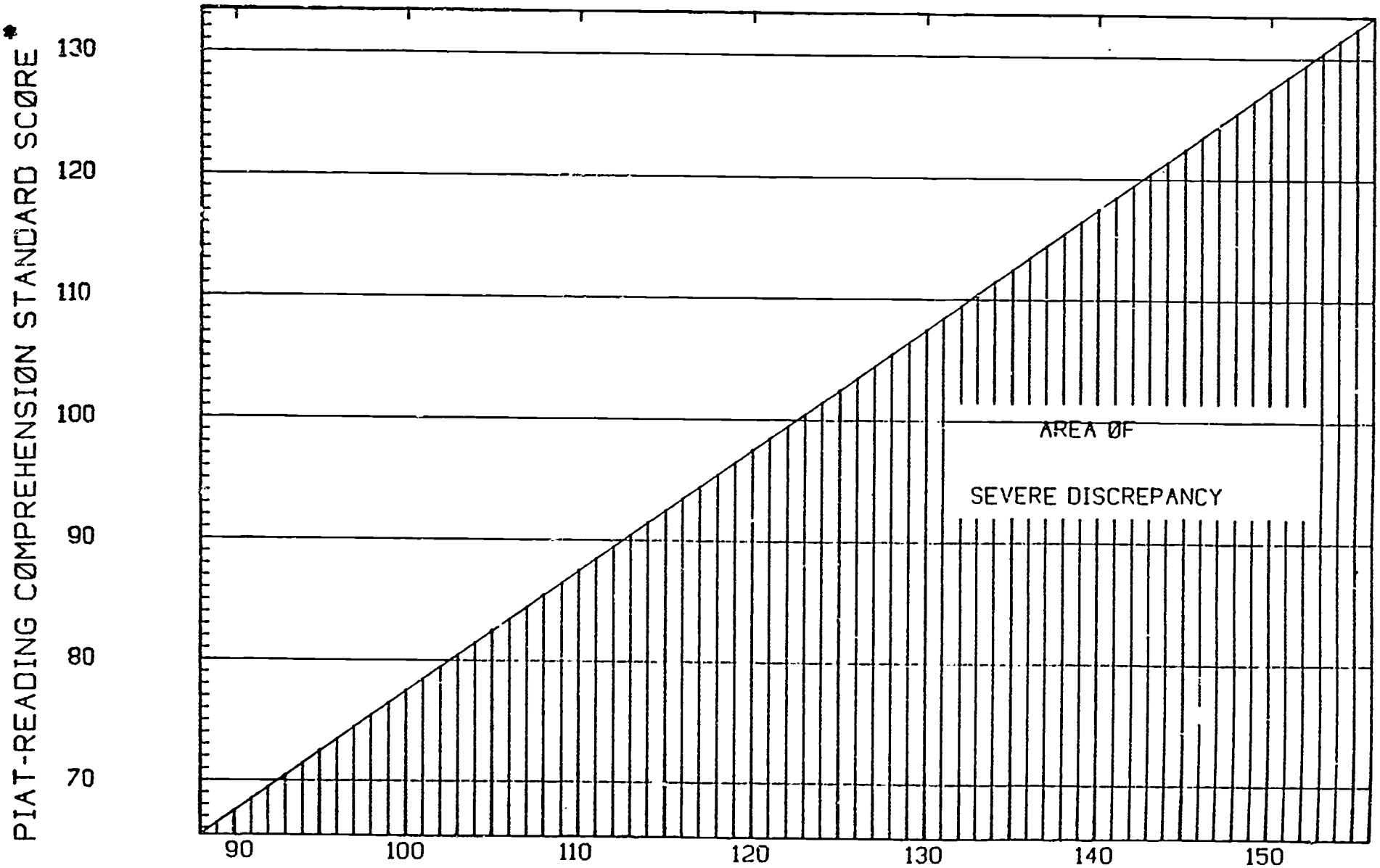
SEVERE DISCREPANCY CRITERIØN CHART 3:
WISC-R VERBAL AND PIAT-READING RECØGNITIØN



*PEABODY INDIVIDUAL ACHIEVEMENT TEST MANUAL, TABLE 19, PAGE 95

SEVERE DISCREPANCY CRITERIØN CHART 4:

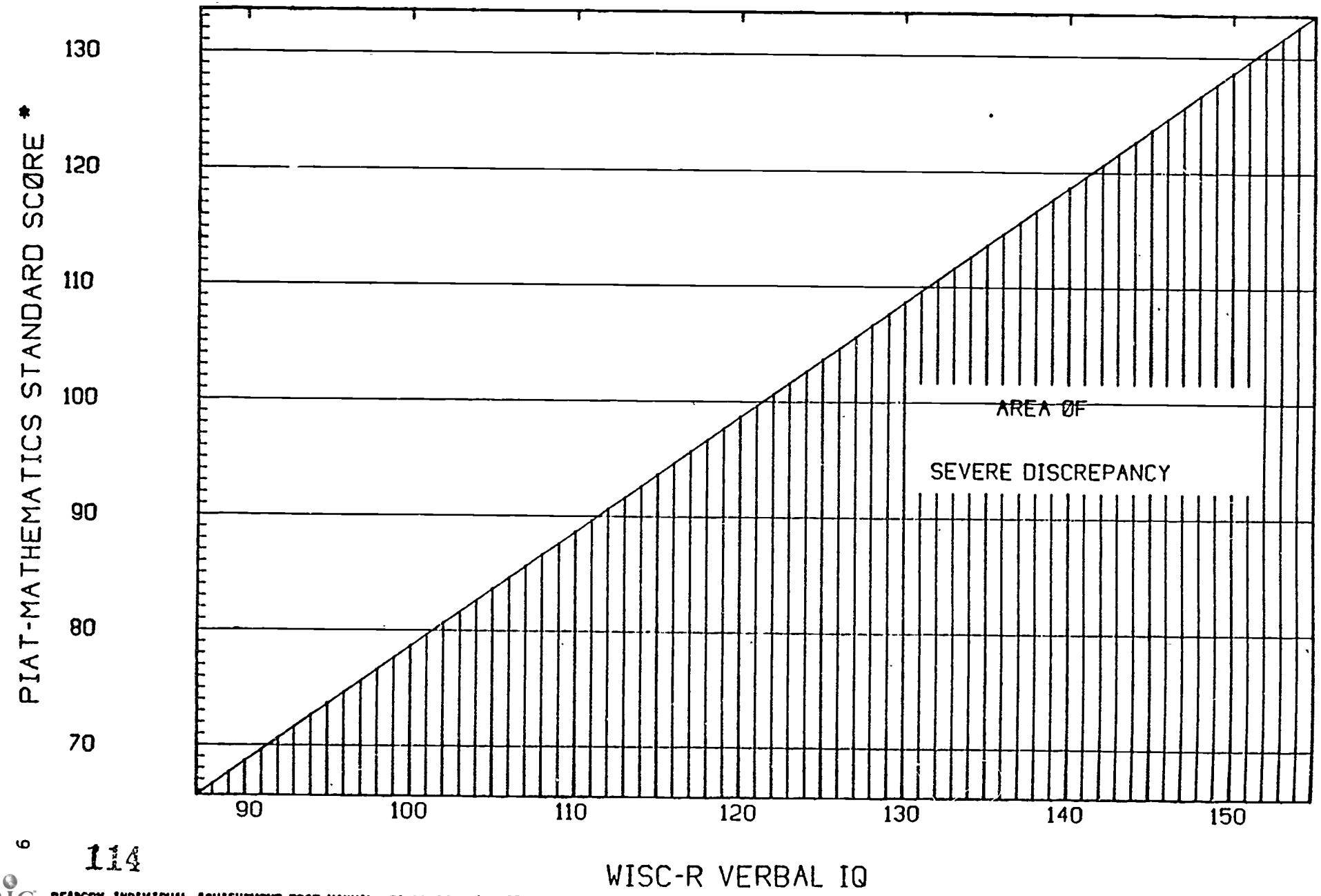
WISC-R VERBAL AND PIAT-READING CØMPREHENSIØN



* PEABODY INDIVIDUAL ACHIEVEMENT TEST MANUAL, TABLE 19, PAGE 95

SEVERE DISCREPANCY CRITERIØN CHART 5:

WISC-R VERBAL AND PIAT-MATHEMATICS

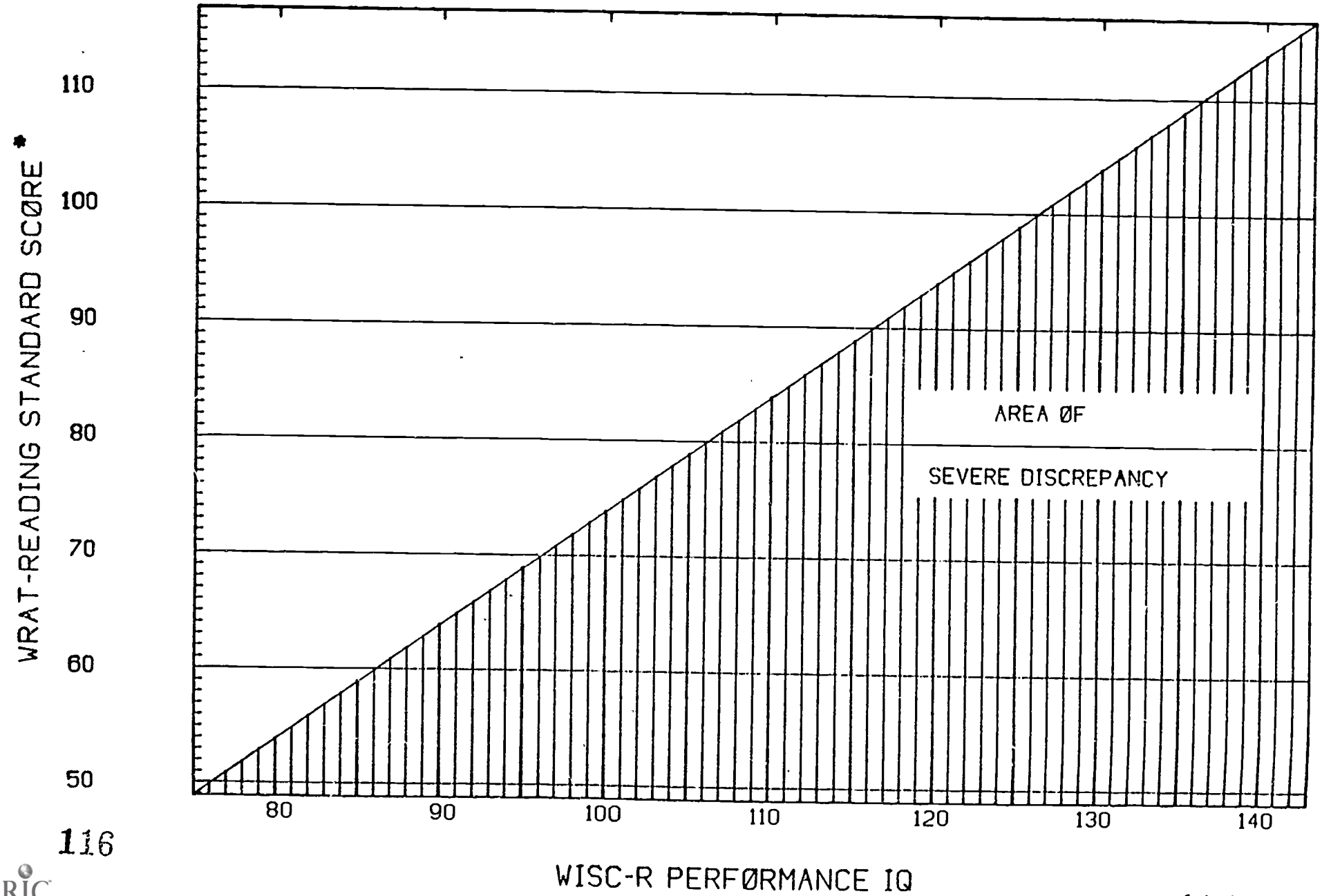


6

PEABODY INDIVIDUAL ACHIEVEMENT TEST MANUAL, TABLE 19, PAGE 95

SEVERE DISCREPANCY CRITERIØN CHART 6:

WISC-R PERFØRMANCE AND WRAT-READING



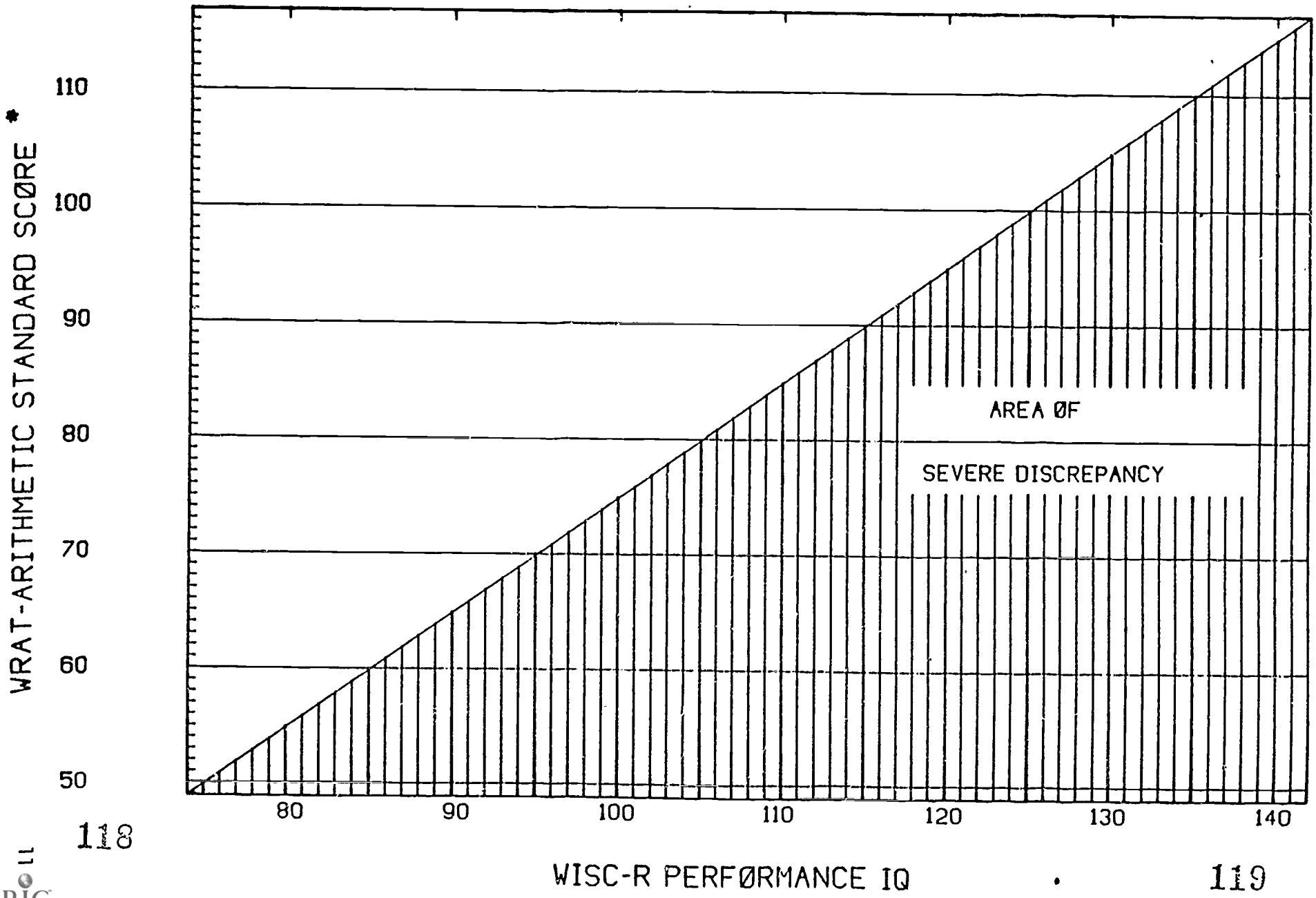
116

JDE RANGE ACHIEVEMENT TEST MANUAL, 1978 EDITION, PAGES 16-38

117

SEVERE DISCREPANCY CRITERIØN CHART 7:

WISC-R PERFØRMANCE AND WRAT-ARITHMETIC



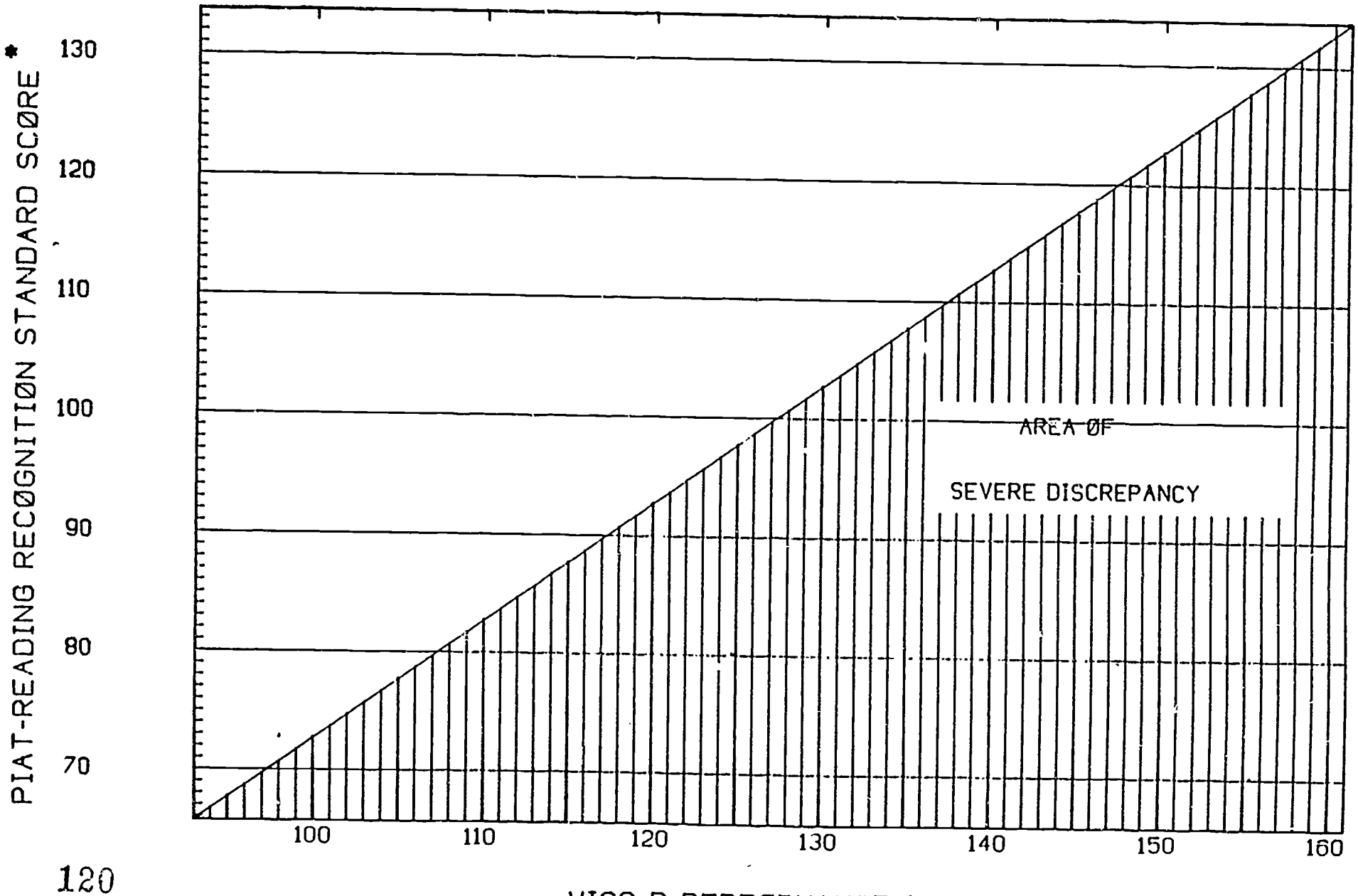
118

119

DE RANGE ACHIEVEMENT TEST MANUAL, 1978 EDITION, PAGES 16-38

SEVERE DISCREPANCY CRITERIØN CHART 8:

WISC-R PERFØRMANCE AND PIAT-READING RECØGNITIØN

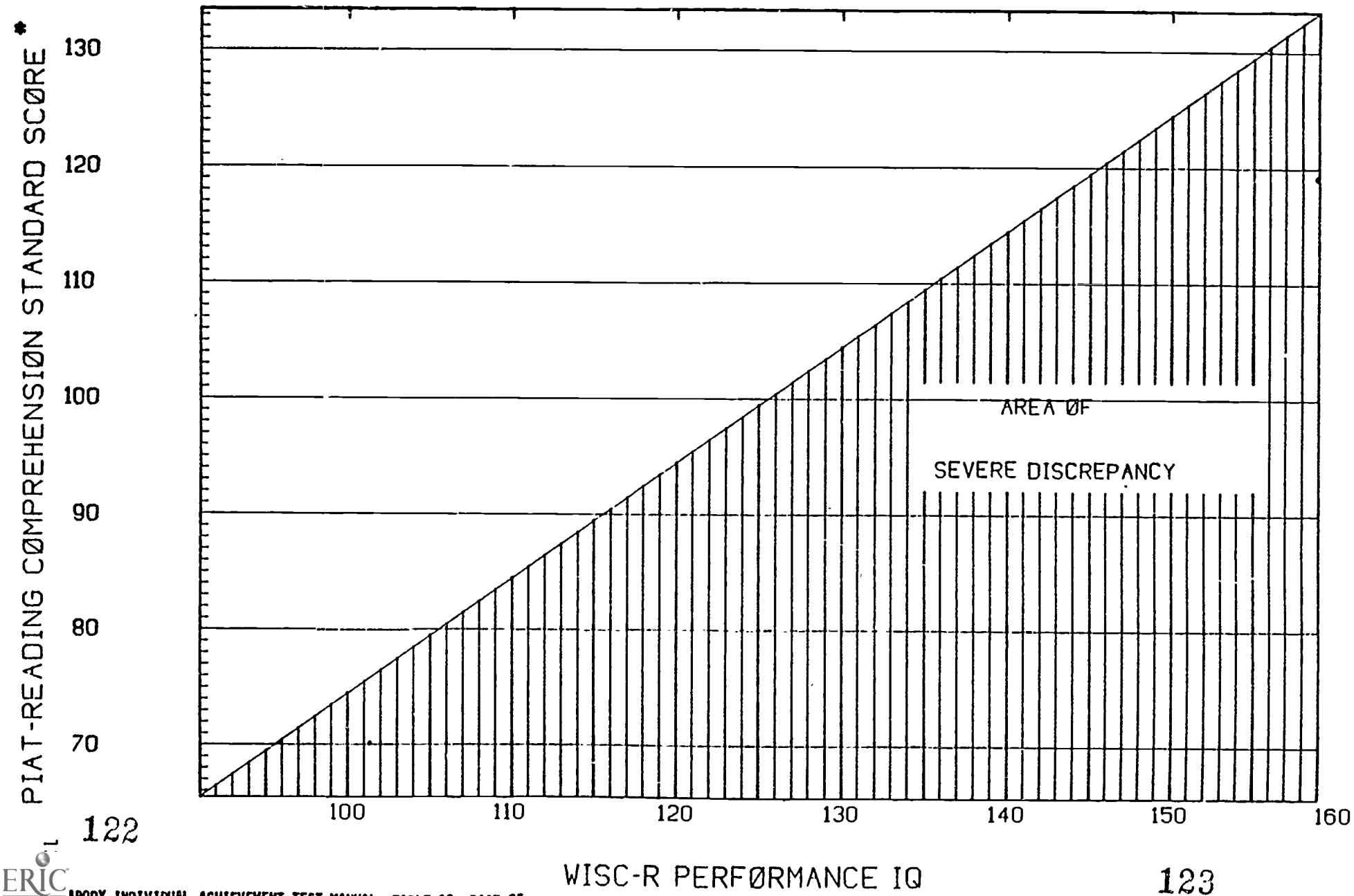


120

PEABODY INDIVIDUAL ACHIEVEMENT TEST MANUAL, TABLE 19, PAGE 95

SEVERE DISCREPANCY CRITERIØN CHART 9:

WISC-R PERFØRMANCE AND PIAT-READING CØMPREHENSIØN



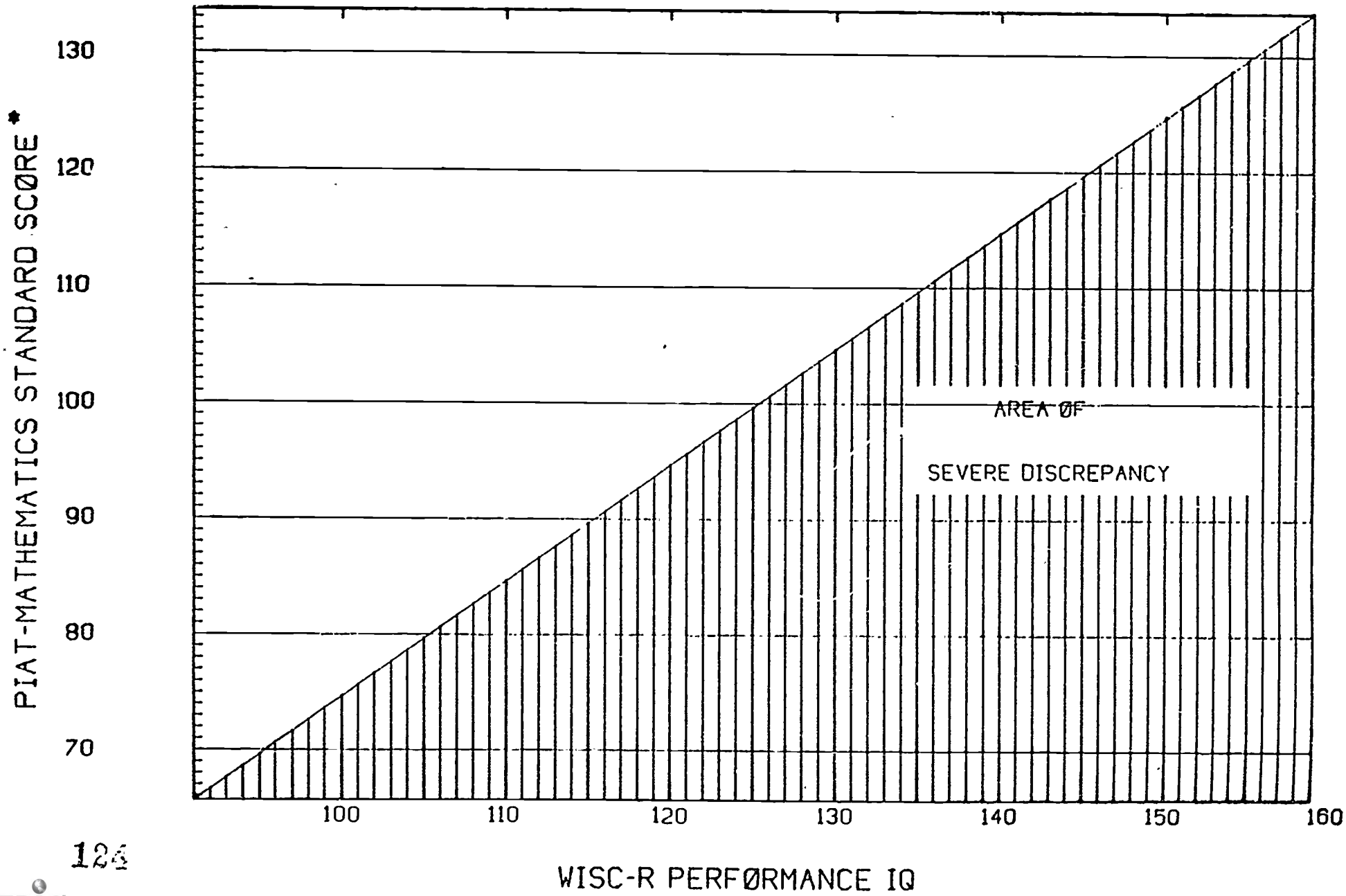
122

123

ABODY INDIVIDUAL ACHIEVEMENT TEST MANUAL, TABLE 19, PAGE 95

SEVERE DISCREPANCY CRITERIØN CHART 10:

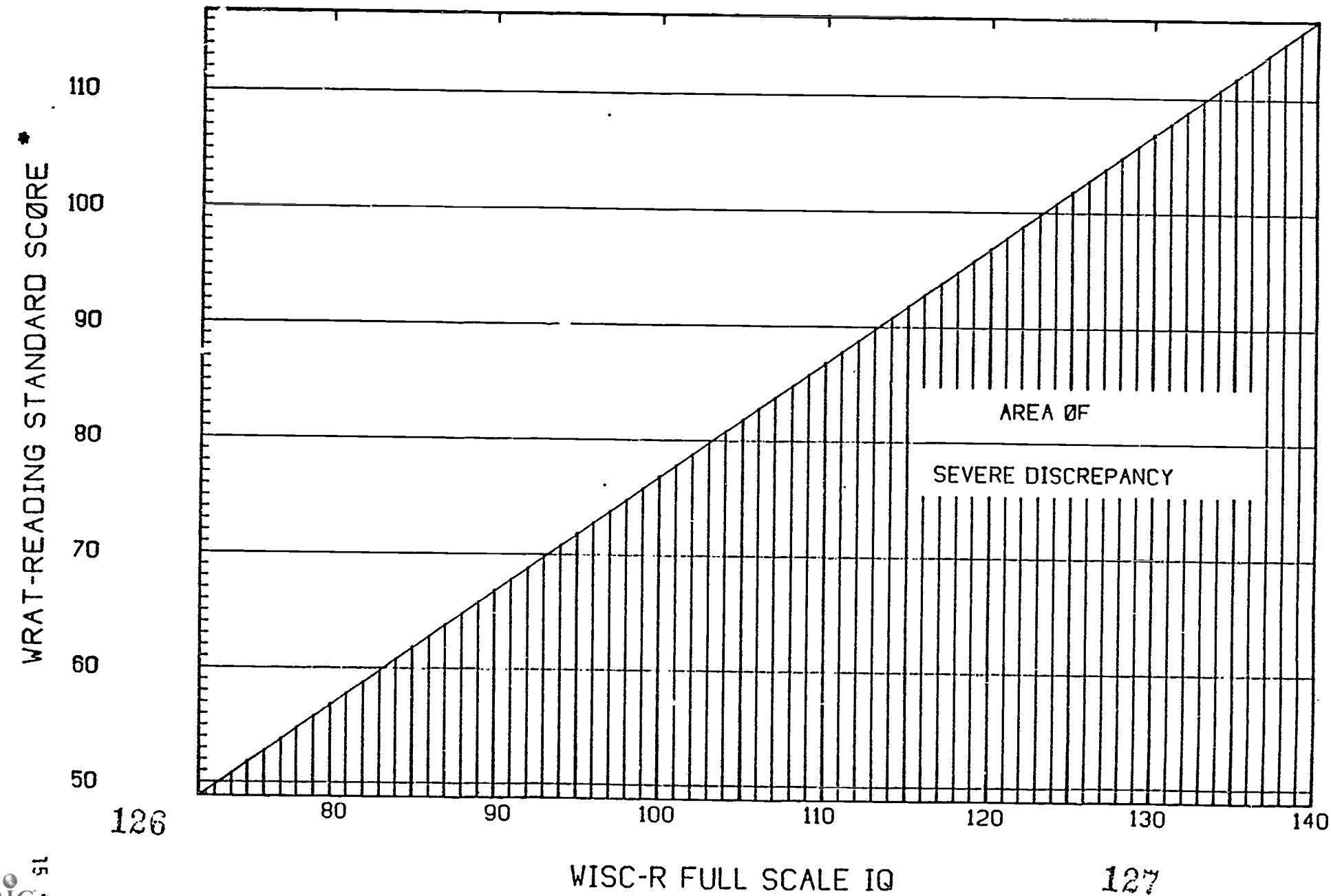
WISC-R PERFØRMANCE AND PIAT-MATHEMATICS



*ABODY INDIVIDUAL ACHIEVEMENT TEST MANUAL, TABLE 19, PAGE 95

SEVERE DISCREPANCY CRITERIØN CHART 11:

WISC-R FULL SCALE AND WRAT-READING



126

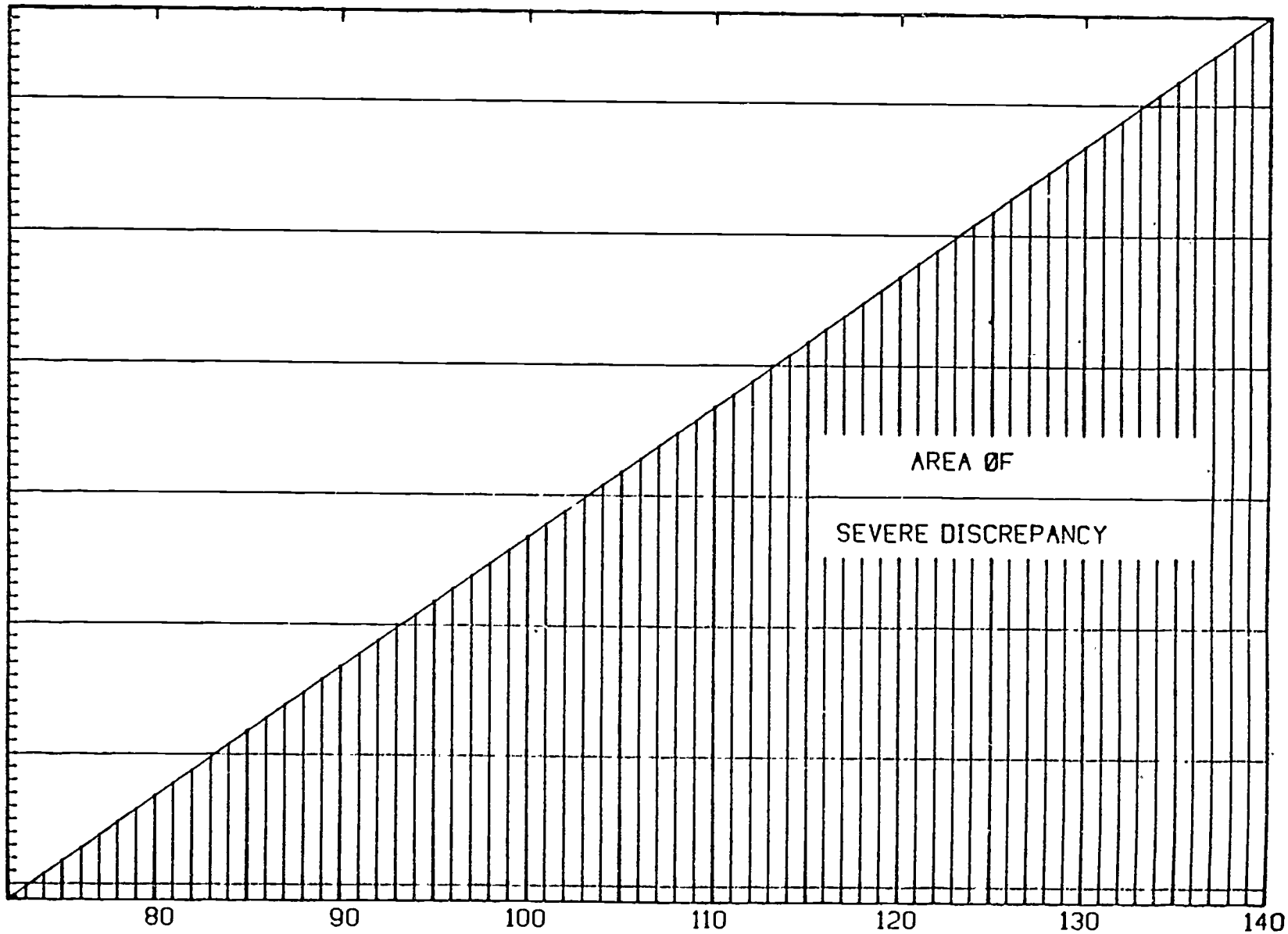
127

15

*WIDE RANGE ACHIEVEMENT TEST MANUAL, 1978 EDITION, PAGES 16-38

SEVERE DISCREPANCY CRITERIØN CHART 12:

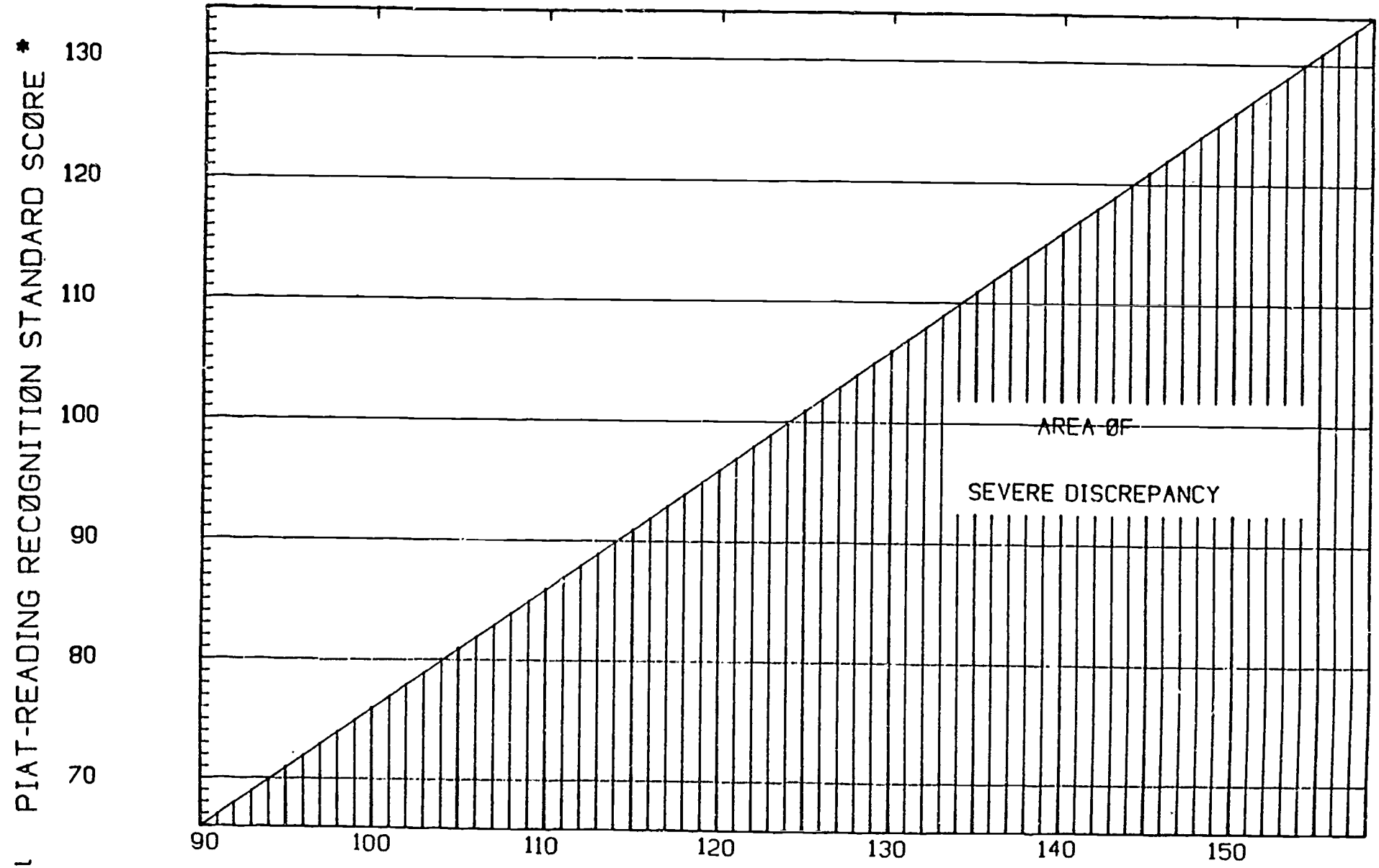
WISC-R FULL SCALE AND WRAT-ARITHMETIC



*WIDE RANGE ACHIEVEMENT TEST MANUAL, 1978 EDITION, PAGES 16-38

SEVERE DISCREPANCY CRITERIØN CHART 13:

WISC-R FULL SCALE AND PIAT-READING RECØGNITIØN



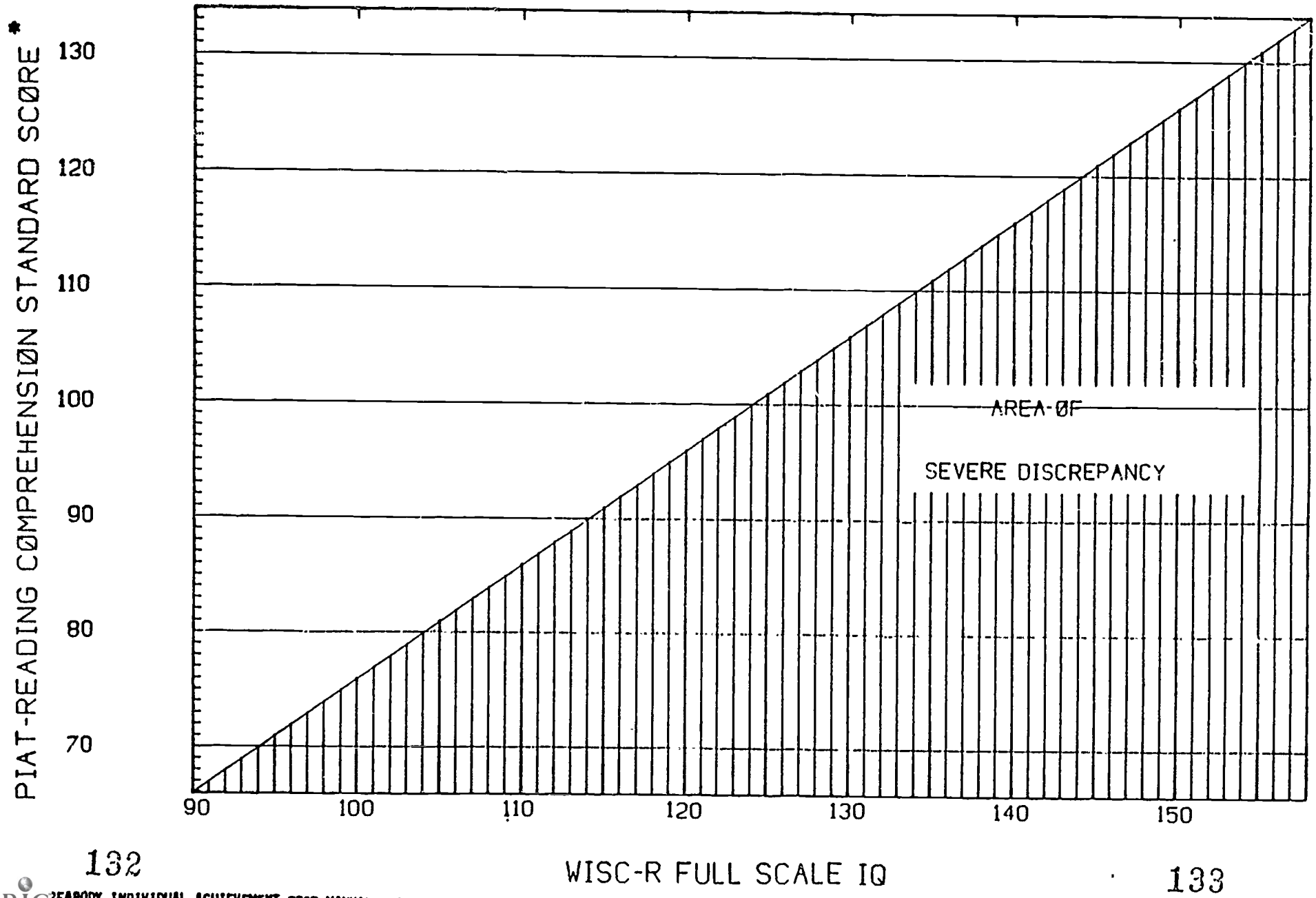
130

131

* PEABODY INDIVIDUAL ACHIEVEMENT TEST MANUAL, TABLE 19, PAGE 95

SEVERE DISCREPANCY CRITERIØN CHART 14:

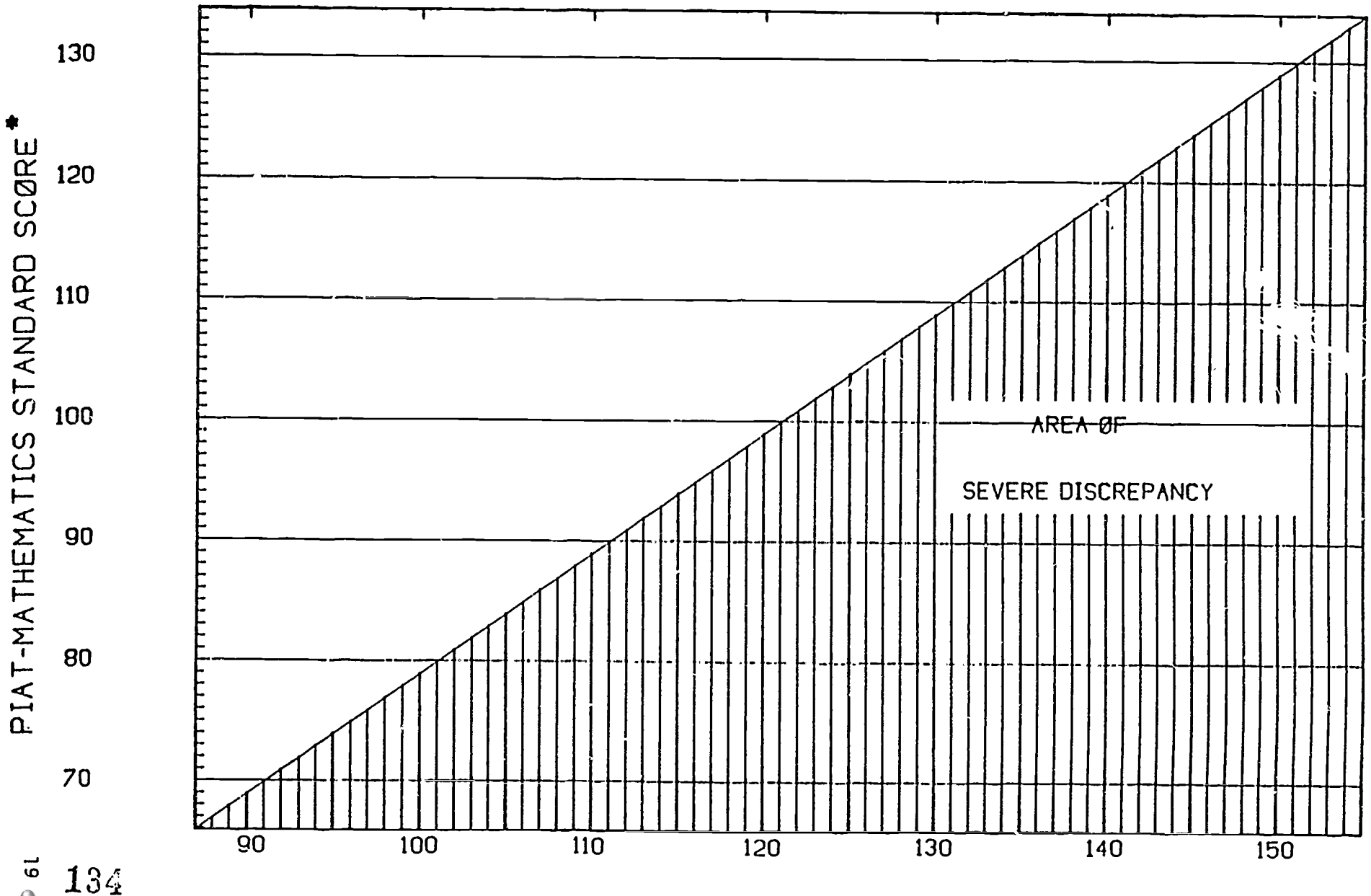
WISC-R FULL SCALE AND PIAT-READING CØMPREHENSIØN



*PEABODY INDIVIDUAL ACHIEVEMENT TEST MANUAL, TABLE 19, PAGE 95

SEVERE DISCREPANCY CRITERIØN CHART 15:

WISC-R FULL SCALE AND PIAT-MATHEMATICS



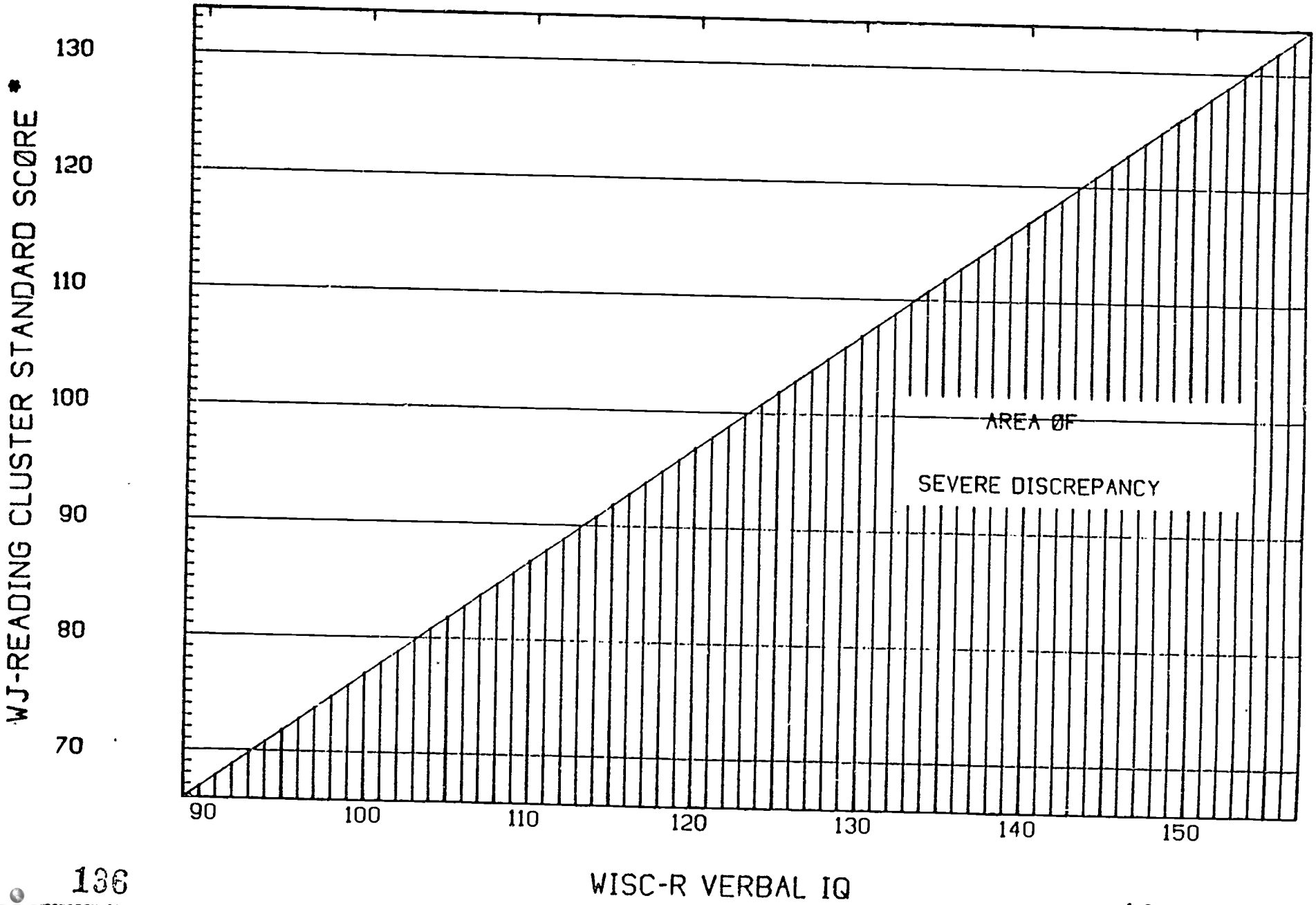
61

134

ABODY INDIVIDUAL ACHIEVEMENT TEST MANUAL, TABLE 19, PAGE 95

135

SEVERE DISCREPANCY CRITERIØN CHART 16:
WISC-R VERBAL AND WJ-READING CLUSTER



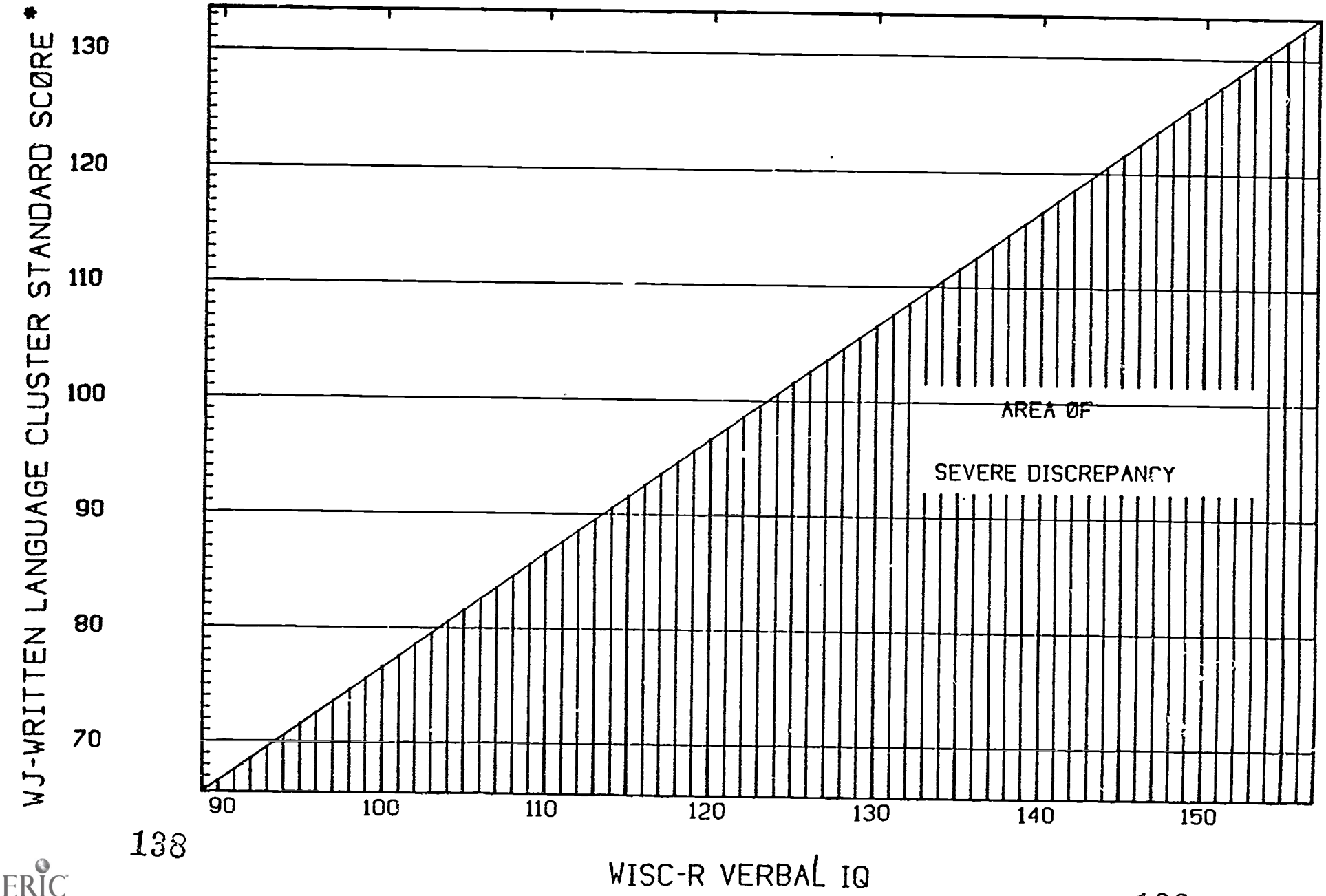
136

DODCOCK-JOHNSON PSYCHO-EDUCATIONAL BATTERY MANUAL, TABLE X, PAGE 407

137

SEVERE DISCREPANCY CRITERIØN CHART 17:

WISC-R VERBAL AND WJ-WRITTEN LANGUAGE CLUSTER

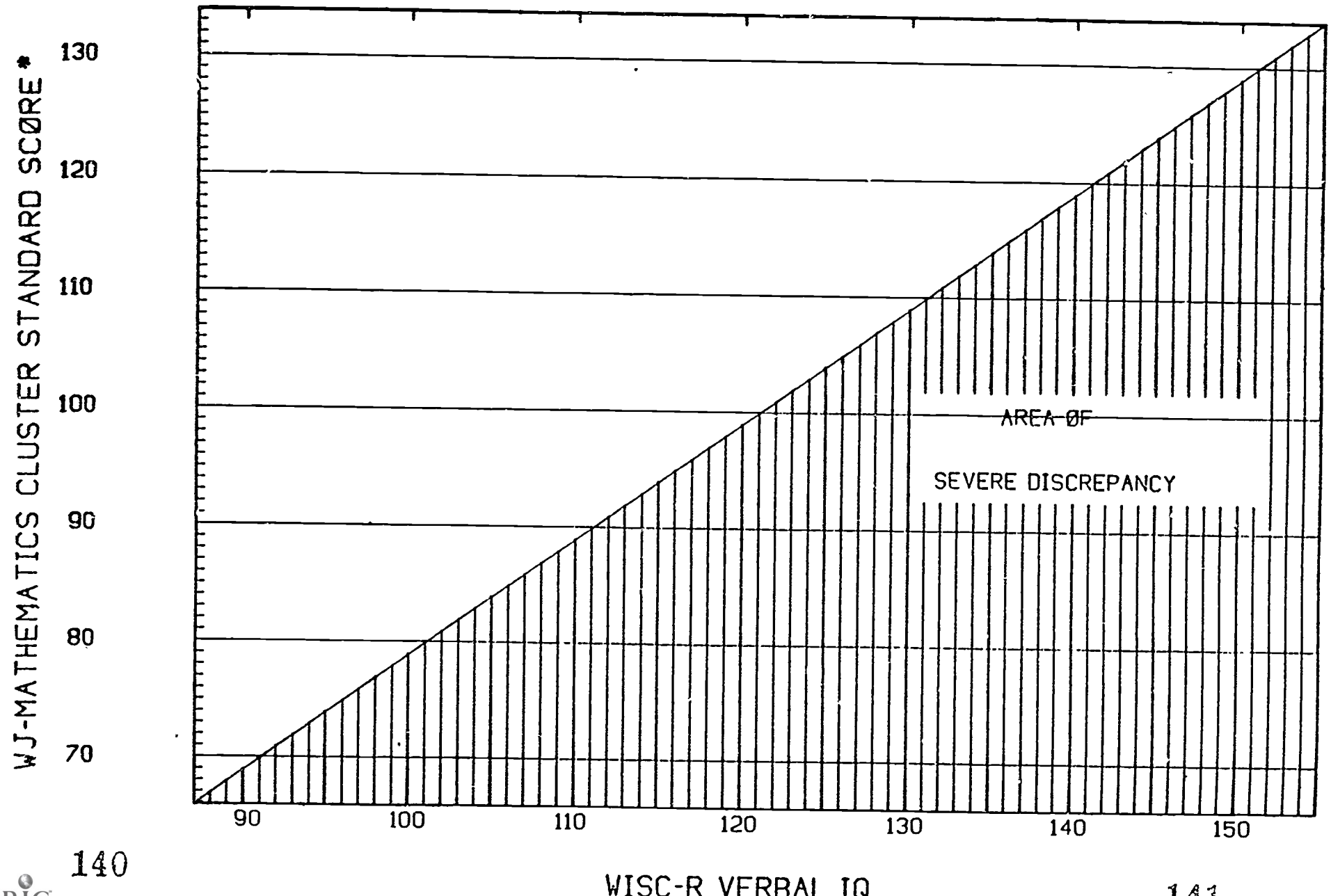


138

* WOODCOCK-JOHNSON PSYCHO-EDUCATIONAL BATTERY MANUAL, TABLE X, PAGE 407

139

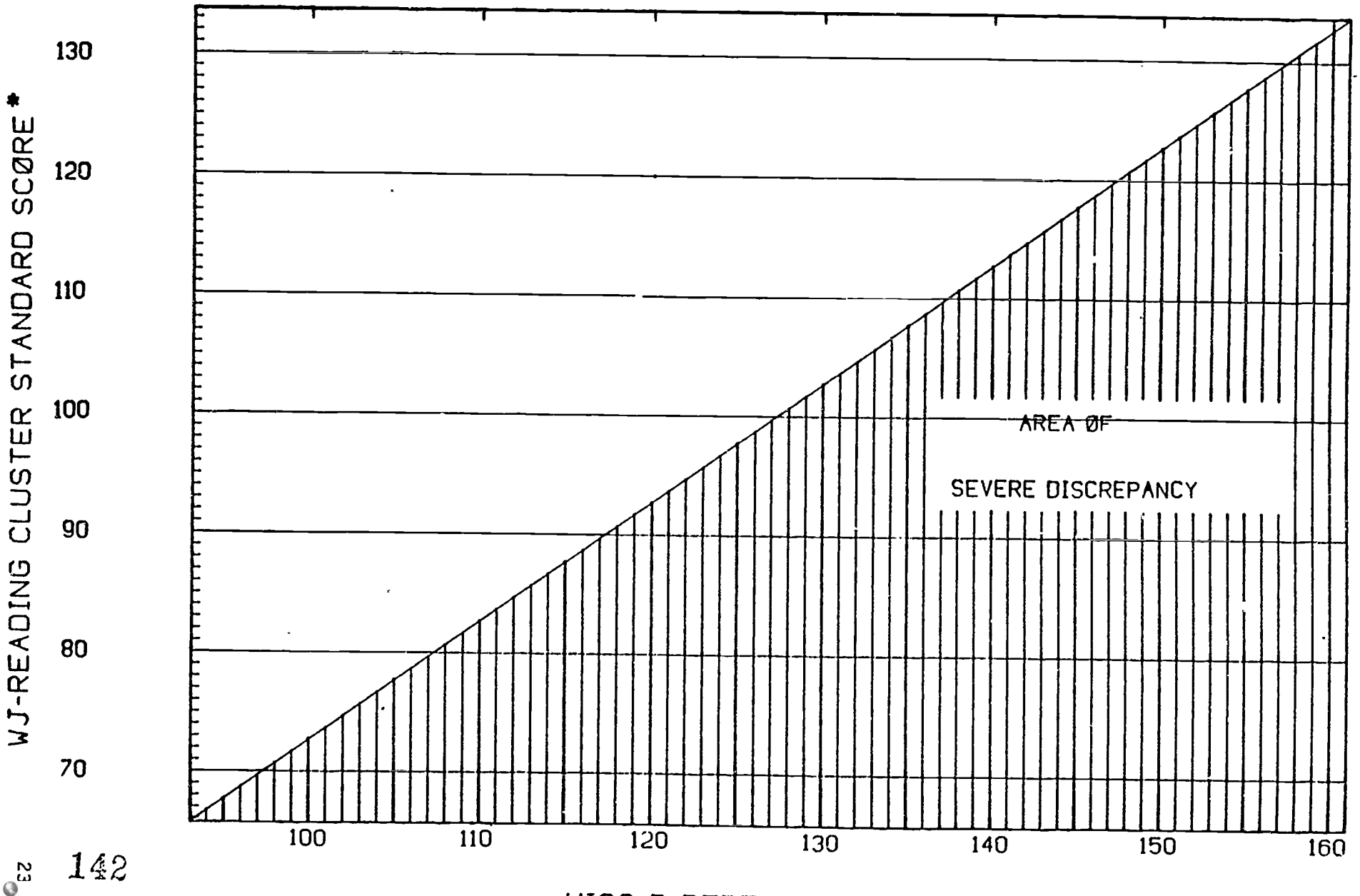
SEVERE DISCREPANCY CRITERIØN CHART 18:
WISC-R VERBAL AND WJ-MATHEMATICS CLUSTER



ODCOCK-JOHNSON PSYCHO-EDUCATIONAL BATTERY MANUAL, TABLE X, PAGE 407

SEVERE DISCREPANCY CRITERIØN CHART 19:

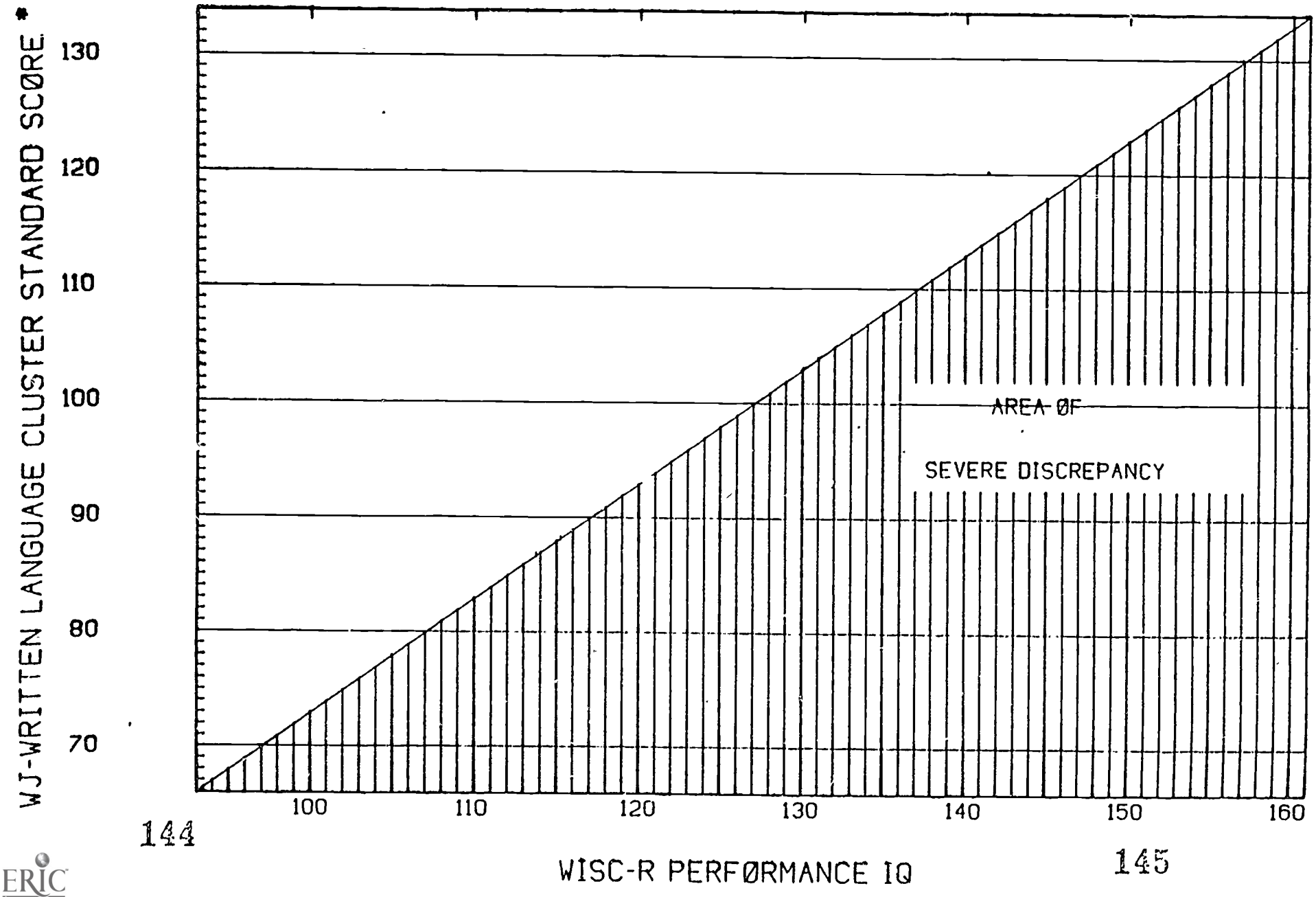
WISC-R PERFØRMANCE AND WJ-READING CLUSTER STD



*WOODCOCK-JOHNSON PSYCHO-EDUCATIONAL BATTERY MANUAL, TABLE X, PAGE 407

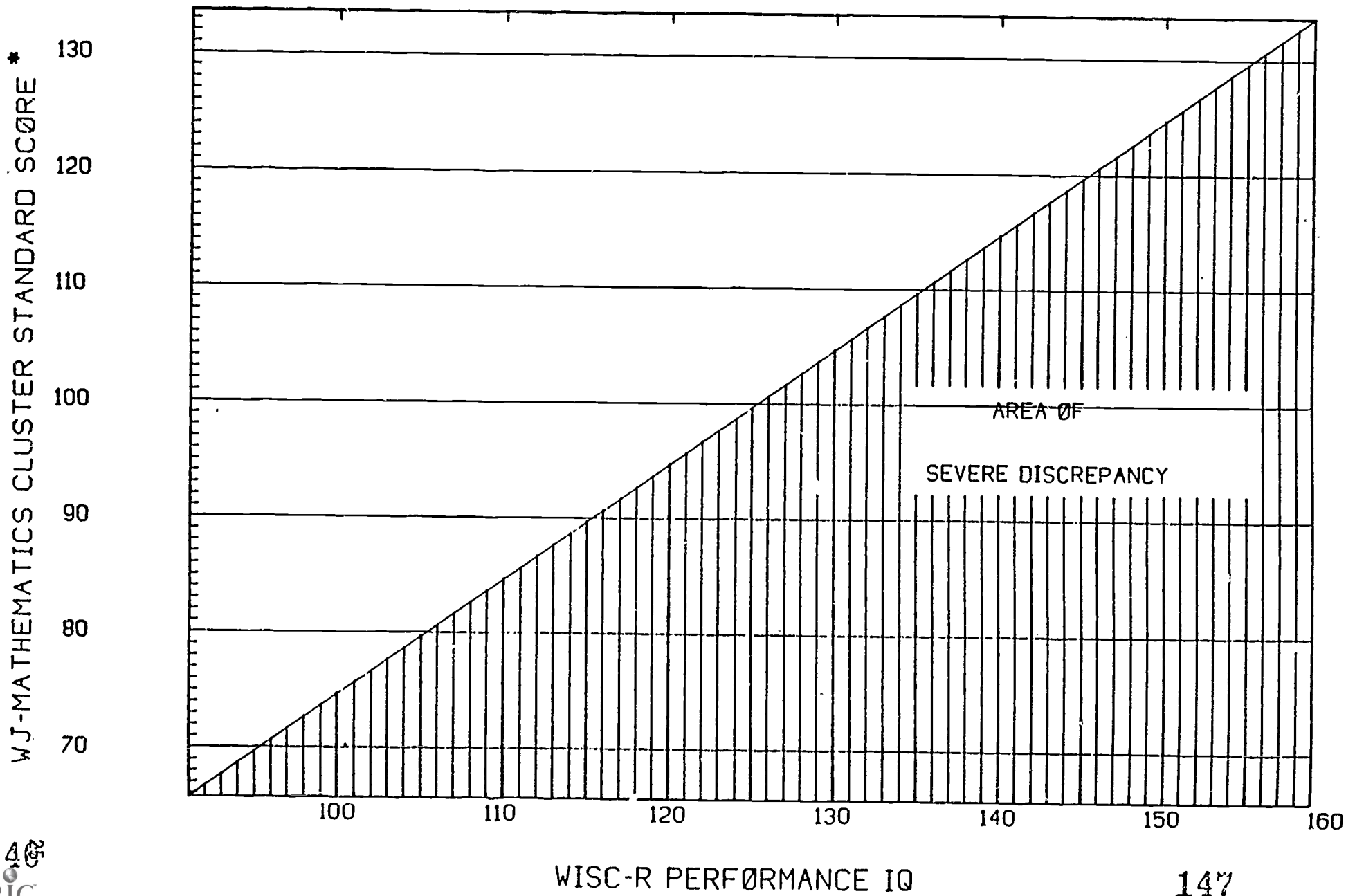
SEVERE DISCREPANCY CRITERION CHART 20:

WISC-R PERFORMANCE AND WJ-WRITTEN LANGUAGE CLUSTER



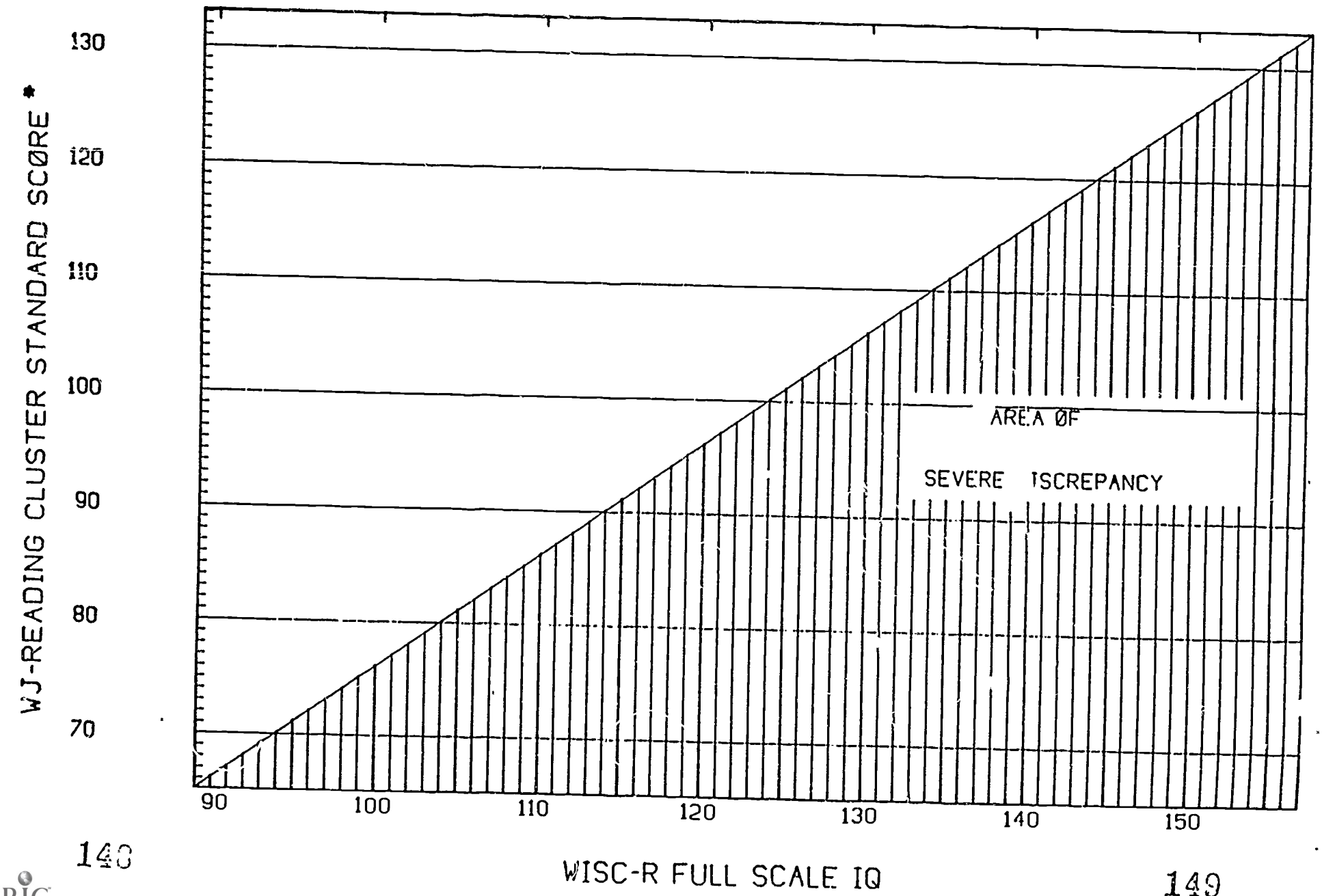
WOODCOCK-JOHNSON PSYCHO-EDUCATIONAL BATTERY MANUAL, TABLE X, PAGE 407

SEVERE DISCREPANCY CRITERIØN CHART 21:
WISC-R PERFØRMANCE AND WJ-MATHEMATICS CLUSTER



WOODCOCK-JOHNSON PSYCHO-EDUCATIONAL BATTERY MANUAL, TABLE X, PAGE 407

SEVERE DISCREPANCY CRITERIØN CHART 22:
WISC-R FULL SCALE AND WJ-READING CLUSTER



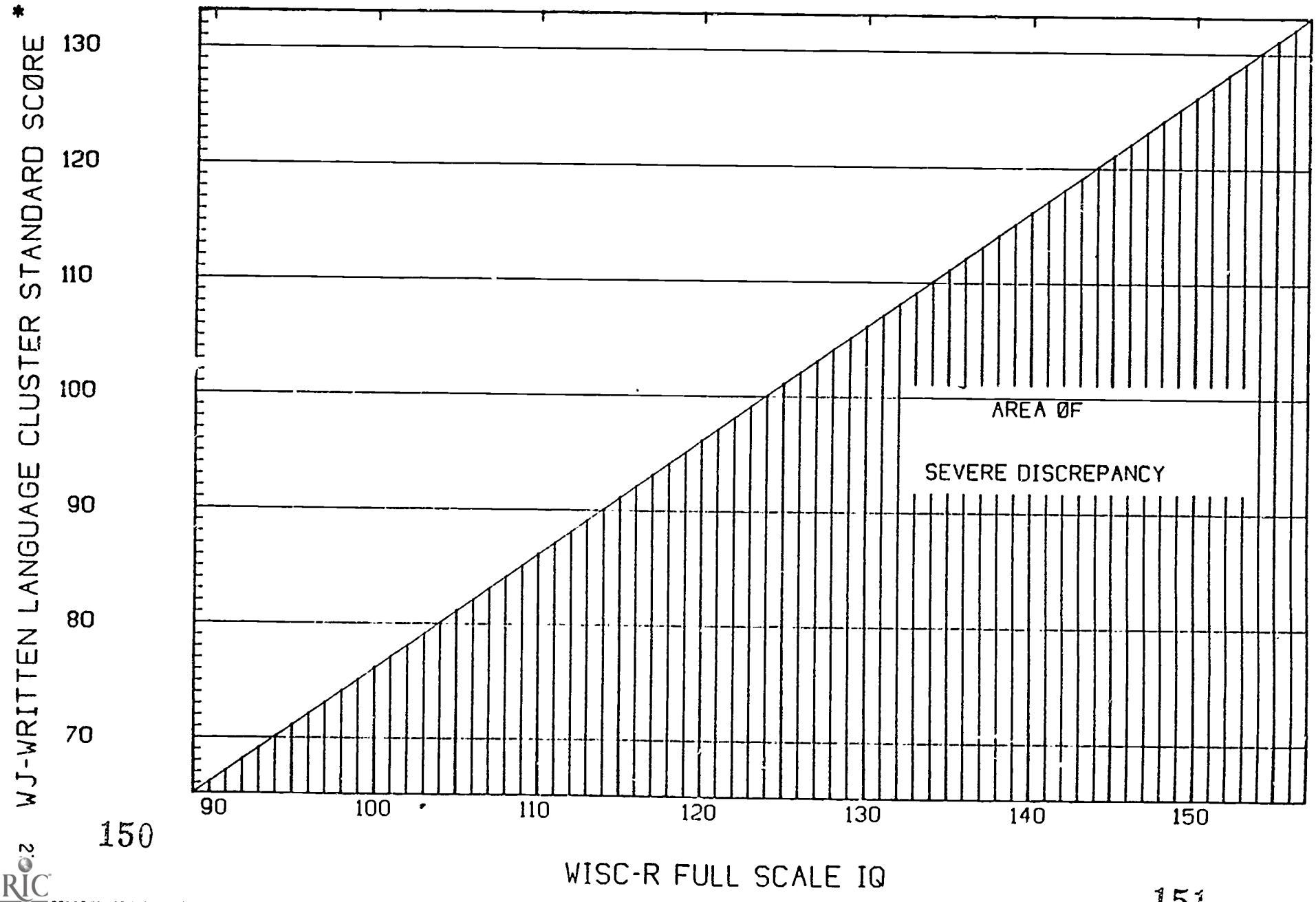
148

149

* WOODCOCK-JOHNSON PSYCHO-EDUCATIONAL BATTERY MANUAL, TABLE X, PAGE 407

SEVERE DISCREPANCY CRITERIØN CHART 23:

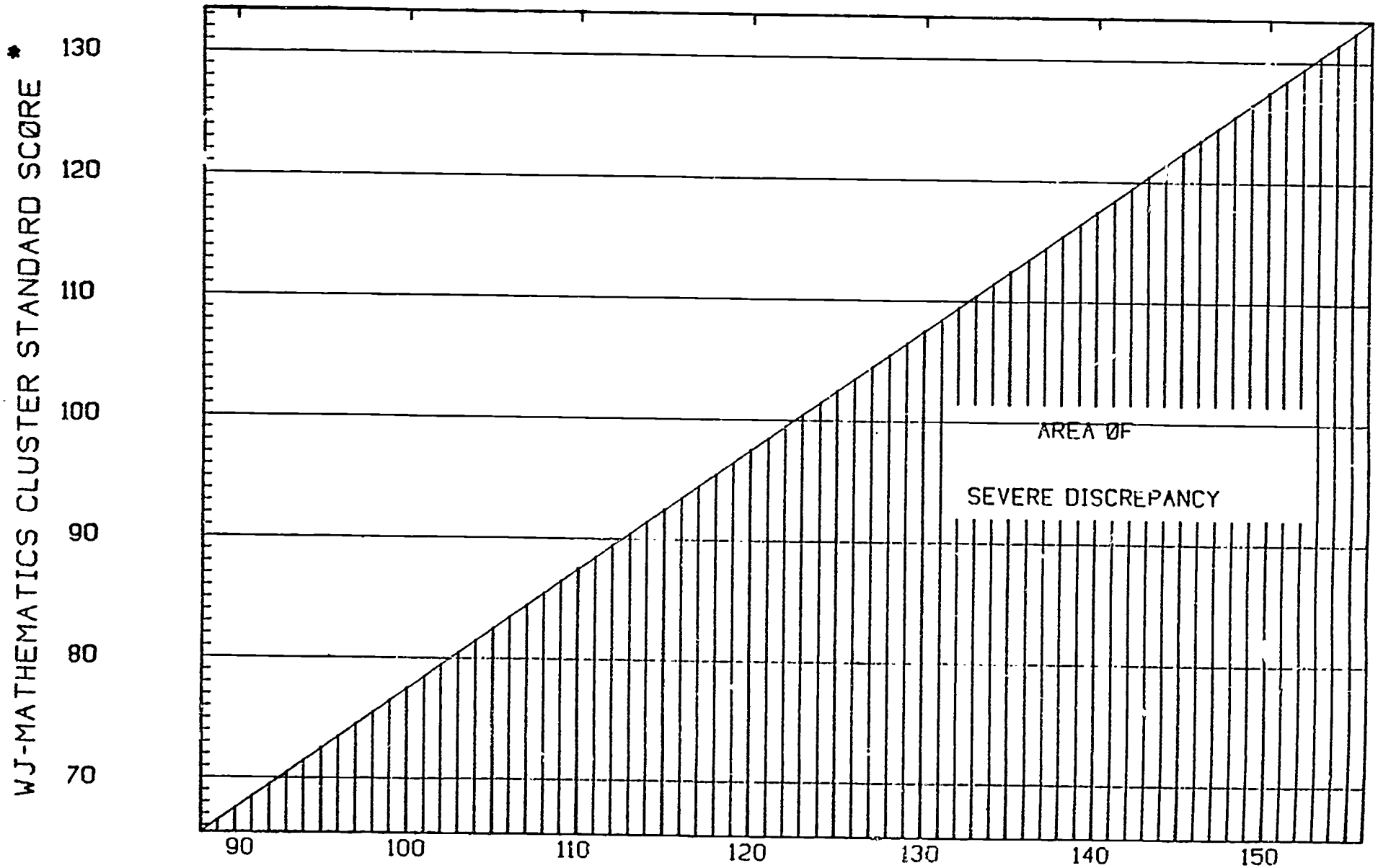
WISC-R FULL SCALE AND WJ-WRITTEN LANGUAGE CLUSTER



150

* DODCOCK-JOHNSON PSYCHO-EDUCATIONAL BATTERY MANUAL, TABLE X, PAGE 407

SEVERE DISCREPANCY CRITERIØN CHART 24:
WISC-R FULL SCALE AND WJ-MATHEMATICS CLUSTER

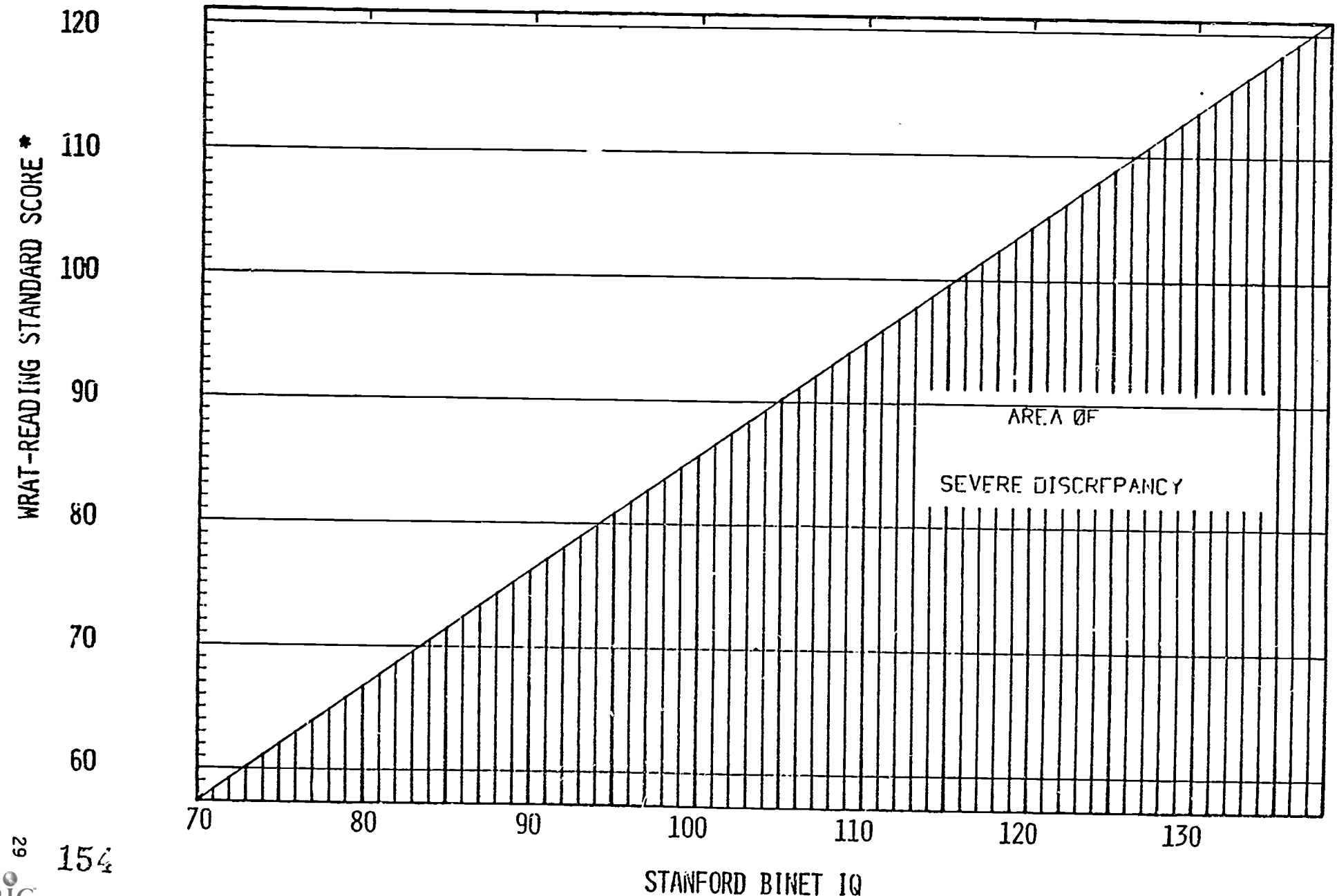


152

153

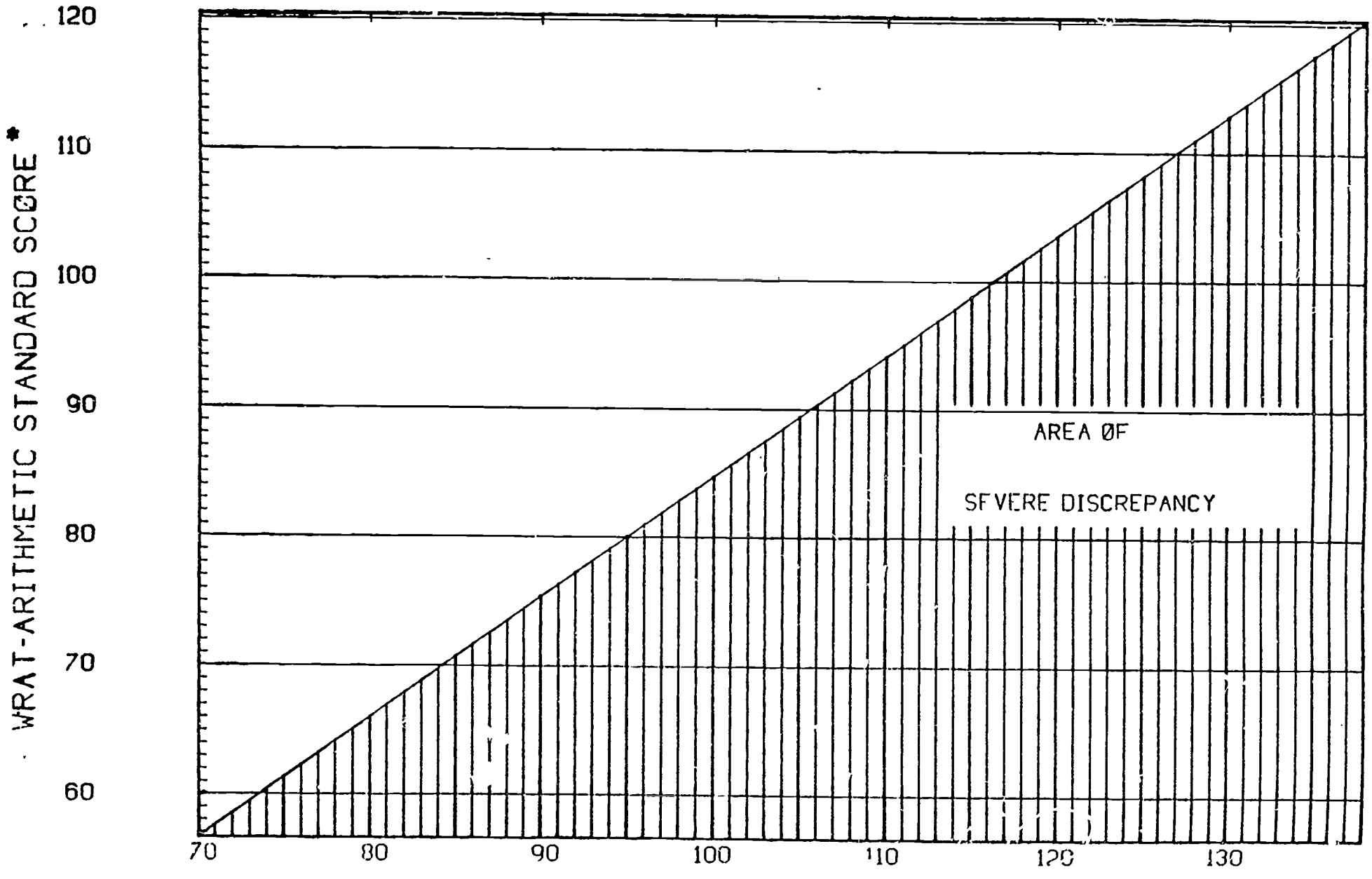
WOODCOCK-JOHNSON PSYCHO-EDUCATIONAL BATTERY MANUAL, TABLE X, PAGE 407

SEVERE DISCREPANCY CRITERION CHART 25:
STANFORD BINET AND WRAT-READING



*WIDE RANGE ACHIEVEMENT TEST MANUAL, 1978 EDITION, PAGES 16-38

SEVERE DISCREPANCY CRITERIØN CHART 26:
STANFØRD BINET AND WRAT-ARITMETIC



156

157

*WIDE RANGE ACHIEVEMENT TEST MANUAL, 1978 EDITION, PAGES 16-38

COMPUTATIONAL PROCEDURE WHEN USING OTHER TESTS

If you use a combination of tests for which charts have not been prepared, you may use the following procedure to determine if a pupil's test scores indicate a severe discrepancy. To use this method, you must obtain: 1) the means and standard deviations of both tests (usually available in the publishers' manuals); and 2) the correlation between the two tests (when available, typically found in professional research journals rather than in publishers' manuals); as well as 3) the pupil's standard scores on the ability and achievement tests.

The steps in the procedure are:

1. Obtain the following technical information for both tests: mean, standard deviation, pupil scores, and the correlation between tests.
   a. If pupil scores on both tests are already expressed on a scale with a mean of 100 and a standard deviation of 15,* go directly to Step 3.
   b. If the pupil score(s) is (are) expressed as percentiles, use Table 1 to obtain the equivalent common standard scores (CSS) which are on a scale having a mean of 100 and a standard deviation of 15. Then go to Step 3.
   c. If pupil scores are on a scale having a mean other than 100 and/or a standard deviation other than 15, go to Step 2.
2. Transform the pupil's standard scores to common standard scores (CSS) on a scale with a mean of 100 and a standard deviation of 15.
3. Obtain the pupil's discrepancy score and the criterion discrepancy score.
4. Compare the pupil's discrepancy score with the criterion discrepancy.

The following pages describe the four steps in greater detail, including sample data. A worksheet is provided on page 39.

---

* For the purposes of this document, scores expressed on a scale with a mean of 100 and a standard deviation of 15 are called common standard scores (CSS).

Step 1: Obtain Needed Technical Data

Procedure: From test manuals and other sources, obtain means, standard deviations, correlation between tests, pupil standard scores.

Ability test: _____

Mean: _____ (M_{ab})

Standard deviation: _____ (SD_{ab})

Pupil's standard score on ability test: _____ (X_{ab})

(must be in same units as mean and standard deviation)

Achievement test: _____

Mean: _____ (M_{ach})

Standard deviation: _____ (SD_{ach})

Pupil's standard score on achievement test: _____ (X_{ach})

(must be in same units as mean and standard deviation)

Correlation between the two tests:

_____ ($r_{ab,\ ach}$)

Example

Ability test: XYZ Ability Test

Mean: 50 (M_{ab})

Standard deviation: 10 (SD_{ab})

Pupil's standard score on ability test: 60 (X_{ab})

Achievement test: Speedy Reading Test

Mean: 30

Standard deviation: 7

Pupil's standard score on achievement test: 16

Correlation: $r_{ab,\ ach} = .39$

Now proceed to step 2.

Step 1(b)

Procedure: Using table 1 below, find the common standard score for the pupil's percentile rank score.

Table 1
Percentile Rank to Standard Score

PERCENTILE RANK	STANDARD SCORE
99	135
98	131
97	128
96	126
95	125
94	123
93	122
92	121
91	120
90	119
89	118
88	118
87	117
86	116
85	116
84	115
83	114
82	114
81	113
80	113
79	112
78	112
77	111
76	111
75	110
74	110
73	109
72	109
71	108
70	108
69	107
68	107
67	107
66	106
65	106
64	105
63	105
62	105
61	104
60	104
59	103
58	103
57	103
56	102
55	102
54	102
53	101
52	101
51	100
50	100
49	100
48	99
47	99
46	98
45	98
44	98
43	97
42	97
41	97
40	96
39	96
38	95
37	95
36	95
35	94
34	94
33	93
32	93
31	93
30	92
29	92
28	91
27	91
26	90
25	90
24	89
23	89
22	88
21	88
20	87
19	87
18	86
17	86
16	85
15	84
14	84
13	83
12	82
11	82
10	81
9	80
8	79
7	78
6	77
5	75
4	74
3	72
2	69
1	65

Step 2: Transformation of Pupil's Test Scores to Common Standard Scores (CSS)

Procedure: Using means, standard deviations, and pupil standard scores, transform the pupil's scores for ability and achievement to common standard scores (mean 100, SD 15).

Example

a. Transform pupil's intellectual ability score to a CSS

Data needed for intellectual ability test:

Pupil's score on the intellectual ability test = X_{ab} — $X_{ab} = 60$

Mean of the intellectual ability test = M_{ab} — $M_{ab} = 50$

Standard deviation of the intellectual ability test = SD_{ab} — $SD_{ab} = 10$

Formula:

$$\text{Common Standard Score}_{ab} = \left[\left(\frac{\text{Pupil ability score - Ability test mean}}{\text{Standard deviation}}\right) \times 15\right] + 100$$

$$CSS_{ab} = \left[\left(\frac{X_{ab} - M_{ab}}{SD_{ab}}\right) \times 15\right] + 100$$

$$CSS_{ab} = \left[\frac{(60 - 50)}{10} \times 15\right] + 100$$

$$= \frac{10}{10} \times 15 + 100$$

$$= 115$$

Example

b. Transform pupil's achievement score to a CSS

Data needed for achievement test:

	Example
Pupil's score on the achievement test = X_{ach}	16
Mean of the achievement test = M_{ach}	30
Standard deviation of the achievement test = SD_{ach}	7

Formula:

$$\text{Common Standard Score}_{ach} = \left[\left(\frac{\text{Pupil's achievement score - Achievement test mean}}{\text{Standard deviation of achievement test}}\right) \times 15\right] + 100$$

$$CSS_{ach} = \left[\left(\frac{X_{ach} - M_{ach}}{SD_{ach}}\right) \times 15\right] + 100$$

$$\left(\frac{16 - 30}{7}\right) \times 15 + 100$$

$$= \left(\frac{-14}{7} \times 15\right) + 100$$

$$= -30 + 100$$

$CSS_{ach} = 70$

Now proceed to step 3.

Step 3: Obtain Pupil's Discrepancy Score
Read Criterion Discrepancy Value

a. Procedure: Substract the pupil's achievement score (CSS_{ach}) from the pupil's ability score (CSS_{ab}) to obtain the pupil's discrepancy score.

		Example	
Pupil's ability score (CSS_{ab})	CSS_{ab}	CSS_{ab}	115
Pupil's achievement score (CSS_{ach})	minus CSS_{ach}	CSS_{ach}	- 70
	discrepancy score		45

b. Procedure: Using the correlation between the tests, find the criterion value from Table 2. Note: The statistical correlation between the ability and achievement tests is given as r=.__ __(two digit decimal). Do not confuse this with a reliability coefficient, which appears similar. A correlation coefficient always involves two tests. A reliability coefficient pertains to a single test.

Now proceed to step 4.

Table 2

Criterion Discrepancy Values for Correlation Coefficients .01 - .99

Correlation	Criterion Value	Correlation	Criterion Value
.01	32	.51	22
.02	32	.52	22
.03	31	.53	22
.04	31	.54	22
.05	31	.55	21
.06	31	.56	21
.07	31	.57	21
.08	31	.58	21
.09	30	.59	20
.10	30	.60	20
.11	30	.61	20
.12	30	.62	20
.13	30	.63	19
.14	30	.64	19
.15	29	.65	19
.16	29	.66	19
.17	29	.67	18
.18	29	.68	18
.19	29	.69	18
.20	28	.70	17
.21	28	.71	17
.22	28	.72	17
.23	28	.73	17
.24	28	.74	16
.25	28	.75	16
.26	27	.76	16
.27	27	.77	15
.28	27	.78	15
.29	27	.79	15
.30	27	.80	14
.31	26	.81	14
.32	26	.82	14
.33	26	.83	13
.34	26	.84	13
.35	26	.85	12
.36	25	.86	12
.37	25	.87	11
.38	25	.88	11
.39	25	.89	11
.40	25	.90	10
.41	24	.91	10
.42	24	.92	9
.43	24	.93	8
.44	24	.94	8
.45	24	.95	7
.46	23	.96	6
.47	23	.97	6
.48	23	.98	5
.49	23	.99	3
.50	23		

Step 4: Compare the Pupil's Discrepancy Score to the Criterion Discrepancy to Determine if the Pupil has a Severe Discrepancy

	Example
Pupil's discrepancy score:	45
Criterion discrepancy score:	25

If the pupil's discrepancy score is equal to or greater than the criterion discrepancy value, a severe discrepancy is indicated.

If the pupil's discrepancy score is less than the criterion discrepancy value, a severe discrepancy is not indicated.

In this example, 45 is greater than 25. A severe discrepancy is indicated.

Summary Work Sheet

Data needed:

Ability Test: Name _____ M_{ab} _____ SD_{ab} _____ Pupil's score X_{ab} _____

Achievement Test: Name _____ M_{ach} _____ SD_{ach} _____ Pupil's score X_{ach} _____

Correlation between tests: $r_{ab,\ ach}$ = _____ Criterion Discrepancy Value (from table 2) _____

Computational Procedures

Step 1

Determine test means and standard deviations, and pupil scores.

Step 2

$$\text{Pupil's common standard score ability} = \frac{\text{Pupil's ability score - Ability test mean}}{\text{Standard deviation of ability test}} \times 15 + 100$$

$$\text{Pupil's } CSS_{ab} = \left[\left(\frac{X_{ab} - M_{ab}}{SD_{ab}}\right) \times 15\right] + 100$$

$$\text{Pupil's common standard score achievement} = \frac{\text{Pupil's achievement score - Achievement test mean}}{\text{Standard deviation of achievement test}} \times 15 + 100$$

$$\text{Pupil's } CSS_{ach} = \left[\left(\frac{X_{ach} - M_{ach}}{SD_{ach}}\right) \times 15\right] + 100$$

Step 3

$$\text{Pupil's discrepancy score} = CSS_{ab} - CSS_{ach}$$

Step 4

Comparison of pupil's discrepancy score to criterion discrepancy value

Criterion _____

Pupil _____

If the pupil's discrepancy score is equal to or greater than the criterion discrepancy value, a severe discrepancy is indicated.

If the pupil's discrepancy score is less than the criterion discrepancy value, a severe discrepancy is not indicated.

WHEN CORRELATIONS ARE NOT AVAILABLE: AN ESTIMATION PROCEDURE

When intellectual ability and achievement are measured by tests for which correlations are not available, the computational procedure cannot be applied. A district may choose to conduct a study to establish the correlation between the tests most commonly used in that district to allow use of the computational procedure. Another alternative is to use the following guidelines, which help to establish the probability that a severe discrepancy does or does not exist.

a. When the difference between a pupil's ability test standard score* and achievement test standard score* is less than 17 standard score points, the probability is low that the 1.5 criterion has been met. Additional evidence should be examined, however, before determining that a severe discrepancy does not exist.

b. When the difference between the pupil's ability test standard score and achievement test standard score is more than 28 standard score points, the probability is high that the 1.5 criterion has been met. Corroborating evidence should be examined, however, before determining that a severe discrepancy does exist.

c. When difference between the pupil's ability test standard score and achievement test standard score is between 17 and 28 standard score points, the probability is moderate that the 1.5 criterion has been met. Further study and examination of the student's performance are necessary to determine if a severe discrepancy exists.

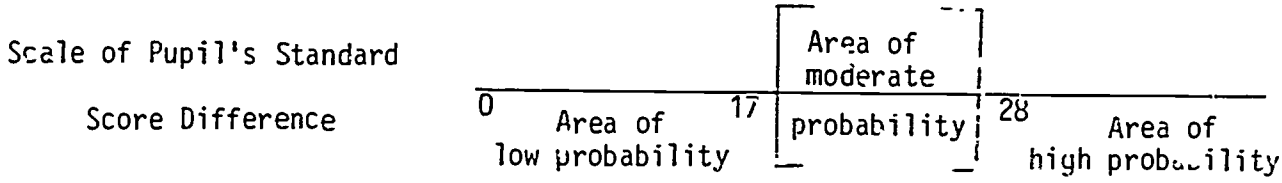
---

* To use this estimation procedure, both test scores must be on a scale with a mean of 100 and a standard deviation of 15. (See Steps 1 or 2 on pp. 32-35.)

Example: A pupil has been tested with the WISC-R and the 123 Math Test. No correlation can be found. The pupil obtained WISC-R scores of FS 105, V 103, P 104. The pupil's score on the 123 Math test (using a scale with mean 100, SD 15) is 85.

WISC-R	105
- 123 Math	85
	20 points

The probability is moderate that the 1.5 criterion has been met. Further study is needed to establish whether there is a severe discrepancy.



82-125 DE12920 3-83 3M

APPENDIX B

Date

Dear parent of_____:

Los Angeles Unified School District is currently participating in a research project with the National Center for Bilingual Research. The research project is concerned with assessment of young children.

Your child has been selected as a potential participant in this project. We would like an opportunity to discuss the project and your participation in it. One of our assistants will be in contact with you to arrange a meeting time.

Thank you for your cooperation and interest in this worthy project.

Sincerely,

_____

Dr. Victor Rodriguez
National Center for Bilingual Research

_____

Los Angeles Unified School District

Fecha

Estimados Padres de Familia de _____:

El Districto Unificado Escolar de Los Angeles (Los Angeles Unified School District) está participando en un proyecto de investigación con el Centro Nacional sbore Investigación BilingÜe (National Center for Bilingual Research). El propósito de este proyecto es estudiar la evaluación de niños jóvenes.

Su hijo/hija ha sido seleccionado(a) para participar en este proyecto. Nos gustaría tener la oportunidad de discutir el proyecto y la participación suya en él. Uno de nuestros asistentes estará en contacto con Usted(es) para arreglar una visita con Usted(es).

Gracias por su cooperación e interés en este proyecto.

Atentamente,

_____
Dr. Víctor Rodríguez
Centro Nacional sobre Investigación BilingÜe

_____
Distrito Escolar Unificado de Los Angeles

APPENDIX C

PERFORMANCE OF SPANISH-SPEAKING MILDLY HANDICAPPED AND AND NON-HANDICAPPED PUPILS ON SPANISH EDITIONS OF THE SOMPA AND KABC

The Research Plan. The Handicapped Minority Research Institute (HMRI), which is part of the National Center for Bilingual Research (NCBR), is studying the assessment of Spanish-speaking pupils. The (Name of school district) and the (Name of the high school) have reviewed the study and agreed to participate. In each school, information will be obtained from Spanish-speaking pupils in special education placements and non-special placements and from their parents. The parent will take part in an interview which lasts approximately one and one half hours. The parent will be asked to complete questions from the Adaptive Behavior Inventory for Children and from the Sociocultural Scales. Once each parent has given informed, written consent for his/her child to participate in the study, the assessment process will begin. Each child in this study will be individually administered the Spanish editions of the System for Multicultural and Pluralistic Assessment (SOMPA) and the Kaufman Assessment Battery for Children (KABC) by bilingual school psychologists.

Purpose of this Study. This study will help children, parents, teachers and educators evaluate the usefulness of Spanish translated assessment instruments. The information derived from this study may assist educators and researchers in identifying nonbiased methods of assessment. In this way children will not be penalized for their unfamiliarity with the English language and culture.

Effect on your Child's Schooling. Participating in this study will help children and their parents learn more about the academic and cognitive functioning of the child. However, being a part of this study is not expected to have any particular effect--good or bad--on your child's performance in school. Schools and teachers will not use any of the information collected in this study for grading or evaluating students, without the parents written permission. The study will not take away from the amount of time that the student is in

school or receiving instruction. All assessment sessions will be scheduled at the convenience of each child and every effort will be made to minimize interference with school demands. Likewise, parent interviews will be arranged to accomodate the parent. A five dollar reimbursement will be given to the parent for his/her time spent in the interview.

Continuing in the Study. This study is scheduled to last for one month. However, parents and children may decide to withdraw from the study at any time, without consequences in or out of school. If for any reason, parents and/or children wish to withdraw from the study, they should notify the HMRI staff member, or one of the persons listed below.

Protecting Privacy. The HMRI will act in every way to respect the privacy of the family and the students. No real names or other specific identifiers, such as family, address, will be used. In general, not even the names of the neighborhoods or the schools will be used. Parents and students have the right to terminate involvement in this study at any time. Only after information from the study has been edited to protect the privacy of participating students and families will data from the study be made available to other researchers and educators for use in the development of nonbiased assessment methods.

School Records. The research plan includes comparing the student's performance on the Spanish editions of the SOMPA and the KABC to other school related assessment. To do this, HMRI staff will record assessment information from the student's school records. This includes tests scores from the most recently administered standardized achievement tests (e.g. CTBS, WRAT) and from language dominance tests. Parent's permission is required for looking at these records. The HMRI will not release any information from these files that can be associated with any particular student.

Scheduling Interviews and Assessment. The HMRI staff member will schedule each parent interview in advance by telephone. The school

psychologist will also schedule assessment sessions with the student in advance. Assessment sessions will always be scheduled at a time that is most convenient for the student and does not interfere with the requirements of school and/or home. Assessment sessions will take place at _____sites. If it is necessary for an HMRI staff member of the school psychologist to make a change in plans, you will be notified.

Questions or Problems. Any questions or problems related to the study may be discussed with the HMRI staff member of School Staff listed below:

Dr. Pauline Mercado

Dr. Robert Rueda

Dr. Victor Rodriguez

(Name and location of sch staff)

PERFORMANCE OF SPANISH-SPEAKING MILDLY HANDICAPPED AND NON-HANDICAPPED PUPILS ON SPANISH EDITIONS OF THE SOMPA AND KABC

1. I have read the attached "Information for Parents" for the study entitled "Performance of Spanish-speaking Mildly Handicapped and Non-handicapped pupils on Spanish Editions of the SOMPA and KABC" conducted by the Handicapped Minority Research Institute (HMRI).

2. I understand that this study will involve the following child from our family: _____

3. I understand that this study is planned to continue for one month, but that I or my child may request to withdraw from the study at any time.

4. I agree that an HMRI staff member may interview me in order to obtain information relative to my child. Also, I agree that my child may be administered the Spanish editions of the SOMPA and KABC.

5. I agree that HMRI staff may have access to the school records previously cited of _____ throughout the course of this study for any purpose related to the study, and direct _____ School to provide access.

6. I understand that HMRI will protect the privacy of our family and our child, and will not release names, address, telephone numbers, or any similar information which could identify individuals or family members.

_____ _____
Signature of Parent/Guardian Date

_____ _____
Printed Name

_____

_____ _____
Address Telephone Number

ACTUACION DE ESTUDIANTES DE HABLA HISPANA LEVEMENTE INCAPACITADOS Y NO-INCAPACITADOS SOBRE EDICIONES EN ESPAÑOL DE SOMPA Y KABC

Plan de Investigación. El Handicapped Minority Research Institute (HMRI - Instituto de Investigación sobre Minorías Incapacitadas) el cual forma parte del National Center for Bilingual Research (NCBR - Centro Nacional sobre Investigación Bilingüe) está estudiando la evaluación de estudiantes de habla hispana. El _____ School District (Distrito Unificado de Escuelas de _____ y el (nombre de la escuela) han revisado el estudio y han acordado en participar. Será obtenida en cada escuela información de padres de familia y de estudiantes de habla hispana que están en clases regulares y estudiantes que están recibiendo educación especial. El padre/madre de familia tomará parte en una entrevista la cual durará aproximadamente una hora y media, para obtener información sobre la niñez del hijo/hija y sobre su desarrollo físico y mental. Una vez que cada padre/madre haya dado su consentimiento por escrito para que su hijo/hija participe en el estudio, se comenzará con el proceso de evaluación. En este estudio, cada niño/niña será administrado individualmente las ediciones en español del Sistema para Evaluación Multicultural y Pluralístico (System for Multicultural and Pluralistic Assessment - SOMPA) y Batería de Evaluación Kaufman para Niños (Kaufman Assessment Battery for Children - KABC) por sicólogos bilingües del distrito escolar de Los Angeles.

Propósito del Estudio. Este estudio ayudará a los niños, padres de familia, maestros y educadores a evaluar la utilidad de estos nuevos instrumentos de evaluación en español. La información derivada de este estudio puede asistir a los educadores e investigadores a identificar métodos de evaluación sin prejuicios. De esta manera, los niños no serán castigados por no conocer el idioma inglés y su cultura.

Efecto sobre la Enseñanza de su Hijo/Hija. La participación en este estudio ayudará a los padres de familia y a los educadores a aprender más sobre el funcionamiento académico e intelectual del niño. Sin embargo, al formar parte de este estudio, no se tendrá ningún efecto

particular--bueno o malo--sobre el desempeño de su hijo en la escuela. Las escuelas y los maestros no usarán la información coleccionada de este estudio para calificar o evaluar a los estudiantes sin el permiso por escrito de los padres de familia. No se tomará tiempo para este estudio mientras el estudiante esté en la escuela recibiendo instrucción. Todas las sesiones de evaluación serán citadas según la conveniencia de cada niño y se hará todo esfuerzo para reducir interrupciones con los deberes escolares. Al mismo tiempo, serán preparadas las entrevistas con los padres de familia según la conveniencia del padre/madre. Se les ofrecerá un donativo a los padres por su tiempo que estarán en la entrevista.

Continuación en el Estudio. Este estudio está programado para durar un mes. Sin embargo, los padres de familia y los niños pueden decidir retirarse del estudio en cualquier tiempo, sin tener consecuencia dentro o fuera de la escuela. Si por algún motivo los padres de familia y/o niños desean retirarse del estudio, ellos notificarán al funcionario del HMRI o a una de las personas detalladas abajo.

Protección de Privacidad. El HMRI actuará de toda manera para respetar completamente la privacidad de la familia y los estudiantes. No se usarán nombres verdaderos u otras formas de identificación, tales como el domicilio de la familia. En general, no se usarán nombres de las vecindades ni de las escuelas. Los padres de familia y los estudiantes tienen el derecho de terminar la participación de este estudio en cualquier tiempo. Sólo después de que haya sido redactada la información del estudio para proteger la privacidad de los estudiantes participantes y sus familias, se hará disponible la información del estudio a otros investigadores y educadores con el fin de usarse en el desarrollo de métodos de evaluación sin prejuicios.

Archivos de Escuelas. El plan de investigación incluye la comparación del funcionamiento del estudiante sobre las ediciones en español del SOMPA y KABC con otras evaluaciones relacionadas con la escuela. Para hacer esto, los funcionarios del HMRI revisarán información de evaluación de los archivos escolares del estudiante. Esto incluye, por

ejemplo, calificaciones de las más recientes pruebas administradas por las escuelas (por ejemplo CTBS, WRAT) y los resultados de las pruebas de dominio del idioma. Se necesita el permiso de los padres de familia para poder ver estos archivos. El HMRI no divulgará ninguna información de estos archivos que pueda ser asociada con cualquier estudiante.

Programar Entrevistas y Evaluación. El funcionario del HMRI programará cada entrevista con el padre de familia por adelantado y por teléfono. El sicólogo de la escuela también programará sesiones de evaluación con el estudiante por adelantado. Las sesiones de evaluación serán siempre programadas cuando sea más conveniente para el estudiante con tal que no interfieran con los requerimientos de la escuela y/o hogar. Las sesiones de evaluación tomarán lugar en los sitios_____. Si el funcionario del HMRI o sicólogo escolar necesite hacer cambio de planes, ustedes serán notificados.

Preguntas o Problemas. Podrá ser discutido cualquier pregunta o problema relacionado con el estudio con un funcionario del HMRI, o funcionario de la escuela detallado a continuación:

Dr. Pauline Mercado
Dr. Robert Rueda
Dr. Víctor Rodríguez

(Nombre y lugar del funcionario de la escuela)

ACTUACION DE ESTUDIANTES DE HABLA HISPANA LEVEMENTE INCAPACITADOS Y NO-INCAPACITADOS SOBRE EDICIONES EN ESPAÑOL DE SOMPA Y KABC

1. Hemos leído la "Información para Padres de Familia" adjunta para el estudio titulado "Actuación de Estudiantes de Habla-Hispana Levemente Incapacitados y No-Incapacitados sobre Ediciones en Español de SOMPA y KABC," y conducido por el Handicapped Minority Research Institute (HMRI - Instituto de Investigación sobre Minorías Incapacitadas).

2. Tenemos entendido que en este estudio participará nuestro hijo/hija:_____

3. Tenemos entendido que este estudio está planeado para continuar por un mes, pero que mi hijo/hija o nosotros podremos solicitar retirarnos del estudio a cualquier momento.

4. Estamos de acuerdo de que un funcionario del HMRI nos entreviste para obtener información en relación con mi hijo/hija. Además, estamos de acuerdo que pueden ser administradas a mi hijo/hija ediciones en español de SOMPA y KABC.

5. Estamos de acuerdo de que los funcionarios del HMRI podrán tener acceso a los archivos escolares de_____ a través del curso de este estudio para cualquier propósito relacionado con el estudio, y dirigimos la escuela_____ para proporcionar acceso.

6. Tenemos entendido de que HMRI protegerá la privacidad de nuestra familia y de nuestro hijo/hija, y que no divulgará nombres, domicilios, números telefónicos o cualquier información similar que pudiera identificar a personas o miembros de nuestra familia.

_____ _____

Firma de Padre de Familia/Guardián Fecha

_____

Nombre con letra de molde

_____

_____ _____

Domicilio Número de Teléfono

APPENDIX D

California State Department of Educat.on
Office of Special Education
Information Concerning Eligibility Criteria

No. 1 Assessing Pupils Suspected of Having a Specific Learning Disability Pursuant to Title 5, CAC, Section 3030(j)

Assessment of pupils for eligibility as individuals with exceptional needs shall be made in accordance with Education Code Sections 56320-56329 and 56338.* The assessment of a pupil suspected of having a specific learning disability requires the determination of a significant discrepancy between intellectual ability and achievement in one or more of the following academic areas: oral expression, listening comprehension, written expression, basic reading skills, reading comprehension, mathematics calculation, or mathematics reasoning.

Further, the discrepancy must be determined to be directly related to a disorder in one of the basic psychological processes which include: attention, visual and auditory processing, sensory-motor skills, and cognitive abilities including association, conceptualization, and expression.

Federal and state statutes require consideration of the following before eligibility can be established:

- The discrepancy cannot be due to environmental, cultural differences or economic disadvantages.
- The discrepancy cannot be due primarily to mental retardation or emotional disturbance.
- The discrepancy cannot be due primarily to visual, hearing, or motor handicaps.
- The discrepancy cannot be corrected through other regular or categorical services offered within the regular instructional program.

School personnel must consider and, where appropriate, utilize the resources of the regular education program before a pupil is referred for special education. Many school districts are using the child study team or guidance council concept as a way of coordinating and monitoring all available services. When a formal referral is initiated, or refused, the due process guarantees and procedural time lines must be followed.

Once a referral has been made, the following steps will provide the information necessary to determine whether a pupil is eligible for special education on the basis of a specific learning disability:

---

*Education Code Section 56338 requires additional consideration beyond that which is required by Education Code Section 56337.

(1) Determine if the pupil has a severe discrepancy between intellectual ability and achievement.

The determination of a severe discrepancy necessitates the use of a standardized achievement test and a test of intellectual ability. The required 1.5 standard deviation difference criterion is the critical threshold that must be corroborated by other assessment data.

The Department has prepared a number of tables to assist in determining if the test results meet the criterion for a "severe discrepancy." Instructions for their use are appended to the tables.

(2) Determine if the pupil has a disorder in one of the basic psychological processes.

State law requires that the evidenced discrepancy be directly related to a processing disorder. This will necessitate that, as part of the multidisciplinary assessment, a psychological processing disorder must be identified in one or more of the following five areas:

(a) Attention
(b) Visual processing
(c) Auditory processing
(d) Sensory-motor skills
(e) Cognitive abilities, including association, conceptualization, and expression

The relationship of the results to the pupil's academic performance should be clearly established by the IEP team.

(3) The discrepancy must not be due to factors of environment, cultural differences or economic disadvantage Also, the discrepancy must not be the result of visual, hearing or motor handicaps, mental retardation, limited school experience or poor attendance.

(4) The discrepancy cannot be corrected through other regular or categorical services offered within the regular instructional program.

Although a pupil must meet the 1.5 standard deviation difference criterion, it is not the only criterion which must be met in order for a child to be eligible for special education. Not all pupils who meet the 1.5 (standard deviation difference) critical threshold and have a disorder in one or more of the basic psychological processes will need to be placed in special education. For example, a student may have a learning disorder, but with support in the form of special materials, teaching strategies, instructional techniques, instructional grouping, or consultative services to the regular program, the student can be appropriately served in the regular classroom without special education placement.

The IEP team must document that the pupil's academic deficits cannot be corrected through modifications of the regular educational program.

Programming and Placement

The assessment results are to be used as the basis for educational programming. Goals and objectives in the IEP must address areas of identified needs. Since there must be a causal relationship between the pupil's academic learning problems and one or more of the processing disorders, specific objectives should address how the identified processing disorders will be ameliorated along with appropriate academic goals and objectives.

Program placement, in one of the options authorized by Education Code Section 56361, should be based on the educational setting that can most appropriately provide for the implementation of the pupil's identified goals and objectives. For example, the determination of placement in the resource specialist program should be made on the basis of the pupil's needs and not on the degree of discrepancy as expressed by the standard deviation of the difference score.

The IEP team should document the reasons for the placement decision so that during required reviews, the placement rationale can be evaluated for its continued applicability.

Use of Standardized Tests

Section 3030(j)(5), Title 5, California Administrative Code states:

> "When standardized tests are considered to be invalid for a specific pupil, the discrepancy shall be measured by alternative means as specified on the assessment plan."

If it is determined that the use of standardized tests would be or is an invalid assessment tool, the assessment personnel will have to use professional judgment, based on such data as the results of informal or criterion-referenced assessments, analysis of pupil work samples, classroom performance and observations to determine the evidence of a severe discrepancy. The need for professional judgment will apply to the areas of written expression and listening comprehension since there are few, if any, standardized tests which measure these skill areas. Special attention should be given in the assessment of pupils whose primary language is other than English, or whose cultural background might mitigate against the use of a certain standardized test.

Summary

The assessment of pupils suspected of having a specific learning disability is highly complex and must be based on a systematic and multidisciplinary assessment. No one procedure can form the basis for identifying a pupil as an individual with exceptional needs.

Standardized test scores must be confirmed by other assessment data and the professional judgment of the assessment team must confirm that the discrepancy between ability and achievement is due to a disorder in one of the basic psychological processes and not due to environment, cultural differences or economic disadvantages; nor due primarily to mental retardation or emotional disturbance; nor due primarily to sensory or motor handicaps or limited school experience or poor school attendance.