This is to certify that the

dissertation entitled

A Comparison of Methods for
Calculating a Severe Discrepancy Between
Ability and Achievement Using the WISC-R,
PIAT, and K-ABC

presented by

Deborah Ellen Bennett

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in School Psychology

Major professor

Date November 12, 1986

A COMPARISON OF METHODS FOR CALCULATING A SEVERE DISCREPANCY BETWEEN ABILITY AND ACHIEVEMENT USING THE WISC-R, PIAT, and K-ABC

By

Deborah Ellen Bennett

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Counseling, Educational Psychology
and Special Education

1986

ABSTRACT

A COMPARISON OF METHODS FOR
CALCULATING A SEVERE DISCREPANCY BETWEEN
ABILITY AND ACHIEVEMENT USING THE WISC-R,
PIAT, AND K-ABC

By

Deborah Ellen Bennett

Four methods for determining a severe discrepancy between ability and achievement for the purposes of diagnosing students as learning disabled were compared using scores from the Wechsler Intelligence Scale for Children-Revised, The Peabody Individual Achievement Test, and the Kaufman Assessment Battery for Children.

Eighty-six students who were referred for academic difficulties were tested with the WISC-R, PIAT, and K-ABC. The score differences between the WISC-R Full Scale IQ and the PIAT subtests as well as between the K-ABC Mental Processing Composite and the K-ABC Achievement subtests were evaluated using four procedures: z-score difference, estimated true score difference, and two regression analysis procedures. The first regression procedure (unadjusted) considered only errors of estimate. The second regression procedure also included an adjustment for test unreliability.

A high degree of agreement in the selection of students was found between the z-score difference and estimated true score difference approaches, particularly for tests with high reliabilities. Considerable agreement was also found between the z-score difference, estimated true score difference, and the adjusted regression procedure when tests with high reliabilities were analyzed. Less agreement was found between the unadjusted regression procedure and the z-score, estimated true score, and adjusted regression approaches. The unadjusted regression method generally identified significantly fewer students than the other three approaches. No significant differences were found between student characteristics: age, sex, grade placement, and IQ across methods.

When comparing the WISC-R/PIAT discrepancies with the K-ABC MPC/Achievement discrepancies, it was found that the two approaches to the determination of a score difference resulted in different populations of students. Percent overlap ranged from 0 to 49.2 percent with 8 out of 16 comparisons resulting in less than 25 percent of the same students identified. The high average PIAT standard scores in combination with lower subtest reliabilities appeared to be the primary source of disparity between the two approaches to the calculation of a discrepancy. The PIAT identified fewer students than the K-ABC in all comparisons and selected less than half the number of students in ten of the sixteen comparisons.

ACKNOWLEDGMENTS

The author would like to thank Dr. Harvey Clarizio and Dr. Susan Phillips for the ideas and advice which have contributed so greatly to the accomplishment of this dissertation. Dr. Herbert Burks has given much time and energy, both in editing this text and in providing guidance throughout my doctoral program. Appreciation is also extended to Dr. Neal Schmitt and Dr. William Mehrens for their valuable comments.

A very special thanks goes to my husband, Vince Bralts, for endless hours of listening, advising, encouraging, and believing. The last note of appreciation is for the children: Natasha, Alisha, and Rebekah who make it all worthwhile.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

WISC-R	Wechsler Intelligence Scale for Children-Revised
K-ABC	Kaufman Assessment Battery for Children
PIAT	Peabody Individual Achievement Test
FSIQ	Full Scale IQ (WISC-R)
VIQ	Verbal IQ (WISC-R)
PIQ	Performance IQ (WISC-R)
MPC	Mental Processing Composite (K-ABC)
M.A.	Mental Age
C.A.	Chronological Age
G.A.	Grade Age
L.D.	Learning Disability

I. INTRODUCTION

The determination of a "severe discrepancy" between ability and achievement is a critical component in the process of diagnosing a learning disability. In order to be considered for Special Education support services under the learning disabilities classification, a student must first exhibit a level of achievement in one or more specified academic areas which is significantly below his or her "expected level of achievement." The expected level of achievement is typically based upon some estimate of ability, such as the student's score on an individually administered intelligence test. The process of quantifying this discrepancy has been an area of increasing concern for school psychologists, measurement specialists, and educational administrators. It has become more obvious that the method of discrepancy determination which is implemented by a school system can have a substantial impact upon the number and characteristics of students who are serviced in learning disability (L.D.) programs.

Unfortunately, there have been few studies which have systematically analyzed the consequences of different discrepancy procedures upon the characteristics of the selected L.D. population. While earlier research has attempted to compare expectancy formulas, little has been done to compare the impact of more recently advocated methods such as the standard score discrepancy (Reynolds, 1981), estimated true score discrepancy (Cone & Wilson, 1981), and regression analysis (Wilson & Cone, 1984) approaches. While considered by most measurement experts to be more acceptable approaches to the quantification of a significant discrepancy, there appears to be a paucity of research which considers the outcome of the actual implementation of each method (Reynolds, 1985; Wilson & Cone, 1984).

One of the major objectives of the present research is to compare the numbers and characteristics of students who demonstrate a severe discrepancy using these three methods. Do all three approaches tend to identify the same students or are there systematic differences in such variables as chronological age, IQ, grade placement, and sex? If there are no observed differences, then ease of application and interpretability of results may be the deciding factors in the selection of a discrepancy procedure. On the other hand, substantial differences in the population which demonstrates a severe discrepancy may warrant a closer examination of the psychometric adequacy of the selected approach as well as consideration of more philosophical issues, such as who should be served by programs for the learning disabled.

Another related issue involves assessing the impact of test characteristics, such as reliability and validity, upon the quantification of a significant discrepancy. To what extent do these critical variables influence the students who are identified by each method; and is there some minimal requirement in terms of technical adequacy for the tests which are used in the determination of a discrepancy? Although it is beyond the scope of this dissertation to recommend minimal standards for tests, such issues are often overlooked once a discrepancy procedure has been endorsed and implemented.

The second major objective of this research is to evaluate the adequacy of the recently published Kaufman Assessment Battery for Children (K-ABC) for use in the determination of a severe discrepancy between a student's ability and achievement. Of particular interest is the comparison of the K-ABC with the well-established and widely used Wechsler Intelligence Scale for Children-Revised (WISC-R) and the Peabody Individual Achievement Test (PIAT). Do the discrepancies found between the Mental Processing Composite (ability component) and the Achievement subtests of the K-ABC correspond with the discrepancies found between the WISC-R Full Scale IQ and the PIAT subtests, or will the selective use of the K-ABC over the WISC-R and PIAT result in the identification of different students?

As with the choice of a method for determining a significant discrepancy, the choice of tests to be administered could have a substantial impact upon students selected for learning disabilities programs. The K-ABC is unique in many ways from the WISC-R and PIAT, both in terms of underlying theory and nature of abilities which are sampled. Because the test authors have purposely attempted to depart from the classical approach to intelligence testing, basic questions need to be answered regarding the application of the K-ABC to the diagnosis of learning disabilities. Again, how will use of the K-ABC in the determination of a discrepancy between ability and achievement affect the size and characteristics of the L.D. population?

II. REVIEW OF LITERATURE

The Review of Literature has been written to analyze and discuss two broad areas of research. The first area of investigation involves the development and analysis of various methods for determining a "severe discrepancy" between ability and achievement. The review starts by examining and evaluating simpler discrepancy methods, such as deviations from grade level, and progresses to the more recently advocated and statistically sophisticated approaches to the problem. An attempt has been made to discuss the benefits and disadvantages of each approach and ultimately provide a rationale for selecting the four methods which have been implemented in the present study.

The second area of investigation focuses on research related to the Kaufman Assessment Battery for Children. Because of its recency in the evolution of individual assessment, both the theoretical structure and the nature of the tasks included in the battery have been outlined and discussed. Studies which have examined the relationship between the K-ABC and other achievement and intelligence tests offer insights regarding how the results of the K-ABC might compare with the WISC-R and the PIAT.

Finally, several critiques of the K-ABC will be summarized. These reviews provide an interesting perspective concerning the accuracy of the K-ABC in reflecting the underlying theoretical constructs selected by its authors. Some basic differences between traditional notions of intelligence and the Kaufmans' concept of "mental processing" are highlighted, and implications of the literature are related to the objectives of the present study.

A. Severe Discrepancy

Cone and Wilson (1981) have provided a useful categorization of available procedures for determining a severe discrepancy between a student's ability and achievement for the purposes of identifying learning disabilities. The categories offer a helpful structure for comparing the many formulas and include: 1. Deviation from grade level, 2. Expectancy formulas, 3. Standard score comparisons, and 4. Regression analyses.

1. Deviation from Grade Level

The deviation from grade level method for determining a severe discrepancy involves the identification of students who score substantially below their grade level on achievement tests. Other sources have referred to this procedure of identifying discrepancy as "years behind" (Forness, Sinclair, & Guthrie, 1983; Erickson, 1975) or simply "low achievement" (Epps, Ysseldyke, & Algozzine, 1985). Cone and Wilson further make the distinction between "constant deviation" and "graduated deviation" procedures. Constant deviation methods require the determination of a fixed discrepancy (e.g., two grade levels) between a student's current grade placement and level of achievement. Graduated deviation procedures allow for variation across grade levels with a smaller deviation required for a "severe discrepancy" at lower grade levels and a larger deviation required at upper grade levels.

Reynolds (1981) discusses the disadvantages of using a constant grade level deviation as a diagnostic criterion for learning disabilities. He explains that "the use of grade equivalent scores at a constant discrepancy level irrespective of actual grade placement produces considerable irregularity and distortion in the magnitude of aptitude/achievement discrepancies required for diagnosis of a learning disability across grade levels" (p. 351). A clear example of this irregularity can be seen by comparing the educational significance of a two year below grade level discrepancy for a third grade student with the educational significance of an eleventh grade student who is achieving at the ninth grade level.

The increasing range of achievement scores which occurs with the progression to higher grade levels was addressed, at least partially, with the introduction of the graduated deviation procedures for determining discrepancy. Several difficulties, however, are still inherent in this approach. Cone and Wilson point out that a graduated deviation formula (as well as a constant deviation formula) discriminates against higher ability students who may exhibit learning disabilities. The brighter the child, the more unlikely it is that a severe discrepancy will be identified. Supportive evidence comes from a study by Fields (1979) which compared the characteristics of students identified using an expectancy formula, a deviation from grade level procedure, and a standard score comparison procedure. Students identified using the expectancy formula and the standard score comparison procedure had significantly higher Full Scale IQs than the group identified using the deviation from grade level method. Reynolds (1984-85) also criticizes these methods as overidentifying children who fall in the "slow learner" range, including many children who are functioning academically at a level which is consistent with their general intellectual ability and age. "Only for children with IQs of precisely 100 will there be no bias in diagnosis with these models" (p. 457).

2. Expectancy Formulas

During the past two decades there has been a proliferation of methods for determining a student's expected level of achievement. Most of these methods utilize some combination of ability (typically IQ) and age or grade level. Cone and Wilson (1981), Danielson and Bauer (1978), Epps, et al. (1985), O'Donnell (1980), and Harris and Sipay (1980) describe and critique several of these formulas. For the purpose of this review, a limited number of the most widely used formulas will be discussed. Many of the difficulties encountered with the use of these selected procedures are also inherent in other expectancy formulas.

One of the simplest and most widely applied procedures for quantifying a student's expected level of achievement is the formula advocated by Kaluger and Kolson (1969):

$$(IQ \times CA)/100 - 5$$

which yields an expected grade level for each student. While useful for students who have entered school at age five and who have not subsequently repeated a grade, the formula becomes more difficult to interpret for students who deviate substantially from this norm.

A second formula which attempts to adjust for the amount of time that a student has spent in school is the revised Bond and Tinker (1973) formula:

$$\text{\# of years in school} \times IQ/100 + 1$$

This formula has been criticized for the same reasons that the deviation from grade level procedures have been criticized, specifically that there is a tendency to overidentify slow learners as learning disabled (Cone & Wilson, 1981). The advantage of this formula over the deviation from grade level procedure is that an attempt is made to incorporate a child's ability level in the determination of a learning discrepancy.

Johnson and Myklebust (1967) propose a third method for the determination of achievement expectancy which includes the child's ability level, chronological age, and "grade age":

$$\text{Expected grade equivalent} = (MA + CA + GA)/3 - 5$$

Cone and Wilson suggest that this formula better accounts for the increasing range and variability of students at higher grade levels than the previously cited formulas.

Finally, the formula which has received the most critical attention in recent years is the modified Harris formula which was proposed by the Bureau of Education for the Handicapped in 1976. The "Federal formula" originated from the 1970 Harris formula for expected grade level:

$$\text{Expected grade equivalent} = (2MA + CA)/3 - 5$$

By substituting a standard score (IQ) for mental age and setting a severe discrepancy at half a child's expected grade equivalent or below, the Harris formula was modified to the following:

$$\text{Severe discrepancy} = CA(IQ/300 + .17) - 2.5$$

Perhaps because of attempts at widespread implementation of this procedure, a number of strong criticisms were leveled at the proponents of the formula. Cone and Wilson (1981) summarize many of the points initially presented by Danielson and Bauer (1978). First, as with several of the previous procedures for the quantification of a severe discrepancy, this formula tends to identify a disproportionate number of slow learners. Secondly, younger children are more likely to be identified as exhibiting a discrepancy between ability and achievement than older children. The formula also will identify more children using certain tests than others because of the failure to consider errors of measurement. This last criticism, however, is applicable to all expectancy formulas. Finally, only 58 percent of children currently receiving support services under the learning disabilities classification would be identified using the Federal formula.

In a study which compared the proportion of students who demonstrated a severe discrepancy using a number of different expectancy formulas, Forness, et al. (1983) found considerable variation in both the incidence of identification and the agreement regarding which children demonstrated a discrepancy. In a sample of 92 children who had been hospitalized for evaluation or short-term treatment of behavior and learning problems, 32 percent were identified as learning disabled using the deviation from grade level procedure, 18.5 percent were identified using the Bond and Tinker formula, 21.7 percent were identified using the Kaluger and Kolson formula, and 24 and 25 percent of the children demonstrated a severe discrepancy using the 1975 Harris formula and the Federal formula, respectively.

Not surprisingly, there was almost perfect agreement between the Federal formula and the Harris formula in identifying children as having a discrepancy between ability and achievement. Somewhat lower agreement was shown between the other procedures, with deviation from grade level being the least consistent procedure among the methods described in this review.

Expectancy formulas have been questioned on a variety of technical grounds, some of which have been discussed in the preceding paragraphs. Wilson and Cone (1984) point out that expectancy formulas ignore the reliabilities of the measures which are incorporated into the calculations, as well as the intercorrelations between the tests. Reynolds (1984-85) objects to expectancy formulas because they attempt "mathematical operations that are not considered appropriate for the types of measurement being employed" (p. 452). The treatment of age and grade information as interval data or ratio data in the calculation of a discrepancy leads to meaningless and misleading results according to Reynolds and others. (More will be said about this concern when discussing standard score procedures for determining a discrepancy.)

Cone and Wilson (1981) echo and amplify Reynolds' concerns regarding the use of expectancy formulas, emphasizing the fact that none of the formulas "addresses errors of measurement, regression toward the mean, norm group comparability, a priori knowledge of incidence, or the increased range and variability of obtained scores for students at higher grade levels" (p. 363). Furthermore, the comparison of grade or age equivalents across tests or, even within tests, is troublesome because of the lack of equal variability of this type of score. As a result of the numerous technical difficulties encountered when using these procedures, most psychometricians patently argue against implementation of the expectancy formulas in determining a discrepancy.

3. Standard Score Comparisons

A number of authors (Erickson, 1975; Reynolds, 1981) have promoted the use of standard score comparisons as a more appropriate method of dealing with the quantification of a severe discrepancy. These procedures typically involve calculating the difference between comparable standard scores (same mean and standard deviation) on an ability test (typically IQ) and on an achievement test. It is expected that the achievement standard score will be in approximately the same relative position in the distribution as the ability estimate.

One form of standard score comparison is the z-score discrepancy procedure outlined by Erickson (1975) in which each test or subtest score is linearly transformed to a z-score by subtracting the group mean from the observed score and dividing the residual by the test standard deviation.

Generally an arbitrary cut-off, with the achievement measure one to two standard score units below the ability measure, is established as a criterion for a severe discrepancy. Reynolds (1981) has attempted to establish the significance of a standard score discrepancy by dividing the z-score difference by the standard error of the difference score. While improving upon the technical quality of the standard score procedure, the statistically significant difference between the z-scores does not always translate into a difference with educational significance. Erickson (1975) stresses the importance of empirically establishing cut-off scores which reflect the characteristics of students in each comparison group.

There are several advantages in using a standard score procedure rather than an expectancy formula in the calculation of a severe discrepancy. Shepard (1980) points out that the Z-score method takes into consideration both ability and grade level when determining a difference. It also adjusts for differences in variability across grades by using the standard deviation for a particular grade. Additionally, standard scores allow for comparisons across tests, subtests, and age levels.

While the standard score procedures appear to offer advantages over the more traditionally used expectancy formulas, there are still difficulties which arise with the use of the Z-score method. Undoubtedly the major problem with this approach is the failure of the standard score procedure to account for the regression between ability and achievement which results from less than perfectly correlated measures. Shepard (1980) explains that "because the actual correlation [between IQ and achievement] is usually more on the order of 0.6, there will be a regression to the mean. It can be shown both mathematically and theoretically that bright children have above average achievement, but their relative position (i.e., Z-score) tends not to be as high as it is in the IQ distribution. Conversely, children with low IQs will, on the average, have relatively higher achievement status than IQ status, although they will still be below the mean" (p. 83).

The effects of the regression phenomenon can be seen quite clearly in a study by Fields (1979) which demonstrated that the z-score procedure identified students with higher IQs than did either the deviation from grade level or Bond and Tinker procedures. Erickson (1975) also presented results which indicate average IQs of 110.3, 92.9, and 88.1 for children identified using the z-score, Bond and Tinker, and deviation from grade level procedures, respectively.

Reynolds (1985) describes the erroneous reasoning of those who call for a discrepancy of one standard deviation between ability and achievement in order to create a pool of eligibility of 16 percent of the student population (one standard deviation below the mean falls at the 16th percentile of a normal distribution). He points out that if two test scores are positively correlated, which ability and achievement test scores invariably are, "the distribution of scores created by subtracting the scores of a set of students on both tests from one another will not be the same as the two univariate distributions" (p. 454). Ultimately the new distribution will have a much smaller standard deviation, serving to identify a smaller number of students as exhibiting a severe discrepancy than originally intended. This problem is further exacerbated by exclusionary practices such as eliminating from consideration children who have IQs less than 85 or children who are not performing below grade level.

On the other hand, Reynolds discusses the impact of multiple comparisons which typically occur when evaluating a child for learning disabilities. The practice of comparing a student's IQ with a number of achievement scores results in a greater incidence of "severe" discrepancies simply on the basis of chance.

Clearly there are a number of difficulties which are inherent in the use of standard score comparisons in the determination of a severe discrepancy. The problems tend to compound depending upon the specific practices and definitions of severe discrepancy which are implemented by a given state or district.

Cone and Wilson (1981) describe an alternative approach which takes into account the measurement errors inherent in the ability and achievement tests that are chosen for the calculation of a standard score discrepancy. Knowing the reliabilities of each test, it is possible to calculate the estimated true score for each observed score using the following formula:

$$\hat{T}_x = r_{xx} (X - \overline{X}) + \overline{X}$$

The estimated true scores, thus, consider errors of measurement resulting from less than perfectly reliable measures. This factor becomes increasingly significant when comparing two tests with substantially different reliabilities, such as the WISC-R Full Scale IQ and the less reliable PIAT achievement subtests (Cone & Wilson, p. 365).

Salvia and Ysseldyke (1978) recommend dividing the difference of the estimated true scores by the standard error of measurement of difference scores in order to determine the significance of the score discrepancy. Cone and Wilson (1981), however, suggest that the more accurate error term for the comparison is the standard error of measurement of the estimated true score difference, described by Stanley (1971), where:

$$SE_{\hat{T}x-\hat{T}y} = \sqrt{r_{xx}^2(1-r_{xx}) + r_{yy}^2 (1-r_{yy})}$$

Cone and Wilson (1981) consider this approach to the determination of a discrepancy to be "the next evolutionary step in the standard score comparison procedure" (p. 365).

Shortcomings of this method, however, include the failure to consider "the main source of regression which is due to the inherent relationship between IQ and achievement" (p. 365).

4. Regression Analysis

In an attempt to address many of the concerns related to the standard score procedures for determining a score discrepancy, Wilson and Cone (1984) and Reynolds (1985) have described a procedure which takes into consideration the regression of aptitude on achievement when assessing a "severe" discrepancy. Using the standard regression or prediction equation, an expected value of the achievement score (Y'), based upon the relationship between the aptitude measure and the achievement measure (r_{xy}), can be calculated:

$$Y' = r_{xy} \frac{S_y}{S_x} (IQ - \overline{X}) + \overline{Y}$$

The observed value for the achievement measure (Y) can thus be subtracted from the expected value (Y'). The discrepancy represented by this residual is considered severe if it equals or exceeds the value:

$$(Z_a)S_y \sqrt{1 - {r_{xy}}^2}$$

After presenting and critiquing several alternatives, Reynolds argues that this model is the only acceptable method for determining a severe discrepancy available at this time. He recommends adjusting Z_a, however, for test unreliability. Both test reliabilities and intercorrelations are therefore included in the calculation of a discrepancy. The following formula represents Reynolds' regression analysis procedure:

$$\hat{Y} - Y_i = (Z_a\ S_y\sqrt{1 - r_{xy}^2}) - 1.65\ SE_{\hat{y}-y_i}$$

$$\text{where } SE_{\hat{y}-y_i} = \sqrt{1-r_{xy}^2} \cdot \sqrt{1- r_{\hat{y}-y_i}}$$

$$\text{and} \quad r_{\hat{y}-y_i} = \frac{r_{yy} + r_{xx}\ r_{xy}^2 - 2r_{xy}^2}{1-r_{xy}^2}$$

When making a single comparison, a z_a of 1.65, corresponding to the traditional 0.05 confidence level for a one-tailed test, is recommended. When making multiple comparisons, a more conservative value of z_a is suggested. Wilson and Cone (1984) summarize the advantages of the regression approach: "Unlike other approaches, it considers regression, measurement errors, and incidence. The expectancy formulas are biased in terms of intelligence and age, and they do not consider measurement errors and regression effects. The variations of the IQ-standard score difference approach consider measurement errors but not regression effects, and they cannot provide a priori estimates of incidence" (p. 107).

6. Grade Equivalents vs. Standard Scores

Before completely dismissing expectancy formulas which utilize grade equivalents in favor of the more psychometrically sound regression procedure, it is useful to examine some of the arguments which support the retention of more descriptive, developmental scores.

Phillips and Clarizio (1986) point out that status standard scores which describe a student's relative standing within a comparison group (either age or grade), do not allow for comparisons across groups. This information is often desirable when determining the appropriate instructional level for children being evaluated for learning problems. It is also suggested that status standard scores fail to consider the increased within-grade variability which occurs at higher grade levels. In other words, the same standard score may, in fact, represent different degrees of learning impairment for children at different grade levels. Descriptive information regarding a student's actual level of performance is lost through attempts to provide normative data regarding the child's relative academic standing.

Hoover (1984) also makes the distinction between status standard scores which compare the student's performance within a single reference group and developmental scores, such as grade and age equivalents, which compare the student's performance with "that of a series of reference groups which differ systematically and developmentally in average achievement" (p. 8). Hoover argues that both types of scores are needed in the interpretation of a child's performance. He cautions against abandoning grade equivalents because of the useful and interpretable information which such scores provide about a student's development in basic skill areas.

Although no specific recommendations are made by either Phillips and Clarizio or Hoover regarding the quantification of a severe discrepancy, it would appear valuable to retain some form of developmental information when considering an educational diagnosis. The obtuse manipulations which accompany regression analysis may serve to obscure the educational meaning of the resultant discrepancy. For this reason, a comparison of methods which provide grade level information with methods which utilize standard scores (status) would seem to be a worthwhile endeavor. It might thereby be possible to evaluate, not only incidence of discrepancy and the agreement among methods, but also the relative gain achieved through use of status indicators versus developmental scores.

For the purposes of this study, however, three approaches to the calculation of a severe discrepancy between ability and achievement have been chosen for comparison. These methods represent the most widely used and endorsed procedures currently recommended by Michigan school districts (from a review of practices submitted by 10 districts). Although the various districts employ different "rules of thumb" in the determination of a severe discrepancy, the basic methods can be characterized as:

1. z-score discrepancy, 2. estimated true score differences, and 3. regression analysis. For the present study, each approach to the quantification of a discrepancy will be implemented using the algorithms provided by the authors reviewed in this section rather than using modified procedures. A more complete description of these algorithms appears in the Methodology section.

B. Description and Background of the K-ABC

The Kaufman Assessment Battery for Children (K-ABC) is a recently developed, individually administered test battery designed for use with children between the ages of two and a half and twelve and a half (Kaufman & Kaufman, 1983a, 1983b). It consists of sixteen subtests, ten of which represent a "mental processing" component, and six of which represent an "achievement" component. The Mental Processing subtests can be further divided into Sequential Processing tasks and Simultaneous Processing tasks. This distinction is based upon neuropsychological research (Luria, 1966), cognitive theory (Neisser, 1967), and factor studies (Das, Kirby, & Jarman, 1975) which support the existence of separate types of mental functioning: sequential (or successive) and simultaneous (or holistic). The distinction between the Mental Processing component and the Achievement component is further based upon the Cattell-Horn model of "fluid" versus "crystallized" intelligence (Cattell, 1968).

The purpose of this section of the literature review is to provide a brief description of the component subtests of the K-ABC and discuss the research which supports the development of the battery. This general background will provide a basis for later examining the studies which consider the relationship between the K-ABC and more traditional tests (section C), as well as critiques which review the extent to which the K-ABC accurately reflects its underlying theoretical models (section D).

1. Description of the K-ABC subtests

Klanderman, Perney, and Kroeschell (1985) have provided a summary of the K-ABC subtests, including a brief description of each task, the ages of administration, and the scale in which each task is included. This summary is included below. A more comprehensive description of each subtest can be found in the Interpretive Manual of the K-ABC.

Sequential Processing Scale

a. Hand Movements (ages 2 1/2 to 12 1/2)
The examinee performs a series of hand movements in the same sequence as the examiner performed them.

b. Number Recall (ages 2 1/2 to 12 1/2)
The examinee repeats a series of digits in the same sequence as the examiner said them.

c. Word Order (ages 4 to 12 1/2)
The examinee touches a series of silhouettes of

common objects in the same sequence that these objects were named orally by the examiner.

Simultaneous Processing Scale

a. Magic Window (ages 2 1/2 to 4 1/2)

The examinee identifies a picture which is rotated behind a narrow window and, hence, only partially visible to the child at any one time.

b. Face Recognition (ages 2 1/2 to 4 1/2)

The examinee selects from a group of people the one or two faces that were just exposed briefly.

c. Gestalt Closure (ages 2 1/2 to 12 1/2)

The examinee names an object or scene pictured in a partially completed "inkblot" drawing.

d. Triangles (ages 4 to 12 1/2)

The examinee assembles two to nine triangles, all identical, into an abstract pattern that matches a model.

e. Spatial Memory (ages 5 to 12 1/2)

The examinee recalls the placement of pictures on a page that was just exposed briefly.

f. Matrix Analogies (ages 5 to 12 1/2)

The examinee selects a concrete picture or an abstract design which best completes a visual analogy.

g. Photo Series (ages 6 to 12 1/2)

The examinee places photographs of an event in chronological order.

Achievement Scale

a. <u>Expressive Vocabulary</u> (ages 2 1/2 to 4 1/2)
The examinee names an object pictured in a photograph.

b. <u>Faces and Places</u> (ages 2 1/2 to 12 1/2)
The examinee names a well-known person, character, or place pictured in a photograph.

c. <u>Arithmetic</u> (ages 3 to 12 1/2)
The examinee demonstrates knowledge of numbers and mathematical concepts, counting and computational skills, and other school-related arithmetic abilities.

d. <u>Riddles</u> (ages 3 to 12 1/2)
The examinee infers the name of a concrete or abstract concept when given a list of its characteristics.

e. <u>Reading/Decoding</u> (ages 5 to 12 1/2)
The examinee identifies letters and words.

f. <u>Reading/Understanding</u> (ages 7 to 12 1/2)
The examinee demonstrates reading comprehension by following commands that are given in sentences. (Klanderman, et al. p. 524)

2. Simultaneous and Successive Processing: Das-Luria Model

Luria (1966) was the first to suggest the distinction between successive and simultaneous modes of integration. Based upon case histories of cortical lesions and subsequent processing dysfunctions, Luria described three functional units of the brain. The first unit or "block" includes the upper and lower brain stem, the reticular formation, and the hippocampus and is associated with arousal or activation of the cortex. In terms of intellectual functioning, the first block is related to motivation and the maintenance of an optimal level of arousal for performing an activity or task. The second block of the brain is located in the posterior region of the neocortex and includes the occipital, temporal, and parietal lobes. This unit is hierarchically structured with a primary projection zone involved in the reception and analysis of information, a secondary projection zone involved in the organization and coding of information, and a tertiary projection zone involved with the integration of information which has already been coded. Moving from the primary zone to the tertiary zone, an increasing degree of lateralization of functions occurs (i.e., the left hemisphere and the right hemisphere of the secondary and tertiary zones become increasingly specialized).

The third block of the brain, as described by Luria (1973), is located in the anterior region of the cortex and is involved with the planning and programming of behavior. "The frontal lobes not only perform the function of synthesis of external stimuli, preparation for action, and formation of programs, but also the function of allowing for the effect of the action carried out and verification that it has taken the proper course" (p. 93).

In examining the consequences of lesions in the occipital-parietal areas (second functional unit), Luria observed that simultaneous organization of stimuli was disturbed. On the other hand, damage to the fronto-temporal region of the cortex interfered with the successive processing of information. This led to the hypothesis that the cortex was involved in two distinctly different modes of integrative activity: simultaneous integration "which has the property of surveyability and refers to any system of relationships" (Das & Molloy, 1975, p. 213) and successive integration which "refers to processing of information in a serial order...a system of cues consecutively activates the components" (Das, Kirby, & Jarman, 1975, p. 89).

Other neuropsychological research has demonstrated differential functioning of the right and left hemispheres of the brain. Sperry (1970) studied the effects of split-brain surgery performed to reduce the symptoms of severe epilepsy. The surgical procedure involves the severing of the corpus callosum, the primary connection between the right and left hemispheres. Using sensitive psychological assessment procedures, Sperry discovered that split-brain surgical patients were unable to integrate information received through the right and left hemispheres: "A patient may see an object in his left visual field, and be able to pick it out with his left hand from a hidden array of objects (both the left visual field and left hand connect directly to the right hemisphere), but be unable to name the object or identify it with his right hand" (Kaufman, 1979, p. 97).

Further research on cerebral specialization has demonstrated that the left hemisphere is more specialized for verbal, analytic, temporal, and digital operations while the right hemisphere is more adept at processing non-verbal, holistic, spatial, analogic, creative and aesthetic information (Bogen, 1969; Bogen, DeZure, Tenhouten & Marsh, 1972; Gazzaniga, 1970; Nebes, 1974, as reported in Kaufman, 1979). Lezak (1976) suggests that differences between the hemispheres exist not only in what is processed but how information is processed. The left hemisphere is described as organizing information on the basis of conceptual similarity while the right hemisphere organizes information on the basis of structural similarity. An apple and a peach are alike to the left brain because they are both fruits, while they are alike to the right brain because they are both round (Kaufman, 1979, p. 98).

The distinctions which have been drawn between the left brain and the right brain are remarkably similar to the dichotomy described by Luria between the frontal-temporal and the occipital-parietal areas of the brain. The split-brain research would suggest that the right brain is responsible for the holistic, simultaneous processing of information while the left brain appears to be more specialized for processing successive, temporally related information. While it is beyond the scope of this review to interpret this discrepancy, it is sufficient to say that research in the area of cortical functioning consistently demonstrates the dual nature of information processing. The existence of a simultaneous-successive dichotomy appears to be well-founded in theory, despite the uncertainty of its neurophysiological etiology. Based upon Luria's neuropsychological research, Das, Kirby, and Jarman (1975) developed an information processing model to explain stimulus integration. It was hypothesized that information integration involves four basic processing units: the input, the sensory register, the central processing unit, and the unit for output. Input is presented to the sensory register in either a parallel (simultaneous) or serial (successive) manner. Once registered, the information is transmitted to the central processing unit. It is speculated that the sensory register accepts complex information in a parallel fashion but "transmits" this information to the central processing unit in a serial manner.

Das, Kirby and Jarman (1979) divide the central processing unit into three components: "that which processes separate information into simultaneous groups, that which processes discrete information into temporally organized successive series, and the decision making and planning component which uses the information so integrated by the other two components" (p. 50). Referring to Luria's functional units of the brain, it can be seen that each of these processing components has an hypothesized relationship to a specific functional area of the cortex. The occipital-parietal area is associated with simultaneous processing, the fronto-temporal area is associated with successive processing, and the anterior lobe is associated with planning and "thinking." Das and his colleagues are quick to point out, however, that an individual has access to both simultaneous and successive modes of processing, and the "decision" to utilize one mode or the other is determined by the individual's habitual mode of processing as well as the nature of the task. Additionally, when responding to a task, the individual may utilize simultaneous or successive processing which is independent of the way in which the information was coded. The example given involves a memory task which could alternatively require serial recall or the recall of categories.

A number of research studies have subsequently confirmed the existence of the simultaneous and sequential processing dichotomy. Das, Kirby, and Jarman (1975) selected a variety of cognitive tasks which they administered to normal and retarded children. A two dimensional, successive-simultaneous, solution resulted from factor analysis. The dimensions were found to be equally descriptive for the normal and retarded children. Molloy (Das, et al., 1975) factor analyzed the scores of six and ten year old children on a similar battery of tests (with the addition of two timed tasks: Word Reading and Color Reading). The results of the analysis yielded a three factor solution: simultaneous integration, successive integration, and speed, which was consistent across age groups. Further investigation by Molloy demonstrated that the three factor solution was also consistent in describing the test performance of fourth grade students from different socioeconomic levels. Das (1973) examined the performance of Canadian children and high-caste Indian children on the battery of tests and found the three factor solution to be consistent across cultures.

A follow-up study by Das and Molloy (1975) added tests of intelligence (Lorge-Thorndike Digit Span and Performance IQ) and measures of reading achievement to the original battery of cognitive tests. For a group of fourth grade students, the factors of simultaneous processing, successive processing, and speed emerged once more in the factor solution. Predictably, Digit Span loaded on the successive dimension while Performance IQ loaded on the simultaneous dimension. Reading achievement emerged as a separate factor with high loadings noted on the vocabulary and comprehension components of the reading achievement test.

A systematic attempt to identify differences between IQ groups (high, normal, low) in the utilization of simultaneous and successive synthesis was undertaken by

Jarman and Das (1977). Factor analysis of the intercorrelations from a number of cognitive tests for the low IQ group yielded three factors which were labeled simultaneous synthesis, successive synthesis and speed, and successive synthesis in performing auditory-visual matching. Analysis of the intercorrelations for the normal IQ group revealed two factors: simultaneous synthesis and successive synthesis, while analysis of the high IQ group intercorrelations yielded three factors: simultaneous synthesis, successive synthesis, and speed. It was concluded that simultaneous synthesis was highly stable across IQ groups. Successive synthesis was also found to be quite stable across IQ groups.

Although not addressed in this paper, an interesting parallel can be seen between the Das, et al. (1975) research and the Jarman and Das (1977) findings. In both cases, the low IQ groups preferred sequential processing for a specific task which is more efficiently solved with a simultaneous approach. Given the observation that the general IQ from school records (Das, et al., 1975) loads on the simultaneous factor, one might speculate that the simultaneous factor represents a higher level of abstraction than the sequential, successive factor. In fact, the successive dimension is described primarily by rote memory tasks (digit span, serial recall, free recall, etc.) across studies. Das (1984) also has observed that sequential processing skills tend to develop earlier than simultaneous skills and predominate at lower age levels.

In a study by Cummins (reported in Das, et al., 1975) an entirely different battery of tests was administered to high school students. The tasks included syllogisms, similarities, paired-associate concrete words, memory span, digit span, paper folding test, and utility test (divergent thinking). Factor analysis yielded a three factor solution: simultaneous processing, successive processing and divergent thinking.

Naglieri, et al. (1981), also concerned about generalizability of the studies by Das and his colleagues, sought to cross-validate the simultaneous and successive processing dimensions using novel tasks. The battery of tests developed for this study was subsequently modified and expanded to form the Mental Processing component of the K-ABC. The tasks chosen for analysis included Concept Formation (requiring the child to distinguish pictures which represented a given concept), Memory for Places (recalling the location of objects presented on a page), Triangles (construction of designs from plastic triangles), Overlapping Designs (construction of hexagons from irregular geometric shapes), Hand Patterning (replication of a series of hand movements), Sequential Hand Movements (similar to Hand Patterning with increasing number of movements), Memory for Words (recollection of a series of common words by pointing to pictures of the words), Raven's Progressive Matrices, and Block Design, Digits Forward, and

Digits Backward from the Wechsler Intelligence Scale for Children-Revised.

A two-factor varimax rotated solution revealed clear-cut simultaneous and sequential processing factors. Progressive Matrices, Block Design, Triangles, Overlapping Design, Concept Formation, and Memory for Places loaded strongly on the simultaneous processing dimension while Digits Forward, Digits Backward, Hand Patterning, Sequential Hand Movements, and Memory for Words all loaded on the sequential factor.

3. Fluid versus Crystallized Intelligence

Horn and Cattell (1966) propose a different, hierachical model of intelligence. The model includes five general factors with the two most important dimensions being fluid intelligence and crystallized intelligence. Both fluid and crystallized intelligence are viewed as components of general intelligence, but fluid intelligence is applied to tasks which are novel to the examinee while crystallized intelligence is believed to reflect education and experience. Other, less significant components include general visualization, general fluency, and general speed.

The Cattell-Horn model (1966) distinguishes between an individual's ability to solve problems which are largely independent of past experiences and formal educational training (fluid intelligence) and the individual's ability to apply aspects of formal training to more familiar problem solving situations (crystallized intelligence). With this distinction in mind, Kaufman and Kaufman have attempted to incorporate "fluid" tasks into the Mental Processing component of the K-ABC while including items which require "crystallized" abilities in the Achievement component.

Of interest for the purpose of this research is the examination of the Cattell-Horn "fluid-crystallized" theory of intelligence and the extent to which the K-ABC Mental Processing and Achievement components correspond to this theoretical dichotomy. Since the distinction between "ability" and "achievement" is implicit in the determination of a severe discrepancy between a student's observed and expected levels of achievement, the extent to which a test or set of tests is capable of separately measuring fluid and crystallized components of intelligence is of interest.

A closer examination of the accuracy of the K-ABC in representing the Das-Luria and Cattell-Horn models will appear in Section D. While the K-ABC grew from the work of researchers and theorists who demonstrated the existence of a simultaneous/sequential information processing dichotomy, it will be seen that the K-ABC Mental Processing subtests may fall short of adequately assessing these two components (Das, 1984). Likewise, critics (Sternberg, 1984; Bracken, 1985) seriously question the application of Cattell's fluid and crystallized components of intelligence to Kaufman's categorization of "mental processing" and "achievement" tasks. Before reviewing critiques of the K-ABC, however, several correlational and factor studies will be examined to determine the relationship of the K-ABC to well-established and widely used intelligence and achievement tests.

C. K-ABC Validity Studies

In attempting to anticipate differences which might be observed between the use of the K-ABC Mental Processing Composite and the Wechsler Full-Scale IQ in the determination of a severe discrepancy between ability and achievement it is useful to review the literature which addresses the comparability of the two instruments. Several validity studies have been conducted which have systematically compared the scores from the K-ABC, the WISC-R, and standardized achievement tests in populations of students referred and not referred for academic problems.

1. Normal Children

Kaufman and Kaufman (1983b) have described a number of correlational studies which were conducted as part of test development and standardization procedures for the K-ABC. Of particular interest is the research which examined the relationship between the K-ABC composite scores and the WISC-R scale scores.

When examining a combination of three samples of normal children (n = 182) the test authors report a correlation of 0.70 between the WISC-R Full Scale IQ and the K-ABC Mental Processing Composite. Somewhat weaker relationships were observed between the Full Scale IQ and the K-ABC Sequential Processing (r = 0.47) and Simultaneous Processing (r = 0.68) standard scores. The strongest score correspondence was reported between the Wechsler Full Scale IQ and the K-ABC Achievement Composite (r = 0.78). Kaufman and Kaufman suggest that this strong relationship is attributable to heavy reliance on verbal skills and factual knowledge in determining the Wechsler Full Scale IQ, skills which are also measured on the achievement subtests of the K-ABC. The authors interpret these correlations as evidence that the Wechsler scales and other traditional measures of intelligence are "to a large extent measures of children's school related accomplishments" (p. 111). The K-ABC Mental Processing tasks are viewed as estimates of more "fluid" abilities while the Achievement tasks (and many of the Wechsler Verbal Scale tasks) represent "crystallized" abilities according to the Kaufmans' interpretation of the Cattell-Horn processing dichotomy.

Additional studies which examined the predictive validity of the K-ABC Achievement Composite through correlations with a number of standardized achievement tests are also reported in the Interpretive Manual of the

K-ABC. Murray and Bracken (Study #28*) found a correlation of 0.52 between the K-ABC Mental Processing Composite and the PIAT Total standard score for a sample of 29 normal children. A predictably higher correlation of 0.72 is reported between the K-ABC Achievement Composite and the PIAT Total standard score with correlations between the MPC and individual PIAT subtest scores ranging from 0.34 (Spelling) to 0.65 (Reading Comprehension).

Other research by Lewis and Swerdlick (Study #25) assessed the relationship between the Iowa Tests of Basic Skills and the K-ABC standard scores for a group of 18 normal children following a six month interval. Correlations were obtained between the ITBS Composite standard score and the K-ABC Mental Processing Composite ($r = 0.58$) and between the ITBS Composite and the K-ABC Achievement standard score ($r = 0.89$). Childers, Durham, and Bolen (Study #10) provide further evidence of the predictive validity of the K-ABC through research which examined the relationship between the K-ABC and the California Achievement Test for 45 normal children. Following a twelve month interval, correlations of 0.65 between the CAT Total Battery and the K-ABC Mental Processing Composite and 0.77 between the CAT Total Battery and the K-ABC Achievement Composite were obtained.

Studies which have examined the concurrent validity of

*Studies which were included in the standardization research (Kaufman & Kaufman, 1983b) are referenced by the study numbers which appear in the K-ABC Interpretive Manual.

the K-ABC are also useful in providing insights into the relationships between the K-ABC and various measures of achievement. Kamphaus (Study #20) evaluated the agreement between the K-ABC MPC standard scores and the Woodcock Reading Mastery Tests Passage Comprehension and the KeyMath Diagnostic Arithmetic Test Written Computation standard scores for a large group of normal children. Coefficients of 0.63 between the K-ABC MPC and the Passage Comprehension subtest and 0.82 between the K-ABC Achievement standard scores and the Passage Comprehension subtest scores were obtained. Somewhat lower correlations were reported between the K-ABC and Written Computation subtests ($r = 0.47$ MPC/Written Computation and $r = 0.59$ Achievement/ Written Computation).

Cummings and McLeskey (Study #13) examined the relationship between the K-ABC Achievement standard score and the PIAT subtest scores for 31 normal children. The results of the study indicate coefficients ranging from 0.63 (Spelling) to 0.82 (Reading Recognition) for PIAT subtests scores. A correlation of 0.89 was obtained between Total PIAT standard score and the K-ABC Achievement standard score. Additional studies reporting the relationship between the K-ABC and Wide Range Achievement Test for normal children show correlations in the range of 0.45 to 0.73.

In a recently reported study, Naglieri (1985b) investigated the relationship between normal children's (n

= 51) performance on the McCarthy Scales, the K-ABC, and the PIAT. A moderate correspondence was found between scores on the McCarthy General Cognitive Index and the K-ABC Mental Processing Composite (r = 0.55). The McCarthy General Cognitive Index was found to be strongly related to academic skills with a correlation of 0.74 between the PIAT Total standard score and the GCI. The correlation coefficient between the K-ABC Mental Processing Composite and the PIAT Total standard score was somewhat lower (r = 0.57) while the coefficient between the K-ABC Achievement standard score and the PIAT Total standard score was 0.86.

In spite of the substantial correlation between the PIAT and the K-ABC Achievement standard scores, the PIAT subtest scores have been reported to be consistently higher than their K-ABC Achievement counterparts. The most extreme example of this mean score difference can be seen when comparing the average K-ABC Reading/Understanding standard score (104.5) with the average PIAT Reading Comprehension standard score (117.7). Naglieri (1985b) suggests a need to determine which subtest scores most closely correspond to the student's actual academic functioning. In spite of these mean score differences, the findings described in this study are very similar to the relationships reported by Murray and Bracken (Study #28 and 1984) and other studies reported in the Interpretive Manual which have evaluated the relationships between the K-ABC Mental Processing and other ability tests, such as the

WISC-R.

The findings of the validity studies for normal children generally provide impressive evidence for construct, predictive, and concurrent validity of the K-ABC. A close correspondence between results obtained from the K-ABC Achievement subtests and PIAT achievement subtests could be anticipated on the basis of the coefficients, although the average standard scores on the PIAT achievement subtests may be somewhat higher than the K-ABC Achievement standard scores. Of additional interest are the reports that the correlations between the K-ABC Mental Processing Composite and Achievement scores tend to be smaller than the correlations between the WISC-R Full Scale IQ and similar measures of achievement. While this finding is consistent throughout many research studies, it should be noted that the Naglieri and Haddad (1984) study failed to demonstrate this pattern of score relationships.

2. Learning Disabled Students

Several studies which examined the performance of identified learning disabled students and learning disabilities referrals on the K-ABC are also reported in the K-ABC Interpretive Manual. As with the reported research with normal subjects, many of these studies have evaluated the relationship of the K-ABC with other well-known and widely used ability and achievement tests. The results of these studies provide clues regarding the impact of the use of the K-ABC on diagnostic decisions. Kaufman and Kaufman (1983b) report the consolidated results of four studies which investigated the relationship between the K-ABC and the WISC-R for learning disabled students (n = 138). The correlations between the Wechsler Scales and the K-ABC Global Scores are very similar to those reported for normal subjects with a coefficient of 0.74 obtained between the K-ABC MPC and the WISC-R Full Scale IQ and a correlation of 0.71 reported between the K-ABC Achievement standard score and the WISC-R Full Scale IQ. The Kaufmans interpret these results as additional evidence of the construct validity of the K-ABC with learning disabled students.

A similar pattern of test intercorrelations was obtained for a group of 60 children who were referred for learning disabilities (Gunnison, Masunaga, Town, & Moffitt, Study #17), although the coefficients were somewhat smaller than the correlations reported for the normal and learning disabled samples. The slightly lower correlations presumably resulted from a restriction of range on the K-ABC rather than from any inherent differences in the variable relationships (Kaufman & Kaufman, 1983b).

Gunnison, et al. (Study #17), Naglieri and Haddad (Study #32), and Klanderman, Kroeschell, and Licht (Study #23) examined the relationships between the K-ABC Achievement subtests and scores obtained by learning disabled students on other individually administered achievement measures. The results of these studies indicate correlations ranging from 0.59 to 0.66 between the PIAT and K-ABC Arithmetic subtests and from 0.40 to 0.71 between the WRAT and K-ABC Arithmetic subtests. Correlations between the WRAT Reading subtest and the K-ABC Reading/Decoding and between the PIAT Reading Recognition and the K-ABC Reading/Decoding ranged from 0.60 to 0.87 (Kaufman & Kaufman, 1983b).

Kaufman and Kaufman further describe the general profiles which typified the learning disabled students who participated in the standardization studies (Gunnison, et al., Study #17; Hooper & Hynd, Study #19). For this group of children, Simultaneous standard scores were, on the average, 2 to 5 points higher than Sequential Processing standard scores. Gestalt Closure represented the strongest Mental Processing subtest for the learning disabled and dyslexic children. Within the Achievement area, the average Riddles subtest score was the highest for this group of students. Subtests which were the most difficult for the learning disabled students included Matrix Analogies, Photo Series and Spatial Memory. It is suggested that learning disabled students experience difficulty on tasks which require the integration of both sequential and simultaneous processing abilities while performing relatively well on tasks which are measures of pure simultaneous processing. Gunnison, et al. (1982) report that 70 percent of the significant discrepancies between Simultaneous and Sequential Processing scores reflected stronger Simultaneous than Sequential abilities for learning disabilities referrals. Studies by Klanderman, et al. (#23), Snyder, et al. (#39), and Naglieri and Haddad (#32) showed an equal proportion of discrepancies in both directions (Kaufman & Kaufman, 1983b, p. 139).

Of particular interest are studies which have investigated the relationships between K-ABC Mental Processing subtest patterns and performance on various measures of achievement. McRae (1981) tested 65 normal children with the PIAT Reading Recognition and Reading Comprehension subtests and, subsequently, identified four groups of students as "high recognition-high comprehension," "high recognition-low comprehension," "low recognition-high comprehension," or "low recognition-low comprehension" on the basis of the PIAT scores. Upon testing these groups with the K-ABC, McRae found a significant relationship between reading comprehension and simultaneous processing, with the group identified as high recognition-low comprehension scoring significantly lower than the other groups on the Simultaneous Processing Scale.

Kaufman and Kaufman (1983b) summarize the findings of the standardization research with learning disabled students, learning disabilities referrals, and dyslexic children, suggesting that these students perform better on Gestalt Closure, Triangles, and Riddles subtests while exhibiting most difficulty with Hand-Movements, Word Order, Faces and Places, Arithmetic, Reading/Decoding and Reading/Understanding subtests. Consistently higher (half a standard deviation) standard scores on Simultaneous than on Achievement Global scores are also reported for this group of handicapped children (pp. 142-143).

In a study of 32 exceptional children (19 learning disabled and 13 educable mentally retarded) Obrzut, Obrzut, and Shaw (1984) examined the relationship of the WISC-R and the K-ABC standard scores. The results of the study demonstrated a stronger relationship between the K-ABC MPC and the WISC-R Full Scale IQ ($r = 0.77$) than between the K-ABC MPC and the K-ABC Achievement scores ($r = 0.54$). For the learning disabled group, the average K-ABC Simultaneous score was eight and a half standard score points higher than the Sequential Processing score, findings which are consistent with results reported in the K-ABC Interpretive Manual.

In addition to the research conducted as part of the K-ABC standardization process, several studies have been undertaken to investigate the generalizability of the three factor (Sequential/Simultaneous/Achievement) structure of the K-ABC for children with learning difficulties.

Keith (1986) performed a principal components factor analysis on the K-ABC subtest scores of 585 children who were referred for psychological evaluation. Seventy-six percent of these students were referred for problems in school achievement. When only the Mental Processing subtests were included in the factor solution, the highest loadings on the first, unrotated factor were achieved with the Photo Series, Triangles, Spatial Memory, and Word Order subtests, suggesting that these subtests are the best measures of general intelligence or "g." Varimax rotation yielded two factors, simultaneous and sequential processing, findings which are consistent with previous studies with normal children. Keith concludes that "the K-ABC MP Scale's factor structure seems quite stable for normal and exceptional children" (p. 243).

When Achievement subtests were included with the MP subtests in the factor analysis, however, a somewhat different structure emerged. The Achievement subtests and the MP Photo Series subtest loaded highest on the first, unrotated factor, suggesting that the Achievement subtests may, in fact, be a better estimate of "g" than most of the Mental Processing subtests. The rotated factor solution yielded a three factor solution which Keith alternatively labels: I. Simultaneous, II. Achievement, III. Sequential or I. General Ability/Reasoning, II. Reading Achievement/Verbal, and III. Verbal Memory. A four factor solution yielded a separate factor for reading with high loadings achieved on Reading Decoding and Reading Understanding subtests.

Keith notes that the factor structure obtained for referred children closely approximates the results of previous studies with normal children (Keith, 1985; Keith &

Dunbar, 1984; Kaufman & Kaufman, 1983b), confirming that the battery is operating quite consistently across groups. The interpretation of the factor structure, however, is not as clear as Kaufman and Kaufman indicate in the K-ABC Interpretive Manual.

In another recent study, Kaufman and McLean (1986) performed a joint factor analysis of the K-ABC and the WISC-R using scores from 198 learning disabled and/or referred students. Citing a previous joint factor study using normal children (Kaufman & McLean, 1985) which demonstrated a merger of each of the K-ABC factors with a previously established WISC-R dimension (i.e., K-ABC Achievement with WISC-R Verbal Comprehension, K-ABC Simultaneous with WISC-R Perceptual Organization, and K-ABC Sequential with WISC-R Freedom from Distractibility), the authors question whether a similar structure would be obtained with children with learning difficulties. The results of the study indicate strong evidence of a characteristic ACID profile on the WISC-R (low scores on Arithmetic, Coding, Information, and Digit Span). In contrast, very little variability was noted among the K-ABC MP subtest scaled scores. Performance IQ was slightly higher than Verbal IQ (VIQ: 93.1, PIQ: 96.4) on the Wechsler while the mean Simultaneous Processing score (93.3) was slightly higher than the mean Sequential Processing score (91.4) on the K-ABC.

Joint factor analysis yielded defensible three and four factor solutions. The three factor solution corresponded to the previous solution obtained with normal children, showing a merger of the WISC-R and K-ABC factors. The four factor solution suggests that the joint structure might be characterized by Simultaneous/Perceptual Organization, Achievement/Verbal, Sequential/Distractibility, and Reading Ability. The four factor solution for L.D. and referred children agrees with the factor solution obtained by Keith (1986), indicating that certain of the subtests (in particular, the K-ABC Achievement subtests) function differently for L.D. children than for normal children. This is not surprising considering that many of the L.D. and referred students were probably experiencing reading difficulties. The authors also argue, however, that the strength of the first three factors in this analysis provides further support for the construct validity of the K-ABC with learning disabled students.

Using the standardization data from the WISC-R, Lawson and Inglis (1985) examined the unrotated factors from a principal components analysis and found two significant factors, the first representing general intelligence or "g" and the second representing a verbal-nonverbal continuum. For learning disabled students, they discovered that "the degree of deficit shown by L.D. children on any particular subtest...is almost exactly proportional to the amount of verbal content of that subtest as expressed by its Factor

II coefficient" (p.80).

In an attempt to replicate this finding with the K-ABC, Inglis and Lawson (1986) submitted intercorrelations from the K-ABC standardization sample to principal components analysis and again obtained two significant unrotated factors: I. a general factor and II. a verbal-nonverbal factor. Number Recall (-.50), Word Order (-.42), Faces and Places (-.08), Riddles (-.08), Reading Decoding (-.26) and Reading Understanding (-.24) all had negative loadings (representing "verbal" subtests) while Hand Movements (.03), Gestalt Closure (.47), Triangles (.40), Matrix Analogies (.21), Spatial Memory (.38), Photo Series (.30), and Arithmetic (.01) had positive loadings (representing "nonverbal" subtests).

To determine the utility of Factor II in explaining the performance of learning disabled students, Inglis and Lawson correlated the Factor II score coefficients with the mean subtest score of the L.D., L.D. referrals, and dyslexic samples from the K-ABC Interpretive Manual. Positive, statistically significant coefficients were obtained in all cases (.55, .64, and .67, respectively), suggesting that, as with the Wechsler subtest scores, the amount of deficit shown by children with learning difficulties tends to be proportional to verbal content. While not preempting the rotated three or four factor solutions offered by other researchers (Keith, 1986; Kaufman & McLean, 1986), Inglis and Lawson suggest that the verbal-nonverbal continuum provides a useful and generalizable structure for interpreting learning difficulties.

Contrasting the Inglis and Lawson study, Hooper and Hynd (1986) performed an analysis of K-ABC subtest score patterns for normal and dyslexic students. Citing previous research (Bannatyne, 1968, 1971; Das, 1984; Das, et al., 1979; Rugel, 1974) which has demonstrated that sequential processing difficulties constitute the primary processing problems for disabled readers, Hooper and Hynd hypothesized that dyslexic children would score significantly lower than normal children on the K-ABC Sequential Processing subtests. Analysis of scores from 55 dyslexic and 30 normal readers (sample from Hooper & Hynd, Study #32) showed that Hand Movements, Number Recall, Word Order, Matrix Analogies, each Achievement subtest, the Sequential Factor, the Achievement Factor, and the Mental Processing Composite all significantly differentiated between the groups in favor of the normal readers. The Matrix Analogies subtest was the only Simultaneous Processing subtest which significantly distinguished between groups. Descriptive discriminant analysis revealed Reading Decoding, Faces and Places, Arithmetic, Number Recall, Matrix Analogies, Reading Understanding, Hand Movements, and Riddles to be significant predictors of group membership, while predictive discriminant analysis revealed that the K-ABC correctly classified 86.7 % of normal and

92.7 % of dyslexic readers.

The authors conclude that the results "implicate sequential processing deficits in the development of reading disability...The study lends additional support to this notion with the significant accuracy with which the K-ABC, particularly the Sequential and Achievement subtests could predict group membership" (p. 201).

Other studies suggest that profile analysis with learning disabled children may be speculative, at best. In a study with 33 learning disabled students, Naglieri and Haddad (1985) demonstrate higher average standard score for the Sequential Processing Scale than for the Simultaneous Processing Scale. Additionally, the PIAT Total standard score is slightly lower than the K-ABC Achievement Composite. Although no individual achievement subtest scores were reported for the K-ABC, this evidence might indicate that the average K-ABC Achievement subtest score is more comparable to the Average PIAT subtest score than suggested by the Naglieri (1985b) study with normal children. Also of note is a "somewhat lower" (5 standard score points) K-ABC Mental Processing Composite than WISC-R Full Scale IQ. This finding is consistent with results reported for learning disabled students in the K-ABC Interpretive Manual and may be attributable to the change in performance of standardization samples over time (Naglieri & Haddad, 1984, p. 54).

Correlations between the K-ABC Mental Processing Composite and the PIAT and WRAT subtests were generally similar to those obtained between the WISC-R Full Scale IQ and the respective achievement subtests. Interestingly, the relationship between the Simultaneous Processing Scale and the PIAT Total Test score was significantly higher (r = 0.71) than the relationship between the Sequential Processing Scale and the PIAT Total (r = 0.30). Naglieri and Haddad suggest that the disparate findings of this research may be attributable to the small sample size, the variable procedures used in identifying the learning disabled children, or the heterogeneous nature of the sample.

Another study which questions the validity of distinctive learning disabilities profiles on the K-ABC was conducted by Naglieri (1985a). Using matched groups of learning disabled (n = 34), borderline mentally retarded (n = 33), and normal children (n = 34), Naglieri failed to find significant differences between the WISC-R Full Scale IQ and the K-ABC Mental Processing Composite across groups. For the learning disabled children, the mean ability estimates from the two measures were 96.8 (Full Scale IQ) and 96.1 (Mental Processing Composite). Even more importantly, the results of the study failed to find significant differences between the mean K-ABC Simultaneous-Sequential discrepancy or between the mean K-ABC Mental Processing-Achievement discrepancy across the three groups of children. Naglieri concludes that "a Simultaneous-Sequential Scale disparity may not typify learning disabled children under the current system of state and federal guidelines" (p. 138).

Through a careful analysis of subtest score profiles for the learning disabled and borderline children, Naglieri discovered that, on both the WISC-R and K-ABC, the exceptional children performed best on tasks which had a simultaneous component (WISC-R: Picture Arrangement and Object Assembly; K-ABC: Gestalt Closure and Photo Series) while demonstrating more difficulty with tasks which were largely sequential or academic in nature (WISC-R: Digit Span, Information, Arithmetic, and Vocabulary; K-ABC: Number Recall, Word Order, Hand Movements, and Reading subtests). Naglieri also found that the exceptional children performed poorly on the Matrix Analogies subtest of the K-ABC. These results are generally consistent with the profiles of learning disabled children reported by Kaufman and Kaufman in the K-ABC Interpretive Manual.

It is suggested that the complexity of the exceptional children's performance on the K-ABC, with both strengths and weaknesses displayed on the sequential and simultaneous processing tasks, contributes to the inadequacy of a simple simultaneous-sequential dichotomy in characterizing the performance of these children.

In examining the intercorrelations between the WISC-R and K-ABC Global Scales standard scores, Naglieri reports a higher correlation between the WISC-R Full Scale IQ and the K-ABC Achievement standard score (r = 0.89) than between the K-ABC Mental Processing Composite and the K-ABC Achievement scores (r = 0.74). He interprets this finding as supportive evidence for the conclusion that the WISC-R Verbal Scale has a substantial achievement component (an argument initially made by Kaufman and Kaufman (1983b), but countered by Keith & Dunbar (1984) and Keith (1985, 1986). It is further argued that it is inappropriate to use two measures of achievement (WISC-R and an individually administered achievement test) when determining an ability/achievement discrepancy for the classification of learning disabled students. The reader is uncertain, at the conclusion of the research report, whether Naglieri is suggesting that a K-ABC Mental Processing/Achievement discrepancy is a more appropriate procedure for classification purposes. If there are no significant differences in the average discrepancies across normal and exceptional groups, this procedure must be seriously questioned, as well.

An additional source of confusion centered on Naglieri's failure to distinguish the direction of the sequential/simultaneous difference when performing his analysis of variance on mean differences across groups. When reviewing his table of mean differences, it is apparent that the average Sequential Scale score is seven and a half points lower than the average Simultaneous

Scale score for children with learning disabilities. This difference is very consistent with Kaufman and Kaufman's report of a Simultaneous Processing score which is a half standard deviation above the Sequential Processing score for this group of exceptional children. For normal children, Naglieri reports a Sequential Processing score which is slightly higher than the average Simultaneous Processing score. The failure to uncover significant Simultaneous/Sequential differences across groups appears to have resulted from the consolidation of all differences, regardless of the direction of difference. The same uncertainty is apparent, though less pronounced, in the results from the Mental Processing/Achievement discrepancy analysis.

Another recent study by Klanderman, Perney and Kroeschell (1985) compares the performance of 44 learning disabled children on the K-ABC, the WISC-R and the Peabody Picture Vocabulary Test (PPVT-R). They found that the Verbal Scale of the WISC-R correlated strongly with the K-ABC Achievement Scale ($r = 0.79$). Additionally, the Simultaneous Scale of the K-ABC showed a strong relationship with the WISC-R Performance scale ($r = 0.82$). When the correlation matrices were factor analyzed, it was found that the subtest factor loadings were somewhat different for this group of children than for normal children, with Hand Movements loading on the "Simultaneous" factor and only Number Recall and Word Order loading on the

"Sequential" factor. The authors interpreted this finding as suggestive that "the learning disabled group processes material somewhat differently than normative populations" with a tendency to "use a combined processing method and slightly more simultaneous processing for the Hand Movements task" (p. 526).

While these results might be indicative of a possible processing distinction between normal and learning disabled children, the study also provides supportive evidence for an alternative factor structure of the K-ABC; one based upon a verbal-nonverbal distinction. Number Recall and Word Order are the only two K-ABC Mental Processing subtests which are primarily dependent on verbal-auditory processing. Das (1984) also notes that the Hand Movements subtest loads on the Simultaneous factor for children above 10 years of age. The fact that Klanderman, et al. do not provide a specific breakdown of age levels of children included in their study, as well as the fact that children up to age 13 years, 2 months were included in their sample (upper age of the K-ABC is 12 1/2 years) creates further concern regarding the validity of the results of this study.

The examination of K-ABC research conducted with learning disabled, dyslexic, and L.D. referred students offers a number of insights regarding the performance of these students on the K-ABC.

It is intuitive that the Achievement subtests would differentiate children with learning problems from normal children (particularly Reading subtests for dyslexic children). More interesting are the patterns of MP subtest scores which were found for the exceptional children. Summarizing the findings from the various studies, Kaufman and Kaufman (1983b) suggest that the Hand Movements and Word Order subtests are the most difficult for learning disabled students while the Gestalt Closure subtest is among the easiest. They conclude that learning disabled students experience difficulty on tasks which require the integration of both simultaneous and sequential processing while performing well on tasks which are measures of pure simultaneous processing. Inglis and Lawson (1986) indicate that Number Recall and Word Order are the most difficult Mental Processing subtests for children with learning problems, reflecting a difficulty of these students in performing tasks with a verbal component. Hooper and Hynd (1986) found Hand Movements, Number Recall, Word Order, and Matrix Analogies to be the most difficult and discriminating K-ABC MP subtests for their sample of dyslexic students, suggesting that there are sequential processing deficits in the development of reading disability. Photo Series was found to be the easiest subtest for the dyslexic children. Naglieri (1985a) reports that his sample of L.D. children achieved the lowest scores on Hand Movements, Number Recall, and Matrix Analogies while performing the best on Gestalt Closure and

Photo Series subtests.

While there are some apparent inconsistencies among these studies, it appears that, in general, the sequential subtests: Hand Movements, Word Order, and Number Recall tend to be more difficult for the children with learning problems. Matrix Analogies is the most difficult of the Simultaneous subtests, while Gestalt Closure (and less consistently, Photo Series) is generally an easier task for the exceptional children.

Several difficulties are encountered when trying to generalize the research findings. To begin, different types of children were included in the research samples. Dyslexic, learning disabled, and learning disabilities referrals were variably represented in the studies. The results may be better generalized to "children with learning difficulties" than specifically to learning disabled students. It will be recalled that McRae (1981) found simultaneous processing deficits to be more closely related to reading comprehension difficulties, results which contrast with those of the research with "dyslexic" children (Hooper & Hynd, 1986). Future research may need to more clearly differentiate the type of learning disability which is being studied. Secondly, several of the studies (Inglis & Lawson; Kaufman & McLean; Hooper & Hynd) make use of data from the K-ABC standardization samples which may contribute to overlapping results. While Kaufman and Kaufman discuss specific studies in the

Interpretation Manual, the samples are often consolidated, making it difficult to pinpoint confounding effects.

In spite of these shortcomings, it appears that the learning disabled students and/or "children with learning difficulties" may have more difficulty with sequential/verbal type tasks than tasks of a predominantly simultaneous nature. Matrix Analogies, which is considered a measure of integrated functioning, may be contributing to the failure of a Sequential/Simultaneous discrepancy to distinguish learning disabled students.

Of note is the close correspondence of the sequential subtests of the K-ABC to the distractibility factor of the WISC-R (Kaufman & McLean, 1986). The three subtests contributing to the distractibility factor (Arithmetic, Coding, Digit Span) have been previously implicated in the performance of learning disabled students. It is possible that the same underlying processes are operating for the Sequential subtests.

Finally, in spite of evidence that reflects different processing strategies for learning disabled students, there is considerable support for Kaufman and Kaufman's original factor structure with learning disabled children. Patterns of test intercorrelations as well as the factor analytic studies provide strong evidence for the generalizability of the two factor structure of the Mental Processing scale. The Achievement scale, quite understandably, may be functioning somewhat differently for the learning disabled students. The debate regarding the best description of the K-ABC factors is still unresolved.

D. Critique of the K-ABC

In spite of Kaufman and Kaufman's careful attempts to adhere to theory in the development of the K-ABC, there have been a number of critical reviews which have questioned their success in accomplishing this goal. Of particular interest has been the extent to which the Das-Luria successive-simultaneous processing distinction has been accurately represented by the Kaufman Sequential and Simultaneous subtests (Bracken, 1985; Das, 1984). There have been additional concerns regarding how well the Mental Processing and Achievement subtests measure "fluid" versus "crystallized" intelligence according to the Cattell-Horn model of intelligence (Bracken, 1985; Sternberg, 1984). Several authors (Jensen, 1984; Keith, 1985) have suggested that the abilities assessed by the K-ABC might, in fact, be characterized by alternative models.

When discussing the match between the K-ABC and the Das-Luria model of information processing, Bracken (1985) questions the ability of only three subtests (Word Order, Number Recall, and Hand Movements) to accurately assess the sequential processing dimension defined by the model. It is suggested that the restricted and specific nature of the K-ABC Sequential tasks "reduce the sequential processing demands of the K-ABC to simple sequential visual and auditory short term memory" (p. 23). This simplistic nonverbal approach to the assessment of sequential abilities contrasts with the Das, Kirby, and Jarman (1979) comprehensive description of sequential processing as "perceptual, conceptual, or mnestic in nature, with conceptual sequential processing representing complex intellectual behaviors (p. 89). Bracken further suggests that the K-ABC Sequential Processing component tasks have very little correspondence with "complex intellectual behaviors and classroom instruction" (p. 25).

Jensen (1984) echoes this concern by describing the K-ABC Sequential tasks as reflective of "Level I" abilities. According to his hierarchical theory of intelligence, Level I abilities are measured by tasks which require rote, short term memory. This contrasts with the more complex cognitive abilities reflected in Level II processing. Jensen further argues that Face Recognition and Spatial Memory subtests of the Simultaneous Processing scale are also reflective of the simpler Level I, short-term memory tasks and that other Simultaneous tasks are inadequate for assessing more complex intellectual abilities.

Das (1984), perhaps the most qualified critic of the correspondence of the K-ABC to the Das-Luria model of intelligence, argues that the Kaufmans have used a very restricted interpretation of Luria's observations regarding brain organization when developing the Mental Processing subtests of the K-ABC. It will be recalled that Luria proposed three functional divisions of the brain, the first associated with arousal, the second associated with coding of information, and the third involved with planning and decision making. Das suggests that the only type of cognitive processing measured by the K-ABC involves Luria's second block or information coding. There is no attempt to assess either arousal or planning, which seriously restricts the utility of the K-ABC as a comprehensive measure of cognitive processing. Not only are there inadequate provisions for assessing all of the components of the Luria model, but the perceptual modality used in evaluating simultaneous information coding is restricted exclusively to visual processing. Das recommends that both simultaneous and successive processes be evaluated through more than one modality, and preferably through auditory, visual, and kinesthetic modalities.

Like Jensen and Bracken, Das observes that all Sequential Processing tasks on the K-ABC are memory tasks. He also notes that, by age 10, the Hand Movements subtest, the only primarily visual sequential task, has a higher factor loading on the Simultaneous dimension, leaving only two verbal subtests to adequately measure sequential processing for older children. This pattern of factor loadings was also discussed in the earlier review (Section C) of the Klanderman, et al. (1985) study with learning disabled children. Das observes that, after age 10, a verbal-nonverbal dichotomy might better characterize the

Mental Processing subtests. It is unclear however, how this dichotomization would correspond with the Wechsler Verbal-Performance distinction.

Keith (1985) has offered an alternative interpretation of the K-ABC based upon independent confirmatory factor analysis. Instead of the Sequential/Simultaneous dichotomy, Keith suggests that the Mental Processing factors might better be characterized as "verbal memory" and "nonverbal reasoning." When both Mental Processing subtests and Achievement subtests are included in the factor solution, Keith finds dimensions which reflect verbal memory, verbal reasoning, nonverbal reasoning, and reading achievement (p.18). While Keith cautions that the factor structure obtained does not disprove the Sequential-Simultaneous processing structure of the K-ABC, he concludes that the results of his analysis "certainly do not support the K-ABC theory as much as the manuals would suggest" (p. 18). Keith's interpretation would correspond quite closely, however, with Jensen's observations that the K-ABC Sequential component is largely represented by Level I, short-term memory skills as well as Das' reinterpretation of the Mental Processing tasks as better represented by a verbal-nonverbal distinction after age 10.

Other critics include Sternberg (1984) who questions the purported ability of the K-ABC to assess a child's problem solving and information processing styles. He suggests that, because the tasks involved on the K-ABC are not pure measures of sequential or simultaneous processing, and because the assessment of information processing style is restricted to tasks rather than to individuals, it is unclear how the K-ABC results contribute to the understanding of a child's approach to problem solving. The observation that the Photo Series subtest was moved from the Sequential to the Simultaneous Processing Scale after it failed to load on the initial (and more intuitively appealing) sequential factor, as well as the inconsistent split loading of the Hand Movements subtest on the sequential and simultaneous factors, would offer support for Sternberg's arguments.

Sternberg also expresses concern over the attempt to limit the assessment of verbal abilities to the Achievement component of the K-ABC. He cites the major emphasis that most major psychometric and information processing theorists place upon verbal abilities in the measurement of intelligence. Evidence offered by Kaufman and Kaufman (1983b) and Naglieri (1985b) which demonstrates the consistently higher correlation between the K-ABC Achievement Scale and other measures of intelligence than between the MPC and traditional IQ is interpreted by Sternberg as lack of convergent-discriminant validity for the K-ABC. Kaufman, on the other hand, argues that the higher correlation between achievement and traditional intelligence measures offers proof of the contamination of intelligence with achievement. The MPC is viewed by its authors as the more "pure" measure of "fluid" intelligence. Others, such as Bracken (1985), counter with observations that "the MPC correlates as well or better than traditional IQ tests with measures of achievement," findings which "indicate that the K-ABC Mental Processing Scale, in spite of efforts to separate intelligence from achievement, is related to achievement as much or more so than the traditional intelligence tests" (p. 25). Bracken's observations are based upon the disparate findings of studies by Naglieri and Haddad (1984) and Murray and Bracken (1984) which demonstrate equal, or higher correlations between the MPC and achievement measures than between the WISC-R Full Scale IQ and similar achievement measures.

Thus, it appears that the attempt to separate fluid from crystallized intelligence according to the Cattell-Horn model of intelligence may have been less than completely successful. Bracken (1985) returns to Cattell's original characterization of fluid and crystallized intelligence, citing Cattell (1969): "Crystallized ability is not identical with scholastic achievement. Many scholastic skills depend largely on rote memory, whereas what factor analysis shows is crystallized ability in that section of school learning involving complex judgement skills that have been acquired by the application of fluid ability" (p. 114). In examining the Achievement subtests of the K-ABC, Bracken would classify only the Riddles subtest as an adequate measure of crystallized ability. He further concludes that "equating the K-ABC Achievement Scale with crystallized intelligence is not an accurate representation of the Cattell-Horn model" (p. 25).

Bracken also finds the MPC of the K-ABC to be a less than convincing representation of Cattell's fluid ability component of intelligence. Like Jensen, who finds an over-representation of lower level, short-term memory tasks in the Sequential Processing subtests, Bracken observes that all of the Sequential Processing subtests are restricted to assessing associative memory skills, and two of the seven Simultaneous subtests are primarily measures of short-term memory. Consequently, only five of seven Simultaneous subtests and only five of ten Mental Processing subtests are measures of non-memory related fluid intelligence (reasoning, judgment, analogies, etc.).

Perhaps the most notable short-coming of the K-ABC identified by its critics is its failure to adequately tap higher level cognitive abilities. Unlike traditional intelligence tests which rely on more higher level mental abilities (what Jensen would refer to as "g"), the K-ABC seems to be over-represented by short-term, associative memory tasks.

The lower correlations between the MPC and K-ABC Achievement Scale than between traditional IQ tests and the K-ABC Achievement Scale have been interpreted by critics as evidence that, in fact, "the K-ABC Achievement Composite measures ability more than Achievement and measures ability better than the MP Scale" (Keith, 1985, p. 10). Keith cites evidence from Swerdlik and Lewis (1983) which demonstrates that the correlation between the WISC-R Verbal Scale and the K-ABC Achievement Scale was larger than the correlation between the K-ABC Achievement Scale and the Composite Achievement score on the Iowa Tests of Basic Skills. If, in fact, the Achievement Scale was measuring achievement, a higher correlation might be anticipated between the two achievement tests. Clearly, there are alternative explanations for the observed relationship between traditional ability tests (particularly verbal intelligence) and the K-ABC Achievement Composite.

When comparing the subtest scores of the K-ABC Mental Processing Composite with the WISC-R subtest results, what might be anticipated in terms of profile differences? Kaufman (1979) published an article on cerebral specialization and intelligence testing in which he discussed the Wechsler Scales in terms of tasks which involved right hemisphere, holistic, simultaneous processing and tasks which required left hemisphere, sequential processing, as well as tasks which appeared to require integrated functioning of the two cerebral hemispheres. Since the K-ABC Mental Processing Composite is based upon this information processing distinction, those Wechsler tasks which appear to require more holistic, simultaneous processing should show a stronger relationship with the Simultaneous Processing Scale of the K-ABC while those which are more dependent on sequential, analytical abilities should correlate more strongly with the Sequential Scale. Regarding holistic processing, Kaufman notes that, of the WISC-R subtests, only Picture Completion and Object Assembly require distinctly simultaneous "right brain" processing. Other Performance subtests are seen as requiring a combination of simultaneous and sequential processing. The Verbal subtests, although primarily characterized as sequential, "left brain" tasks, in most cases demand an integration of visual-spatial and temporal-language skills. Noted examples include the mental arithmetic (Arithmetic) and digits backward (Digit Span) tasks of the Verbal Scale. It is also observed that the heavy reliance on verbal instructions on all but the Object Assembly and Picture Completion subtests of the Wechsler increases the necessity for "left brain" involvement for efficient and effective task solution.

Kaufman describes research with split-brain patients which shows clearly the need for the integration of simultaneous and sequential processes for the accurate solution of Picture Arrangement and Block Design subtest items. Indeed, most academic tasks such as reading require not only a sequential, left brain component, but also the perceptual functions of the right brain (letter and pattern recognition). Without the interaction of the hemispheres and the respective specialized functions, the child will likely experience some degree of learning difficulty. Several reseachers (Dencla, 1974; Witelson, 1977) have hypothesized that learning difficulties such as developmental dyslexia have resulted from faulty interhemispheric integration. Witelson (1977) has speculated that the individual with developmental dyslexia has "two right hemispheres and none left." Das (1984) observes that successive or sequential processes appear to develop earlier than simultaneous abilities. It might be expected, therefore, that until the age of six a child will have more difficulty solving simultaneous tasks than sequential tasks. By grade three, however, Das, Snart, and Mulcahy (1982, as reported in Kaufman & Kaufman, 1983b) have shown that the performance of learning disabled children is depressed on both simultaneous and sequential processing tasks.

It appears, then, that depressed scores on tasks which require either simultaneous or sequential or integrated processing (most WISC-R tasks) could show a relationship to learning disabilities. On the other hand, intact simultaneous and sequential abilities would not necessarily predict learning success if there is faulty integration of the two processes.

The WISC-R might be faulted for its inability to isolate simultaneous from successive processing difficulties. On the other hand the K-ABC may be viewed as deficit in tapping more complex skills which require integrated processing and which are more representative of abilities required for school learning. Yet, even as argued previously, the K-ABC Mental Processing subtests are not pure measures of a single processing style and, in fact, load to varying degrees on both the primary and secondary factors (especially Hand Movements). One is still left with the question regarding the extent to which the solution of a sequential task such as Hand Movements has been mediated by simultaneous or successive processing.

E. Summary and Implications for the Present Study

In reviewing the research related to the present study, four general areas of literature have been discussed: 1. Methods for determining a severe discrepancy between ability and achievement, 2. Description and background of the Kaufman Assessment Battery for Children (K-ABC), 3. K-ABC validity studies, and 4. Critical reviews of the K-ABC.

The first section of the review addressed problems involved in the determination of a severe discrepancy between ability and achievement when considering a student for eligibility under the Special Education Learning Disabilities classification. Several approaches to the calculation of a severe discrepancy were presented and critiqued on the basis of Cone and Wilson's (1981) classification of methods: 1. Deviation from grade level, 2. Expectancy formulas, 3. Standard score procedures, and

4. Regression analysis. Advantages and disadvantages of each approach were discussed. Because of the current widespread use and statistical advantages of the z-score discrepancy, estimated true score difference, and regression analysis procedures, these three methods have been selected for comparison in the present study. By contrasting these procedures, it will be possible to examine the effect of considering various sources of error in the calculation of a significant score discrepancy. Among the sources of error are errors of measurement resulting from the less than perfect reliabilities of selected tests and errors in predicting achievement from ability due to less than perfect correlations between tests.

The second section of the literature review was concerned with examining the theoretical background, development, and structure of the Kaufman Assessment Battery for Children. The Das-Luria model of simultaneous and successive mental processing and the Cattell-Horn model of fluid and crystallized intelligence were presented and related to the tasks which were selected for the K-ABC Simultaneous, Sequential, and Achievement Scales.

A summary of correlational studies which examined the relationship between the K-ABC, WISC-R, and PIAT was presented in section three of the review. General findings which would support the comparability of discrepancies calculated using the WISC-R Full Scale IQ and PIAT

Achievement subtests with the K-ABC Mental Processing Composite (MPC) and K-ABC Achievement subtests include: 1. The high, positive correlations between the Full Scale IQ and the Mental Processing composite and 2. The high, positive correlations between the PIAT subtests and the K-ABC Achievement subtests. The evidence which demonstrates a similar pattern and magnitude of test intercorrelations for normal and learning disabled students offers support for the generalizability of results from students with normal children to learning disabled populations.

Results which might predict different outcomes from WISC-R Full Scale IQ/PIAT comparisons and K-ABC MPC/K-ABC Achievement comparisons include the reported tendency for PIAT standard scores to be higher than the respective K-ABC Achievement scores, as well as the tendency for the WISC-R Full Scale IQ to be slightly higher than the K-ABC MPC. The finding that the correlation between the K-ABC MPC and achievement subtests tends to be lower than the correlation between the WISC-R Full Scale IQ and achievement subtests could also influence discrepancies which are obtained using methods which consider test intercorrelations (i.e., regression analysis procedures). Finally, although the reliabilities of the WISC-R Full Scale IQ and the K-ABC MPC are comparable, the reliabilities of the PIAT subtests are consistently lower than the respective K-ABC Achievement subtests. The lower PIAT subtest reliabilties could serve to decrease the relative incidence of significant discrepancies when using methods which consider this source of error (i.e., z-score discrepancy and estimated true score difference approaches).

The last section of the review involved a critical examination of the underlying structure of the K-ABC, as well as a closer look at the relationship of this structure to the Wechsler Intelligence Scales. The discussion raised questions regarding the accuracy of the K-ABC in representing its underlying theoretical models. It was generally concluded that, relative to most traditional tests of intelligence, the K-ABC is over-represented by less complex, short-term memory tasks. The attempt by Kaufman and Kaufman to minimize the verbal component of the K-ABC may have further reduced the diagnostic utility of the battery in school settings. In spite of efforts to represent a simultaneous/sequential processing dichotomy through task selection, it was shown that certain subtests, such as Hand Movements, have loadings on both simultaneous and sequential factors, and that the magnitude of these loadings varied across age groups. Independent confirmatory factor analysis demonstrated that an equally plausible model based upon verbal memory and nonverbal reasoning dimensions could be applied to the K-ABC Mental Processing subtests.

It is difficult to determine the full impact of test content on the resulting performance of students with learning difficulties. The conclusion that profile analysis with learning disabled students is speculative, at best, makes attempts at subtest score prediction even more questionable. It may be useful, however, to return to studies which demonstrate divergent patterns of subtest score performance for learning disabled students when interpreting differences which might emerge between performance on the K-ABC and WISC-R. Among the reported trends are Kaufman and Kaufman's (1983) findings that L.D. children tend to perform somewhat better on the K-ABC Simultaneous Processing tasks than on Sequential Processing tasks, as well as Naglieri's (1985b) findings that learning disabled children tend to earn higher scores on the WISC-R Picture Arrangement and Object Assembly subtests and on the K-ABC Gestalt Closure and Photo Series subtests.

In conclusion, the studies reviewed present a wide variety of factors which could potentially influence the discrepancies which are obtained when comparing the WISC-R Full Scale IQ and PIAT subtests and the K-ABC MPC and K-ABC Achievement subtests across selected methods for calculating a significant discrepancy. These factors will be considered when interpreting the results of the present study.

III. METHODOLOGY

A. Rationale

In attempting to answer the questions which were generated in the previous sections, two general areas of investigation were pursued. The first area of research addresses the issues surrounding the adequacy of selected procedures in quantifying a "severe discrepancy" between ability and achievement for purposes of identifying students under the special education learning disabilities classification. A systematic analysis of several techniques which are currently endorsed and utilized by diagnostic teams will allow a comparison of the impact of each method on the size and characteristics of the population identified as learning disabled.

While previous studies have focused on the comparison of the proportion of students who demonstrate a severe discrepancy between ability and achievement using divergent criteria, this study will additionally attempt to describe characteristics, such as IQ, sex, and chronological age, of students who are systematically included or excluded from a learning disabilities classification based upon the severe discrepancy criterion which has been implemented. Also, unlike previous studies which have examined differences between single types of discrepancy calculation procedures, for example, differences between expectancy formulas or differences between standard score methods, this study will compare two standard score discrepancy formulas and two approaches to the more recently advocated regression analysis procedure. The comparison of discrepancies obtained through each of these procedures should help to clarify some of the practical, as well as theoretical, benefits and disadvantages of using various approaches when calculating discrepancies.

Another major difference between this study and most of the prior research in the area is the nature of the sample of students which has been selected. Many studies have chosen to evaluate test information from students who have been previously classified as learning disabled. The number of students who would be classified or declassified by various discrepancy procedures is then determined. Often such studies involve a consolidation of achievement test information which disguises the criteria upon which initial decisions were based. Also, in retrospect, it is nearly impossible to identify the discrepancy formula which was implemented in the original diagnosis of such students. This study will examine a group of students who have been referred for evaluation because of learning difficulties. As a result, it will be possible to determine the extent to which a wider range of student characteristics interact with various discrepancy criteria. Also retained in the analysis will be scores from a number of achievement subtests. The inclusion of these scores will allow an analysis of the numbers and characteristics of students who demonstrate a severe discrepancy in different academic areas under the selected definitions.

The second part of the study has been designed to compare the recently developed Kaufman Assessment Battery for Children (K-ABC) with the more established and widely used Wechsler Intelligence Scale for Children-Revised (WISC-R) and Peabody Individual Achievement Test (PIAT) to determine the adequacy of the new battery in the determination of a severe discrepancy for learning disabilities classification. Since the K-ABC consists of a Mental Processing component which is analogous to the WISC-R Full Scale IQ and an Achievement component which can be considered analogous to the PIAT, it is of interest to evaluate differences in a discrepancy determination which might result through the use of the K-ABC. With the increased usage of the K-ABC by practitioners, the convenient comparison of the Mental Processing Composite with the various achievement subtests for determination of discrepancy seems a natural extension of assessment procedures. The importance of determining the validity and the impact of this diagnostic strategy is obvious.

It will be recalled that, in a study which examined differences in K-ABC score profiles, Naglieri (1985) found no significant differences between the mean Mental

Processing-Achievement discrepancy for learning disabled, mentally retarded and normal groups. Again, the consolidation of achievement subtest scores can serve to conceal differences in specific academic areas. Additionally, the evaluation of mean differences between test scores can hide the direction of the differences. With learning disabled students, it is anticipated that ability estimates will exceed achievement estimates. Normal and mentally retarded students, on the other hand, might exhibit differences in either direction. A purpose of this study will be to evaluate both the magnitude and the direction of the discrepancy between the Mental Processing Composite and the K-ABC Achievement subtests and contrast these differences with the discrepancies obtained in WISC-R Full Scale IQ-PIAT achievement subtest comparisons.

B. Goals

The goals of the proposed research can be summarized as follows:

Part I

1. Compare the proportion of students who demonstrate a significant discrepancy between ability and achievement using four procedures:

   a. Z-score discrepancy (Method 1)
   b. Estimated true score differences (Method 2)
   c. Regression analysis I: Wilson & Cone (Method 3)
   d. Regression analysis II: Reynolds (Method 4).

2. Determine the extent of agreement between each pair of methods in the identification of students who demonstrate a significant discrepancy between ability and achievement.

3. Determine systematic variation in characteristics: IQ, chronological age, sex, and grade placement for students demonstrating a severe discrepancy between ability and achievement using the four selected procedures.

Part II

Compare the following WISC-R Full Scale IQ/PIAT subtest discrepancies with corresponding K-ABC Mental Processing/Achievement subtest discrepancies:

a. FSIQ-PIAT Arithmetic with MPC-KABC Mathematics
b. FSIQ-PIAT Reading Recognition with MPC-KABC Reading/Decoding
c. FSIQ-PIAT Reading Comprehension with MPC-KABC Reading/Understanding
d. FSIQ-PIAT Total with MPC-KABC Achievement

1. Determine the proportion of students who demonstrate a discrepancy between ability and achievement.
2. Determine the extent to which there is agreement between pairs of tests in identifying students who demonstrate a significant discrepancy between ability and achievement.

C. Definitions and Formulas

1. Method 1: Z-score discrepancy

In order to compare a standard score for ability measures with the standard scores for achievement measures, the procedure recommended by Reynolds (1981) for the determination of a significant standard score difference will be applied. For each ability-achievement comparison, the following calculations will be made:

a. Calculate Z for the ability measure

$$Z_X = \frac{X - \overline{X}}{\sigma_X}$$

where X = WISC-R Full Scale IQ or K-ABC Mental Processing Composite

$\overline{X}$ = mean standard score for FSIQ or MPC

σ_X = standard deviation of scores for FSIQ or MPC

b. Calculate Z for the achievement measure

$$Z_Y = \frac{Y - \overline{Y}}{\sigma_Y}$$

where Y = PIAT or KABC achievement subtest score

$\overline{Y}$ = mean standard score for PIAT or K-ABC achievement subtest score

σ_Y = standard deviation of scores for PIAT or K-ABC achievement subtest

c. Calculate the difference in Z scores for the two tests

$$D = Z_x - Z_y$$

d. Divide the difference by the standard error of the difference scores

$$Z_{dif} = \frac{D}{\sqrt{(1-r_{xx}) + (1-r_{yy})}}$$

where r_{xx} = reliability of FSIQ or MPC (internal consistency)

r_{yy} = reliability of PIAT or K-ABC achievement subtest (internal consistency)

e. Compare Z_{dif} to $Z_a = Z_{.05} = 1.65$ from the standard normal table.

For the purposes of this study, the .05 level of significance has been chosen with a corresponding Z_a of 1.65. A one-tailed test was selected because the comparisons of interest involve discrepancies in which ability scores exceed achievement scores.

2. Method 2: Estimated true score differences

Following the procedure outlined by Stanley (1971) and Cone and Wilson (1981), true scores will be estimated for each observed ability and achievement test score and the significance of the difference in true scores will be calculated:

a. Calculate estimated true Z scores for ability and achievement measures

$$Z_{\hat{T}x} = r_{xx}Z_x$$

$$Z_{\hat{T}y} = r_{yy}Z_y$$

where r_{xx}, r_{yy}, Z_x, and Z_y are previously defined

b. Calculate difference between estimated true Z scores

$$D_{\hat{T}x-\hat{T}y} = Z_{\hat{T}x} - Z_{\hat{T}y}$$

c. Calculate standard error of the estimated true score difference

$$SE_{\hat{T}x-\hat{T}y} = \sqrt{r_{xx}^2(1-r_{xx}) + r_{yy}^2\ (1-r_{yy})}$$

where r_{xx} and r_{yy} are previously defined

d. Divide the true Z score difference by the standard error of estimated true score difference

$$Z_{dif} = \frac{D_{\hat{T}x-\hat{T}y}}{SE_{\hat{T}x-\hat{T}y}}$$

e. Compare Z_{dif} to $Z_a = Z_{.05} = 1.65$ from the standard normal table.

3. <u>Method 3: Regression Analysis</u>

Method 3 employs the regression analysis procedure which was outlined by Wilson and Cone (1984). For purposes of comparison with other standard score methods, it will be assumed that the mean standard score for the population being sampled is 100 and the standard deviation is 15 for

each test under consideration (WISC-R, K-ABC, PIAT). Intercorrelations between tests have been drawn from either test manuals or other published sources which have reported the relationships. Correlations between the K-ABC Mental Processing Composite and the K-ABC Achievement subtests were obtained from the intercorrelation matrices for seven age groups (6 to 12 1/2 years) which were presented in the K-ABC Interpretive Manual. The correlations were z-transformed (Fisher z transformation), averaged, and converted back into correlations which were then included in the discrepancy equations. The correlations between the WISC-R Full Scale IQ and the PIAT subtests were obtained from a research study by Wikoff (1979) which investigated the WISC-R as a predictor of achievement.

Wilson and Cone advise determining appropriate means, standard deviations, and intercorrelations for each population under consideration. Therefore, it should be cautioned that the statistics borrowed from the broader standardization group may vary to some extent for the population of students which are currently being studied. These variations could influence the number of students demonstrating a discrepancy between ability and achievement. Reynolds (1984-85) advises using the statistics from "a large, stratified, random sample of normally functioning children" (p. 472) which offers some support for the choice of statistics from the standardization group for these analyses (Methods 1 through 4).

The following steps are involved in the calculation of a severe discrepancy for Method 3:

a. Calculate the predicted Z score for achievement

$$Z_{\hat{y}} = r_{xy}(Z_x)$$

where r_{xy} is the correlation between the ability test and the achievement subtest and Z_x is previously defined

b. Subtract the observed Z_y from the predicted $Z_{\hat{y}}$

$$D_{Z\hat{y}Zy} = Z_{\hat{y}} - Z_y$$

where Z_y is previously defined

c. Divide the difference by the standard error of estimate

$$Z_{dif} = \frac{D_{Z\hat{y}Zy}}{\sqrt{1 - r_{xy}^2}}$$

d. Compare Z_{dif} to $Z_a = Z_{.05} = 1.65$ from the standard normal table.

4. Method 4: Regression Analysis II

Method 4 is based upon the regression analysis formula recommended by Reynolds (1984-85). Unlike the previous regression analysis approach in which Z_{dif} is compared to $Z_a = 1.65$, Reynolds recommends comparing Z_{dif} to a modified Z_a. The z_{mod} is an adjustment of Z_a for the unreliability of the subtests which are being compared. With perfectly reliable tests, Z_{mod} will equal Z_a, and score differences which are significant for Method 4 will also be significant for Method 3.

Reynolds suggests using the following equation for calculating a discrepancy:

$$Y - Y_i = (2\ SD\sqrt{1 - r_{xy}^2}) - 1.65\ SE_{\hat{y}-y_i}$$

$$\text{where } SE_{\hat{y}-y_i} = \sqrt{1-r_{xy}^2} \cdot \sqrt{1- r_{\hat{y}-y_i}}$$

$$\text{and } \quad r_{\hat{y}-y_i} = \frac{r_{yy} + r_{xx}\, r_{xy}^2 - 2r_{xy}^2}{1-r_{xy}^2}$$

S. Phillips (personal communication) has suggested that the correct form of this equation for standard score comparisons is:

$$Z_{dif} = Z_x\,(r_{xy}) - Z_y > (Z_a - 1.65\ SE_{z\hat{y}-zy_i})\,(\sqrt{1-r_{xy}^2})$$

where Z_a would equal 1.65 for one-tailed comparisons

and $SE_{z\hat{y}-zy_i} = \sqrt{1- r_{\hat{y}-y_i}}$

Using the latter equation, the following steps would be taken in the calculation of a discrepancy:

a. Calculate Z_{dif} as in Method 3, above

$$Z_{dif} = \frac{D_{Z\hat{y}Zy}}{\sqrt{1 - r_{xy}^2}}$$

b. Calculate Z_{mod} (note: Z_{mod} is Z_a corrected for the unreliability of the tests)

$$Z_{mod} = Z_a - 1.65\ SE_{z\hat{y}-zy_i}$$

where $Z_a = 1.65$

and $SE_{z\hat{y}-zy_i} = \sqrt{1- r_{\hat{y}-y_i}}$

c. Compare Z_{dif} to Z_{mod}.

5. Comparison of Discrepancy Methods

In order to compare the four standard score procedures for determining discrepancy (Methods 1, 2, 3, and 4), the formulas were solved for $Z_a = 1.65$ and plotted. Sample statistics were selected to demonstrate the relationships between the methods and the impact of test reliabilities and intercorrelations on the obtained discrepancies. Two graphs were developed representing the following situations:

Graph 1: $r_{xx} = .96$; $r_{yy} = .74$; $r_{xy} = .58$

Graph 2: $r_{xx} = .94$; $r_{yy} = .87$; $r_{xy} = .71$

Linear equations were developed for each method in the following manner:

a. Method 1: Z score difference

$$\frac{Z_x - Z_y}{\sqrt{1 - r_{xx} + 1 - r_{yy}}} = 1.65$$

substituting

$$\frac{Z_x - Z_y}{\sqrt{1 - .96 + 1 - .74}} = 1.65$$

reducing and rearranging

$$Z_y = -.90 + Z_x$$

b. Method 2: Estimated True Scores

$$\frac{r_{xx}Z_x - r_{yy}Z_y}{\sqrt{r_{xx}^2(1 - r_{xx}) + r_{yy}^2(1 - r_{yy})}} = 1.65$$

substituting

$$\frac{.96\ Z_x - .74\ Z_y}{\sqrt{.96^2\ (1 - .96) + .74^2\ (1 - .74)}} = 1.65$$

reducing and rearranging

$$Z_y = -.95 + 1.30\ Z_x$$

c. Method 3: Regression Analysis (Wilson & Cone)

$$\frac{r_{xy}(Z_x) - Z_y}{\sqrt{1 - r_{xy}^2}} = 1.65$$

substituting

$$\frac{.58\ (Z_x) - Z_y}{\sqrt{1 - .58^2}} = 1.65$$

reducing and rearranging

$$Z_y = -1.35 + .58\ Z_x$$

d. Method 4: Regression Analysis (Reynolds)

$$Z_x\ (r_{xy}) - Z_y > (Z_a - 1.65\ SE_{z\hat{y}-zy_i})\ (\sqrt{1-r_{xy}^2})$$

substituting

$$\frac{Z_x(.58) - Z_y}{\sqrt{1 - .58^2}} = (1.65 - 1.65\sqrt{1 - \frac{.74 + .96(.58)^2 - 2(.58)^2}{1 - .58^2}})$$

reducing and rearranging

$$Z_y = -.482 + .58Z_x$$

e. The four linear equations are graphed for the ability and achievement range of interest for each of the two conditions above (See Figures 1 and 2).

It can be seen that the number of students identified will vary depending upon the test reliabilities and intercorrelations as well as the actual distribution of ability and achievement scores for the students sampled. The graphs provide a tool for generating hypotheses regarding the numbers of students who will be identified by different methods based upon selected test characteristics and intercorrelations.

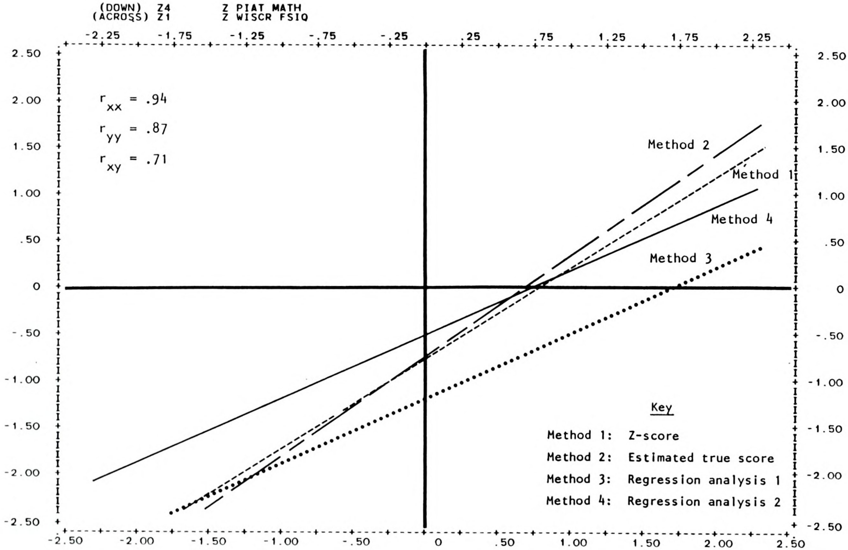


Figure 1: An example comparison of four methods for determining a severe discrepancy: WISC-R/PIAT Mathematics

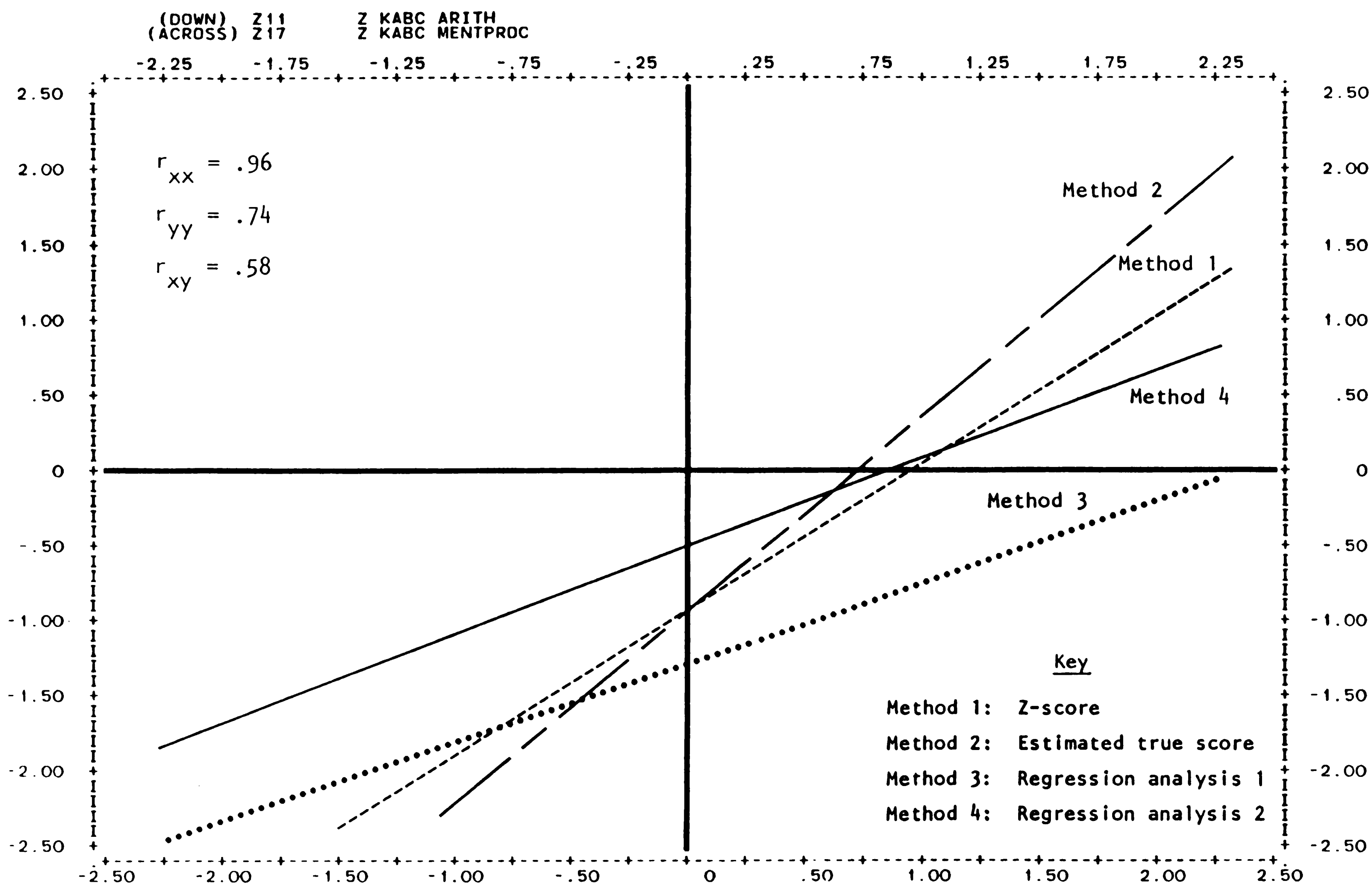


Figure 2: An example comparison of four methods for determining a severe discrepancy: K-ABC MPC/Arithmetic

D. Hypotheses

Based upon these observations and considerations, the following hypotheses, which correspond with the previously stated research objectives, have been generated for the first part of this study.

Part I

1. There will be no significant differences between the proportion of students who demonstrate a significant discrepancy between ability and achievement using the four proposed methods (Objective 1).

2. The percent of agreement in the identification of students who demonstrate a discrepancy will be significantly higher when comparing Method 1 with Method 2 than when comparing Method 1 or Method 2 with Method 3 or Method 4 (Objective 2).

3. The correlations between pairs of discrepancies obtained using Methods 1 and 2 will be higher than correlations obtained using Methods 1 and 3, Methods 2 and 3, Methods 1 and 4, or Methods 2 and 4 (Objective 2).

$$r_{12} > r_{13}, r_{23}, r_{14}, r_{24}$$

4. Methods 3 and 4 will include significantly fewer high ability students and significantly more low ability students than either Method 1 or 2 (Objective 3).

5. There will be no significant differences in the mean age, grade placement, or sex ratio of identified students across the four methods (Objective 3).

Part II

Because of the large number of factors (such as test reliability, subtest intercorrelations, floor and ceiling effects, subtest content, and mode of presentation) which can potentially influence obtained discrepancies, as well as the conflicting patterns of test intercorrelation which have been reported in the literature, the following hypotheses regarding the comparison of K-ABC, WISC-R, and PIAT subtests scores are suggested:

1. There will be no significant differences between the proportions of students who demonstrate a severe discrepancy between ability and achievement using K-ABC Mental Processing/Achievement comparisons and WISC-R/PIAT comparisons across the four methods for calculating a discrepancy (Objective 1).

2. There will be a significant correspondence (percent of agreement and intercorrelation) between the following pairs of test scores for each of the four discrepancy methods:

   a. FSIQ-PIAT Arithmetic with MPC-KABC Mathematics
   b. FSIQ-PIAT Reading Recognition with MPC-KABC Reading Decoding

c. FSIQ-PIAT Reading Comprehension with MPC-KABC Reading Understanding

d. FSIQ-PIAT General Information with MPC-KABC Faces and Places

e. FSIQ-PIAT Total with MPC-KABC Achievement (Objective 2).

E. Subjects

The sample of students included in this study consists of children in the State of Michigan who had been referred for psychoeducational evaluation during the Fall of 1983 and who were subsequently diagnosed using the WISC-R, the K-ABC, and PIAT. Table 1 shows the descriptive statistics for this group of students. It can be seen that urban students represent 12.8 percent of the sample while the remaining 87.2 percent of the sample is approximately equally divided between rural and suburban students. 1980 Census data indicate that 27.9 percent of the population is urban, while 43.8 percent is suburban and 28.3 percent is rural. Males account for more than twice the number of students as females, a trend which is consistent with other studies of learning disabled students and learning disabilities referrals.

White students represent almost 90 percent of the children in the sample, while Blacks and other minorities constituted 10.5 percent of the sample. This contrasts

Table 1

Student Characteristics (n = 86)

Educational Setting	Freq.	Pct.
1. Urban	11	12.8
2. Rural	38	44.2
3. Suburban	37	43.0

Sex	Freq.	Pct.
1. Male	61	70.9
2. Female	25	29.1

Race	Freq.	Pct.
1. White	77	89.5
2. Black/Other	9	10.5

Grade Placement	Freq.	Pct.
1. Kindergarten	3	3.5
2. First	14	16.3
3. Second	26	30.2
4. Third	14	16.3
5. Fourth	15	17.4
6. Fifth	6	7.0
7. Sixth	8	9.3

Mean: 2.86 S.D.: 1.6

Chronological Age Mean: 8.83 S.D.: 1.7

Range: 6.0 to 12.3 years

Parental Education	Freq.	Pct.
1. Less than 12 yrs.	23	26.7
2. 12 years	38	44.2
3. 13-15 years	11	12.8
4. 16 or more years	10	11.6
5. Not reported	4	4.7

with 1980 U.S. Census data which shows that Blacks and other minorities make up 26.9 percent of the total population (Kaufman & Kaufman, 1983b; 1980 Census of Population). The under-representation of minority students is quite likely related to the disproportionately small number of urban students in the sample.

Further examination of Table 1 shows that grade placement is concentrated on the second grade level with relatively few students at the kindergarten, fifth, and sixth grade levels. Chronological age statistics also show a concentration of students in the seven to ten year age range. Parental education is 12 years or less in more than 70 percent of the cases. Approximately 24 percent of the parents have some college or advanced training.

All of the students in the sample were evaluated by certified school psychologists or supervised school psychology interns. The average years of experience for the school psychologists/interns who contributed data to this study was 7.78 years (range 0 to 22 years). 32.6 percent were male and 67.4 percent were female.

To be included in the study, the following requirements were established:

1. Age at the time of the evaluation between 6 years, 0 months and 12 years, 6 months
2. WISC-R Full Scale IQs ranging from 70 to 125
3. Students referred for learning problems
4. Diagnosis conducted using WISC-R, PIAT, and K-ABC

5. Evaluation completed by certified school psychologist or school psychology intern who was supervised by a certified school psychologist
6. English as a primary language
7. Student free of physical/motor handicap

Students with complete data sets who met these criteria were selected from a larger sample of 171 referrals. Children whose ability estimates fell into the Educable Mentally Impaired Range (I.Q < 70) or who were referred primarily for behavioral or emotional difficulties (suspected Emotionally Impaired) were not included in the study. Gifted referrals were also eliminated.

F. Instruments

1. Wechsler Intelligence Scale for Children-Revised

The Wechsler Intelligence Scale for Children-Revised (WISC-R) is an individually administered intelligence test developed for children between the ages of 6 and 16. The WISC-R provides estimates of verbal abilities (Verbal IQ), visual-perceptual abilities (Performance IQ), and a measure of general ability (Full Scale IQ). A more complete description of the subtests, standardization group, and application of the WISC-R is included in Appendix A: Description of Tests.

2. Kaufman Assessment Battery for Children

The K-ABC has been described in detail in the literature review (see Review of Literature: B. Description and Background of the K-ABC; C. K-ABC Validity Studies; D. Critique of the K-ABC). The reader is referred to those sections for a more complete discussion of the K-ABC.

3. Peabody Individual Achievement Test

The PIAT is an individually administered achievement test designed for children between the ages of 5 years, 8 months and 18 years, 3 months. It includes five subtests: Mathematics, Reading Recognition, Reading Comprehension, Spelling, and General Information. The battery is based largely on a multiple-choice format where the student is presented with a test plate (page) and must select a correct response from among four possible answers. A more complete description of the subtests is included in Appendix A: Description of Tests.

G. Procedures

The data for this study were submitted by school psychologists who participated in a course on the Kaufman Assessment Battery for Children which was offered through Michigan State University and the Michigan Association of School Psychologists in Fall of 1983. Each school psychologist was requested to administer the three selected instruments to five referred students as well as collect behavioral ratings and personal data for each child. The data collection instruments appear in Appendix B.
Also included were reason for referral and order of test administration. Complete protocols were returned by 86.5

percent of the course participants.

H. Data Analysis

In order to test the hypotheses which have been generated, the following statistical techniques will be used.

1. Kappa Statistic: The Kappa statistic described by Cohen (1960) for measuring nominal agreement among raters will be used to measure the agreement between pairs of formulas in identifying the same students as exhibiting a severe discrepancy between ability and achievement. (See Appendix F for the formula for calculating kappa). This procedure will be used to test Hypothesis 2, Part 1 and Hypothesis 2, Part 2. The percent overlap of students identified by the various pairs of methods and tests will also be reported. The overlap statistic will provide descriptive evidence regarding the agreement of the four methods for determining a discrepancy (Part I) and the agreement between the K-ABC MPC/Achievement and WISC-R FSIQ/PIAT subtest discrepancies (Part II) in identifying students.

2. Spearman Rank Order Correlation Coefficient: Spearman's rho will also be used to measure the agreement between formulas in calculating discrepancy. The rank order based on the magnitude of the discrepancy calculated for each student using a particular method or pair of tests

will be correlated with the rank order of discrepancies obtained using other methods or tests. This procedure will also be used to test Hypothesis 2, Part I and Hypothesis 2, Part 2.

3. Analysis of Variance: Analysis of variance procedures will be used to compare mean ability estimates, mean chronological age, and mean grade placement differences across methods. Planned comparisons will be used to compare the hypothesized differences proposed in Hypothesis 3, Part 1 while post hoc comparisons will be used to compare means for hypotheses which propose no significant differences (Hypotheses 3, Part 2).

4. Friedman two-way analysis of variance by ranks/ chi-square: The Friedman two-way analysis of variance will be used to test hypotheses which involve the comparison of the proportions of students identified by different methods (Hypothesis 1, Part 1). Chi-square will be applied when comparing the proportions of students who demonstrate a severe discrepancy across tests. When analyzing proportions which are correlated (Hypothesis 1, Part 2), McNemar's test for correlated proportions will be used to estimate chi-square (see Appendix F for a description of procedures).

I. Limitations

Perhaps the most significant limitation of the proposed study is the non-random nature of the sample of students who are being investigated. Although students represent diverse educational settings, the generalization of many of the results of this study to areas outside the State of Michigan is questionable. The generalization is especially limited in terms of the actual numbers of referred students who exhibit a discrepancy using various methods. Certainly substantial differences exist between states, and even between school districts, in the number and type of students who are referred for evaluation. Less variable are the results obtained from investigating the characteristics and relative proportions of students identified using the selected discrepancy procedures and selected test comparisons.

A second limitation of this research involves the procedure which was used by each psychologist in selecting subjects to be evaluated using the three instruments and submitted once the evaluations were complete. Although students were evaluated within a circumscribed period of time, the cases received from each psychologist may have been selected on a non-random basis. Cases with missing data were also eliminated from consideration, which may have biased the final data set in some additional manner.

Third, it was not possible to consider discrepancies in the areas of oral expression, listening comprehension, and written expression, three academic areas which are also used to qualify students as learning disabled. Since no standard scores were available in the three areas, it was impossible to indicate the additional number of students who would have shown a severe discrepancy in these achievement areas. Additionally, when examining the criteria for qualifying students as learning disabled, it becomes apparent that a severe discrepancy is a necessary but not sufficient condition in classification decisions. In addition to demonstrating a discrepancy, it must be shown that the student's learning difficulties are not a result of visual, hearing, or motor handicaps, of mental retardation, or of environmental, cultural or economic disadvantage (P.L. 94-142,121a.5 b(9)). Because it was not possible to control for all of these factors, the direct translation of the observed discrepancies into "learning disabilities" is strongly cautioned against.

Another important distinction which should be made when interpreting the discrepancies which result from various methods is the difference between statistical significance and educational significance of obtained test scores. While a small raw score difference between two highly reliable instruments might result in a statistically significant discrepancy at the .05 level, the difference may not constitute an educationally meaningful difference.

Since the data obtained for this study were collected as part of a training course in the K-ABC, it is expected that the protocols which were received represented some of the first experiences of psychologists with this new instrument. While all psychologists are required to have familiarity with individually administered, standardized tests, it is difficult to asses the impact that lack of experience may have had on the resultant K-ABC test scores.

Finally, the test data which were collected for this study consisted principally of face sheets from the K-ABC, PIAT, and WISC-R. Therefore, it was impossible to perform a random check to determine the accuracy with which the various tests were scored. The examination of eight complete protocols indicated that proper determination of basal and ceiling had occurred. Calculations and table consultation were also accurate for this limited sample of protocols. Again, this does not insure against improper administration or scoring of tests for the larger sample of students.

IV. RESULTS AND DISCUSSION

Before testing the hypotheses which were set forth in the previous section, it is useful to examine the statistics which describe the test scores of the students included in this study. Table 2 illustrates the means and standard deviations of the scale scores and standard scores of the K-ABC, WISC-R, and PIAT. Of interest is the consistent tendency of the K-ABC Achievement subtest scores to be lower than the respective PIAT subtest scores. The mean K-ABC Reading Decoding subtest score is 6.52 standard score points below the mean PIAT Reading Recognition subtest score, while the mean K-ABC Reading Understanding subtest score is 6.10 standard score points below the mean PIAT Reading Comprehension subtest. It will be recalled that Naglieri (1985b), in a study of normal children, found an average PIAT Reading Comprehension standard score of 117.7 and an average K-ABC Reading/Understanding standard score of 104.5. The disparity in scores, therefore, appears to have some substantiation in previous comparative studies. This trend may have a substantial impact on the number of students who exhibit a severe discrepancy between ability

Table 2

Means and Standard Deviations of WISC-R, K-ABC, and PIAT Subtests and Score Composites

K-ABC Subtests	Mean	S.D.
1. Hand Movements	9.14	2.4
2. Gestalt Closure	9.71	3.0
3. Number Recall	9.13	3.0
4. Triangles	9.52	3.2
5. Word Order	9.11	3.1
6. Matrix Analogies	9.67	2.3
7. Spatial Memory	9.01	2.7
8. Photo Series	9.80	2.5
9. Faces & Places	87.65	12.7
10. Arithmetic	89.14	11.7
11. Riddles	95.94	11.3
12. Reading Decoding	84.75	12.5
13. Reading Understanding	85.22	12.3
14. Sequential	95.23	14.9
15. Simultaneous	97.30	11.9
16. Mental Processing	95.83	12.8
17. Achievement	87.42	11.1

WISC-R Subtests	Mean	S.D.
1. Information	7.97	2.7
2. Similarities	9.97	3.0
3. Arithmetic	7.94	2.4
4. Vocabulary	9.42	2.9
5. Comprehension	10.04	2.8
6. Digit Span	7.52	2.8
7. Picture Completion	10.20	2.5
8. Picture Arrangement	9.99	2.6
9. Block Design	9.29	2.5
10. Object Assembly	9.86	2.6
11. Coding	8.44	2.8
12. Verbal IQ	94.11	12.7
13. Performance IQ	96.86	12.3
14. Full Scale IQ	94.94	11.1

PIAT Subtests	Mean	S.D.
1. Mathematics	92.93	10.6
2. Reading Recognition	91.27	9.1
3. Reading Comprehension	91.32	10.3
4. Spelling	90.20	10.0
5. General Information	94.84	11.8
6. Total	89.74	10.0

and reading achievement. A higher average PIAT standard score also was achieved in the area of arithmetic, although the disparity is less pronounced than in reading achievement. The difference between the average PIAT Total standard score and the average K-ABC Achievement Scale score is the smallest (2.32 standard score points favoring the PIAT). It might be anticipated that the average standard score difference will be larger with MPC/K-ABC Achievement comparisons than with WISC-R/PIAT subtest comparisons, particularly in the area of reading.

It is interesting to note that the K-ABC Mental Processing subtests show very little variability, with the average scale scores ranging from 9.01 on the Spatial Memory subtest to 9.80 on the Photo Series subtest. This flat profile of average subtest scores contrasts with previous studies which suggest that learning disabled students, learning disabilities referrals, and dyslexic students perform more poorly on Hand Movements, Word Order, Number Recall, and Matrix Analogies subtests (Kaufman & Kaufman, 1983b). It is consistent, however, with the Kaufman & McLean (1986) study of learning disability referrals and children with diagnosed learning disabilities which showed little variability among the K-ABC Mental Processing subtest scores.

Sequential Processing scores were slightly lower (2 points) than Simultaneous Processing scores, a finding which is consistent with the standardization studies with learning disabled students (#17 and #19) which showed Sequential Processing standard scores to be 2 to 5 points below Simultaneous Processing standard scores.

The average WISC-R subtest scores, however, reflected a characteristic "ACID" profile with depressed Arithmetic, Coding, Information, and Digit Span scores. This pattern of scores has been identified with learning disabled students in previous literature (Kaufman & McLean, 1986), although its utility and generalizability has been seriously questioned by other authors (Clarizio & Veres, 1983).

The WISC-R Verbal IQ is 2.75 points lower than the average Performance IQ, while there is less than a point difference between the average K-ABC MPC (95.83) and the WISC-R Full Scale IQ (94.94). This finding is somewhat inconsistent with the K-ABC standardization studies with learning disabilities referrals which showed a lower K-ABC MPC (90.8) than WISC-R FSIQ (94.7) (Study #17). The average scores are more consistent, however, with the scores from Naglieri's (1985a) sample of learning disabled students (FSIQ: 96.8; MPC: 96.1).

B. Testing of Hypotheses: <u>Part I</u>

The first hypothesis in Part I states: "There will be no significant differences between the proportions of students who demonstrate a severe discrepancy between ability and achievement using the four proposed methods for calculating a discrepancy."

In order to test this hypothesis, z_{dif} was calculated for each student using the four methods described in the previous section: Method 1, Z-score differences; Method 2, Estimated true score differences; Method 3, Regression analysis I; Method 4, Regression analysis II. If z_{dif} exceeded 1.65 for Methods 1, 2, or 3, the student was considered to have a severe discrepancy between ability and achievement for that method. If z_{dif} exceeded z_{mod} for Method 4, the student was determined to have a severe discrepancy for that Method. (Recall that z_{mod} adjusted z_a for unreliability of the ability and achievement tests).

The proportion of students demonstrating a discrepancy for each method across each WISC-R/PIAT and K-ABC MPC/Achievement comparison appears in Table 3. A Friedman two-way analysis of variance by ranks was conducted to determine if significant differences existed between proportions of students identified by each method. For purposes of this analysis, χr^2 was calculated and compared to critical values of chi-square for 3 degrees of freedom (see Appendix F for the formula for calculating χr^2 and justification for using the chi-square approximation).

It can be seen that χr^2 is significant, suggesting that the methods for determining a severe discrepancy do, in fact, tend to select different proportions of students. A closer examination of Table 3 reveals that Methods 1 and 2 select similar proportions of students while Method 3 generally identifies smaller numbers of students than

Table 3

Proportion of Students Showing a Severe Discrepancy Between Ability and Achievement Using Four Methods

	METHOD			
TEST	METH1	METH2	METH3	METH4
PIAT MATH	8.1	10.5	2.3	31.4
PIAT RREC	27.9	24.4	4.7	23.3
PIAT RCOMP	16.9	16.9	7.0	62.0
PIAT TOT	32.1	29.6	3.7	33.3
KABC ARITH	46.5	40.7	17.4	53.5
KABC RDEC	53.5	53.5	24.4	52.3
KABC RUND	57.8	57.8	28.9	57.9
KABC ACHIEV	53.5	52.3	31.4	44.2

$\chi_r^2 = 15.19$ $df = 3$ $p<.01$

Methods 1 or 2. Method 4 selects similar proportions of students as Methods 1 and 2 for PIAT Reading Recognition, PIAT Total Achievement, and each of the K-ABC subtest comparisons. In two instances (PIAT Mathematics and PIAT Reading Comprehension), Method 4 selects considerably more students that the other three methods (more than three times as many students as Methods 1 or 2). The reliabilities are low for both the PIAT Mathematics and the PIAT Reading Comprehension (.74 and .64, respectively), suggesting that, when using tests with low reliabilities, Method 4 will tend to select more students than other methods.

The analysis presented in Table 3 does not support the first hypothesis. There are significant differences between the proportion of students who demonstrate a severe discrepancy across the selected methods. Therefore, it can be anticipated that, in a group of referred children, the method chosen will have a significant impact on the numbers of students who are qualified as learning disabled.

2. The second hypothesis in Part I states: "The agreement in identification of students who demonstrate a discrepancy will be significantly higher when comparing Method 1 with Method 2 than when comparing Methods 1 or 2 with Methods 3 or 4."

Two approaches were taken to the testing of this hypothesis. First, the percent overlap between methods was calculated (number of students who are identified by two methods divided by the total number of students identified by both methods). The percent overlap indicates the extent of agreement among the pool of eligible students identified by two methods. Secondly, kappa, the extent of nominal agreement among methods, was calculated. Kappa reflects the consistency of decisions made across two methods after the effects of chance have been removed (see Appendix F). The results of these comparisons appear in Tables 4 through 7.

Table 4 shows a high degree of correspondence between Method 1 and Method 2 according to both percent overlap and kappa. The lowest agreement occurred with the WISC-R/PIAT Mathematics comparison (k=.587). Examination of the z_{dif} s for discrepant cases (see Appendix E for a listing of z_{dif}'s for each comparison) shows that lack of agreement occurred on cases for which the z_{dif} for one method was slightly below 1.65 (e.g., 1.58) while the z_{dif} for the other method was slightly above 1.65 (e.g., 1.68). Only seven students were identified with Method 1 while nine were identified with Method 2. It is possible that the relatively low reliability of the PIAT Mathematics subtest, (r_{xx} =.74), coupled with the highest mean achievement score (92.93) of the subtests in the study resulted in a very low number of students identified (and, consequently, the lower degree of agreement between methods). It should be noted, however, that nearly perfect agreement was achieved

Table 4

Agreement between Method 1 and Method 2 in the Selection of Students

	Number with Discrepancy				
TEST	Method 1	Method 2	%Overlap	Kappa	Sig
PIAT MATH	7	9	45.5	.587	$p<.01$
PIAT RREC	24	21	87.5	.910	$p<.001$
PIAT RCOMP	12	12	92.3	.949	$p<.001$
PIAT TOTAL	26	24	92.3	.947	$p<.001$
KABC ARITH	40	35	87.5	.882	$p<.001$
KABC RDEC	46	46	100.0	1.00	$p<.001$
KABC RUND	44	44	100.0	1.00	$p<.001$
KABC ACHIEV	46	45	97.8	.977	$p<.001$

between Methods 1 and 2 for PIAT Reading Comprehension, which had the lowest subtest reliability (.64) and a high average standard score score (91.32).

In general, the close correspondence revealed through these analyses (kappa is greater than .87 in seven of the eight comparisons) indicates that Method 1 and Method 2 are highly comparable. When using achievement tests with high reliabilities, the methods are nearly interchangeable.

Method 1 (Z-score differences) and Method 3 (Regression analysis 1) are compared in Table 5. It is apparent that the percent overlap between these methods is considerably lower than achieved between Methods 1 and 2. Method 1 identifies six times the number of students for the PIAT Reading Recognition subtest and nearly nine times the number of students for the PIAT Total comparisons. In each case, all students identified by Method 3 are also identified by Method 1. The percent overlap ranges from 11.5 percent for comparisons involving the PIAT Total standard scores to 58.7 percent for comparisons involving the K-ABC Achievement Composite. Kappa is significant in all K-ABC subtest comparisons, but in only one of the PIAT comparisons. The low agreement appears to be attributable to the limited number of students identified by Method 3, particularly for the PIAT subtests for which the average standard scores are higher.

Table 5

Agreement between Method 1 and Method 3 in the Selection of Students

TEST	Number with Discrepancy Method 1	Method 3	%Overlap	Kappa	Sig
PIAT MATH	7	2	28.6	.423	N.S.
PIAT RREC	24	4	16.7	.224	N.S.
PIAT RCOMP	12	5	41.7	.543	p<.01
PIAT TOTAL	26	3	11.5	.150	N.S.
KABC ARITH	40	15	37.5	.391	p<.001
KABC RDEC	46	21	45.7	.439	p<.001
KABC RUND	44	22	50.0	.466	p<.001
KABC ACHIEV	46	27	58.7	.569	p<.001

Table 6 shows the relationship between Method 1 (Z-score discrepancy) and Method 4 (Regression analysis 2) The percent overlap varies from 25.9 percent for comparisons involving the PIAT Mathematics subtest to 91.3 percent for comparisons involving the K-ABC Reading Understanding subtest. Kappa is significant at the .05 level for comparisons involving the PIAT Mathematics and the PIAT Reading Comprehension scores. Significance at the .001 level was achieved for each of the remaining comparisons. Similar numbers of students were identified by the two methods in six out of eight comparisons (PIAT Reading Recognition, PIAT Total, and the K-ABC subtest comparisons). For the two comparisons involving subtests with lower reliabilities (PIAT Mathematics and PIAT Reading Comprehension), Method 4 identified nearly four times as many students as Method 1. A high degree of consistency (kappa=.79 to .89) was achieved for each of the comparisons involving the K-ABC subtests, while moderate kappas (.57 to .58) were achieved for comparisons involving the PIAT Mathematics and PIAT Reading Comprehension subtests.

Method 2 (Estimated true score differences) and Method 3 (Regression analysis 1) are compared in Table 7. The percent overlap ranges from 10 percent for PIAT Mathematics comparisons to 60 percent for K-ABC Achievement Composite comparisons. Kappa is significant for all K-ABC subtest comparisons but is consistently lower for the PIAT subtests. This trend is due, again, to the very limited

Table 6

Agreement between Method 1 and Method 4 in the Selection of Students

TEST	Number with Discrepancy Method 1	Method 4	%Overlap	Kappa	Sig
PIAT MATH	7	27	25.9	.324	p<.05
PIAT RREC	24	20	51.7	.574	p<.001
PIAT RCOMP	12	44	27.3	.222	p<.05
PIAT TOTAL	26	27	55.9	.579	p<.001
KABC ARITH	40	46	87.0	.861	p<.001
KABC RDEC	46	45	82.0	.790	p<.001
KABC RUND	44	44	91.3	.892	p<.001
KABC ACHIEV	46	38	82.6	.815	p<.001

Table 7

Agreement between Method 2 and Method 3
in the Selection of Students

	Number with Discrepancy				
TEST	Method 2	Method 3	%Overlap	Kappa	Sig
PIAT MATH	9	2	10.0	.149	N.S.
PIAT RREC	21	4	19.0	.262	N.S.
PIAT RCOMP	12	5	31.0	.412	p<.05
PIAT TOTAL	24	3	12.5	.167	N.S.
KABC ARITH	35	15	42.9	.471	p<.001
KABC RDEC	46	21	45.7	.439	p<.001
KABC RUND	44	22	50.0	.466	p<.001
KABC ACHIEV	45	27	60.0	.588	p<.001

number of students identified by Method 3, particularly using the PIAT subtests. The comparisons involving the K-ABC subtests, however, reflect more consistency, with all students identified by Method 3 also identified by Method 2 and overlap ranging from 42.9% to 60.0%.

Table 8 illustrates the comparison of Method 2 (Estimated true score differences) with Method 4 (Regression analysis 2). It can be seen that percent overlap ranges from 27.3 percent for PIAT Reading Comprehension comparisons to 89.1 percent for K-ABC Reading Understanding comparisons. Kappa is significant at the .05 level for the PIAT Mathematics, and Reading Comprehension subtest comparisons. The .01 level of significance was achieved with the other PIAT subtest comparisons (Reading Recognition and Total), while the remainder of the kappas were significant at the .001 level. Again, a higher degree of correspondence was found between methods for subtests with higher reliabilities.

The two regression procedures, Method 3 and Method 4, are compared in Table 9. Percent overlap varies from 7.4 percent for the PIAT Mathematics subtest to 71.1 percent for the K-ABC Achievement comparisons. Kappa was nonsignificant for all of the PIAT subtests, with Method 4 identifying from five to nine times as many students as Method 3. Significance at the .01 level was achieved for comparisons involving the K-ABC Arithmetic subtest, while significance at the .001 level was obtained for the

Table 8

Agreement between Method 2 and Method 4 in the Selection of Students

TEST	Number with Discrepancy Method 2	Method 4	%Overlap	Kappa	Sig
PIAT MATH	9	27	28.6	.341	p<.05
PIAT RREC	21	20	44.4	.456	p<.01
PIAT RCOMP	12	44	27.3	.222	p<.05
PIAT TOTAL	24	27	41.7	.400	p<.01
KABC ARITH	35	46	76.1	.747	p<.001
KABC RDEC	46	45	85.7	.837	p<.001
KABC RUND	44	44	89.1	.892	p<.001
KABC ACHIEV	45	38	84.4	.838	p<.001

Table 9

Agreement between Method 3 and Method 4 in the Selection of Students

TEST	Number with Discrepancy Method 3	Method 4	%Overlap	Kappa	Sig
PIAT MATH	2	27	7.4	.099	N.S.
PIAT RREC	4	20	20.0	.277	N.S.
PIAT RCOMP	5	44	11.4	.089	N.S.
PIAT TOTAL	3	27	11.1	.143	N.S.
KABC ARITH	15	46	32.6	.310	$p<.01$
KABC RDEC	21	45	46.7	.455	$p<.001$
KABC RUND	22	44	50.0	.457	$p<.001$
KABC ACHIEV	27	38	71.1	.733	$p<.001$

remaining K-ABC subtest comparisons.

While kappa reached significance in many of the above comparisons, the statistic should be interpreted in light of the corresponding percent overlap, particularly when the placement of students is being considered based upon the resulting severe discrepancies. It is not unlikely, for example, to achieve a highly significant kappa with less than fifty percent of the same students being identified using two methods. Consequently, in many comparisons, quite a different population of students would show severe discrepancies depending on the method selected. In discussing the testing of kappa for significance, Cohen (1960) indicates, "...it is generally of as little value to test kappa for significance as it is for any other reliability coefficient-to know merely that kappa is beyond chance is trivial since one usually expects much more than this in the way of reliability in psychological measurement. It may, however, serve as a minimum demand in some applications" (p. 44). With this in mind, the greatest agreement appears to be achieved between Method 1 and Method 2, with perfect agreement resulting in two of the eight comparisons and kappa greater than .87 in seven of the eight comparisons.

Method 4 corresponds closely with Methods 1 and 2 when comparing the K-ABC subtests. With subtests with lower reliabilities (PIAT Mathematics and PIAT Reading Comprehension), however, Method 4 tends to identify

considerably more students than Methods 1 and 2.
The lowest agreement appears between comparisons involving Method 3 (Regression analysis 1), primarily because of the very limited number of students identified using this approach. The agreement, when using the PIAT subtests, tends to be lower than agreement obtained with K-ABC subtests across methods.

Figures 3 through 10 graphically illustrate the relationship between the discrepancy methods. The symbols below each line represent students who were selected by the respective methods. The number of students falling into the selection area corresponds to the entries in the preceding tables.

WISC-R/PIAT Mathematics comparisons appear in Figure 3. The test reliabilities and intercorrelations are listed at the top of the graph. It can be seen that Method 3, Regression analysis I, is the most restrictive of the Methods, selecting only two students. The regression procedure cutoff lines have equal slopes (r_{xy}) and vary only in the intercepts. Because of the substantial correction for subtest unreliability, the intercept for Method 4 is larger than the intercept for Method 3. Consequently, more students are identified with Method 4 than with Method 3.

The low percent overlap between Methods 1 and 2 is explained both by the diverging lines (resulting from the

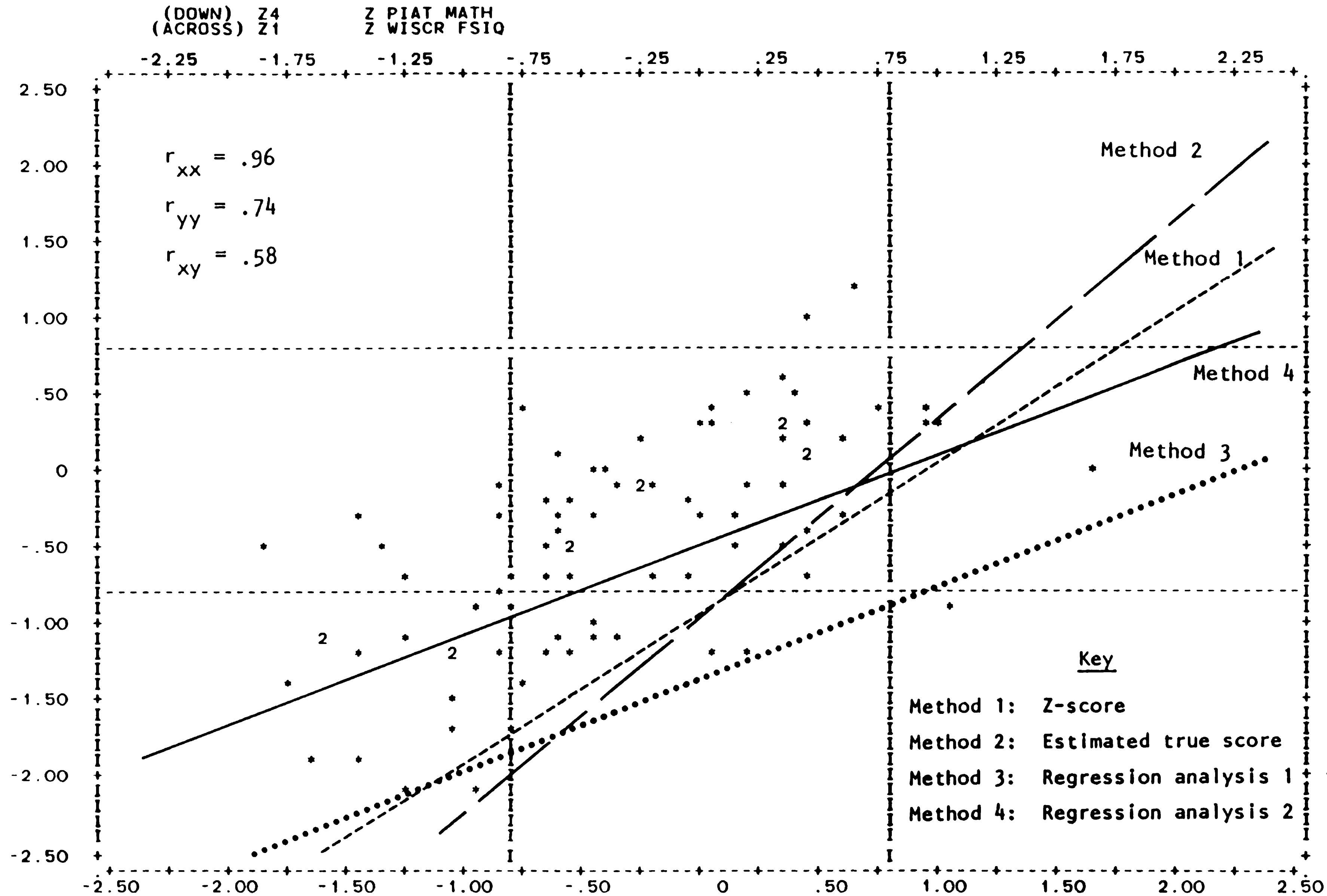


Figure 3: A comparison of four methods for determining a severe discrepancy: WISC-R/PIAT Mathematics

lower test reliabilities) as well as the several students who fall near the lines (close to the 1.65 cutoff). The relatively high average PIAT Mathematics subtest score is also illustrated by the large number of scores falling above $z_y = 0$ (standard score = 100).

Figure 4 shows the relationship between methods for the WISC-R/PIAT Reading Recognition subtest. The high degree of agreement between Methods 1 and 2 results from the high reliabilities of the Full Scale IQ (.96) and the PIAT Reading Recognition subtest (.89). It can be seen that Methods 1, 2, and 4 include all of the students who are selected by Method 3. Once again, the correction for unreliability creates a higher intercept for Method 4 than for Method 3. The correction is somewhat less than on Figure 3, however.

The WISC-R/PIAT Reading Comprehension comparisons appear in Figure 5. The relationship between Method 1 and Method 2 is more discrepant than in Figure 4, and the disparity can be attributed directly to the lower PIAT Reading Comprehension subtest reliability ($r_{yy} = .64$). Because of the distribution of scores, however, the agreement between the two methods is still relatively high (overlap = 92.3 %). The lines representing the two regression methods are again parallel, with all of the students selected by Method 3 also identified by Method 4. As with Figure 3, a large correction for unreliability resulted in Method 4 selecting substantially more students

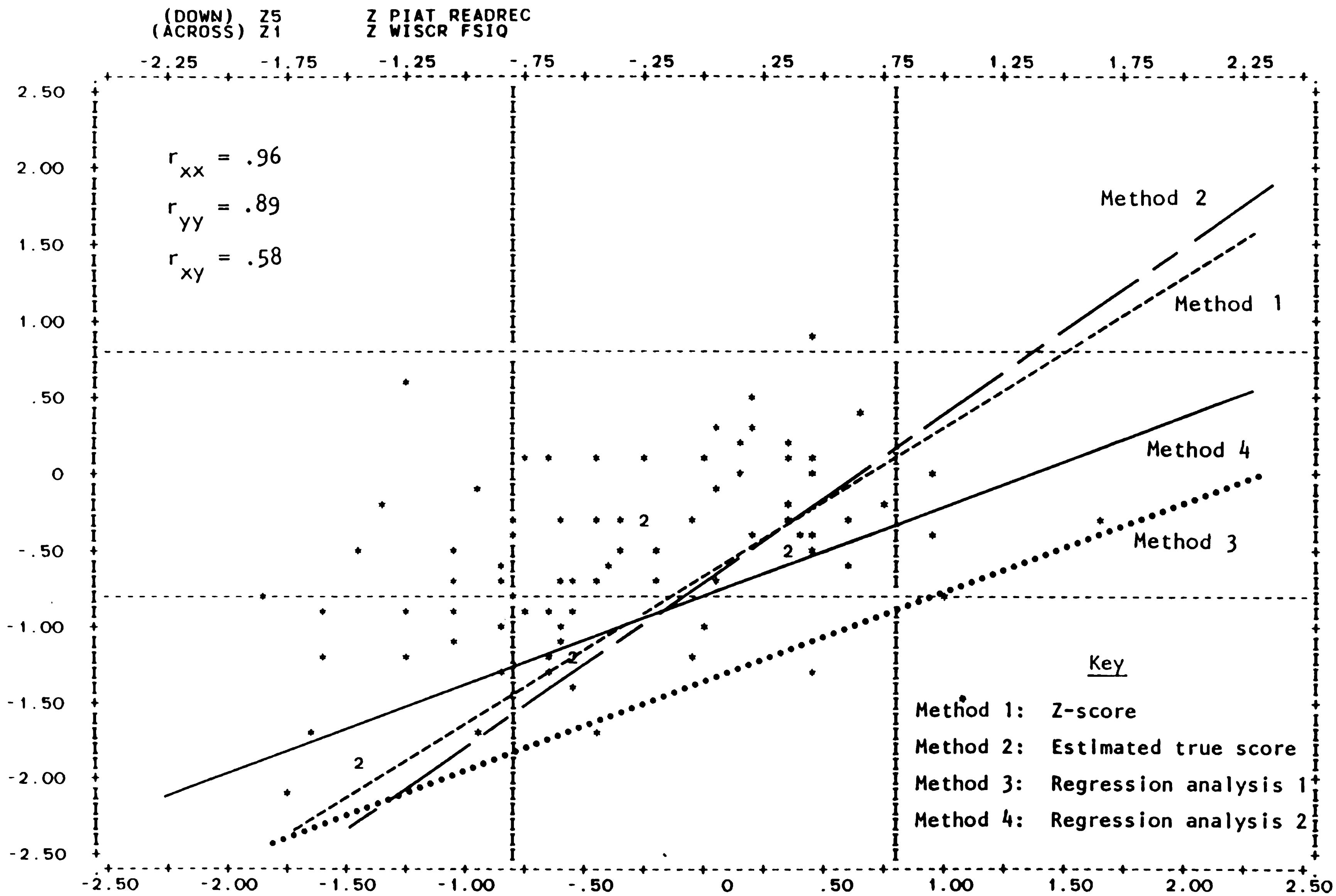


Figure 4: A comparison of four methods for determining a severe discrepancy: WISC-R/PIAT Reading Recognition

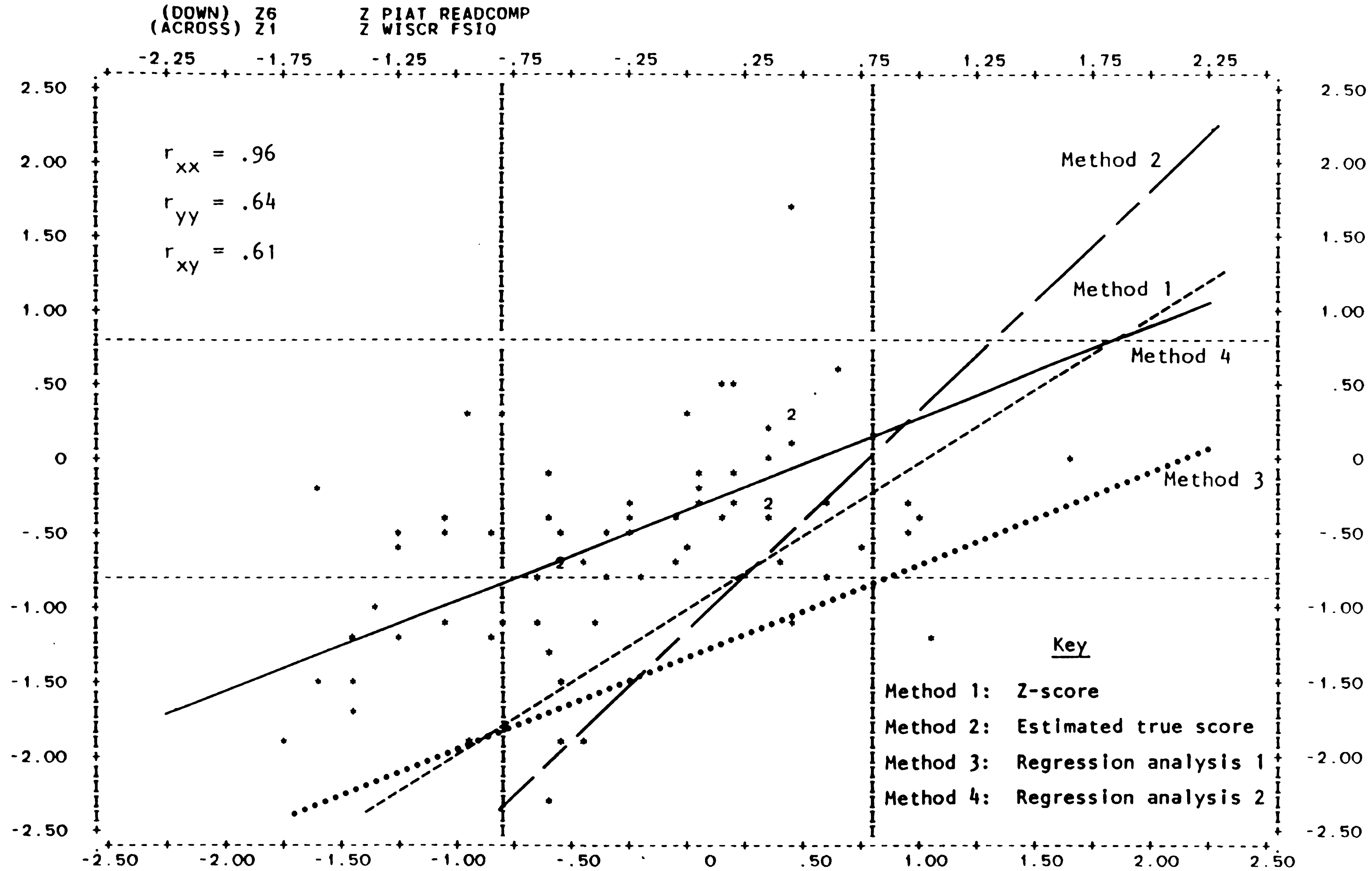


Figure 5: A comparison of four methods for determining a severe discrepancy: WISC-R/PIAT Reading Comprehension

than Method 3. The distance between the intercepts for Method 3 and Method 4 is greater in Figure 5 than in any of the other Figures, 3 through 10. This difference is attributable, again, to the low PIAT Reading Comprehension subtest reliability.

The WISC-R/PIAT Total comparisons appear in Figure 6. A high degree of correspondence between Methods 1 and 2 is illustrated in this graph. As with Figure 4, the reliabilities of the two tests are high (FSIQ: r_{xx} = .96; PIAT Total: r_{yy} = .89) and the small number (2) of students who were not selected by both methods are represented by symbols which lay close to both lines. Methods 3 and 4 are illustrated by parallel lines, with Method 4 inclusive of Method 3. It can be seen that Method 3 selects only 3 students, while Method 4 identifies considerably more students (27).

The K-ABC comparisons are represented by Figures 7 through 10. Figures 7 through 10 show virturally identical lines for Methods 1 and 2. The K-ABC MPC as well as all of the K-ABC achievement subtests have high reliabilities (.87 to .94) which explains the high degree of correspondence noted in the tables describing percent overlap (Tables 5 through 10), as well as the similarities in slopes and intercepts illustrated here.

Of additional interest is the relationship between Methods 3 and 4 in Figures 7 through 10. Consistent with previous figures, Method 4 is inclusive of Method 3 with

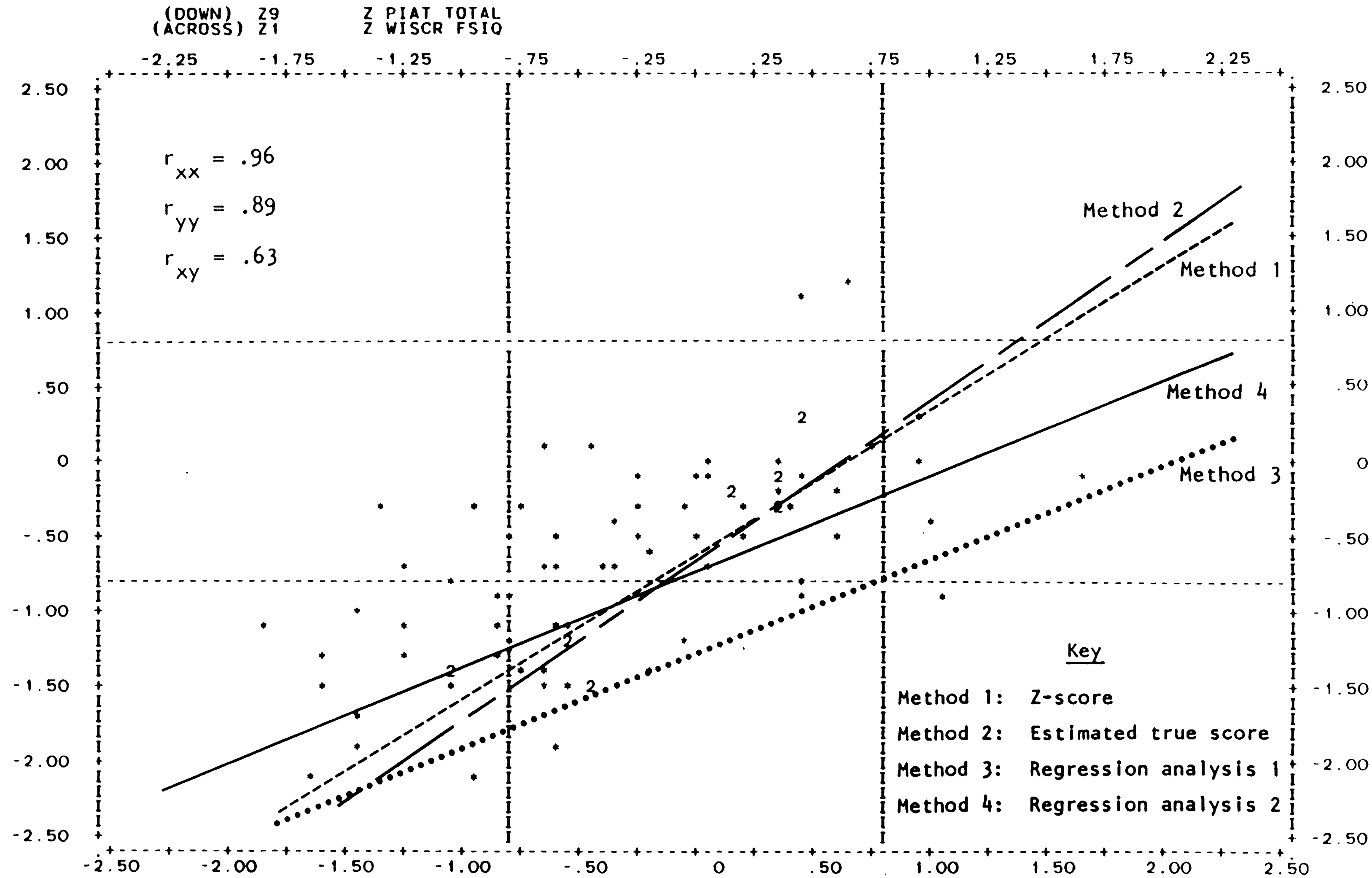


Figure 6: A comparison of four methods for determining a severe discrepancy: WISC-R/PIAT Total

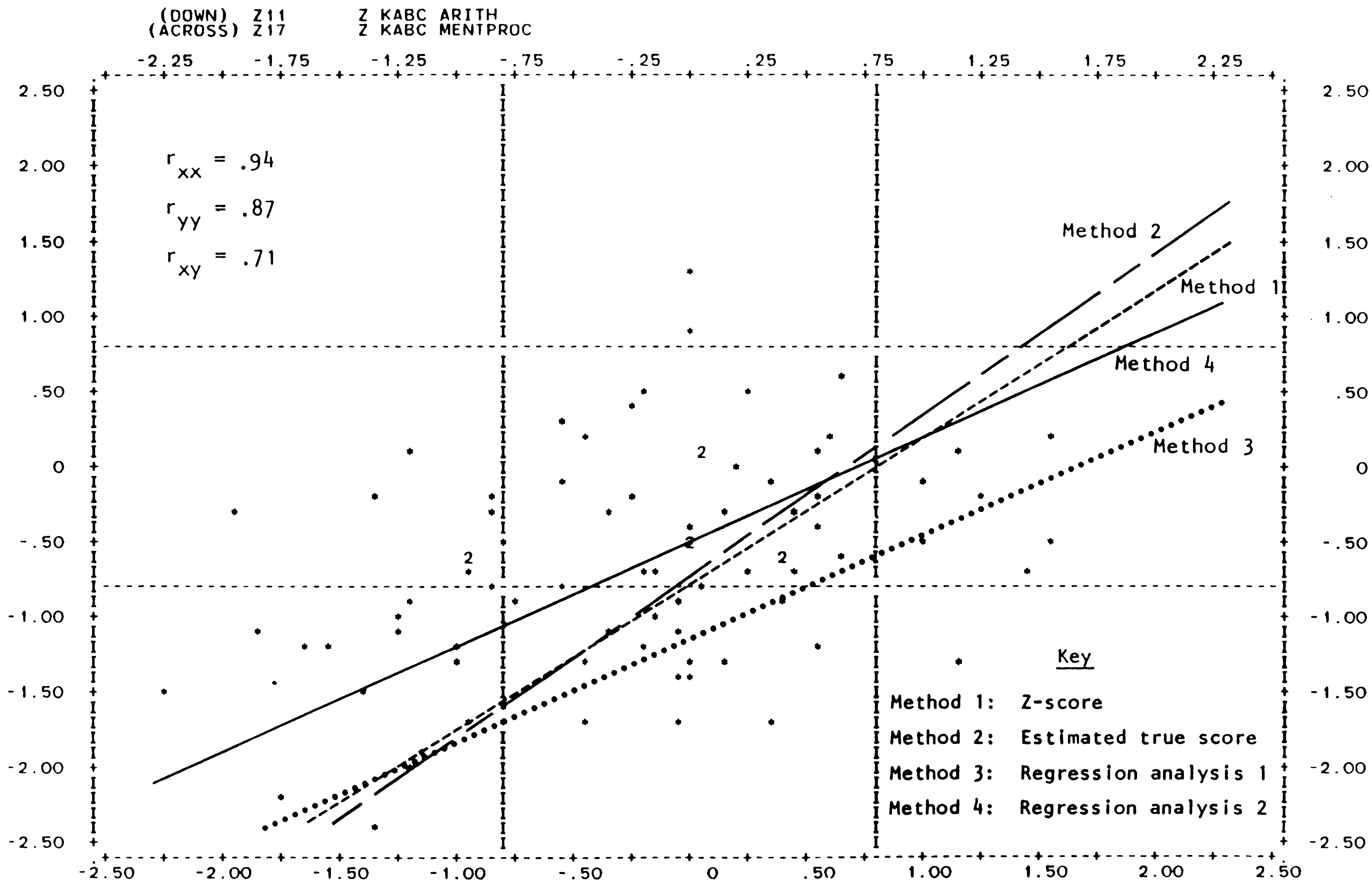


Figure 7: A comparison of four methods for determining a severe discrepancy: K-ABC MPC/Arithmetic

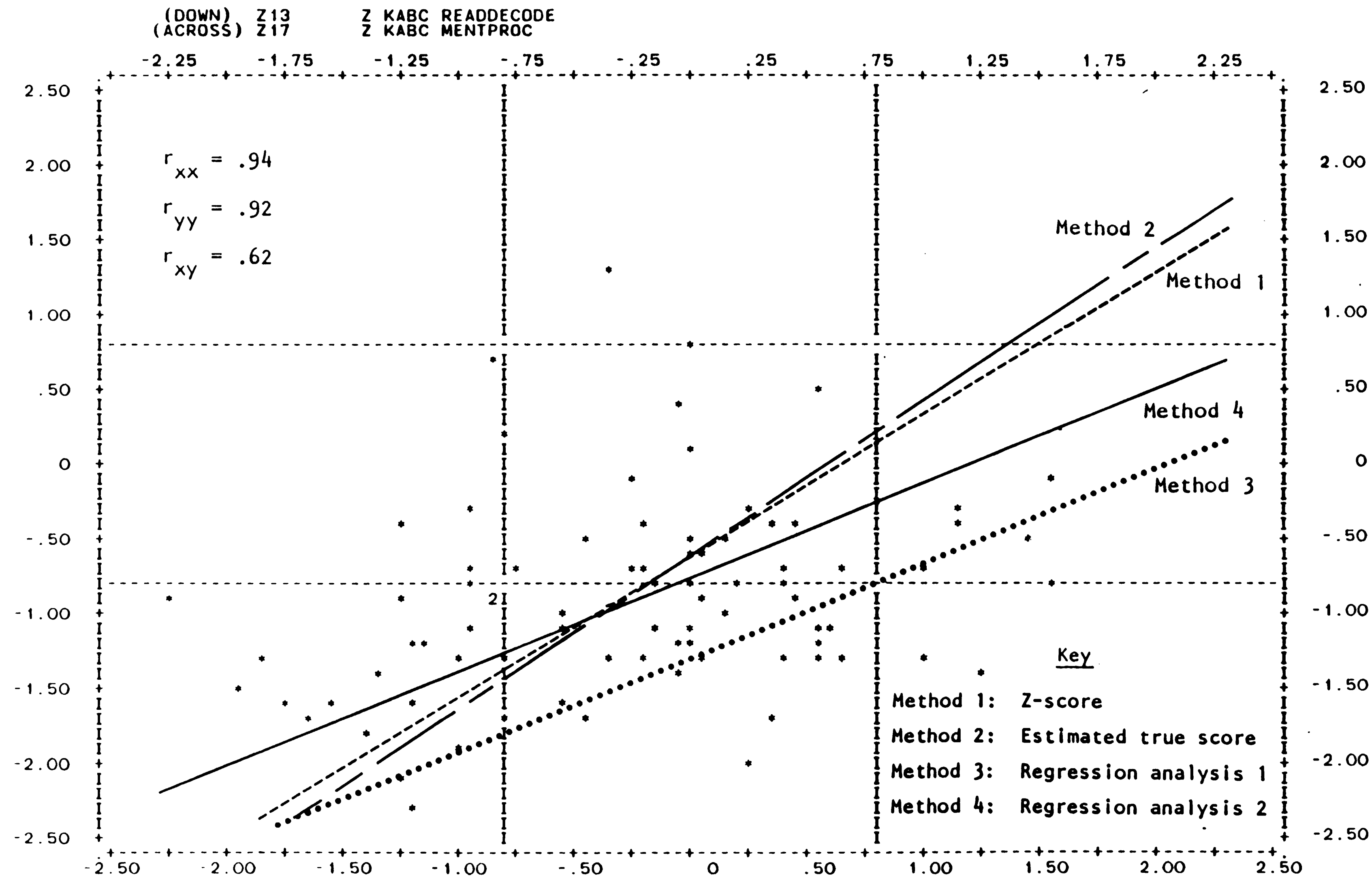


Figure 8: A comparison of four methods for determining a severe discrepancy: K-ABC MPC/Reading Decoding

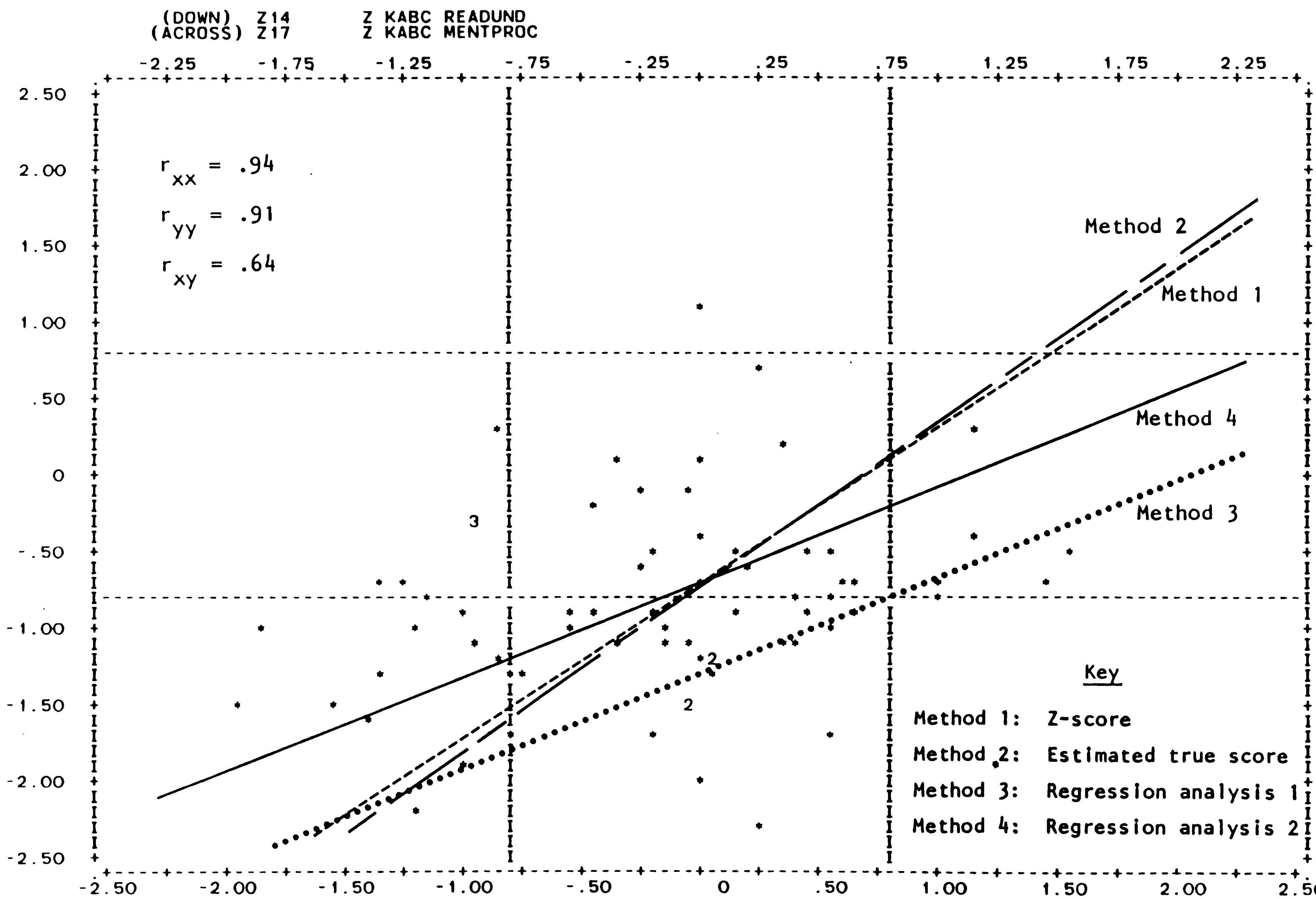


Figure 9: A comparison of four methods for determining a severe discrepancy: K-ABC MPC/Reading Understanding

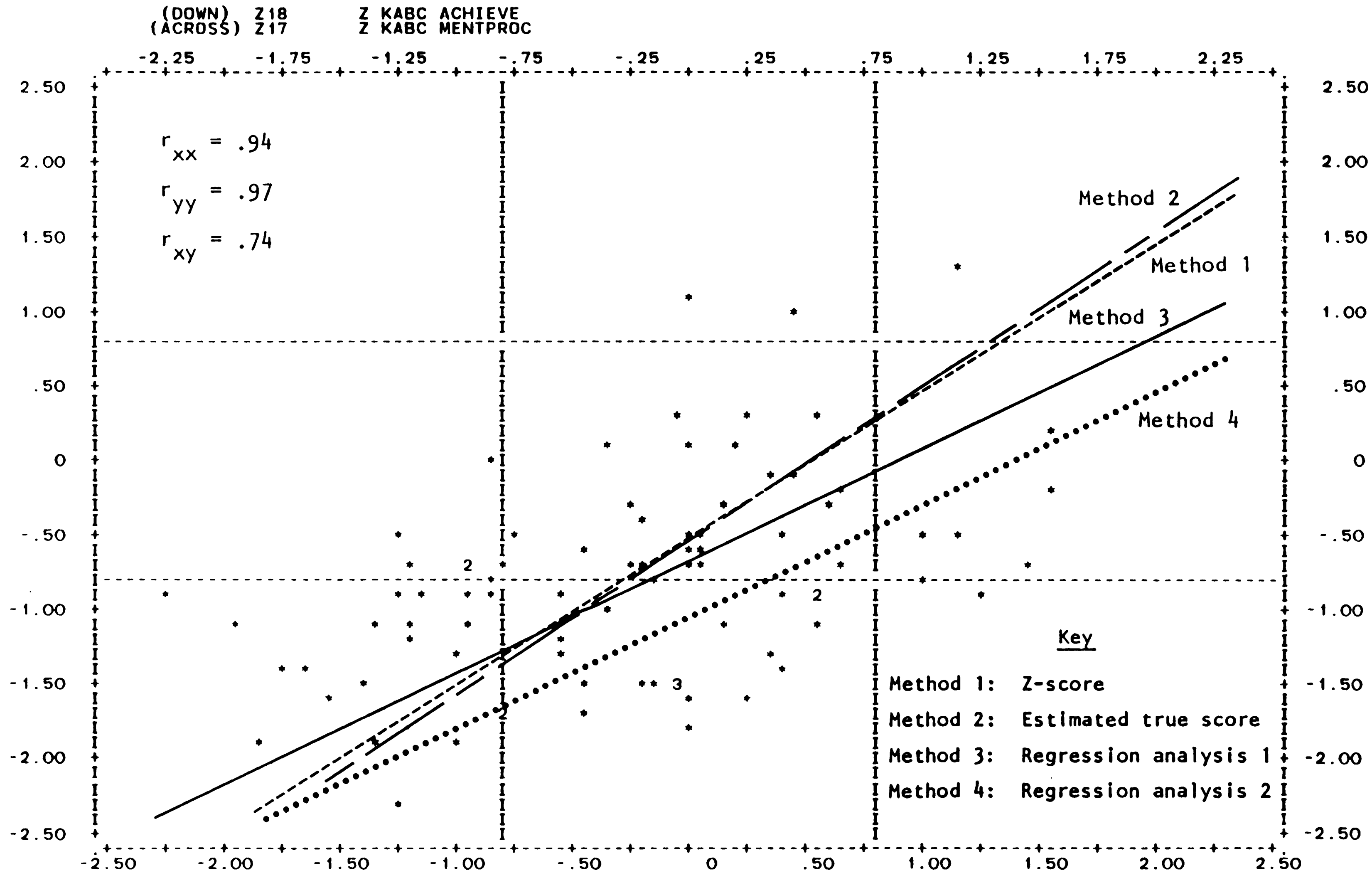


Figure 10: A comparison of four methods for determining a severe discrepancy: K-ABC MPC/Achievement Composite

the intercept of Method 4 varying as a function of the achievement subtest reliability. As the reliability increases, the intercepts for Method 3 and 4 become closer. With perfectly reliable tests, Method 1 will equal Method 2. Figure 10 illustrates the comparison of two reliable tests (r_{xx} = .94; r_{yy} = .97). In this comparison, the intercepts for the two regression procedures are the closest, resulting in the higher degree of overlap observed between Methods 3 and 4 in Table 10 (kappa = .733; percent overlap = 71.1).

It can be concluded from this set of analyses that Methods 1 and 2 might be used interchangeably in the calculation of a discrepancy, particularly when using tests with high reliabilities (.85 or greater). In such cases Method 1 might be preferred to Method 2 because of the relative simplicity of the discrepancy calculation. Even with lower test reliabilities, however, there is considerable overlap between methods.

Because Method 4 corrects for unreliability, the extent of agreement between Methods 3 and 4 is dependent on the reliabilities of the subtests. When using tests with high reliabilities, there will tend to be greater agreement between the two methods. However, even when using subtests with high reliabilities, such as the K-ABC Reading Decoding subtest (r_{yy} = .92), the overlap between Methods 3 and 4 may be only moderate (.455%) due to the distribution of test

scores.

When using tests with low reliability, it can be expected that Method 4 will select appreciably more students than Method 3. Although Method 4 creates a less restrictive cutoff than Method 3, errors of measurement are still present. The probability of selecting a student who does not have a true discrepancy is greater. Therefore, it is advised that tests with lower reliabilities be avoided.

A much less consistent relationship is evidenced when comparing Methods 1 and 2 with Methods 3 and 4. Since the slope of the regression cutoff line is equal to the correlation between tests (r_{xy}) and the slopes for Methods 1 and 2 are greater than or equal to 1.0, it might be expected that, as r_{xy} increases, the regression methods would be more consistent with Methods 1 and 2. This, however, does not appear to be the case. The impact of intercept differences and score distributions appear to be more important in determining the extent of agreement between methods than slope changes.

Method 4 tends to correspond more closely with Methods 1 and 2 than does Method 3, again, because the intercept of the Method 4 cutoff line is generally closer to the intercepts of lines for Methods 1 and 2. This is particularly apparent when comparing the more reliable subtests (Figures 4, 6, and 7 through 10). It might be concluded, therefore, that subtest reliabilities are more important in determining the agreement between methods than test intercorrelations.

3. Hypothesis 3 of Part I states: "The correlation between pairs of discrepancies obtained using Method 1 and 2 will be higher than correlations obtained when comparing Method 1 with Methods 3 and 4 or when comparing Method 2 with Methods 3 and 4."

Table 10 shows the rank order correlations between methods across all subtests. Since these correlations take into consideration the entire range of z_{dif}'s, there is considerable agreement demonstrated between each comparison. Methods 3 and 4 correlate perfectly because both use the same numerator (slope) in the discrepancy formula. Setting aside this artifact, it can be seen that Method 1 and Method 2 correlate almost perfectly, with coefficients ranging from .944 for the PIAT Reading Comprehension subtest comparisons to .999 for the K-ABC Achievement Composite comparisons. All correlations involving the K-ABC subtests exceed .93, while the correlations involving the PIAT subtests are generally somewhat lower. Each correlation is significant at the .001 level.

While it is of interest to examine the correlations among methods, it should be noted that the coefficients are spuriously high because of the range of z_{dif}'s which are considered, as well as the inclusion of similar elements in certain of the discrepancy formulas. Because of these confounding effects, tests of the significance of

Table 10

Agreement between Methods for the Range of Z_{dif} 's

	METHODS (M) COMPARED					
TEST	M1-M2	M1-M3	M1-M4	M2-M3	M2-M4	M3-M4
PIAT MATH	.954	.856	.856	.676	.676	1.00
PIAT RREC	.997	.890	.890	.855	.855	1.00
PIAT RCOMP	.944	.878	.878	.712	.712	1.00
PIAT TOTAL	.994	.880	.880	.836	.836	1.00
KABC ARITH	.998	.956	.956	.940	.940	1.00
KABC RDEC	.999	.944	.944	.938	.938	1.00
KABC RUND	.999	.961	.961	.957	.957	1.00
KABC ACHIEV	.999	.975	.975	.979	.979	1.00

differences in the correlation coefficients were not undertaken.

4. Hypothesis 4 of Part I states: "Methods 3 and 4 will include significantly fewer high ability students and significantly more low ability students than either Method 1 or 2."

Before testing this hypothesis, the data from several test comparisons were consolidated. It was reasoned that a student could exhibit a severe discrepancy on either the Mathematics, Reading Recognition, or Reading Comprehension subtest of the PIAT and be identified as having a discrepancy between ability and achievement. Likewise, the student could show a discrepancy on the Arithmetic, Reading Decoding, or Reading Understanding subtest of the K-ABC and be identified as having a discrepancy between ability and achievement. With this consideration, students exhibiting a severe discrepancy in any of the three achievement areas were selected using each method. It is acknowledged that, when using multiple comparisons, a more conservative z_a would be recommended [see Reynolds (1984-85) for guidelines in establishing a cutoff when using multiple test comparisons]. For purposes of simplifying the analysis, however, $z_a = 1.65$ was retained. It might be expected that a somewhat smaller population with slightly higher average FSIQ and MPC would be identified using a larger z_a.

The WISC-R Full Scale IQ scores for students showing a discrepancy on PIAT subtests were analyzed using a one-way analysis of variance. The analysis appears in Table 11. The resulting F ratio of 1.915 is not significant, suggesting that there are no significant differences in the mean IQ of students identified by the four methods. It should be noted, however, that the pattern of mean scores for the four methods follows the pattern of scores predicted in hypothesis four. That is, the mean I.Q. scores for the regression approaches (Methods 3 and 4) are lower than the mean I.Q. scores for Methods 1 and 2. The limited number of students identified by Method 3 may have served to reduce the sensitivity of this analysis. While not undertaken in the present study, it would appear worthwhile to further examine the patterns of mean Verbal and Performance I.Q.'s across methods.

A similar analysis of variance was performed on the Mental Processing Composite of the K-ABC across methods. The results appear in Table 12. As with the previous analysis of the WISC-R Full Scale IQ, the analysis of variance on the K-ABC MPC yielded no significant differences between methods. In fact, the mean MPC's are remarkably similar for students identified by each method. Hypothesis 4, therefore, appears to be unsupported by these analyses. As a result, it can be concluded that the four methods would not select populations with significantly different ability estimates.

Table 11

Analysis of Variance for Full Scale IQ's of Identified Students across Methods for Determining a Discrepancy

METHOD			
METH1	METH2	METH3	METH4
$\bar{X}$ =98.10	$\bar{X}$ =98.95	$\bar{X}$ =93.00	$\bar{X}$ =94.86
SD=11.77	SD=11.95	SD=11.15	SD=11.37
SS=5267.6	SS=5143.9	SS=2612.0	SS=8796.55
n=39	n=37	n=22	n=69

$s_b^2 = 256.36$

$s_w^2 = 133.87$

F = 1.915 df= 3, 163 N.S.

Table 12

Analysis of Variance of K-ABC MPC's of Identified Students across Methods for Determining a Discrepancy

METHOD			
METH1	METH2	METH3	METH4
$\bar{X}$ =98.82	$\bar{X}$ =99.31	$\bar{X}$ =98.33	$\bar{X}$ =98.00
SD=12.95	SD=12.88	SD=14.49	SD=13.10
SS=10055.0	SS=9626.51	SS=8821.44	SS=10646.0
n=61	n=59	n=43	n=63

$s_b^2 = 19.45$

$s_w^2 = 176.34$

F = .110 df = 3, 222 N.S.

5. Hypothesis 5, Part I states: "There will be no significant differences in the mean age, grade placement or sex ratio of identified students across the four methods for determining a discrepancy."

In order to test differences in mean chronological age across methods, a one-way analysis of variance was performed on the ages of students identified by any of the three WISC-R/PIAT subtest comparisons. Table 13 illustrates this analysis.

The F ratio of 2.42 with 3 and 163 degrees of freedom is non-significant. The analysis suggests that there are no significant differences in mean chronological age of students identified as having a severe discrepancy across methods when using WISC-R/PIAT comparisons.

Table 14 shows a one-way analysis of variance performed on the ages of students identified by any of the three K-ABC MPC/Achievement subtest comparisons across methods for calculating a discrepancy.

Again, a non-significant F ratio indicates that there are no significant differences in mean chronological age of students identified with MPC/K-ABC Achievement subtest comparisons. The first part of Hypothesis 5 appears to be supported.

To test the second part of Hypothesis 5, similar analysis of variance tables were developed to examine mean grade placement. The results appear in Tables 15 and 16. Non-significant F ratios resulting from both analyses

Table 13

Analysis of Variance of Chronological Age of Students Identified by WISC-R/PIAT Comparisons across Methods for Determining a Discrepancy

METHOD 1	METHOD 2	METHOD 3	METHOD 4
$\bar{X}$ =8.38	$\bar{X}$ =8.18	$\bar{X}$ =7.76	$\bar{X}$ =8.88
SD =1.80	SD =1.63	SD =1.73	SD =1.73
SS =123.21	SS =95.17	SS =63.16	SS =199.95
N =39	N =37	N =22	N =68

s_b^2 = 7.20

s_w^2 = 2.97

F = 2.42 d.f. = 3, 162 N.S.

Table 14

Analysis of Variance of Chronological Age of Students Identified by MPC/K-ABC Achievement Comparisons across Methods for Determining a Discrepancy

METHOD 1	METHOD 2	METHOD 3	METHOD 4
$\bar{X}$ =8.76	$\bar{X}$ =8.71	$\bar{X}$ =8.62	$\bar{X}$ =8.77
SD =1.82	SD =1.80	SD =1.97	SD =1.76
SS =198.51	SS =187.8	SS =162.38	SS =192.98
N =61	N =59	N =43	N =63

s_b^2 = .225

s_w^2 = 3.34

F = .067 df = 3, 222 N.S.

Table 15

Analysis of Variance of Grade Placement of Students Identified by WISC-R/PIAT Comparisons across Methods for Determining a Discrepancy

METHOD 1	METHOD 2	METHOD 3	METHOD 4
$\bar{X}$ =2.38	$\bar{X}$ =2.24	$\bar{X}$ =1.82	$\bar{X}$ =2.84
SD =1.60	SD =1.49	SD =1.50	SD =1.63
SS =97.23	SS =80.81	SS =47.27	SS =181.25
N =39	N =37	N =22	N =69

$s_b^2 = 6.33$

$s_w^2 = 2.49$

$F = 2.54$ df = 3, 163 N.S.

Table 16

Analysis of Variance of Grade Placement of Students Identified by MPC/K-ABC Achievement Comparisons across Methods for Determining a Discrepancy

METHOD 1	METHOD 2	METHOD 3	METHOD 4
$\bar{X}$ =2.80	$\bar{X}$ =2.75	$\bar{X}$ =2.58	$\bar{X}$ =2.76
SD =1.62	SD =1.59	SD =1.72	SD =1.59
SS =157.64	SS =149.19	SS =124.47	SS =157.43
N =61	N =59	N =43	N =63

$s_b^2 = .608$

$s_w^2 = 2.65$

$F = .229$ df = 3, 222 N.S.

support the hypothesis of no significant differences in mean grade placement for students identified as having a severe discrepancy across methods.

Finally, the difference in the proportions of male and female students identified as having a severe discrepancy across methods was examined through contingency tables and chi-square analyses. Tables 17 and 18 illustrate the proportions of males and females showing a severe discrepancy for WISC-R/PIAT and K-ABC MPC/Achievement Composite comparisons, respectively. The non-significant chi-square in each case supports the hypothesis of no significant differences in the proportions of male and female students identified across methods.

The testing of Hypothesis 5 has indicated that there are no significant differences in the characteristics (sex, age, grade placement) of students who are identified by the various methods for determining a severe discrepancy. It might be anticipated that students with similar demographic characteristics will be selected, regardless of the formula chosen for calculating a discrepancy. It should be cautioned, however, that the original distribution of students was not even with regard to age or grade placement. A concentration of students in the seven to ten year age range along with an over-representation of second grade students may have limited the sensitivity of the analyses with regard to these characteristics. A more definitive analysis would result from testing an equal

Table 17

Proportions of Students Showing a Severe Discrepancy Using WISC-R/PIAT Comparisons by Sex for Methods of Determining a Discrepancy

METHOD	Males	Females
METH1	.492	.360
METH2	.492	.280
METH3	.295	.160
METH4	.787	.840

$\chi^2 = 2.45$ N.S.

Table 18
Proportions of Students Showing a Severe Discrepancy Using MPC/KABC Achievement Comparisons by Sex for Methods for Determining a Discrepancy

METHOD	Males	Females
METH1	.721	.680
METH2	.689	.680
METH3	.508	.480
METH4	.754	.680

$\chi^2 = .097$ N.S.

number of students at each age or grade level.

In spite of the indication that the choice of a method for calculating a discrepancy can have a substantial impact on the population of students identified as learning disabled, there has been no attempt to determine the adequacy of the methods in selecting learning disabled students. In an effort to determine the validity of each method in identifying students, the z_{dif}'s which were calculated using the four formulas were correlated with the classroom teacher's perception of the severity of the child's difficulty in each of the achievement areas.

Appendix B, page 3, shows the teacher rating form which was completed concurrently with the administration of the K-ABC, WISC-R, and PIAT. Each student's teacher rated the severity of the child's academic problems from 1 to 5, where a rating of 1 indicated that the child did not have a problem and a rating of 5 indicated that the child had a very serious problem. Items 35, 36, 39, and 47 of the rating scale correspond with the areas of academic concern. Ratings on these items were correlated directly with the z_{dif}'s which were obtained using the respective achievement tests. Tables 19 through 22 illustrate the pattern of correlations.

Table 19 reflects the relationships between the teachers' ratings of the students' difficulty in the area of reading decoding and the z_{dif}'s which resulted from

Table 19

Correlations between Teacher Ratings of Students' Academic Difficulty and Calculated Discrepancies: Reading Recognition

	TEST	
	PIAT RREC	K-ABC RDEC
METHOD 1	.084	.371***
METHOD 2	.065	.368***
METHOD 3	.211*	.425***
METHOD 4	.211*	.425***

Table 20

Correlations between Teacher Ratings of Students' Academic Difficulty and Calculated Discrepancies: Reading Comprehension

	TEST	
	PIAT RCOMP	K-ABC RUND
METHOD 1	.095	.330**
METHOD 2	.017	.323**
METHOD 3	.182*	.405***
METHOD 4	.182*	.405***

* $p < .05$; ** $p < .01$; *** $p < .001$

comparisons using the PIAT Reading Recognition and K-ABC Reading Decoding subtests. It can be seen that the regression analysis procedures result in discrepancies which correspond the most closely with teachers' perceptions, although the relationships are far from perfect. For the WISC-R/ PIAT Reading Recognition comparisons, the correlations between the teacher ratings and the four methods range from r =.065 for Method 2 to r = .211 for Methods 3 and 4. For the K-ABC MPC/Reading Decoding comparisons, the correlations ranged from r = .368 (Method 2) to r = .425 (Methods 3 and 4). It would appear that the K-ABC Reading Decoding subtest relates more closely to teachers' perceptions of achievement deficit than the PIAT Reading Recognition test. Additionally, the regression methods appear somewhat more consistent with teachers' ratings than Methods 1 and 2.

Table 20 shows the correlations between teachers' ratings and discrepancies in the area of reading comprehension. The results are quite similar to the correlations found in Table 19, with the K-ABC MPC/Reading Understanding comparisons relating more closely to the teacher ratings (r = .330 to .405) than the WISC-R/PIAT Reading Comprehension comparisons (r = .017 to .182). Again, there is a tendency for the regression analysis discrepancies to relate slightly better to the teacher ratings than Method 1 and 2 discrepancies.

The correlations in the area of arithmetic show a similar pattern, although the relationships are not as strong as in the reading areas. Table 21 illustrates the correlations between teacher ratings of the students' difficulty in arithmetic and the discrepancies achieved for WISC-R/PIAT Mathematics and K-ABC MPC/Arithmetic comparisons. The correlations reached significance at the .05 level for discrepancies obtained using the regression procedures (Methods 3 and 4). Although significant, the correlations are low, accounting for a very small proportion of the total variance between the measures.

Table 22 illustrates the correlations between the teacher ratings of a student's overall achievement difficulty and the discrepancies achieved using WISC-R/PIAT Total and K-ABC MPC/Achievement test scores. It is apparent that the pattern of correlations is different for the total achievement comparisons. All of the correlations are low and nonsignificant, with the correlations for the regression analysis procedures the lowest of all. This finding is not surprising in light of the fact that the total achievement scores reflect achievement in a number of areas. Often children referred for learning disabilities show a deficit in only one area of achievement and may excel in other areas.

While teacher ratings may not offer an ideal external criterion for quantifying learning deficits, the ratings do provide one source by which to judge the adequacy of the

Table 21

Correlations between Teacher Ratings of Students' Academic Difficulty and Calculated Discrepancies: Arithmetic

	TEST	
	PIAT MATH	K-ABC ARITH
METHOD 1	-.050	.075
METHOD 2	-.166	.047
METHOD 3	.185*	.193*
METHOD 4	.185*	.193*

Table 22

Correlations between Teacher Ratings of Students' Academic Difficulty and Calculated Discrepancies: Total Achievement

	TEST	
	PIAT TOTAL	K-ABC ACHIEV
METHOD 1	.041	.147
METHOD 2	.071	.140
METHOD 3	-.137	.077
METHOD 4	-.137	.077

* $p < .05$; ** $p < .01$; *** $p < .001$

various methods used in determining a discrepancy as well as the adequacy of the achievement subtests. It should be emphasized that in each analysis the correlations are low: less than 20 percent of the variance in Z_{dif}'s can be predicted from the ratings. These low correlations would suggest the need for a better defined external criterion and, perhaps, the need for multiple rating sources.

With this limitation in mind, however, it would appear that the K-ABC Reading subtests agree more closely with teachers' perceptions than do the PIAT Reading subtests. The consistent tendency of the regression analysis approaches to correlate higher with teacher ratings also suggests that these procedures may be slightly more valid while representing the more statistically sound approach to the determination of a discrepancy.

C. Hypotheses: Part II

1. The first hypothesis proposed in Part II states: "There will be no significant differences between the proportions of students who demonstrate a severe discrepancy between ability and achievement using the K-ABC Mental Processing/Achievement comparisons and the WISC-R/PIAT comparisons across the four methods for calculating a discrepancy."

In order to test this hypothesis, Tables 22 through 25 were developed to illustrate the proportions of students who were identified as having a severe discrepancy, first

using a WISC-R/PIAT comparison and, then, the K-ABC MPC/Achievement counterpart (e.g., FSIQ/PIAT Mathematics with MPC/K-ABC Arithmetic). Chi square was then calculated using McNemar's test for correlated proportions (see Appendix F for the formula for calculating this estimate of chi-square) to determine if significant differences existed between the tabled proportions.

Table 23 shows the analysis for the PIAT Mathematics and K-ABC Arithmetic comparisons. Significant chi-squares were achieved for all methods, indicating that there are significant differences in the proportion of students selected. Closer examination reveals that the K-ABC Arithmetic identified from one and a half to seven times more students than the PIAT Mathematics across methods. The differences were highly significant ($p<.001$) for Methods 1 and 2. For comparisons involving Methods 3 and 4, the differences were significant at the .01 level.

Table 24 demonstrates the differences in proportions of students selected using the PIAT Reading Recognition and the K-ABC Reading Decoding subtests. Significant chi squares were achieved in all four analyses. Once again, from two to five times more students were identified with the K-ABC Achievement subtest than with the respective PIAT subtest.

The PIAT Reading Comprehension subtest is compared with the K-ABC Reading Understanding subtest in Table 25. Chi squares were significant for three of the four methods

Table 23

Numbers of Students Showing a Severe Discrepancy between Ability and Achievement Using WISC-R/PIAT Mathematics and K-ABC MPC/Arithmetic Comparisons

METHOD	Discrep	PIAT MATH	KABC ARITH	χ^2	df	sig
METH1	yes	7	40	27.67	1	$p<.001$
	no	79	46			
METH2	yes	9	35	19.53	1	$p<.001$
	no	77	51			
METH3	yes	2	15	8.47	1	$p<.01$
	no	84	71			
METH4	yes	27	46	8.75	1	$p<.01$
	no	59	40			

Table 24

Numbers of Students Showing a Severe Discrepancy between Ability and Achievement Using WISC-R/PIAT Reading Recognition and MPC/KABC Reading Decoding Comparisons

METHOD	Discrep	PIAT RREC	KABC RDEC	χ^2	df	sig
METH1	yes	24	46	12.97	1	p<.001
	no	62	40			
METH2	yes	21	46	16.46	1	p<.001
	no	65	40			
METH3	yes	4	21	12.19	1	p<.001
	no	82	65			
METH4	yes	20	45	17.45	1	P<.001
	no	66	41			

Table 25

Numbers of Students Showing a Severe Discrepancy between Ability and Achievement Using WISC-R/PIAT Reading Comprehension and MPC/KABC Reading Understanding Comparisons

METHOD	Discrep	PIAT RCOMP	KABC RUND	χ^2	df	sig
METH1	yes	12	44	30.12	1	p<.001
	no	59	33			
METH2	yes	12	44	30.12	1	p<.001
	no	59	33			
METH3	yes	5	22	13.47	1	p<.001
	no	66	55			
METH4	yes	44	44	.03	1	N.S.
	no	27	33			

for determining a discrepancy. The only non-significant chi-square was achieved in comparing the proportions of students identified using Method 4. Reconfirming the trend of the previous analyses, the PIAT subtest identified significantly fewer students than the K-ABC Achievement subtest for Methods 1 through 3 (less than 1/3 in each instance).

Table 26 shows the last in this series of analyses. This table illustrates the comparison of the PIAT Total with the K-ABC Achievement Composite in the identification of students. Chi square is significant for Methods 1 through 3, with two to nine times more students identified using the K-ABC Achievement score than using the PIAT Total score. The proportions of students identified using Method 4 were not significantly different across the two achievement subtests.

It appears, at least for Methods 1, 2 and 3, that the WISC-R/PIAT subtest comparisons identify significantly fewer students as showing a severe discrepancy than do the K-ABC MPC/Achievement comparisons. The difference appears to be attributable, in part, to the higher mean PIAT standard scores noted earlier in this section. With higher achievement standard scores it will be more difficult to establish a discrepancy between ability and achievement.

Of interest in these analyses is the finding that no significant differences between the proportions of students

Table 26

Numbers of Students Showing a Severe Discrepancy between Ability and Achievement Using WISC-R/PIAT Total and K-ABC MPC/Achievement Comparisons

METHOD	Discrep	PIAT TOTAL	KABC ACHIEV	χ^2	df	sig
METH1	yes	26	46	11.38	1	p<.001
	no	55	40			
METH2	yes	24	45	11.43	1	p<.001
	no	57	41			
METH3	yes	3	27	22.04	1	p<.001
	no	78	59			
METH4	yes	27	38	3.23	1	N.S.
	no	54	48			

identified with the PIAT and K-ABC were found in two of the four subtest comparisons using Method 4 (PIAT Reading Comprehension and PIAT Total). In the case of the PIAT Reading Comprehension comparisons, it appears that the adjustment for unreliability in Method 4 was substantial due to the low reliability of this subtest. As a result, more students were identified by this adjustment. The numbers identified were, thereby, closer to the numbers identified with the K-ABC Reading Understanding subtest. It is unclear why differences did not emerge for the total achievement comparisons.

The results of the testing of Hypothesis 1, therefore, appear to be divided. When considering Methods 1, 2, and 3 it can be expected that there will be significant differences in the proportion of students who are identified with the K-ABC and PIAT subtests: the PIAT will identify significantly fewer students. When using Method 4, however, no significant differences in proportions are anticipated between the K-ABC and PIAT in the areas of reading comprehension and in the achievement composite. In the area of reading recognition and mathematics, it might be expected that the PIAT will select significantly fewer students than the K-ABC.

2. The second hypothesis in Part II predicts: "There will be a significant correspondence between the following pairs of test scores across the four methods for

determining a discrepancy:

a. FSIQ/PIAT Mathematics with MPC/K-ABC Arithmetic

b. FSIQ/PIAT Reading Rec. with MPC/K-ABC Reading Dec.

c. FSIQ/PIAT Reading Comp. with MPC/K-ABC Reading Und.

d. FSIQ/PIAT Total with MPC/K-ABC Achievement

Percent overlap, kappa, and Spearman's rank order coefficient were calculated for each of the pairs of discrepancies across the four methods.

Table 29 shows the extent of agreement in decisions made using the FSIQ/PIAT Mathematics and MPC/K-ABC Arithmetic. Kappa is low and non-significant for each of the four comparisons. Rho is also low, accounting for less than 5 percent of the variance between the pairs of comparisons for each method used.

The comparison of FSIQ/PIAT Reading Recognition with MPC/K-ABC Reading Decoding is illustrated in Table 28. Percent overlap ranges from 8.7 percent for comparisons using Method 3 to 34.6 percent using Method 1. Again, kappa is generally low across methods, reaching significance at the .05 level for Methods 1, 2, and 4. Rho ranges from .35 for Methods 1 and 2 to .41 for Methods 3 and 4.

Table 29 shows the agreement between FSIQ/PIAT Reading Comprehension and MPC/K-ABC Reading Understanding comparisons. The percent overlap between the ability/achievement comparisons ranges from 17.4 for Method 3 to 49.2 for Method 4. Kappa is non-significant in each

Table 27

Agreement between WISC-R/PIAT Mathematics and MPC/K-ABC Arithmetic in Identifying Students

	Number with Discrepancy				
METHOD	PIAT MATH	KABC ARITH	%Overlap	Kappa	Rho
METHOD1	7	40	11.9	.086	.199
METHOD2	9	35	15.8	.128	.179
METHOD3	2	15	0.0	-.043	.253
METHOD4	27	46	32.7	.161	.253

* P<.05, ** P<.01, ***P<.001

Table 28

Agreement between WISC-R/PIAT Reading Recognition and MPC/K-ABC Reading Decoding in Identifying Students

	Number with Discrepancy				
METHOD	PIAT RREC	KABC RDEC	%Overlap	Kappa	Rho
METHOD1	24	46	34.6	.234*	.352
METHOD2	21	46	31.4	.214*	.353
METHOD3	4	21	8.7	.089	.414
METHOD4	20	45	32.7	.251*	.414

* P<.05, ** P<.01, ***P<.001

Table 29

Agreement between WISC-R/PIAT Reading Comprehension and MPC/K-ABC Reading Understanding in Identifying Students

	Number with Discrepancy				
METHOD	PIAT RCOMP	KABC RUND	%Overlap	Kappa	Rho
METHOD1	12	44	24.4	.173	.491
METHOD2	12	44	24.4	.173	.471
METHOD3	5	22	17.4	.205	.516
METHOD4	44	44	49.2	.103	.516

* P<.05, ** P<.01, ***P<.001

of the four comparisons. Rho ranges from .47 for Method 2 to .52 for Methods 3 and 4.

The comparison of FSIQ/PIAT Total and MPC/K-ABC Achievement appears in Table 30. The percent overlap ranges from 11.1 percent for Method 3 to 38.5 for comparisons involving Method 1. Rho indicates similar levels of agreement across methods with coefficients ranging from about .33 for Methods 1 and 2 to .38 for Methods 3 and 4.

The low degree of agreement between tests which is revealed in Tables 27 through 30 is apparent. For each comparison, the percent overlap never exceeds fifty percent. That is to say, less than half of the same students would be identified by both a WISC-R/PIAT comparison and a K-ABC/Achievement comparison, even when the same method for calculating a discrepancy is implemented. In half the comparisons, the overlap is less than 25 percent. Reflecting the same lack of concurrence between tests, kappa is consistently low, reaching significance in only four of the sixteen comparisons.

The lack of agreement between tests in identifying students with a severe discrepancy between ability and achievement suggests that the tests chosen will have a significant impact on any decision making process: the approaches to the determination of a discrepancy are not interchangeable. The WISC-R/PIAT comparisons consistently identify different students than the K-ABC MPC/Achievement subtest comparisons. The difference in selection can be

Table 30

Agreement between WISC-R/PIAT Total and MPC/K-ABC Achievement in Identifying Students

	Number with Discrepancy				
METHOD	PIAT TOT	KABC ACHIEV	%Overlap	Kappa	Rho
METHOD1	26	46	38.5	.246**	.332
METHOD2	24	45	32.7	.173	.331
METHOD3	3	27	11.1	.143	.379
METHOD4	27	38	35.4	.218	.379

* P<.05, ** P<.01, ***P<.001

attributed, in part, to the tendency of the PIAT to select fewer students than the K-ABC. However, even in the instances using Method 4 where no significant differences in proportions of students identified emerged (arithmetic and reading recognition) the PIAT and the K-ABC did not select the same students.

Therefore, with the exception of four comparisons where kappa reached significance (reading recognition: Methods 1, 2, and 4, total achievement: Method 1), Hypothesis 2 is rejected. It should be noted, however, that in the cases where kappa was significant the percent overlap was still under 50 percent.

Discussion

The preceding analyses have highlighted several important points which should be considered when evaluating children who have been referred for possible learning disabilities. Perhaps the most significant finding involves the differential functioning of the PIAT and the K-ABC achievement tests. When using the PIAT, the higher average standard scores in combination with lower test reliabilities make it difficult to obtain reliable differences between tests. The PIAT identified significantly fewer students than the K-ABC in 14 of 16 comparisons.

Method 4 adjusts Z_a for unreliability of the subtests, resulting in a less restrictive cutoff criterion.

The only two comparisons for which there were no significant differences involved comparing the two PIAT subtests which have the lowest reliabilities with their K-ABC counterparts using Method 4. In both cases, the application of the adjustment for unreliability resulted in a substantial modification of the regression cutoff line. In nearly every analysis, the WISC-R/PIAT comparisons identified significantly fewer students than the respective K-ABC MPC/Achievement comparisons. The graphs (Figures 3 through 10), which illustrate the bivariate distribution of students scores with cutoff lines representing the four methods for determining a discrepancy, reflect the impact of the higher PIAT standard scores on the number of students qualified using any of the four methods. If a single over-riding factor could be identified which affected the discrepancy analyses between tests, it would be the higher PIAT scores.

Also important in the comparisons has been the influence of subtest reliabilities. The lower subtest reliabilities for PIAT Mathematics and Reading Comprehension resulted in the diverging of selection cut-off lines in Figures 3 and 5.

The consideration of different average standard scores and subtest reliabilities does not account for all of the differential functioning of the two test batteries, however. For example, the total achievement scores for the PIAT and the K-ABC have the most comparable average standard scores (PIAT Total: 89.74; K-ABC Achievement Composite: 87.42) as well as high reliabilities (PIAT Total: r_{yy}=.89; K-ABC Achievement Composite: r_{yy}=.97); yet the two achievement measures selected a significantly different population of students (percent overlap: 11.1% to 38.5% across methods). It would appear that other factors (test content, test format, etc.) are contributing to the observed differences. Since such factors were not controlled in the present study, it is impossible to assess the impact of these variables.

When looking at how the various methods function in selecting students, it appears that the z-score and estimated true score methods operate very similarly. This is especially true with tests with high reliabilities. With tests with reliabilities above .85, the extra calculations required to compute the estimated true scores do not appear to be necessary. The straight z-score procedure would appear adequate in determining the differences. It has been argued that tests with low reliabilities should not be used in diagnostic situations. For this reason, it is advised that the PIAT Mathematics and Reading Comprehension subtests be avoided in the determination of a discrepancy for learning disabilities. Even when an adjustment for subtest unreliability is made using Reynolds' regression formula, errors of measurement are still present. While useful in adjusting the cutoff for comparisons involving tests with acceptable reliabilities (e.g. > .85), when extrapolated to poor quality tests, the results are difficult to interpret.

The two regression approaches to the calculation of a discrepancy are perfectly correlated but result in the selection of different populations of students. Method 4 corrects for lack of subtest reliability, making it much easier to establish a severe discrepancy. The less reliable the subtest, the larger the difference between the numbers of students identified with Method 3 and Method 4. The most extreme cases were illustrated with the PIAT Mathematics subtest (r_{xx} = .74) and the PIAT Reading Understanding subtest (r_{xx} = .64). For the comparisons involving the PIAT Mathematics, Method 3 identified 2 students, while Method 4 identified 27 students. For the PIAT Reading Comprehension, Method 3 identified 5 students, while Method 4 identified 44 students.

It would appear that the correction for unreliability, which is included in Reynold's regression analysis procedure, is an appropriate adjustment. When this adjustment is not made, it is very difficult to obtain a significant score difference, even at the .05 level.

In looking at Figures 3 through 10, it becomes apparent that Methods 1 and 2 are generally functioning quite differently than Methods 3 and 4. One then might ask, which set of procedures is preferred? The correlations of discrepancies obtained by each method with teacher ratings of the severity of a child's problem in the academic area of concern revealed that the regression procedures correlated somewhat better with the ratings than the z-score difference or the estimated true score difference. A regression procedure, therefore may be a more valid approach to the determination of a discrepancy. The question of which intercept to use with the regression cutoff line may be a point of debate, however.

Figure 11 illustrates how several approaches to regression analysis compare. Several school districts have recommended the use of a straight regressed score difference (e.g. a 1.0 or 1.5 z-score difference) which takes the form:

$$z_x (r_{xy}) - z_y > 1.0 \text{ (or } 1.5)$$

These lines are plotted along with the cutoff lines for Reynolds' regression analysis (Method 4) to demonstrate the impact of selecting one procedure over another. It is clear that the choice of an alpha level (or more simply, the magnitude of the regressed score-observed score difference) will have an influence on the size of the population which will qualify as learning disabled. It would seem appropriate for an individual school district to determine the means and standard deviations of the ability and achievement scores of the general population of students, determine the percentage of that population which can effectively be served in programs for learning disabled, and, then set a cutoff which corresponds to that

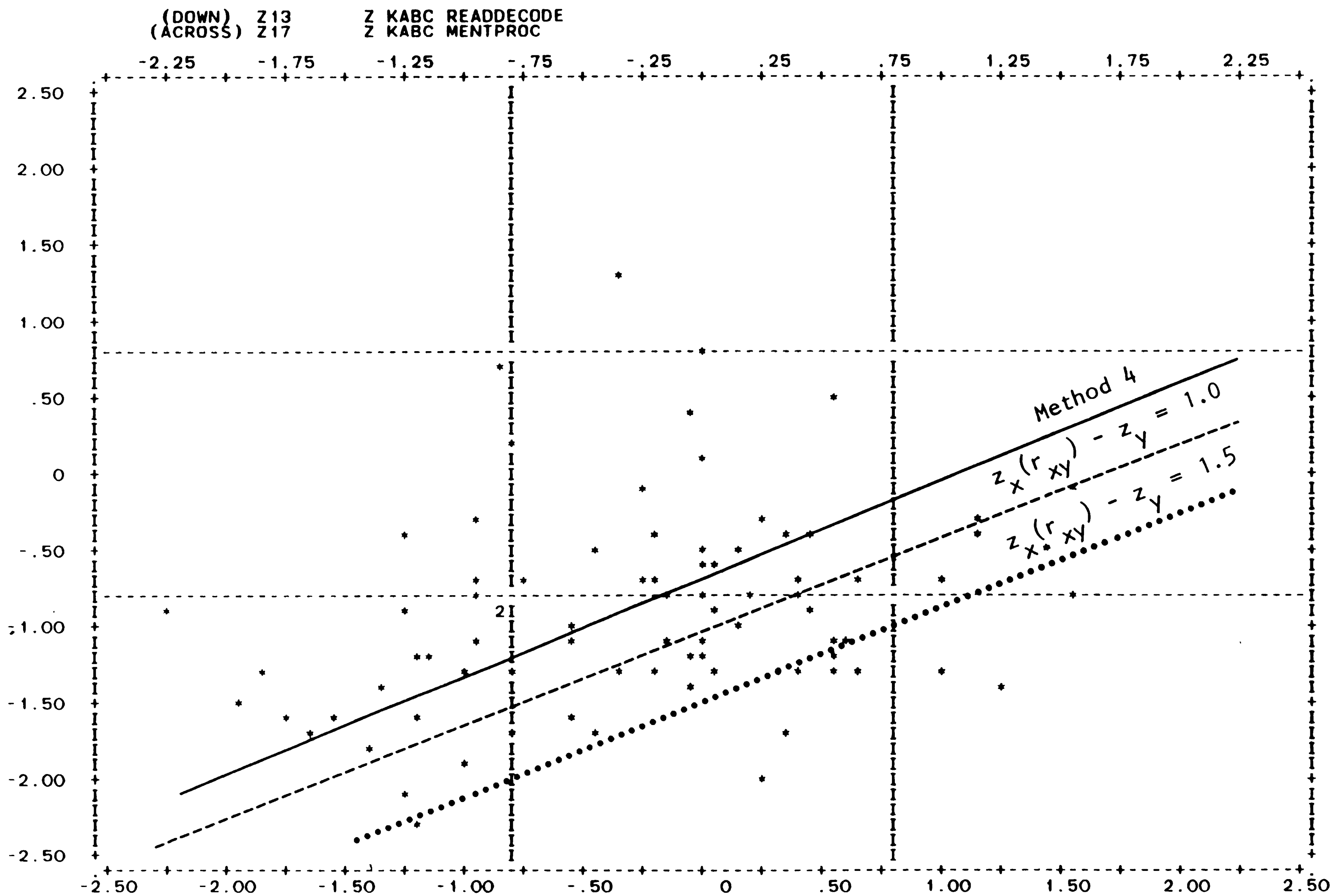


Figure 11: A comparison of three cutoffs for the regression analysis approach to determining a severe discrepancy

number of students. The reader is referred to Reynolds (1984-85) for recommendations and considerations in establishing a cutoff. Difficulties which might be encountered when attempting to set a cutoff include the effects of multiple test score comparisons, the consideration of exclusionary practices (e.g. the discrepancy cannot be primarily a result of cultural or environmental factors), the possible low referral rate of students with certain characteristics (e.g. high ability), and changes in the population characteristics from year to year.

Once a cutoff line has been established, however, it would be relatively easy for school psychologists and other diagnosticians to plot the standard scores for a given student on a graph, such as appears in Figure 12. If the point corresponding to the ability and achievement scores falls below the cutoff line, a severe discrepancy would be established. Several examples are included in Figure 12.

The results of the comparisons of the PIAT and K-ABC Achievement subtests illustrate the difficulty in utilizing tests with different average standard scores. Even if a predetermined cutoff is established, variations such as the much higher PIAT standard scores will significantly influence the numbers of students who exhibit a "severe discrepancy." While it is not always feasible to determine the distributions of every subtest

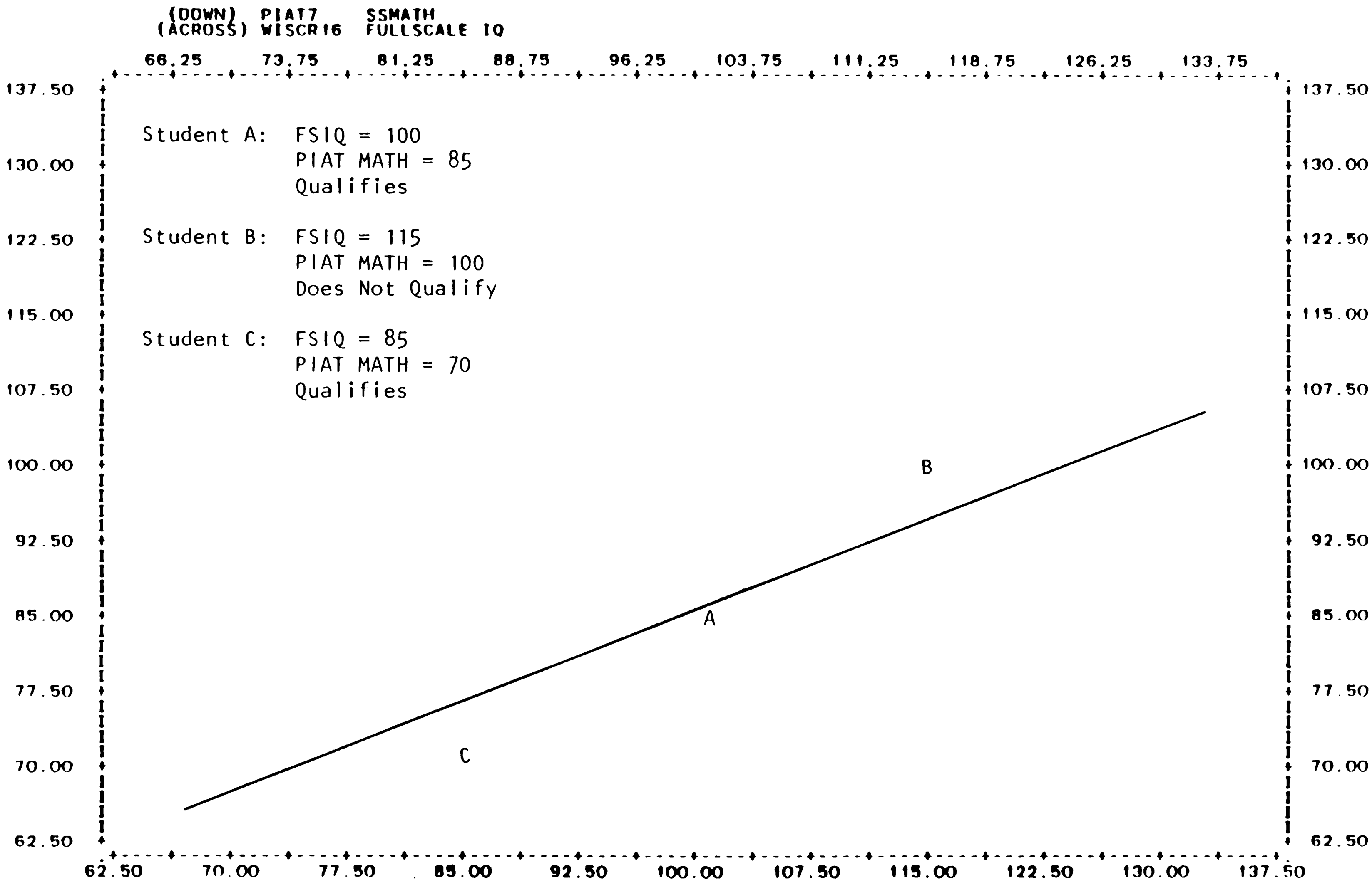


Figure 12: Utilization of a regression analysis cutoff line: an example

which is currently being used, it would seem advisable to avoid tests, such as the PIAT, which have not been re-normed in many years and which have notoriously low reliabilities for several subtests.

The verdict is still not in on whether the K-ABC MPC comparisons are preferred to the WISC-R IQ comparisons in the diagnosis of learning disabilities. The flat profile of K-ABC Mental Processing scores suggests that, for this sample of students, there is not a great deal of variability in subtest performance. Whether this finding will hold for the referred students who are actually diagnosed as learning disabled remains to be seen (Becht, 1986, is currently investigating this question).

The high reliabilities of the K-ABC subtests, in addition to strong evidence of the construct validity of the battery, makes the use of this test quite appealing. Keith's (1985) question regarding the K-ABC, "What does it measure?" must still be addressed, however. If, in fact, the dimensions which are measured by the K-ABC are better characterized along a verbal-nonverbal continuum, it could argued that the WISC-R measures these abilities and measures them better (particularly since the Kaufmans have attempted to minimize the verbal component of the K-ABC).

From a practical perspective, one can seriously question the utility of the Sequential Processing Scale in estimating a student's ability. The observation that all of these tasks measure short-term memory, as well as the

finding that the Sequential subtests load on the "distractibility" factor in joint factor analysis with the WISC-R, suggests that this dimension may have diagnostic utility only in the assessment of attention and immediate recall rather than any higher level, integrative abilities.

Thus, while the K-ABC MPC/Achievement comparisons appear to be preferred over the WISC-R/PIAT comparisons in the determination of a reliable difference between scores, it has not been determined that the K-ABC would be preferred to the WISC-R in the diagnosis of students with learning difficulties. The PIAT has been shown to be the weaker link in the WISC-R/Achievement comparisons. It is quite possible that a more reliable and recently normed achievement battery used with the WISC-R would provide both a technically sound <u>and</u> meaningful strategy for the diagnosis of learning disabilities.

VI. SUMMARY AND RECOMMENDATIONS

This dissertation has compared four methods for calculating a severe discrepancy between ability and achievement when diagnosing a student for a learning disabilities classification. Also of interest has been the comparison of score discrepancies obtained using the WISC-R Full Scale IQ and PIAT subtests with the score discrepancies obtained using the K-ABC Mental Processing Composite and K-ABC Achievement subtests.

In order to perform these comparisons, 86 students between the ages of 6 and 12 1/2 who were referred for academic difficulties were tested with the WISC-R, PIAT, and K-ABC. The score differences between the WISC-R Full Scale IQ and the PIAT subtests (Mathematics, Reading Recognition, Reading Comprehension, and Total) were contrasted with the differences obtained between the K-ABC Mental Processing Composite and the K-ABC Achievement subtests (Arithmetic, Reading/Decoding, Reading/Understanding, and Achievement Composite). The differences were evaluated using four standard score procedures for calculating a severe discrepancy: z-score difference (Method 1), estimated true score difference (Method 2), and two regression procedures. The first regression procedure

(Method 3) considered only errors of estimate, while the second regression procedure (Method 4) also included an adjustment for test unreliability.

A high degree of agreement in the selection of students was found between the z-score difference and estimated true score difference approaches to determining a discrepancy. The agreement was especially apparent for tests with high reliabilities (K-ABC Achievement subtests and PIAT Reading Recognition and Total). Considerable agreement was also found between the z-score difference, estimated true score difference, and the adjusted regression analysis procedure for K-ABC subtests. Less agreement was obtained when comparing the unadjusted regression procedure with the other three methods, primarily because of the limited number of students identified with the unadjusted regression formula.

When comparing the WISC-R/PIAT discrepancies with the K-ABC MPC/Achievement discrepancies, it was found that the two approaches to the determination of a score difference resulted in different populations of students. Percent overlap ranged from 0 to 49.2 percent with 8 out of 16 comparisons resulting in less than 25 percent of the same students identified. The high average PIAT standard scores in combination with lower subtest reliabilities appeared to be a primary source of disparity between the K-ABC and PIAT comparisons. The PIAT identified fewer students than the K-ABC in all comparisons and selected less than half the number of students in ten of the sixteen comparisons.

In an attempt to determine the adequacy of approaches to determining a discrepancy between ability and achievement, teacher ratings of the severity of a child's academic deficit were correlated with the magnitude of the discrepancy obtained using the different tests and methods. Higher correlations were achieved between the teacher ratings and the K-ABC subtests, particularly in the area of reading. Slightly higher correlations were obtained for the regression analysis methods than for the z-score and estimated true score approaches in the areas of reading and arithmetic.

Analysis of variance and chi square analyses were performed to determine if there were significant differences between student characteristics (sex, age, grade placement, and ability) for students who were identified by the four methods. No significant differences were found, suggesting that the four methods would select students with similar characteristics.

Based upon the literature review and data analysis, it was concluded that a regression analysis approach to the determination of a discrepancy would be preferred to either the straight z-score or estimated true score methods. Several lines representing varying cutoffs were graphed, illustrating the impact of the choice of an alpha level. It was recommended that the distribution of ability and achievement scores of the students in a given school district, as well as the size of the population which can be served in learning disabilities programs, be considered when setting a cutoff (intercept) for the regression analysis line. While the Reynolds adjusted regression analysis procedure appears to be the more technically correct approach to the calculation of a discrepancy, the establishment of an alpha level will be based as much on pragmatic considerations as on statistical significance. For this reason, it is recommended that the ability and achievement test intercorrelation (r_{xy}) be used to set the slope of a regression analysis cutoff line (such as appears in the figures in the Section IV). The intercept of the cutoff line can be estimated by considering the numbers of students who can be served by support programs in a given school district. It may be necessary to adjust this intercept to correspond with variations in referral and diagnostic practices, as well as general population characteristics.

The analysis of discrepancies obtained when using the PIAT reflected the difficulties which would be encountered if this achievement battery were used in diagnosing learning disabilities. High average subtest scores and lower subtest reliabilities made it difficult to obtain reliable differences between scores. When using any approach to the determination of a severe discrepancy, it was recommended that only recently normed tests with high reliabilities be used.

together with more than rating source, may provide a better estimate of the adequacy of each discrepancy procedure.

Also of interest is the extent to which different methods for determining a discrepancy would serve to displace students who are currently served under the Learning Disability classification. Which procedures are the most consistent with current practices? Although diagnosis as "learning disabled" was purposely avoided in the present study, the extent to which implementation of an alternative discrepancy formula will change a population of classified students is a critical consideration when modifying evaluation practices.

It has been recommended that the actual probabilities associated with each discrepancy formula be determined using population estimates of test means, variances, and intercorrelations. This is a valuable endeavor which was not undertaken in the current study. Of interest is the comparison of these probabilities with obtained proportions of students who are referred and shown to exhibit discrepancies. One might expect, for example, that high ability students who are functioning near the mean on achievement tests (and, thus exhibit a severe discrepancy between ability and achievement) would be under-represented in the population of students who are referred for learning disabilities.

Also of interest is the comparison of the proportion of students who would be expected to have a significant discrepancy with the proportion of students who are actually classified as learning disabled. It must also be shown that the student's learning difficulties are not a result of visual, hearing, ot motor handicaps, of mental retardation, or of environmental, cultural, or economic disadvantage. Because of exclusionary considerations, it might be expected that more students would exhibit a severe discrepancy than would ultimately be classified as learning disabled.

Another consideration which has been alluded to in previous discussion has been the impact of multiple comparisons on the probability of obtaining a significant difference between scores. Reynolds (1984-85) has discussed this issue, recommending adjusting z_a when more than one comparison is being made. He has suggested delineating "the precise method of accounting for multiple comparisons in a way that can be easily applied" (p. 469). This would appear to be an important direction for future study, as well.

Replication of this study should involve a random sample of referred students, equally representing age or grade levels. A greater representation of minority students would also allow the determination of differential selection of students by race/ethnicity. It is acknowledged that the sample of students used in the present study was not randomly selected from a well-defined population. The generalizability of the results is limited by this constraint.

APPENDICES

APPENDIX A

DESCRIPTION OF TESTS: WISC-R
PIAT

Wechsler Intelligence Scale for Children-Revised

The Wechsler Intelligence Scale for Children-Revised by David Wechsler (1974) is a norm-referenced test of general intellectual performance. It assesses intellectual functioning by sampling performance on several different types of test tasks, including verbal comprehension, arithmetic reasoning, and visual-motor tasks. Results may be used to gain an overall estimate of present intellectual performance and to gather information about possible strengths and weaknesses in various aptitude areas.

Characteristics of the WISC-R

The WISC-R is an individual test that requires 50 to 75 minutes for administration. Norms are available for students aged 6-0 to 16-11. The test kit resembles a briefcase and contains the manual, record forms, and needed testing materials. The tester must furnish only a stopwatch and writing implements.

Ten required and two supplementary subtests are included on the WISC-R, and subtests are categorized as Verbal or Performance.

Verbal	Performance
Information	Picture Completion
Similarities	Picture Arrangement
Arithmetic	Block Design
Vocabulary	Object Assembly
Comprehension	Coding
(Digit Span)	(Mazes)

During test administration, Verbal and Performance subtests are alternated, and all students take all the required subtests.

On the Verbal tests, students listen to questions and reply orally. Each of the Verbal subtests is described briefly below. To preserve confidentiality of WISC-R test items, the examples provided here are similar but not identical to actual test items.

- *Information*—The student responds orally to general information questions, such as "How many eyes do you have?" and "From what animal do we get hamburger?"
- *Similarities*—The student is asked to describe how two items, such as a pony and a cow or a car and an airplane, are alike.
- *Arithmetic*—The student listens to arithmetic problems and attempts to solve the problems without pencil and paper.
- *Vocabulary*—The student is asked to tell the meaning of several words. The tester asks, "What is a ________?" or "What does ______ mean?"
- *Comprehension*—The student answers questions that require social reasoning, such as "What are some reasons why we need firemen?"
- *Digit Span* (Supplementary)—The tester reads a series of numbers to the student at the rate of one digit per second. On the first portion of the test, the student attempts to repeat the numbers exactly as the tester presented them. On the second portion of the test, the student attempts to repeat the number series backward.

On the Performance tests, students listen to oral directions and then perform some visual-motor task. Following is a brief description of each of the Performance subtests. For each item of these subtests, a time limit is given.

- *Picture Completion*—The student is shown a drawing of an object or scene with an important part missing. The student either points to or tells what is missing.
- *Picture Arrangement*—The student is shown several cards with pictures on them. The student arranges the pictures to tell a story.
- *Block Design*—The student is given several colored cubes or blocks. The student is then shown a picture of a design and must make an identical design using the blocks.
- *Object Assembly*—The student is given several puzzle pieces which must be assembled into the correct shape.
- *Coding*—The student is given a worksheet that presents a code, such as specific geometric shapes for each of the digits. The student

From: McLoughlin, J.A., & Lewis, R.B. (1981). Assessing special students. Columbus: Charles E. Merrill, pp. 226-230, 200-205.

is then presented with several lines of digits and must write the correct geometric shape for each digit in order.

- *Mazes* (Supplementary)—The student is presented with several mazes. Using a pencil, the student must solve each maze.

Appropriate student population

The WISC-R was standardized on a large sample (approximately 2,200) selected to resemble the United States as described in the 1970 Census in terms of race (white and nonwhite), geographic region, occupation of head of household, and urban-rural residence. The standardization sample appears to be representative of the United States population. However, in the case of race, while the percentage of nonwhite individuals in the standardization sample is the same as in the general population (15%), only 330 nonwhite individuals were included, 305 classified as "black" and 25 to represent the groups of American Indian, Oriental, Puerto Rican, and Chicano.

Students in the sample ranged in age from 6½ to 16½. Equal numbers of male and female students were included at each age level. Only "normal" students were selected. The manual states that institutionalized mentally retarded children, children with severe emotional problems, and bilingual children who could not speak and understand English were excluded from the sample. Thus, the WISC-R appears most appropriate for elementary and secondary students from ages 6-0 to 16-11 who are white, speak English, are not in institutions for the mentally retarded, and do not have severe emotional problems.

Skill requirements for students

The WISC-R is appropriate for students who are able to sustain attention for extended periods of time. Facility in the English language is necessary since students must comprehend instructions and questions and answer questions in English. Both spoken and motor responses are required; students must point, manipulate small objects, and write with a pencil. No reading skills are needed, but students must answer simple arithmetic questions.

Skill requirements for tester

According to the manual, the WISC-R "should be administered and scored by a competent, trained examiner" (Wechsler, 1974, p. 53). In most states, individual intelligence tests such as the WISC-R may be administered in school settings only by licensed school psychologists. However, even if special educators are generally not responsible for giving and scoring this test, they should be knowledgeable about it so that they can analyze test results.

Results

The WISC-R yields global IQ scores and scaled scores for each separate subtest. Three IQ scores are obtained—Verbal IQ, Performance IQ, and Full Scale IQ. The Verbal IQ is a standard score that indicates performance on the five Verbal tests, the Performance IQ is a standard score that indicates performance on the five Performance tests, and the Full Scale IQ is a standard score that indicates performance on all 10 of the required subtests.

Each of these global standard scores is distributed normally with a mean of 100 and a standard deviation of 15. Thus, approximately 95% of the population falls within the range of IQ 70 to IQ 130; that is, from two standard deviations below the mean to two standard deviations above the mean. Significantly subaverage performance on the WISC-R, according to the AAMD definition of mental retardation, would be indicated by a score of IQ 69 and below. The average standard errors of measurement for the Verbal IQ, Performance IQ, and Full Scale IQ (across all age levels) are 3.60, 4.66, and 3.19 IQ points respectively. In comparing a student's performance on the Verbal and Performance scales, Wechsler (1974) recommends that the difference be at least 15 IQ points to consider it important enough to warrant further investigation.

Standard scores, called *Scaled Scores*, are obtained for each subtest. Scaled Scores are distributed normally with a mean of 10 and a standard deviation of 3. Thus, approximately 95% of the population falls within a range of SS 4 to SS 16; that is, from two standard deviations below the mean to two standard deviations above the mean. The standard error of measurement for Scaled Scores, averaged across age levels, ranges from 1.15 to 1.70 Scaled Score points depending on the specific subtest. In comparing a student's performance on two subtests, Wechsler (1974) recommends that the difference be at least 3 Scaled Score points to be considered significant. Subtest Scaled Scores are plotted on a profile which appears on the front of the record form. It is also possible to calculate Test Age scores for each of the WISC-R subtests.

Quality

The quality of a test is judged both by its reliability and its validity. The suggested minimum criterion for adequate reliability and validity correlation coefficients is .80. Reliability of the WISC-R was studied via the split-half procedure, except with the Coding and Digit Span subtests (for which the test-retest method was used). Verbal, Performance, and Full Scale IQs have average reliabilities of .94, .90, and .96 respectively across all ages. Average reliability coefficients range from .77 to .86 for Verbal subtests and from .70 to .85 for Performance subtests.

Validity of the WISC-R was studied by correlation with other measures of intellectual performance. For example, WISC-R results were compared with results from the Stanford-Binet (1972 norms). Average correlations of the WISC-R Verbal, Performance, and Full Scale IQs with the Stanford-Binet IQ were .71, .60, and .73 respectively. Reliability and validity of the WISC-R appear to approach the criteria for adequacy. However, Anastasi (1976) points to the need for validation of the WISC-R with measures of factors such as school performance rather than solely with other measures of intellectual performance.

The WISC-R and assessment questions

The WISC-R is one of the individual tests of intellectual performance most often used in educational decision making. Results from this test, along with information from other sources, are used to answer the assessment question "What is the level of intellectual performance?" The WISC-R is well-suited to answer this question, since it yields data on current student functioning in several aptitude areas as well as an overall summary.

In making decisions about the existence of intellectual performance problems, the team must set criteria for below average performance. According to the American Association on Mental Deficiency's definition of *mental retardation* (Grossman, 1973), significantly subaverage general intellectual functioning "refers to performance which is more than two standard deviations from the mean" (p. 11). On the WISC-R then, two types of standard scores are available, IQ scores and scaled scores; and these scores may be classified using the following system.

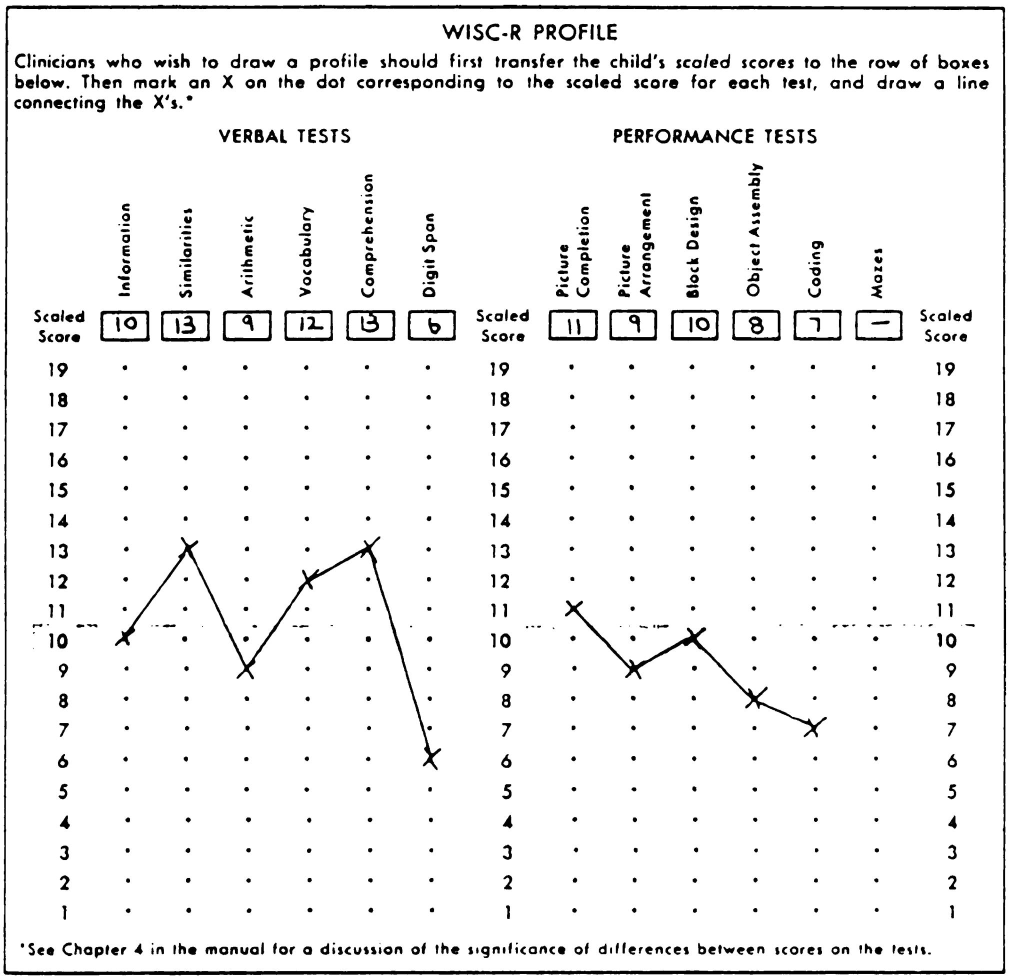
WISC-R IQ Range	WISC-R Scaled Score Range
(*a*) IQ 69 and below	SS 3 and below
(*b*) IQ 70 to IQ 84	SS 4 to SS 6
(*c*) IQ 85 to 115	SS 7 to SS 13
(*d*) IQ 116 to IQ 130	SS 14 to SS 16
(*e*) IQ 131 and above	SS 17 and above

Classification

(*a*) Below Average performance
(*b*) Low Average performance
(*c*) Average performance
(*d*) High Average Performance
(*e*) Above Average Performance

In analyzing WISC-R results, the team can use this classification system. The Full Scale IQ provides information about the overall level of intellectual performance. Likewise, the Verbal and Performance IQs allow the team to determine levels of performance in each of these areas. In addition, subtest scaled scores may be analyzed to determine possible areas of strengths and weaknesses. Standard errors of measurement must, of course, be considered. Wechsler presents specific guidelines for interpreting significant differences between Verbal and Performance IQs (minimum of 15 IQ points) and between subtest scaled scores (minimum of 3 scaled score points).

Figure 9-1 presents WISC-R results for a



	Raw Score	Scaled Score
VERBAL TESTS		
Information	13	10
Similarities	16	13
Arithmetic	11	9
Vocabulary	32	12
Comprehension	22	13
(Digit Span)	(8)	(6)
Verbal Score		67
PERFORMANCE TESTS		
Picture Completion	19	11
Picture Arrangement	21	9
Block Design	25	10
Object Assembly	16	8
Coding	32	7
(Mazes)	(—)	(—)
Performance Score		45

	Scaled Score	IQ
Verbal Score	67 •	108
Performance Score	45 •	92
Full Scale Score	102	101

*Prorated from 4 tests, if necessary

Figure 9-1 Sample WISC-R Profile

Reproduced from the Wechsler Intelligence Scale for Children—Revised, by D. Wechsler. Copyright 1971, 1974 by The Psychological Corporation. All rights reserved. Reproduced by special permission.

student aged 10-0. Scaled scores are plotted on the profile. The following statements may be made about this profile.

- **Overall Test Performance**
  - **Results:** The student obtained a Verbal IQ in the range 104 to 112, a Performance IQ in the range 87 to 97, and a Full Scale IQ in the range 98 to 104. The chances that these ranges of scores include the true scores of the student are about 68 out of 100.
  - **Comments:** These results indicate that the student is presently functioning within

the average range in Verbal IQ, Performance IQ, and Full Scale IQ.

- **Strengths and Weaknesses**
  - **Results:** On the verbal tests, the student scored within one standard deviation of the mean on all tests except Digit Span. On Digit Span, the student scored between one and two standard deviations below the mean. On the Performance Tests, the student scored within one standard deviation of the mean on all tests.
  - **Comments:** These results indicate that the student is presently performing in the average range on all WISC-R subtests except Digit Span. On the Digit Span subtest, the student's present performance is in the low average range. This may indicate a possible weakness in short-term auditory memory for digits.

Besides analysis of WISC-R results for current levels of intellectual performance, many experts advocate analysis of performance on various subtests to locate patterns of strengths and weaknesses (Lerner, 1976). For example, Bannatyne (1974) suggests grouping subtest results into several categories and analyzing student performance in each of these categories.

- *Spatial Ability*—Picture Completion, Block Design, Object Assembly
- *Verbal Conceptualization Ability*—Comprehension, Similarities, Vocabulary
- *Sequencing Ability*—Digit Span, Arithmetic, Coding
- *Acquired Knowledge*—Information, Arithmetic, Vocabulary

Anastasi (1976), in reviewing factor analytic studies of the WISC, notes that four categories or factors generally emerge: general, verbal comprehension, perceptual-spatial, and memory (or freedom from distractibility). However, Wechsler (1974) warns that results of research are mixed in their support of subtest analysis.

The WISC-R, then, is an individual norm-referenced standardized test of intellectual performance. It may be used with both elementary and secondary students, and it provides information about current overall performance in intellectual functioning. Results from this test also allow analysis of performance in Verbal and Performance areas, as well as in the many areas tapped by the WISC-R subtests.

Peabody Individual Achievement Test

The Peabody Individual Achievement Test or PIAT (Dunn & Markwardt, 1970) is a norm-referenced standardized test of several areas of academic achievement. According to the manual, the PIAT provides "a wide range, screening measure of achievement in the areas of mathematics, reading, spelling, and general information" (p. 1). The test authors suggest that PIAT results be used as an overview of present academic achievement and as a guide to further assessment.

Characteristics of the PIAT

The PIAT requires approximately 30 to 40 minutes to administer and score. Norms are available for students in Kindergarten to grade 12 and for ages 5-3 to 18-3. Test equipment includes two test notebooks, the manual, and Individual Record Booklets. The test note-

books are in an easel format and contain the test items and directions for administration. The notebooks stand by themselves; while the student sees the test item, the tester sees instructions for administration. No other materials are required except writing implements for the tester.

The PIAT includes five subtests: Mathematics, Reading Recognition, Reading Comprehension, Spelling, and General Information. Most subtests use a multiple-choice response format. The student is shown a test plate with four possible responses and must select the correct answer. Each subtest (except Reading Comprehension) contains 84 items, arranged in order of difficulty. Following is a brief description of each of the subtests.

- *Mathematics Subtest*—The student listens to a question while viewing four possible responses. The student responds by pointing to the answer, saying the answer, or saying the number of the response. Test items include many types of problems: matching numerals; word problems requiring addition, subtraction, multiplication, and division; time, money, and measurement; quantitative vocabulary and use of mathematics symbols; algebra, geometry, and trigonometry; and so forth.
- *Reading Recognition Subtest*—The first few subtest items are in multiple-choice format and require matching of objects, letters, and words. Beginning with item 10, the student must look at a stimulus and respond orally, either by naming letters or by reading isolated words.
- *Reading Comprehension Subtests*—Test items require the student to read a sentence silently; then, from a set of four pictures, select the one that best depicts the sentence. The student may not read the sentence more than once and may not look at the sentence once the pictures have been exposed.
- *Spelling Subtest*—Multiple-choice items begin with exercises requiring the student to locate the one stimulus in four that is different from the rest. Next, the student must select the correct letter from four choices, given the letter name and sound; or the student must locate the correct word of four after hearing the word and its initial sound. Beginning with item 15, after listening to the tester read a word and a sentence containing the word, the student must select the correct spelling of that word from four choices.
- *General Information Subtest*—Questions are read to the student by the tester. No visual stimuli are presented, and the student must respond orally to each question. The test items sample the student's knowledge in many areas: science, social studies, art, music, sports, and so forth.

Appropriate student population

The PIAT was standardized in 1969 on a group of 2,889 students enrolled in the " 'mainstream of education'; namely attending regular classrooms of the public day schools operated by local school systems" (p. 26). No students enrolled in special classes were included in the standardization sample. The manual does not state whether all students in the sample were English-speaking.

The standardization sample included students in grades Kindergarten through 12 (ages 4 to 21) with an approximately equal mixture of males and females. The sample was selected to resemble the United States (as described by 1967 U. S. Bureau of the Census data) in geographical distribution, type of community (urban, suburban, and rural), parental occupation, and race. A majority of the PIAT sample (84.4%) was classified as "white," 11.3% as "Negro," and 4.3% or 123 cases as "other" (Mexican American, Puerto Rican, American Indian, and Oriental). The final sample differed from the United States population on several variables (race, geographical distribution, parental occupation); however, "it was considered more important to retain all usable cases than fit the census figures exactly" (p. 31).

The PIAT is most appropriate for students in grades Kindergarten through 12 who are enrolled in regular classes. But this does not mean that this test is not appropriate for students receiving special education services. The PIAT was standardized more than a decade ago, and it is difficult to equate the regular standard deviation of 15. Thus, approximately 95% of the population falls within the range of standard score 70 to standard score 130 (that is, from two standard deviations below the mean to two standard deviations above the mean).

Three types of "grade" scores may be calculated from subtest and total test raw scores: grade equivalents, percentile ranks for the student's grade placement, and normalized standard scores. Again, standard scores are distributed with a mean of 100 and a standard deviation of 15.

The standard error of measurement is reported in the PIAT manual for *raw scores* only. The median standard errors of measurement, across ages, for the five subtests are 5, 3, 6, 5, and 4 raw score points respectively. The mean standard error of measurement for the total test across ages is 12 raw score points. In comparing a student's performance on two subtests, the manual recommends a difference of at least 8 raw score points in Kindergarten and grade 1, at least 13 raw score points in grades 2 and 3, and at least 15 raw score points in grades 4 to 12 for interpretation of a "true" difference with a 95% level of confidence.

PIAT results may be plotted on a profile provided on the front of the protocol. If age norms are used, then either age equivalents or percentile ranks by age may be plotted. If grade norms are used, then either grade equivalents or percentile ranks by grade may be plotted. The manual recommends comparison of a student's PIAT performance with either Mental Age or IQ. However, the manual offers no suggestions for what may be considered a significant discrepancy between intellectual functioning and a student's achievement on the PIAT.

Quality

The quality of a test is judged both by its reliability and its validity. The suggested minimum criterion for adequate reliability and validity correlation coefficients is .80. Reliability of the PIAT was studied by the test-retest method. Reliability coefficients for all subtests for all ages ranged from .42 to .94. Total test reliability ranged from .82 (at Kindergarten) to .92 (at grade 12). Across ages, reliability was lowest for Reading Comprehension (.64) and Spelling (.65). Thus, reliability of the PIAT appears better for older students and better for some subtests than others.

Concurrent validity of the PIAT was studied by correlation with the Peabody Picture Vocabulary Test or PPVT (Dunn, 1959), which the PIAT manual calls "a measure of scholastic aptitude" (p. 50). The PIAT General Information Subtest correlated best with PPVT results (.68). The median correlation for all ages and all subtests was .57.

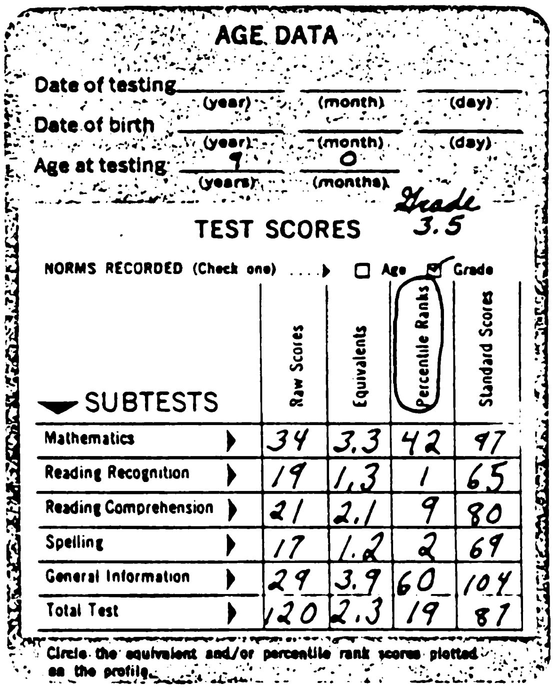
The manual reports only one study which investigated the relationship of the PIAT to other measures of academic achievement. This research (Sitlington, 1970) found substantial correlation between results of the WRAT and the PIAT, especially on reading and spelling subtests. However, the sample size was relatively small (n = 46), and the subjects represented a special population (mildly retarded adolescents).

Validity of the PIAT is not well-substantiated by the data presented in the manual. No extensive study of the relationship of the PIAT to other measures of academic performance is cited. Further investigation is needed regarding the validity of this test, the reliability of some of its subtests, and its reliability with young students.

The PIAT and assessment questions

The PIAT is one of the individual achievement tests most frequently used in educational decision making. Results from this test, along with information from other sources, are used to answer the key assessment question: "**Is there a school performance problem?**" The PIAT is well suited to provide answers to this question since it yields data on current student performance in several academic areas plus an overall summary score.

In making decisions about the existence of school performance problems, the team must set criteria for below average school performance. PIAT results include many types of scores, such as age and grade equivalents, percentile ranks, and normalized standard scores. Age and grade scores are not appropriate for this purpose. Either percentile ranks or standard scores, either for grade or age, should be



AGE DATA

Date of testing ____ (year) ____ (month) ____ (day)

Date of birth ____ (year) ____ (month) ____ (day)

Age at testing 9 (years) 0 (months)

Grade 3.5

TEST SCORES

NORMS RECORDED (Check one) .... ☐ Age ☑ Grade

SUBTESTS	Raw Scores	Equivalents	Percentile Ranks	Standard Scores
Mathematics	34	3.3	42	97
Reading Recognition	19	1.3	1	65
Reading Comprehension	21	2.1	9	80
Spelling	17	1.2	2	69
General Information	29	3.9	60	104
Total Test	120	2.3	19	87

Circle the equivalent and/or percentile rank scores plotted on the profile.

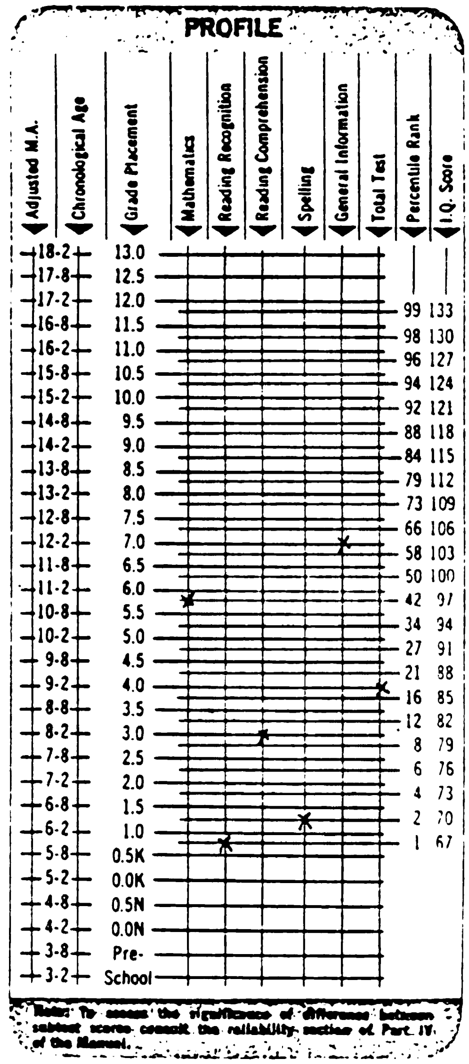


Figure 8-1 Sample PIAT Profile

From Peabody Individual Achievement Test by L. M. Dunn and F. C. Markwardt. Circle Pines, Minn.: American Guidance Service, 1970. Copyright © 1970 by American Guidance Service, Inc. Reprinted by permission.

used to determine below average performance. Chapter 6 of this book presents a classification system for standard score ranges. For the PIAT, this would be

PIAT Standard Score Range	Classification
69 and below	Below average performance
70 to SS 84	Low average performance
85 to SS 115	Average performance
116 to SS 130	High average performance
131 and above	Above average performance

In analyzing PIAT results, the team can use this classification system to answer the assessment question: "**What is the level of academic achievement and strengths and weaknesses in school learning?**" The total test standard score presents information about the overall level of academic achievement. Comparison of subtest scores permits the team to identify strengths and weaknesses in school performance. Of course, it is necessary in interpreting results of this test to consider both the standard error of measurement and the standard error of measurement of differences.

Figure 8-1 presents a PIAT protocol for a student aged 9-0 in grade 3.5. Percentile ranks by grade are plotted on the profile. The following statements may be made about this profile.

- **Overall Test Performance**
  - **Results:** The student obtained a total test standard score in the range 79 to 90.* The chances that this range of scores includes the true score of the student are about 68 out of 100.
  - **Comments:** These results indicate that the

* The standard error of measurement for Total Test is 12 raw score points. The Total Test raw score range is thus 108 to 132, which is equivalent to the standard score range 79 to 90.

student is presently functioning in the low average to average range in academic achievement.

- **Strengths and Weaknesses**
  - **Results:** The student scored within one standard deviation of the mean on the Mathematics and General Information subtests. The student scored between one and two standard deviations below the mean on the Reading Comprehension subtest. The student scored greater than two standard deviations below the mean on the Reading Recognition and Spelling subtests.
  - **Comments:** These results indicate that the student is presently functioning in the average range on measures of Mathematics and General Information. The student performed in the low average range in Reading Comprehension and in the below average range in Reading Recognition and Spelling. This student appears to have strengths in Mathematics and General Information. Possible areas of need are Reading Recognition, Reading Comprehension, and Spelling.

PIAT results are useful for determining present levels of academic achievement in several curriculum areas. These levels aid the team in identifying academic subjects in which the student may be performing in the below average range and in determining the student's strengths and weaknesses in school learning. Thus, PIAT results may aid the team in establishing the student's eligibility for special education services and may also point to further assessment needs. The team considers the academic areas in which the student has shown below average and low average performance and prioritizes these according to the severity of the skill deficit and the functional relevance of the skill. Then, further assessment may be done in high priority areas of the curriculum which are determined by the team to be areas of educational need.

It may also be useful to analyze the student's responses to PIAT test questions. However, on the PIAT, most subtests use a multiple-choice format, and thus analysis of student responses may not be very informative. Reading Recognition and General Information are not multiple-choice, and it may certainly be worthwhile to do a work sample analysis of the student's attempts to respond to items on these subtests.

Further assessment of specific skills in areas of educational need is essential since global scores such as percentile ranks and standard scores do not facilitate program planning. Also, some PIAT subtests require test tasks that are not representative of typical classroom demands. For example, classroom reading comprehension tasks usually do not require matching a sentence with a picture. And classroom spelling tasks usually require writing correctly spelled words rather than merely picking the word that is correctly spelled from several distractors.

The PIAT, then, is an individual norm-referenced standardized achievement test that provides information about present levels of school performance in several academic areas. It may be used with both elementary and secondary students; results aid in determining possible areas of educational need and strengths and weaknesses in school learning. Further assessment in specific skill areas is necessary before planning instructional programs.

APPENDIX B

DATA COLLECTION FORMS

Dr. Harvey F. Clarizio
461 Erickson Hall
Michigan State University
East Lansing, MI 48824
(517) 355-8524
Office hours: T & Th 1-3

CEP 822
Kaufman Assessment Battery for Children

Following completion of this course the student should be able to (1) administer and score the K-ABC, (2) to discuss the theoretical rationale underlying the K-ABC, (3) to interpret the K-ABC from an empirical or psychometric standpoint, (4) to interpret the K-ABC from a clinical standpoint, (5) conduct via the K-ABC a psychological assessment of minority youth and a wide variety of exceptional children, and (6) be better able to compare the K-ABC vis-a-vis the WISC-R.

COURSE PREREQUISITE: All students must be either practicing school psychologists or school psychology interns with prior courses in tests and measurements and individual mental assessment.

COURSE REQUIREMENTS: The course requirements are twofold:

(1) Required Readings: The K-ABC Administration and Scoring Manual;
The K-ABC Interpretive Manual

(2) Field Experience: Each participant is to administer and interpret 5 K-ABC's. To insure that experience is acquired with a variety of children, one K-ABC and the WISC-R should be given to each of the following, if at all possible:
a.) a learning disabled child
b.) an emotionally impaired youngster
c.) a black child
d.) an EMI student
e.) and a student who was referred for psychological evaluation but who was found not eligible for special education.

Each case should be an initial referral. The order in which the K-ABC and WISC-R are given should be alternated from case to case. For example, with your first case, if you give the WISC-R first and then the K-ABC, in your second case you should give the K-ABC first and then the WISC-R. The first 2 tests should be observed by a colleague in the course or by someone trained in the K-ABC. Use school-age children between the ages of 5-0 and 12-5. Do not prorate.

Please use the PIAT and the Teacher Rating Scale with each case. These measures will give you an idea of how well the K-ABC relates to school success.

For the field based experience portion of the course, you should complete:
a.) the data form for each child (delete student's name)
b.) the clinical rating sheet
c.) the test evaluation protocols (K-ABC, WISC-R and PIAT)
d.) and the teacher rating scale.

Both the K-ABC and WISC-R test booklets are to be turned in by December 1. The tests will be returned to you when we meet for our final session on December 6, 7-10 p.m. on the MSU campus in Room C103 Wells Hall.

During the first part of the final session, you will observe a video tape of a K-ABC administration and you will be asked to score it. So bring a K-ABC record booklet and manual. During the second portion of our final meeting, we will discuss your observations and judgments regarding the clinical usefulness of this scale.

GRADING is on a Pass/Fail basis

CEP 882

DATA FORM

Subject # Sex: ______ Grade: ______

Date of Birth: __________ CA: ________ Race: ______________________________
(White, Black, Hispanic, Other)

SES (Parental Education)(✓):

Less than High School	High School Grad or GED	One or More Years College or Technical School	College Graduate
____ (0-11 years)	____ (12 years)	____ (13-15 years)	____ (16 years +)

Reason for Referral: ______________________________

Order of Administration: (✓) _____ **K-ABC first** _____ **WISC-R first**

Examiner: __________________ **Race:** __________________ **Sex:** ________

Examiner's School Setting: _____ **Urban** _____ **Rural** _____ **Suburban**

Years of Experience as a School Psychologist: __________________

Please rate every item on the following scale:

very high	high	average	low	very low
1	2	3	4	5

K-ABC WISC-R

1. _______ interest level of the materials _______

Comments:

2. _______ clarity of instructions and teaching examples _______

Comments:

3. _______ emphasis on language ability required by the tasks _______

Comments:

4. _______ precision of floor levels of the various subtests for different age groups _______

Comments:

5. _______ clarity of scoring procedures _______

Comments:

CLASSROOM ADJUSTMENT RATING SCALE

Child's name: Teacher's name: Date: 11-24-83

Section I: Please rate *every* item on the following scale: 1 = not a problem, 2 = very mild problem, 3 = moderate problem, 4 = serious problem, 5 = very serious problem

Child's classroom behavior:

1. 1 disruptive in class
2. 1 fidgety, hyperactive, can't stay in seat
3. 2 talks out of turn, disturbs others while they are working
4. 1 constantly seeks attention, "clowns around"
5. 1 overly aggressive to peers (fights, is overbearing, belligerent)
6. 1 defiant, obstinate, stubborn
7. 1 impulsive, is unable to delay
8. 1 withdrawn
9. 1 shy, timid
10. 1 does not make friends
11. 1 overconforms to rules
12. 1 daydreams, is preoccupied, "off in another world"
13. 1 unable to express feelings
14. 1 anxious
15. 1 worried, frightened, tense
16. 1 depressed
17. 1 cries easily, pouts, sulks
18. 1 does not trust others
19. 1 shows other signs of "nervousness," specify: ____
20. 1 specific fears, specify: ____

Other behaviors:

21. 2 lacks self-confidence
22. 2 overly sensitive to criticism
23. 1 reacts poorly to disappointment
24. 1 depends too much on others
25. 1 pretends to be ill
26. ___ other, specify: ____
27. 1 poor grooming or personal hygiene

Child's academic performance:

28. 5 underachieving (not working up to potential)
29. 1 poorly motivated to achieve
30. 1 poor work habits
31. 4 difficulty following directions
32. 2 poor concentration, limited attention span
33. 4 motor coordination problem
34. 1 other, specify: verbal

Child has specific academic problems in:

35. 3 reading recognition
36. 4 reading comprehension
37. 5 spelling
38. 1 general information (listening–verbal)
39. 4 math
40. 1 numbers
41. 4 writing
42. 1 colors
43. 1 concepts
44. 5 language skills problems, specify: speech

45. Considering your knowledge of this child's work please assign a letter grade to best represent the child's current functioning level in each academic area below:

C reading recognition
D reading comprehension
D spelling
D math
C general knowledge

Section II: From your experiences with this child, please check (✓) where he or she would lie on the following dimensions taking into account the direction of each item:

46. Know child well — 1 2 3 4 5 6 7 — Barely know child

47. Child has significant school adjustment problems — 1 2 3 4 5 6 7 — Child has no school adjustment problems

Note: Adapted & Modified.
For Research Only: from R. P. Lorion, E. L. Cowen, and R. A. Caldwell, 1975, Plenium Publishing Corporation.

APPENDIX C

SPSS PROGRAM

VOGELBACK COMPUTING CENTER
NORTHWESTERN UNIVERSITY

S P S S - - STATISTICAL PACKAGE FOR THE SOCIAL SCIENCES

VERSION 9.0 , MSU , DEC 11 , 1984

100000 CM MAXIMUM FIELD LENGTH REQUEST

```
RUN NAME        KABC LIST CASES
FILE NAME       KABCDATA
VARIABLE LIST   DEM1 TO DEM16, DIS,ADJ1 TO ADJ51,
                PIATDAT1 TO PIATDAT3,PIAT1 TO PIAT12,KABCDAT1
                TO KABCDAT3,KABC1 TO KABC38,WISCDAT1 TO WISCDAT3,
                WISCR1 TO WISCR16,GRADE1 TO GRADE6
INPUT FORMAT    FIXED(F2.0,F3.0,1X,3F1.0,F2.0,2F1.0,5F2.0,3F1.0,27X,
                27F1.0/25F1.0,9F2.0,6F3.0,9F2.0,1X/10F2.0,5F3.0,5F2.0,
                4F3.0,4F2.0,4F1.0,5F2.0,1X/11F2.0,3F3.0/3X,6F3.1)
```

ACCORDING TO YOUR INPUT FORMAT, VARIABLES ARE TO BE READ AS FOLLOWS

VARIABLE	FORMAT	RECORD	COLUMNS
DEM1	F 2. 0	1	1- 2
DEM2	F 3. 0	1	3- 5
DEM3	F 1. 0	1	7- 7
DEM4	F 1. 0	1	8- 8
DEM5	F 1. 0	1	9- 9
DEM6	F 2. 0	1	10- 11
DEM7	F 1. 0	1	12- 12
DEM8	F 1. 0	1	13- 13
DEM9	F 2. 0	1	14- 15
DEM10	F 2. 0	1	16- 17
DEM11	F 2. 0	1	18- 19
DEM12	F 2. 0	1	20- 21
DEM13	F 2. 0	1	22- 23
DEM14	F 1. 0	1	24- 24
DEM15	F 1. 0	1	25- 25
DEM16	F 1. 0	1	26- 26
DIS	F 1. 0	1	54- 54
ADJ1	F 1. 0	1	55- 55
ADJ2	F 1. 0	1	56- 56
ADJ3	F 1. 0	1	57- 57
ADJ4	F 1. 0	1	58- 58
ADJ5	F 1. 0	1	59- 59
ADJ6	F 1. 0	1	60- 60
ADJ7	F 1. 0	1	61- 61
ADJ8	F 1. 0	1	62- 62
ADJ9	F 1. 0	1	63- 63
ADJ10	F 1. 0	1	64- 64
ADJ11	F 1. 0	1	65- 65
ADJ12	F 1. 0	1	66- 66
ADJ13	F 1. 0	1	67- 67
ADJ14	F 1. 0	1	68- 68
ADJ15	F 1. 0	1	69- 69
ADJ16	F 1. 0	1	70- 70
ADJ17	F 1. 0	1	71- 71
ADJ18	F 1. 0	1	72- 72
ADJ19	F 1. 0	1	73- 73
ADJ20	F 1. 0	1	74- 74
ADJ21	F 1. 0	1	75- 75
ADJ22	F 1. 0	1	76- 76
ADJ23	F 1. 0	1	77- 77
ADJ24	F 1. 0	1	78- 78
ADJ25	F 1. 0	1	79- 79
ADJ26	F 1. 0	1	80- 80
ADJ27	F 1. 0	2	1- 1
ADJ28	F 1. 0	2	2- 2

ACCORDING TO YOUR INPUT FORMAT, VARIABLES ARE TO BE READ AS FOLLOWS

VARIABLE	FORMAT	RECORD	COLUMNS
KABC26	F 2. 0	3	44- 45
KABC27	F 3. 0	3	46- 48
KABC28	F 3. 0	3	49- 51
KABC29	F 3. 0	3	52- 54
KABC30	F 3. 0	3	55- 57
KABC31	F 2. 0	3	58- 59
KABC32	F 2. 0	3	60- 61
KABC33	F 2. 0	3	62- 63
KABC34	F 2. 0	3	64- 65
KABC35	F 1. 0	3	66- 66
KABC36	F 1. 0	3	67- 67
KABC37	F 1. 0	3	68- 68
KABC38	F 1. 0	3	69- 69
WISCDAT1	F 2. 0	3	70- 71
WISCDAT2	F 2. 0	3	72- 73
WISCDAT3	F 2. 0	3	74- 75
WISCR1	F 2. 0	3	76- 77
WISCR2	F 2. 0	3	78- 79
WISCR3	F 2. 0	4	1- 2
WISCR4	F 2. 0	4	3- 4
WISCR5	F 2. 0	4	5- 6
WISCR6	F 2. 0	4	7- 8
WISCR7	F 2. 0	4	9- 10
WISCR8	F 2. 0	4	11- 12
WISCR9	F 2. 0	4	13- 14
WISCR10	F 2. 0	4	15- 16
WISCR11	F 2. 0	4	17- 18
WISCR12	F 2. 0	4	19- 20
WISCR13	F 2. 0	4	21- 22
WISCR14	F 3. 0	4	23- 25
WISCR15	F 3. 0	4	26- 28
WISCR16	F 3. 0	4	29- 31
GRADE1	F 3. 1	5	4- 6
GRADE2	F 3. 1	5	7- 9
GRADE3	F 3. 1	5	10- 12
GRADE4	F 3. 1	5	13- 15
GRADE5	F 3. 1	5	16- 18
GRADE6	F 3. 1	5	19- 21

THE INPUT FORMAT PROVIDES FOR 149 VARIABLES. 149 WILL BE READ.
IT PROVIDES FOR 5 RECORDS ("CARDS") PER CASE.
A MAXIMUM OF 80 "COLUMNS" ARE USED ON A RECORD.

```
MISSING VALUES DEM3 TO DEM5,DEM7,DEM8,DEM14 TO DEM15,
               DIS,ADJ1 TO ADJ51,KABC35 TO KABC38(9,BLANKS)/DEM6,
               DEM9 TO DEM13,PIATDAT1 TO PIATDAT3,PIAT1 TO PIAT6,
               KABCDAT1 TO KABCDAT3,KABC1 TO KABC8,
               WISCDAT1 TO WISCDAT3,WISCR1 TO
               WISCR13(BLANKS,99)/PIAT7 TO PIAT12,KABC17 TO KABC21,
               KABC27 TO KABC30,WISCR14 TO WISCR16 (BLANKS,0,999)/
               PIAT1 TO PIAT6,KABC9 TO KABC16,KABC22 TO KABC26,
               KABC31 TO KABC34(BLANKS)/
               DEM16(0,9,BLANKS)/GRADE1 TO GRADE6(0,BLANKS)
VAR LABELS     DEM1,EXAMINER ID/DEM2,STUDENT ID/DEM3,RACE EXAM/DEM4,
               SEX EXAM/DEM5,SETTING/DEM6,YRS EXP/DEM7,RACE STUDENT/
               DEM8,SEX STUDENT/DEM9,MONTH BIRTH/DEM[illegible],DAY BIRTH/
               DEM11,YEAR BIRTH/DEM12,CA YRS/DEM13,[illegible] MONTHS/
               DEM14,GR PLACE/DEM15,REPEAT/DEM16,SES [illegible]ARENT/,
               DIS,CHILD DISAB/
               ADJ1,DISRUPTIVE/ADJ2,FIDGETY/ADJ3,TAL[illegible] OUT/ADJ4,
               SEEKS ATTENT/ADJ5,AGGRESSIVE/ADJ6,DEF[illegible]NT/ADJ7,
               IMPULSIVE/ADJ8,WITHDRAWN/ADJ9,SHY/ADJ[illegible],NO FRIENDS/ADJ11,
               OVERCONFORMS/ADJ12,DAYDREAMS/ADJ13,CAN[illegible] EXPR FEEL/
               ADJ14,ANXIOUS/ADJ15,WORRIED/ADJ16,DEPRESSED/ADJ17,
               CRIES POUTS/ADJ18,NO TRUST/ADJ19,NERVOUS/ADJ20,
               FEARS/ADJ21,NO SELF CONF/ADJ22,OVER SENS/ADJ23,
               DISAPPOINTED/ADJ24,OVER DEPEND/ADJ25,FEIGNS ILL/ADJ26,
               OTHER/ADJ27,POOR GROOM/ADJ28,UNDERACHIEVE/ADJ29,POOR MOT/
               ADJ30,POOR WORKHAB/ADJ31,DIFF DIRECT/ADJ32,POOR CONCENT/
               ADJ33,POOR COORD/ADJ34,OTHER/ADJ35,READ REC/ADJ36
                READ COMP/ADJ37,SPELLING/ADJ38,GEN INFO/ADJ3[illegible] ADJ40,
               NUMBERS/ADJ41,WRITING/ADJ42,COLORS/ADJ43,CONC[illegible]
```

```
                 ADJ44.LANG SKILLS/ADJ45.READ REC/ADJ46.READ COMP/ADJ47.
                 SPELLING/ADJ48.MATH/ADJ49.GEN KNOW/ADJ50. KNOW CHILD/
                 ADJ51.ADJUST PROB/PIATDAT1. MONTH TEST/PIATDAT2.
                 DAY TEST/PIATDAT3.YEAR TEST/PIAT1.PCTMATH/PIAT2.
                 PCTREADREC/PIAT3.PCTREADCOMP/PIAT4.PCTSPELLING/PIAT5.
                 PCTGENINFO/PIAT6.PCTTOTAL/PIAT7.SSMATH/PIAT8.SSREADREC/
                 PIAT9.SSREADCOMP/PIAT10.SSSPELLING/PIAT11.SSGENINFO/
                 PIAT12.SSTOTAL/KABCDAT1.MONTH TEST/KABCDAT2.DAT TEST/
                 KABCDAT3.YEAR TEST/KABC1.SCSHANDMOVE/KABC2.SCSGESTCLOS/
                 KABC3.SCSNUMBREC/KABC4.SCSTRIANGLES/KABC5.SCSWORDORD/
                 KABC6.SCSMATRIXAN/KABC7.SCSSPATMEM/KABC8.SCSPHOTOSER/
                 KABC9.PCTHANDMOV/KABC10.PCTGESTCLOS/KABC11.PCTNUMBREC/
                 KABC12.PCTTRIANGLES/KABC13.PCTWORDORD/KABC14.PCTMATRIXAN/
                 KABC15.PCTSPATMEM/KABC16.PCTPHOTOSER/KABC17.SSFACEPLACE/
                 KABC18.SSARITH/KABC19.SSRIDDLES/KABC20.SSREADDECODE/
                 KABC21.SSREADUND/KABC22.PCTFACEPLACE/KABC23.PCTARITH/
                  KABC24.PCTRIDDLES/KABC25.PCTREADDECODE/KABC26.
                 PCTREADUNDERSTAND/KABC27.SSSEQUENTIAL/KABC28.
                 SSSIMULANEOUS/KABC29.SSMENTPROCESS/KABC30.SSACHIEVE/
                 KABC31.PCTSEQUENTIAL/KABC32.PCTSIMULTANEOUS/KABC33.
                 PCTMENTPROCESS/KABC34.PCTACHIEVE/KABC35.SEQ TO SIM/
                 KABC36. SEQ TO ACHIEVE/KABC37.SIM TO ACHIEVE/KABC38.
                 MPC TO ACHIEVE/WISCDAT1. MONTH TEST/WISCDAT2. DATY TEST/
                 WISCDAT3.YEAR TEST/WISCR1. SCSINFORMATION/WISCR2.
                 SCSSIMILARITIES/WISCR3.SCSARITHMETIC/WISCR4.SCSVOCAB/
                 WISCR5.SCSCOMP/WISCR6.SCSDIGSPAN/WISCR7.TOTSCSVERBAL/
                 WISCR8.SCSPICTCOMPLET/WISCR9.SCSPICTARRANGE/WISCR10.
                 SCSBLOCKDES/WISCR11.SCSOBJASSEMBLY/WISCR12.SCSCODING/
                 WISCR13.TOTSCSPERFORM/WISCR14.VERBAL IQ/WISCR15.
                 PERFORMANCE IQ/WISCR16.FULLSCALE IQ
VALUE LABELS     DEM3.DEM7(1)WHITE (2)BLACK (3)HISPANIC/DEM4.
                 DEM8 (1)MALE (2)FEMALE/DEM16 (1)0-11 YEARS (2)12 YEARS
                 (3)13-15 YEARS (4)16+ YEARS/
                 DIS (1)EMI (2)LD (3)EI (4)REG (5)OTHER/
                 ADJ1 TO ADJ44(1)NOT PROBLEM (2)VERY MILD (3)MODERATE
                 (4)SERIOUS (5)VERYSERIOUS/ADJ45 TO ADJ49 (1)A (2)B
                 (3)C (4)D (5)E/ADJ50 (1)WELL (7)BARELY KNOW/ADJ51
                 (1)SIGNIF PROB (7)NO PROBLEM/KABC35 TO KABC38
                 (1)MORE THAN  (2)LESS THAN (3)ABOUT SAME
COMPUTE           Z1=(WISCR16-100)/15
COMPUTE          Z2=(WISCR14-100)/15
COMPUTE          Z3=(WISCR15-100)/15
IF               (PIAT7 GT 0) Z4=(PIAT7-100)/15
IF               (PIAT8 GT 0) Z5=(PIAT8-100)/15
IF               (PIAT9 GT 0) Z6=(PIAT9-100)/15
IF               (PIAT10 GT 0) Z7=(PIAT10-100)/15
IF               (PIAT11 GT 0) Z8=(PIAT11-100)/15
IF               (PIAT12 GT 0) Z9=(PIAT12-100)/15
IF               (KABC17 GT 0) Z10=(KABC17-100)/15
IF               (KABC18 GT 0) Z11=(KABC18-100)/15
IF               (KABC19 GT 0) Z12=(KABC19-100)/15
IF               (KABC20 GT 0) Z13=(KABC20-100)/15
IF               (KABC21 GT 0) Z14=(KABC21-100)/15
IF               (KABC27 GT 0) Z15=(KABC27-100)/15
IF               (KABC28 GT 0) Z16=(KABC28-100)/15
IF               (KABC29 GT 0) Z17=(KABC29-100)/15
IF               (KABC30 GT 0) Z18=(KABC30-100)/15
ASSIGN MISSING   Z1 TO Z18 (9)
COMPUTE          DIF1=Z1-Z4
COMPUTE          DIF2=Z1-Z5
COMPUTE          DIF3=Z1-Z6
COMPUTE          DIF4=Z1-Z7
COMPUTE          DIF5=Z1-Z8
COMPUTE          DIF6=Z1-Z9
COMPUTE          DIF7=Z1-Z10
COMPUTE          DIF8=Z1-Z11
COMPUTE          DIF9=Z1-Z12
COMPUTE          DIF10=Z1-Z13
COMPUTE          DIF11=Z1-Z14
COMPUTE          DIF12=Z17-Z4
COMPUTE          DIF13=Z17-Z5
COMPUTE          DIF14=Z17-Z6
COMPUTE          DIF15=Z17-Z7
COMPUTE          DIF16=Z17-Z8
COMPUTE          DIF17=Z17-Z9
COMPUTE          DIF18=Z17-Z10
COMPUTE          DIF19=Z17-Z11
```

```
•               COMPUTE         DIF20=Z17-Z12                                                                   116
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
KABC LIST CASES                                           10/25/86     SPSS V9.0    .15.05.23.     PAGE    5
•               COMPUTE         DIF21=Z17-Z13
•               COMPUTE         DIF22=Z17-Z14
•               COMPUTE         DIF23=Z1-Z17
•               COMPUTE         DIF24=Z2-Z3
•               COMPUTE         DIF25=Z15-Z16
•               COMPUTE         DIF26=Z17-Z18
•               ASSIGN MISSING  DIF1 TO DIF26 (9)
•               IF              (DIF1 GE 1.0) REYN1=1
•               IF              (DIF2 GE 1.0) REYN2=1
•               IF              (DIF3 GE 1.0) REYN3=1
•               IF              (DIF4 GE 1.0) REYN4=1
•               IF              (DIF5 GE 1.0) REYN5=1
•               IF              (DIF6 GE 1.0) REYN6=1
•               IF              (DIF18 GE 1.0) REYN7=1
•               IF              (DIF19 GE 1.0) REYN8=1
•               IF              (DIF20 GE 1.0) REYN9=1
•               IF              (DIF21 GE 1.0) REYN10=1
•               IF              (DIF22 GE 1.0) REYN11=1
•               IF              (DIF26 GE 1.0) REYN12=1
•               ASSIGN MISSING  REYN1 TO REYN12(0)
•               IF              (REYN1 EQ 1 OR REYN2 EQ 1 OR REYN3 EQ 1)
•                               QUAL1=1
•               IF              (REYN8 EQ 1 OR REYN10 EQ 1 OR REYN11 EQ 1)
•                               QUAL2=1
•               ASSIGN MISSING  QUAL1,QUAL2 (0)
•               VAR LABELS      Z1,Z WISCR FSIQ/Z2,Z WISCR VIQ/Z3,
•                               Z WISCR PIQ/Z4,Z PIAT MATH/Z5, Z PIAT READREC/
•                               Z6,Z PIAT READCOMP/Z7,Z PIAT SPELLING/Z8,
•                               Z PIAT GENINFO/Z9,Z PIAT TOTAL/Z10,
•                               Z KABC FACEPLACE/Z11, Z KABC ARITH/Z12,
•                               Z KABC RIDDLES/Z13,Z KABC READDECODE/Z14,
•                               Z KABC READUND/Z15,Z KABC SEQUENTIAL/Z16,
•                               Z KABC SIMULTANEOUS/Z17, Z KABC MENTPROC/Z18,
•                               Z KABC ACHIEVE
•               VAR LABELS      DIF1,FSIQ TO PIATMATH/DIF2,FSIQ TO PIATREADREC/
•                               DIF3,FSIQ TO PIATREADCOMP/DIF4,FSIQ TO PIATSPELL/
•                               DIF5,FSIQ TO PIATGENINFO/DIF6,FSIQ TO PIATTOTAL/
•                               DIF7,FSIQ TO KABCFACEPLACE/DIF8,FSIQ TO KABCARITH/
•                               DIF9,FSIQ TO KABCRIDDLES/DIF10,FSIQ TO KABCREADDEC/
•                               DIF11,FSIQ TO KABCREADUND/DIF12,MPC TO PIATARITH/
•                               DIF13,MPC TO PIATREADREC/DIF14,MPC TO PIATREADCOMP/
•                               DIF15,MPC TO PIATSPELL/DIF16,MPC TO PIATGENINFO/
•                               DIF17,MPC TO PIATTOTAL/DIF18,MPC TO KABCFACEPLAC/
•                               DIF19,MPC TO KABCARITH/DIF20,MPC TO KABCRIDDLES/
•                               DIF21,MPC TO KABCREADDEC/DIF22,MPC TO KABCREADUND/
•                               DIF23,FSIQ TO MPC/DIF24,VIQ TO PIQ/DIF25,SEQ TO SIM
•               PRINT FORMATS   DIF1 TO DIF26 (3)
•               VAR LABELS      GRADE1,GEPIATMATH/GRADE2,GEPIATREADRED/GRADE3,
•                               GEPIATREADCOMP/GRADE4,GEPIATSPELL/GRADE5,GEPIATGI/
•                               GRADE6,GEPIATTOTAL
•               PRINT FORMATS   GRADE1 TO GRADE6(1)
•               COMPUTE         R1=.96
•               COMPUTE         R2=.93
•               COMPUTE         R3=.91
•               COMPUTE         R4=.74
•               COMPUTE         R5=.89
•               COMPUTE         R6=.64
•               COMPUTE         R7=.65
•               COMPUTE         R8=.76
•               COMPUTE         R9=.89
•               COMPUTE         R10=.84
•               COMPUTE         R11=.87
•               COMPUTE         R12=.86
•               COMPUTE         R13=.92
•               COMPUTE         R14=.91
•               COMPUTE         R15=.89
•               COMPUTE         R16=.93
•               COMPUTE         R17=.94
•               COMPUTE         R18=.97
•               COMPUTE         ZDIF1=DIF1/SQRT((1-R1)+(1-R4))
•               COMPUTE         ZDIF2=DIF2/SQRT((1-R1)+(1-R5))
•               COMPUTE         ZDIF3=DIF3/SQRT((1-R1)+(1-R6))
•               COMPUTE         ZDIF4=DIF4/SQRT((1-R1)+(1-R7))
```

```
COMPUTE          ZDIF5=DIF5/SQRT((1-R1)+(1-R8))
COMPUTE          ZDIF6=DIF6/SQRT((1-R1)+(1-R9))
COMPUTE          ZDIF7=DIF18/SQRT((1-R17)+(1-R10))
COMPUTE          ZDIF8=DIF19/SQRT((1-R17)+(1-R11))
COMPUTE          ZDIF9=DIF20/SQRT((1-R17)+(1-R12))
```

- -


```
COMPUTE          ZDIF10=DIF21/SQRT((1-R17)+(1-R13))
COMPUTE          ZDIF11=DIF22/SQRT((1-R17)+(1-R14))
COMPUTE          ZDIF12=DIF26/SQRT((1-R17)+(1-R18))
ASSIGN MISSING   ZDIF1 TO ZDIF12 (9)
PRINT FORMATS    ZDIF1 TO ZDIF12 (3)
COMPUTE          TRUE1=R1*Z1
COMPUTE          TRUE2=R2*Z2
COMPUTE          TRUE3=R3*Z3
COMPUTE          TRUE4=R4*Z4
COMPUTE          TRUE5=R5*Z5
COMPUTE          TRUE6=R6*Z6
COMPUTE          TRUE7=R7*Z7
COMPUTE          TRUE8=R8*Z8
COMPUTE          TRUE9=R9*Z9
COMPUTE          TRUE10=R10*Z10
COMPUTE          TRUE11=R11*Z11
COMPUTE          TRUE12=R12*Z12
COMPUTE          TRUE13=R13*Z13
COMPUTE          TRUE14=R14*Z14
COMPUTE          TRUE15=R15*Z15
COMPUTE          TRUE16=R16*Z16
COMPUTE          TRUE17=R17*Z17
COMPUTE          TRUE18=R18*Z18
ASSIGN MISSING   TRUE1 TO TRUE18(9)
COMPUTE          SETDIF1=SQRT(((R1**2)*(1-R1))+((R4**2)*(1-R4)))
COMPUTE          SETDIF2=SQRT(((R1**2)*(1-R1))+((R5**2)*(1-R5)))
COMPUTE          SETDIF3=SQRT(((R1**2)*(1-R1))+((R6**2)*(1-R6)))
COMPUTE          SETDIF4=SQRT(((R1**2)*(1-R1))+((R7**2)*(1-R7)))
COMPUTE          SETDIF5=SQRT(((R1**2)*(1-R1))+((R8**2)*(1-R8)))
COMPUTE          SETDIF6=SQRT(((R1**2)*(1-R1))+((R9**2)*(1-R9)))
COMPUTE           SETDIF7=SQRT(((R17**2)*(1-R17))+((R10**2)*(1-R10)))
COMPUTE          SETDIF8=SQRT(((R17**2)*(1-R17))+((R11**2)*(1-R11)))
COMPUTE          SETDIF9=SQRT(((R17**2)*(1-R17))+((R12**2)*(1-R12)))
COMPUTE          SETDIF10=SQRT(((R17**2)*(1-R17))+((R13**2)*(1-R13)))
COMPUTE          SETDIF11=SQRT(((R17**2)*(1-R17))+((R14**2)*(1-R14)))
COMPUTE          SETDIF12=SQRT(((R17**2)*(1-R17))+((R18**2)*(1-R18)))
ASSIGN MISSING   SETDIF1 TO SETDIF12(9)
COMPUTE          ZTDIF1=(TRUE1-TRUE4)/SETDIF1
COMPUTE          ZTDIF2=(TRUE1-TRUE5)/SETDIF2
COMPUTE          ZTDIF3=(TRUE1-TRUE6)/SETDIF3
COMPUTE          ZTDIF4=(TRUE1-TRUE7)/SETDIF4
COMPUTE          ZTDIF5=(TRUE1-TRUE8)/SETDIF5
COMPUTE          ZTDIF6=(TRUE1-TRUE9)/SETDIF6
COMPUTE          ZTDIF7=(TRUE17-TRUE10)/SETDIF7
COMPUTE          ZTDIF8=(TRUE17-TRUE11)/SETDIF8
COMPUTE          ZTDIF9=(TRUE17-TRUE12)/SETDIF9
COMPUTE          ZTDIF10=(TRUE17-TRUE13)/SETDIF10
COMPUTE          ZTDIF11=(TRUE17-TRUE14)/SETDIF11
COMPUTE          ZTDIF12=(TRUE17-TRUE18)/SETDIF12
ASSIGN MISSING   ZTDIF1 TO ZTDIF12(9)
PRINT FORMATS    TRUE1 TO TRUE18,SETDIF1 TO SETDIF12,
                 ZTDIF1 TO ZTDIF12(3)
IF               ((TRUE1-TRUE4) GE 1.0) WAYNE1=1
IF               ((TRUE1-TRUE5) GE 1.0) WAYNE2=1
IF               ((TRUE1-TRUE6) GE 1.0) WAYNE3=1
IF               ((TRUE1-TRUE7) GE 1.0) WAYNE4=1
IF               ((TRUE1-TRUE8) GE 1.0) WAYNE5=1
IF               ((TRUE1-TRUE9) GE 1.0) WAYNE6 1
IF                 ((TRUE17-TRUE10) GE 1.0)WAYNE7=1
IF               ((TRUE17-TRUE11)GE 1.0)WAYNE8=1
IF               ((TRUE17-TRUE12)GE 1.0)WAYNE9=1
IF               ((TRUE17-TRUE13)GE 1.0)WAYNE10=1
IF               ((TRUE17-TRUE14)GE 1.0)WAYNE11=1
IF               ((TRUE17-TRUE18)GE 1.0)WAYNE12=1
ASSIGN MISS'NG   WAYNE1 TO WAYNE12 (0)
IF               (WAYNE1 EQ 1 OR WAYNE2 EQ 1 OR WAYNE3 EQ 1)
                 QUAL3=1
IF               (WAYNE8 EQ 1 OR WAYNE10 EQ 1 OR WAYNE11 EQ 1)
                 QUAL4=1
```

```
ASSIGN MISSING  QUAL3,QUAL4 (O)
COMPUTE         CA=DEM12+(DEM13/12)
ASSIGN MISSING  CA(9)
PRINT FORMATS   CA(1)
COMPUTE         ELA=(CA*WISCR16)/100 - 5
COMPUTE         GRDIF1=ELA-GRADE1
COMPUTE         GRDIF3=ELA-GRADE3
COMPUTE         GRDIF4=ELA-GRADE4
COMPUTE         GRDIF5=ELA-GRADE5
```

- -


```
COMPUTE         GRDIF6=ELA-GRADE6
ASSIGN MISSING  ELA,GRDIF1 TO GRDIF6(99)
PRINT FORMATS   ELA,GRDIF1 TO GRDIF6(2)
COMPUTE         RR1=.58
COMPUTE         RR2=.58
COMPUTE         RR3=.61
COMPUTE         RR4=.43
COMPUTE         RR5=.73
COMPUTE         RR6=.63
COMPUTE         RR7=.53
COMPUTE         RR8=.71
COMPUTE         RR9=.64
COMPUTE         RR10=.62
COMPUTE         RR11=.64
COMPUTE         RR12=.74
COMPUTE         REGRES1=(Z1*(RR1)-Z4)/SQRT(1-(RR1**2))
COMPUTE         REGRES2=(Z1*(RR2)-Z5)/SQRT(1-(RR2**2))
COMPUTE         REGRES3=(Z1*(RR3)-Z6)/SQRT(1-(RR3**2))
COMPUTE         REGRES4=(Z1*(RR4)-Z7)/SQRT(1-(RR4**2))
COMPUTE         REGRES5=(Z1*(RR5)-Z8)/SQRT(1-(RR5**2))
COMPUTE         REGRES6=(Z1*(RR6)-Z9)/SQRT(1-(RR6**2))
COMPUTE         REGRES7=(Z17*(RR7)-Z10)/SQRT(1-(RR7**2))
COMPUTE         REGRES8=(Z17*(RR8)-Z11)/SQRT(1-(RR8**2))
COMPUTE         REGRES9=(Z17*(RR9)-Z12)/SQRT(1-(RR9**2))
COMPUTE         REGRES10=(Z17*(RR10)-Z13)/SQRT(1-(RR10**2))
COMPUTE         REGRES11=(Z17*(RR11)-Z14)/SQRT(1-(RR11**2))
COMPUTE         REGRES12=(Z17*(RR12)-Z18)/SQRT(1-(RR12**2))
ASSIGN MISSING  REGRES1 TO REGRES12(9)
PRINT FORMATS   REGRES1 TO REGRES12(3)
IF              ((Z1*(RR1)-Z4) GE 1.0) MONROE1=1
IF              ((Z1*(RR2)-Z5) GE 1.0) MONROE2=1
IF              ((Z1*(RR3)-Z6) GE 1.0) MONROE3=1
IF              ((Z1*(RR4)-Z7) GE 1.0) MONROE4=1
IF              ((Z1*(RR5)-Z8) GE 1.0) MONROE5=1
IF              ((Z1*(RR6)-Z9) GE 1.0) MONROE6=1
IF              ((Z17*(RR7)-Z10) GE 1.0)MONROE7=1
IF              ((Z17*(RR8)-Z11) GE 1.0)MONROE8=1
IF              ((Z17*(RR9)-Z12) GE 1.0)MONROE9=1
IF              ((Z17*(RR10)-Z13) GE 1.0)MONROE10=1
IF              ((Z17*(RR11)-Z14) GE 1.0)MONROE11=1
IF              ((Z17*(RR12)-Z18) GE 1.0)MONROE12=1
ASSIGN MISSING  MONROE1 TO MONROE11 (0)
IF              (MONROE1 EQ 1 OR MONROE2 EQ 1 OR MONROE3 EQ 1)
                QUAL5=1
IF              (MONROE8 EQ 1 OR MONROE10 EQ 1 OR MONROE11 EQ 1)
                QUAL6=1
ASSIGN MISSING  QUAL5, QUAL6 (0)
COMPUTE         REYREG1=(Z1*(RR1)-Z4)/((1-
                (SQRT(1-((R4+(R1*(RR1**2))-2*(RR1**2))/
                (1-(RR1**2))))))*(SQRT(1-RR1**2)))
COMPUTE         REYREG2=(Z1*(RR2)-Z5)/((1-
                (SQRT(1-((R5+(R1*(RR2**2))-2*(RR2**2))/
                (1-(RR2**2))))))*(SQRT(1-RR2**2)))
COMPUTE         REYREG3=(Z1*(RR3)-Z6)/((1-
                (SQRT(1-((R6+(R1*(RR3**2))-2*(RR3**2))/
                (1-(RR3**2))))))*(SQRT(1-RR3**2)))
COMPUTE         REYREG4=(Z1*(RR4)-Z7)/((1-
                (SQRT(1-((R7+(R1*(RR4**2))-2*(RR4**2))/
                (1-(RR4**2))))))*(SQRT(1-RR4**2)))
COMPUTE         REYREG5=(Z1*(RR5)-Z8)/((1-
                (SQRT(1-((R9+(R1*(RR5**2))-2*(RR5**2))/
                (1-(RR5**2))))))*(SQRT(1-RR5**2)))
COMPUTE         REYREG6=(Z1*(RR6)-Z9)/((1-
                (SQRT(1-((  +(R1*(RR6**2))-2*(RR6**2))/
                (1-(RR6**  ))))))*(SQRT(1-RR6**2)))
```

```
.                COMPUTE          REYREG7=(Z17*(RR7)-Z10)/((1-
.                                 (SQRT(1-((R10+(R17*(RR7**2))-2*(RR7**2))/
.                                  (1-(RR7**2))))))*(SQRT(1-RR7**2)))
.                COMPUTE          REYREG8=(Z17*(RR8)-Z11)/((1-
.                                 (SQRT(1-((R11+(R17*(RR8**2))-2*(RR8**2))/
.                                  (1-(RR8**2))))))*(SQRT(1-RR8**2)))
.                COMPUTE          REYREG9=(Z17*(RR9)-Z12)/((1-
.                                 (SQRT(1-((R12+(R17*(RR9**2))-2*(RR9**2))/
.                                  (1-(RR9**2))))))*(SQRT(1-RR9**2)))
.                COMPUTE          REYREG10=(Z17*(RR10)-Z13)/((1-
.                                 (SQRT(1-((R13+(R17*(RR10**2))-2*(RR10**2))/
.                                  (1-(RR10**2))))))*(SQRT(1-RR10**2)))
.                COMPUTE          REYREG11=(Z17*(RR11)-Z14)/((1-

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

KABC LIST CASES                                            10/25/86     SPSS V9.0    .15.05.23.     PAGE    8

.                                 (SQRT(1-((R14+(R17*(RR11**2))-2*(RR11**2))/
.                                  (1-(RR11**2))))))*(SQRT(1-RR11**2)))
.                COMPUTE          REYREG12=(Z17*(RR12)-Z18)/((1-
.                                 (SQRT(1-((R18+(R17*(RR12**2))-2*(RR12**2))/
.                                  (1-(RR12**2))))))*(SQRT(1-RR12**2)))
.                ASSIGN MISSING   REYREG1 TO REYREG12(9)
.                PRINT FORMATS       REYREG1 TO REYREG12(3)
.                IF               (ZDIF1 GE 1.65 OR ZDIF2 GE 1.65 OR ZDIF3 GE
.                                  1.65) DISCREP1=1
.                IF               (DISCREP1 NE 1)DISCREP1=0
.                IF               ( ZDIF8 GE 1.65 OR ZDIF10 GE 1.65
.                                 OR  ZDIF11 GE 1.65) DISCREP2=1
.                IF               (DISCREP2 NE 1)DISCREP2=0
.                IF               (ZTDIF1 GE 1.65 OR ZTDIF2 GE 1.65 OR ZTDIF3 GE
.                                  1.65)DISCREP3=1
.                IF               (DISCREP3 NE 1)DISCREP3=0
.                IF               (ZTDIF8 GE 1.65 OR ZTDIF10 GE 1.65 OR
.                                 ZTDIF11 GE 1.65) DISCREP4=1
.                IF               (DISCREP4 NE 1)DISCREP4=0
.                IF               (REGRES1 GE 1.65 OR REGRES2 GE 1.65 OR REGRES3
.                                 GE 1.65)DISCREP5=1
.                IF               (DISCREP5 NE 1)DISCREP5=0
.                IF               (REGRES8 GE 1.65 OR REGRES10 GE 1.65 OR REGRES11
.                                 GE 1.65) DISCREP6=1
.                IF               (DISCREP6 NE 1)DISCREP6=0
.                IF               (REYREG1 GE 1.65 OR REYREG2 GE 1.65 OR REYREG3
.                                 GE 1.65)DISCREP7=1
.                IF               (DISCREP7 NE 1)DISCREP7=0
.                IF               (REYREG8 GE 1.65 OR REYREG10 GE 1.65 OR REYREG11
.                                 GE 1.65)DISCREP8=1
.                IF               (DISCREP8 NE 1)DISCREP8
.                IF                (ZDIF6 GE 1.65)IQ1=WIS(?16
.                IF               (ZTDIF6 GE 1.65)IQ2=WISCR16
.                IF               (REGRES6 GE 1.65)IQ3=WISCR16
.                IF               (REYREG6 GE 1.65)IQ4=WISCR16
.                IF               (ZDIF12 GE 1.65)MPC5=KABC29
.                IF               (ZTDIF12 GE 1.65)MPC6=KABC29
.                IF               (REGRES12 GE 1.65)MPC7=KABC29
.                IF               (REYREG12 GE 1.65)MPC8=KABC29
.                ASSIGN MISSING   IQ1 TO IQ4(O)
.                ASSIGN MISSING    MPC5 TO MPC8 (O)
.                LIST CASES          CASES=86/VARIABLES=DEM2,KABC29,WISCR16 ?DIF6,
.                                 ZDIF12,ZTDIF6,ZTDIF12,REGRES6,REGRES12,RE?REG6,REYREG12

CPU TIME REQUIRED..       2.415 SECONDS

.                SCATTERGRAM        PIAT8(62.5,137.5) WIT? WISCR16(62.5,137.5)
.                OPTIONS            2
.                READ INPUT DATA

          GIVEN    2 VARIABLES. INITIAL CM ALLOWS FOR   2254 CASES
                                MAXIMUM CM ALLOWS FOR   4494 CASES

          OPTION - 2
          LISTWISE DELETION OF MISSING DATA
```

APPENDIX D

TEST INTERCORRELATIONS

	PIAT7	PIAT8	PIAT9	PIAT10	PIAT11	PIAT12
WISCR1	.4428 (86) P= .001	.4201 (86) P= .001	.5407 (71) P= .001	.2243 (82) P= .021	.5766 (81) P= .001	.5158 (81) P= .001
WISCR2	.4767 (86) P= .001	.3720 (86) P= .001	.4083 (71) P= .001	.4207 (82) P= .001	.5395 (81) P= .001	.4808 (81) P= .001
WISCR3	.5613 (86) P= .001	.3200 (86) P= .001	.4345 (71) P= .001	.3194 (82) P= .002	.3985 (81) P= .001	.4914 (81) P= .001
WISCR4	.5187 (86) P= .001	.5074 (86) P= .001	.5065 (71) P= .001	.4021 (82) P= .001	.6046 (81) P= .001	.6182 (81) P= .001
WISCR5	.4414 (86) P= .001	.4590 (86) P= .001	.4293 (71) P= .001	.3610 (82) P= .001	.5170 (81) P= .001	.5132 (81) P= .001
WISCR6	.4431 (68) P= .001	.3457 (68) P= .002	.4540 (55) P= .001	.3662 (64) P= .001	.3189 (64) P= .005	.5021 (63) P= .001
WISCR7	.4941 (86) P= .001	.4707 (86) P= .001	.5072 (71) P= .001	.3241 (82) P= .001	.6912 (81) P= .001	.5916 (81) P= .001
WISCR8	.3266 (86) P= .001	.3656 (86) P= .001	.3744 (71) P= .001	.2724 (82) P= .007	.3828 (81) P= .001	.4581 (81) P= .001
WISCR9	.2704 (86) P= .006	.1140 (86) P= .148	.1357 (71) P= .130	.0622 (82) P= .289	.2255 (81) P= .021	.2698 (81) P= .007
WISCR10	.3563 (86) P= .001	.1948 (86) P= .036	.1009 (71) P= .201	.1593 (82) P= .076	.3009 (81) P= .003	.2138 (81) P= .028
WISCR11	.1981 (84) P= .035	.1199 (84) P= .139	-.0037 (69) P= .488	.1377 (80) P= .112	.0669 (79) P= .279	.1417 (80) P= .105
WISCR12	.2624 (84) P= .008	.0898 (84) P= .208	.1445 (70) P= .116	.1872 (80) P= .048	.1765 (79) P= .060	.1876 (79) P= .049
WISCR13	.4034 (85) P= .001	.1980 (85) P= .035	.1842 (70) P= .063	.1828 (81) P= .051	.3548 (81) P= .001	.3509 (80) P= .001
WISCR14	.6337 (86) P= .001	.5303 (86) P= .001	.5988 (71) P= .001	.4385 (82) P= .001	.7103 (81) P= .001	.6983 (81) P= .001
WISCR15	.4210 (86) P= .001	.2093 (86) P= .027	.1767 (71) P= .070	.1891 (82) P= .044	.3666 (81) P= .001	.3743 (81) P= .001
WISCR16	.6390 (86) P= .001	.4480 (86) P= .001	.4840 (71) P= .001	.3872 (82) P= .001	.6588 (81) P= .001	.6476 (81) P= .001

	KABC1	KABC2	KABC3	KABC4	KABC5	KABC6	KABC7	KABC8	KABC17	KABC18
WISCR1	.0907 (86) P= .203	.2041 (86) P= .030	.1888 (86) P= .041	.0943 (86) P= .194	.1374 (86) P= .104	.2552 (86) P= .009	.1379 (86) P= .103	.1486 (86) P= .086	.4775 (86) P= .001	.4777 (86) P= .001
WISCR2	.1112 (86) P= .154	.2050 (86) P= .029	.0502 (86) P= .323	.1226 (86) P= .130	.0229 (86) P= .417	.2629 (86) P= .007	.1290 (86) P= .118	.0824 (86) P= .225	.3314 (86) P= .001	.4217 (86) P= .001
WISCR3	.2698 (86) P= .006	.2349 (86) P= .015	.2220 (86) P= .020	.2661 (86) P= .007	.3991 (86) P= .001	.2161 (86) P= .023	.2308 (86) P= .016	.3874 (86) P= .001	.1573 (86) P= .074	.6277 (86) P= .001
WISCR4	.0697 (86) P= .262	.1199 (86) P= .136	.0512 (86) P= .320	.0242 (86) P= .412	.1207 (86) P= .134	.2498 (86) P= .010	.0454 (86) P= .339	.1764 (86) P= .052	.4707 (86) P= .001	.3664 (86) P= .001
WISCR5	.0642 (86) P= .279	.1421 (86) P= .096	.1054 (86) P= .167	.0397 (86) P= .358	.1197 (86) P= .136	.3038 (86) P= .002	.0537 (86) P= .312	.1453 (86) P= .091	.3332 (86) P= .001	.3525 (86) P= .001
WISCR6	.3425 (68) P= .002	.2061 (68) P= .046	.4825 (68) P= .001	.3030 (68) P= .006	.4625 (68) P= .001	.1676 (68) P= .086	.3952 (68) P= .001	.1965 (68) P= .054	.3931 (68) P= .001	.3955 (68) P= .001
WISCR7	.1229 (86) P= .130	.2099 (86) P= .026	.1982 (86) P= .034	.0968 (86) P= .188	.2155 (86) P= .023	.2485 (86) P= .011	.0763 (86) P= .242	.1766 (86) P= .052	.4812 (86) P= .001	.4725 (86) P= .001
WISCR8	.3241 (86) P= .001	.4251 (86) P= .001	.2452 (86) P= .011	.3923 (86) P= .001	.3482 (86) P= .001	.1134 (86) P= .149	.3517 (86) P= .001	.2417 (86) P= .012	.2613 (86) P= .008	.3419 (86) P= .001
WISCR9	.2047 (86) P= .029	.4005 (86) P= .001	.1498 (86) P= .084	.4368 (86) P= .001	.3859 (86) P= .001	.3114 (86) P= .002	.2178 (86) P= .022	.2729 (86) P= .006	.2020 (86) P= .031	.3552 (86) P= .001
WISCR10	.2315 (86) P= .016	.3481 (86) P= .001	.2320 (86) P= .016	.5926 (86) P= .001	.2538 (86) P= .009	.3680 (86) P= .001	.2876 (86) P= .004	.3316 (86) P= .001	.0660 (86) P= .273	.2672 (86) P= .006
WISCR11	.3099 (84) P= .002	.4136 (84) P= .001	.1822 (84) P= .049	.5087 (84) P= .001	.3057 (84) P= .002	.1164 (84) P= .146	.3101 (84) P= .002	.2809 (84) P= .005	.0521 (84) P= .319	.1032 (84) P= .175
WISCR12	.1071 (84) P= .166	.1094 (84) P= .161	-.0050 (84) P= .482	.2575 (84) P= .009	.1880 (84) P= .043	.1963 (84) P= .037	.2737 (84) P= .006	.0256 (84) P= .409	.1700 (84) P= .061	.2296 (84) P= .018
WISCR13	.4022 (85) P= .001	.5476 (85) P= .001	.2526 (85) P= .010	.6474 (85) P= .001	.4794 (85) P= .001	.2968 (85) P= .003	.4641 (85) P= .001	.3360 (85) P= .001	.2310 (85) P= .017	.3864 (85) P= .001
WISCR14	.1980 (86) P= .034	.2716 (86) P= .006	.1716 (86) P= .057	.1521 (86) P= .081	.2245 (86) P= .019	.3178 (86) P= .001	.1730 (86) P= .056	.2531 (86) P= .009	.4781 (86) P= .001	.5973 (86) P= .001
WISCR15	.4030 (86) P= .001	.5494 (86) P= .001	.2728 (86) P= .006	.6636 (86) P= .001	.4936 (86) P= .001	.3093 (86) P= .002	.4488 (86) P= .001	.3483 (86) P= .001	.2328 (86) P= .016	.3890 (86) P= .001
WISCR16	.3606 (86) P= .001	.4838 (86) P= .001	.2615 (86) P= .008	.4659 (86) P= .001	.4277 (86) P= .001	.3792 (86) P= .001	.3538 (86) P= .001	.3443 (86) P= .001	.4423 (86) P= .001	.6058 (86) P= .001

	KABC19	KABC20	KABC21	KABC22	KABC27	KABC28	KABC29	KABC30
WISCR1	.5700 (86) P= .001	.4208 (86) P= .001	.4213 (76) P= .001	.4719 (81) P= .001	.1782 (86) P= .050	.2848 (86) P= .004	.2666 (86) P= .007	.5504 (86) P= .001
WISCR2	.4833 (86) P= .001	.3239 (86) P= .001	.2092 (76) P= .035	.2278 (81) P= .020	.0615 (86) P= .287	.2347 (86) P= .015	.1872 (86) P= .042	.3949 (86) P= .001
WISCR3	.3938 (86) P= .001	.4158 (86) P= .001	.2160 (76) P= .030	.2630 (81) P= .009	.3325 (86) P= .001	.3880 (86) P= .001	.4228 (86) P= .001	.4186 (86) P= .001
WISCR4	.6446 (86) P= .001	.4165 (86) P= .001	.3988 (76) P= .001	.4945 (81) P= .001	.0999 (86) P= .180	.1694 (86) P= .059	.1605 (86) P= .070	.5262 (86) P= .001
WISCR5	.6706 (86) P= .001	.3965 (86) P= .001	.3226 (76) P= .002	.2907 (81) P= .004	.1266 (86) P= .123	.2179 (86) P= .022	.2025 (86) P= .031	.4919 (86) P= .001
WISCR6	.2625 (68) P= .015	.2423 (68) P= .023	.3070 (60) P= .009	.3902 (63) P= .001	.5578 (68) P= .001	.4073 (68) P= .001	.5577 (68) P= .001	.2851 (68) P= .009
WISCR7	.7224 (86) P= .001	.4709 (86) P= .001	.3899 (76) P= .001	.4632 (81) P= .001	.2122 (86) P= .025	.2429 (86) P= .012	.2658 (86) P= .007	.5796 (86) P= .001
WISCR8	.4206 (86) P= .001	.3058 (86) P= .002	.3396 (76) P= .001	.2342 (81) P= .018	.3764 (86) P= .001	.5194 (86) P= .001	.5294 (86) P= .001	.3711 (86) P= .001
WISCR9	.3926 (86) P= .001	.2067 (86) P= .028	.2289 (76) P= .023	.2109 (81) P= .029	.2672 (86) P= .006	.4789 (86) P= .001	.4445 (86) P= .001	.1766 (86) P= .052
WISCR10	.1945 (86) P= .036	.1269 (86) P= .122	.0727 (76) P= .266	.0268 (81) P= .406	.2695 (86) P= .006	.6274 (86) P= .001	.5428 (86) P= .001	.2192 (86) P= .021
WISCR11	.2090 (84) P= .028	.1193 (84) P= .140	.0210 (74) P= .430	.0731 (79) P= .261	.2694 (84) P= .007	.5109 (84) P= .001	.4669 (84) P= .001	.0418 (84) P= .353
WISCR12	.1435 (84) P= .096	.2217 (84) P= .021	.1861 (75) P= .055	.2506 (79) P= .013	.1599 (84) P= .073	.2953 (84) P= .003	.2752 (84) P= .006	.1647 (84) P= .067
WISCR13	.4279 (85) P= .001	.2634 (85) P= .007	.2188 (75) P= .030	.2437 (80) P= .015	.4394 (85) P= .001	.7419 (85) P= .001	.7032 (85) P= .001	.2782 (85) P= .005
WISCR14	.7588 (86) P= .001	.5168 (86) P= .001	.4134 (76) P= .001	.4596 (81) P= .001	.2354 (86) P= .015	.3557 (86) P= .001	.3484 (86) P= .001	.6346 (86) P= .001
WISCR15	.4325 (86) P= .001	.2633 (86) P= .007	.2101 (76) P= .034	.2417 (81) P= .015	.4526 (86) P= .001	.7460 (86) P= .001	.7138 (86) P= .001	.2845 (86) P= .004
WISCR16	.7295 (86) P= .001	.4790 (86) P= .001	.3815 (76) P= .001	.4411 (81) P= .001	.4114 (86) P= .001	.6404 (86) P= .001	.6243 (86) P= .001	.5688 (86) P= .001

	KABC1	KABC2	KABC3	KABC4	KABC5	KABC6	KABC7	KABC8	KABC17	KABC18
PIAT7	.4578 (86) P= .001	.1738 (86) P= .055	.2782 (86) P= .005	.2826 (86) P= .004	.4261 (86) P= .001	.2606 (86) P= .008	.4165 (86) P= .001	.2887 (86) P= .004	.3504 (86) P= .001	.6168 (86) P= .001
PIAT8	.2111 (86) P= .026	.1352 (86) P= .107	.2884 (86) P= .004	.1800 (86) P= .049	.2272 (86) P= .018	.2593 (86) P= .008	.0878 (86) P= .211	.1692 (86) P= .060	.3487 (86) P= .001	.4155 (86) P= .001
PIAT9	.0922 (71) P= .222	.1983 (71) P= .049	.1356 (71) P= .130	.1270 (71) P= .146	.0869 (71) P= .236	.3021 (71) P= .005	.0502 (71) P= .339	.0745 (71) P= .268	.3152 (71) P= .004	.4195 (71) P= .001
PIAT10	.2064 (82) P= .031	.1132 (82) P= .156	.2294 (82) P= .019	.1933 (82) P= .041	.2723 (82) P= .007	.3112 (82) P= .002	.1607 (82) P= .075	.1754 (82) P= .058	.3570 (82) P= .001	.4560 (82) P= .001
PIAT11	.0849 (81) P= .226	.3085 (81) P= .003	.1521 (81) P= .088	.2199 (81) P= .024	.1587 (81) P= .079	.2396 (81) P= .016	.3348 (81) P= .001	.1410 (81) P= .105	.5656 (81) P= .001	.3962 (81) P= .001
PIAT12	.2886 (81) P= .004	.2785 (81) P= .006	.2465 (81) P= .013	.2530 (81) P= .011	.3646 (81) P= .001	.3416 (81) P= .001	.3250 (81) P= .002	.2997 (81) P= .003	.5167 (81) P= .001	.6009 (81) P= .001

- - - - - - - - - - - - P E A R S O N   C O R R E L A T I O N   C O E F F I C I E N T S - - -

	KABC19	KABC20	KABC21	KABC22	KABC27	KABC28	KABC29	KABC30
PIAT7	.4721 (86) P= .001	.4667 (86) P= .001	.3416 (76) P= .001	.3339 (81) P= .001	.4111 (86) P= .001	.3713 (86) P= .001	.4514 (86) P= .001	.4946 (86) P= .001
PIAT8	.5154 (86) P= .001	.6545 (86) P= .001	.6079 (76) P= .001	.3116 (81) P= .002	.3365 (86) P= .001	.3184 (86) P= .001	.3788 (86) P= .001	.6791 (86) P= .001
PIAT9	.4028 (71) P= .001	.6432 (71) P= .001	.7148 (69) P= .001	.3729 (66) P= .001	.1414 (71) P= .120	.2633 (71) P= .013	.2518 (71) P= .017	.5859 (71) P= .001
PIAT10	.3780 (82) P= .001	.5439 (82) P= .001	.2835 (74) P= .007	.3570 (77) P= .001	.3269 (82) P= .001	.3102 (82) P= .002	.3730 (82) P= .001	.5299 (82) P= .001
PIAT11	.6828 (81) P= .001	.3214 (81) P= .002	.2050 (73) P= .041	.5333 (76) P= .001	.1871 (81) P= .047	.4129 (81) P= .001	.3682 (81) P= .001	.5062 (81) P= .001
PIAT12	.6479 (81) P= .001	.6675 (81) P= .001	.5063 (73) P= .001	.5138 (76) P= .001	.3791 (81) P= .001	.4686 (81) P= .001	.5065 (81) P= .001	.7214 (81) P= .001

- - - - - - - - - - - - P E A R S O N   C O R R E L A T I O N   C O E F F I C I E N T S - - - - - - - - -

	KABC27	KABC28	KABC29	KABC30
KABC17	.2730	.2081	.2784	.5124
	(86)	(86)	(86)	(86)
	P= .005	P= .027	P= .005	P= .001
KABC18	.3667	.3801	.4402	.6616
	(86)	(86)	(86)	(86)
	P= .001	P= .001	P= .001	P= .001
KABC19	.3474	.4208	.4486	.6629
	(86)	(86)	(86)	(86)
	P= .001	P= .001	P= .001	P= .001
KABC20	.2499	.3117	.3341	.7337
	(86)	(86)	(86)	(86)
	P= .010	P= .002	P= .001	P= .001
KABC21	.1594	.2513	.2463	.6432
	(76)	(76)	(76)	(76)
	P= .085	P= .014	P= .016	P= .001
KABC22	.2322	.2283	.2729	.5049
	(81)	(81)	(81)	(81)
	P= .018	P= .020	P= .007	P= .001

(COEFFICIENT / CASES / SIGNFICANCE)   (99.0000 MEANS UNCOMPUTABLE)

APPENDIX E

LISTING OF Z_{DIF}'S

CASE-NO	DEM2	KABC29	WISCR16	ZDIF1	ZDIF8	ZTDIF1	ZTDIF8	REGRES1	REGRES8	REYREG1	REYREG8
1	1.	71.	81.	-1.095	-3.671	-1.707	-3.925	-.083	-1.476	-.233	-3.420
2	5.	82.	93.	1.217	-3.059	.923	-3.197	1.059	-1.399	2.958	-3.242
3	6.	100.	78.	.730	3.059	-.063	2.981	1.247	1.893	3.483	4.387
4	7.	87.	92.	.243	-1.376	-.044	-1.497	.439	-.495	1.225	-1.147
5	9.	100.	107.	-.974	-2.141	-.690	-2.087	-.895	-1.325	-2.500	-3.071
6	11.	79.	78.	-.487	.153	-1.228	-.103	.429	.671	1.198	1.555
7	12.	106.	96.	-.243	2.906	-.372	2.904	-.026	1.634	-.073	3.786
8	13.	93.	87.	-.122	1.988	-.567	1.854	.365	1.423	1.019	3.297
9	15.	97.	103.	.487	-1.682	.570	-1.676	.224	-.959	.626	-2.222
10	16.	86.	84.	1.217	1.835	.611	1.621	1.368	1.520	3.821	3.523
11	18.	98.	91.	-.365	1.224	-.661	1.168	.064	.812	.178	1.882
12	19.	108.	87.	-1.339	3.977	-1.732	3.971	-.453	2.242	-1.266	5.195
13	21.	99.	89.	1.217	3.059	.784	2.969	1.196	1.921	3.341	4.451
14	22.	95.	101.	2.313	1.682	2.249	1.580	1.521	1.179	4.247	2.731
15	27.	119.	116.	3.530	3.365	3.934	3.507	1.823	1.561	5.092	3.617
16	28.	101.	107.	2.191	0	2.340	.012	1.232	-.027	3.442	-.064
17	29.	109.	109.	.730	.918	1.011	1.002	.182	.321	.507	.744
18	33.	75.	90.	0	-1.071	-.346	-1.343	.344	.024	.960	.055
19	35.	106.	89.	-2.069	2.294	-2.362	2.308	-1.013	1.255	-2.830	2.909
20	46.	77.	84.	.243	-.765	-.321	-1.021	.714	.158	1.993	.366
21	47.	96.	102.	.852	-.153	.885	-.197	.504	.015	1.408	.035
22	48.	85.	86.	2.069	.459	1.496	.267	1.872	.696	5.229	1.612
23	49.	108.	96.	-.852	1.682	-.954	1.736	-.435	.822	-1.216	1.904
24	50.	122.	101.	-.365	4.894	-.315	5.034	-.280	2.425	-.782	5.620
25	51.	107.	109.	1.582	1.682	1.827	1.724	.755	.849	2.107	1.968
26	52.	92.	97.	-.243	-1.988	-.337	-2.034	-.061	-1.011	-.169	-2.343
27	53.	110.	111.	.609	.153	.964	.269	.031	-.180	.087	-.417
28	55.	82.	87.	.609	-.765	.132	-.961	.856	.021	2.391	.048
29	56.	101.	97.	.974	-.153	.828	-.137	.758	-.122	2.116	-.283
30	58.	100.	96.	-.243	-2.906	-.372	-2.832	-.026	-1.799	-.073	-4.168
31	60.	100.	105.	-.487	1.224	-.293	1.192	-.499	.757	-1.394	1.755
32	62.	101.	105.	1.582	1.988	1.688	1.950	.892	1.203	2.491	2.788
33	63.	75.	93.	.974	2.600	.690	2.234	.895	2.296	2.500	5.320
34	65.	104.	92.	-.609	2.141	-.860	2.135	-.134	1.216	-.375	2.817
35	66.	97.	84.	.730	2.294	.145	2.200	1.041	1.502	2.907	3.481
36	67.	106.	125.	3.043	2.294	3.779	2.308	1.187	1.255	3.314	2.909
37	68.	89.	107.	.609	.306	.825	.166	.169	.491	.471	1.139
38	69.	66.	84.	.243	-1.835	-.321	-2.196	.714	-.203	1.993	-.469
39	72.	88.	81.	-.243	1.835	-.891	1.645	.489	1.465	1.367	3.396
40	73.	108.	110.	-.974	1.071	-.586	1.139	-.998	.443	-2.788	1.027
41	75.	100.	94.	-.730	.918	-.907	.894	-.285	.568	-.795	1.316
42	76.	93.	91.	.852	2.906	.504	2.748	.882	1.991	2.464	4.613
43	77.	117.	81.	1.461	5.659	.740	5.719	1.635	3.036	4.567	7.035
44	80.	83.	88.	1.704	1.682	1.216	1.436	1.558	1.508	4.352	3.495
45	82.	81.	90.	-.852	-.459	-1.162	-.675	-.229	.238	-.640	.551
46	83.	86.	88.	-.122	-.765	-.532	-.913	.331	-.089	.923	-.206
47	84.	93.	105.	.243	-1.529	.406	-1.575	-.008	-.755	-.023	-1.748
48	85.	99.	92.	1.217	2.447	.888	2.373	1.093	1.542	3.054	3.574
49	86.	80.	78.	-2.191	-2.600	-2.860	-2.774	-.717	-1.060	-2.002	-2.457
50	87.	115.	102.	1.095	3.365	1.118	3.459	.668	1.671	1.865	3.872
51	88.	103.	101.	-.609	.459	-.548	.483	-.444	.202	-1.239	.467
52	91.	72.	76.	-.974	-1.682	-1.764	-1.975	.170	-.273	.475	-.632
53	92.	99.	103.	2.556	1.835	2.551	1.777	1.615	1.163	4.512	2.696
54	93.	81.	75.	.365	3.059	-.516	2.753	1.105	2.415	3.086	5.596
55	96.	80.	76.	-.852	2.447	-1.647	2.145	.252	2.064	.704	4.782
56	97.	99.	91.	-.487	3.824	-.778	3.714	-.018	2.394	-.050	5.548
57	98.	86.	90.	-.243	-.765	-.579	-.913	.180	-.089	.503	-.206
58	99.	92.	99.	1.217	-.918	1.131	-.990	.853	-.348	2.382	-.807
59	100.	100.	92.	0	3.212	-.277	3.130	.275	1.988	.768	4.607
60	101.	88.	74.	-.609	2.141	-1.483	1.943	.484	1.655	1.353	3.835
61	103.	87.	88.	.122	-.153	-.299	-.305	.494	.262	1.380	.608
62	104.	115.	115.	1.217	2.447	1.685	2.565	.303	1.103	.846	2.556
63	106.	102.	91.	-1.339	3.365	-1.594	3.303	-.591	2.028	-1.650	4.699
64	107.	98.	100.	.487	1.988	.466	1.914	.327	1.286	.914	2.979
65	108.	105.	90.	.974	4.741	.586	4.681	.998	2.797	2.788	6.483
66	111.	82.	72.	-2.556	1.835	-3.417	1.573	-.756	1.630	-2.112	3.778
67	113.	85.	86.	-.122	.612	-.602	.416	.399	.790	1.115	1.832
68	114.	110.	107.	.365	2.906	.592	2.952	.005	1.524	.014	3.532
69	115.	107.	114.	.974	2.600	1.417	2.618	.173	1.417	.485	3.284
70	117.	97.	93.	-.852	1.071	-1.058	1.007	-.332	.745	-.928	1.726
71	120.	108.	92.	-.122	2.141	-.394	2.183	.193	1.106	.539	2.562

CASE-NO	DEM2	KABC29	WISCR16	ZDIF1	ZDIF8	ZTDIF1	ZTDIF8	REGRES1	REGRES8	REYREG1	REYREG8
72	121.	100.	106.	-.122	1.224	.091	1.192	-.288	.757	-.805	1.755
73	123.	86.	95.	-.365	-.612	-.523	-.764	-.074	.006	-.206	.013
74	124.	96.	99.	.243	-1.529	.198	-1.539	.198	-.837	.553	-1.939
75	133.	88.	80.	-1.582	-.612	-2.208	-.740	-.376	-.049	-1.051	-.114
76	135.	123.	107.	.609	4.741	.825	4.897	.169	2.303	.471	5.337
77	138.	104.	105.	.122	-.459	.290	-.399	-.090	-.394	-.251	-.913
78	139.	95.	105.	.852	0	.989	-.060	.401	.137	1.120	.318
79	142.	123.	114.	1.217	3.059	1.650	3.257	.337	1.262	.942	2.924
80	143.	117.	105.	0	2.294	.173	2.440	-.172	.953	-.480	2.209
81	144.	105.	100.	-.487	.918	-.466	.954	-.327	.431	-.914	.998
82	148.	74.	93.	-.243	1.071	-.476	.732	.077	1.376	.215	3.190
83	149.	87.	103.	-.609	-1.529	-.479	-1.646	-.512	-.590	-1.431	-1.367
84	150.	92.	87.	-1.095	.612	-1.499	.500	-.290	.598	-.809	1.386
85	151.	81.	95.	1.461	-.612	1.225	-.824	1.154	.143	3.223	.331
86	157.	102.	107.	1.582	1.071	1.757	1.067	.823	.608	2.299	1.408

CASE-NO	DEM2	KABC29	WISCR16	ZDIF2	ZDIF10	ZTDIF2	ZTDIF10	REGRES2	REGRES10	REYREG2	REYREG10
1	1.	71.	81.	-.172	-1.069	-.420	-1.170	.571	.427	1.004	.722
2	5.	82.	93.	3.271	1.069	3.109	.990	1.796	1.091	3.157	1.847
3	6.	100.	78.	1.033	3.207	.719	3.177	1.247	1.529	2.193	2.589
4	7.	87.	92.	.861	.178	.736	.127	.684	.505	1.203	.854
5	9.	100.	107.	-1.033	-2.138	-.918	-2.118	-.732	-1.020	-1.287	-1.726
6	11.	79.	78.	1.033	1.069	.719	.979	1.247	1.188	2.193	2.011
7	12.	106.	96.	0	3.029	-.053	3.024	.137	1.251	.242	2.117
8	13.	93.	87.	.344	7.127	.165	7.034	.611	3.625	1.074	6.135
9	15.	97.	103.	1.549	1.425	1.556	1.401	.633	.777	1.114	1.314
10	16.	86.	84.	-1.033	-.535	-1.223	-.583	.059	.197	.104	.334
11	18.	98.	91.	.172	1.782	.049	1.758	.391	.914	.688	1.547
12	19.	108.	87.	1.033	4.276	.839	4.267	.938	1.781	1.649	3.014
13	21.	99.	89.	.516	6.592	.360	6.527	.624	3.176	1.097	5.376
14	22.	95.	101.	1.893	2.673	1.867	2.629	.866	1.436	1.523	2.430
15	27.	119.	116.	6.541	7.127	6.615	7.134	2.560	2.785	4.501	4.714
16	28.	101.	107.	4.648	3.563	4.642	3.534	1.969	1.667	3.462	2.822
17	29.	109.	109.	3.098	4.454	3.152	4.448	1.164	1.834	2.046	3.103
18	33.	75.	90.	1.721	0	1.552	-.096	1.162	.807	2.044	1.366
19	35.	106.	89.	-2.238	4.454	-2.336	4.436	-.686	1.930	-1.206	3.267
20	46.	77.	84.	0	.178	-.212	.088	.550	.828	.967	1.401
21	47.	96.	102.	-.172	-.356	-.142	-.368	-.151	-.041	-.265	-.069
22	48.	85.	86.	2.066	2.316	1.836	2.237	1.463	1.589	2.573	2.689
23	49.	108.	96.	.172	4.633	.115	4.620	.219	1.951	.386	3.302
24	50.	122.	101.	-.689	5.345	-.661	5.380	-.362	1.839	-.636	3.112
25	51.	107.	109.	2.238	3.563	2.310	3.557	.755	1.473	1.327	2.494
26	52.	92.	97.	.861	1.425	.803	1.381	.512	.938	.901	1.588
27	53.	110.	111.	2.410	3.742	2.505	3.745	.768	1.461	1.350	2.474
28	55.	82.	87.	-.344	0	-.509	-.069	.283	.581	.498	.984
29	56.	101.	97.	1.377	2.673	1.308	2.652	.758	1.242	1.333	2.103
30	58.	100.	96.	-.861	-.178	-.896	-.177	-.272	-.085	-.478	-.144
31	60.	100.	105.	2.238	2.138	2.257	2.118	.892	1.020	1.569	1.726
32	62.	101.	105.	1.721	1.782	1.751	1.769	.647	.817	1.137	1.383
33	63.	75.	93.	.689	3.742	.581	3.611	.568	2.592	.999	4.386
34	65.	104.	92.	2.238	6.058	2.084	6.017	1.339	2.760	2.354	4.671
35	66.	97.	84.	-1.549	3.029	-1.729	2.989	-.187	1.541	-.328	2.609
36	67.	106.	125.	5.164	3.207	5.386	3.200	1.596	1.336	2.806	2.261
37	68.	89.	107.	2.582	0	2.620	-.042	.987	.355	1.736	.601
38	69.	66.	84.	-.344	-3.563	-.549	-3.661	.386	-.602	.679	-1.018
39	72.	88.	81.	-.861	1.247	-1.094	1.190	.244	.982	.429	1.662
40	73.	108.	110.	.689	.178	.807	.207	-.016	-.173	-.029	-.293
41	75.	100.	94.	.516	1.425	.426	1.412	.452	.680	.794	1.151
42	76.	93.	91.	1.377	3.385	1.229	3.327	.964	1.840	1.695	3.115
43	77.	117.	81.	-4.820	3.920	-4.970	3.949	-1.638	1.320	-2.881	2.235
44	80.	83.	88.	0	.178	-.159	.111	.412	.634	.725	1.073
45	82.	81.	90.	-2.066	-2.316	-2.155	-2.368	-.638	-.491	-1.122	-.831
46	83.	86.	88.	-1.377	-1.782	-1.507	-1.819	-.242	-.398	-.426	-.673
47	84.	93.	105.	2.238	.178	2.257	.150	.892	.311	1.569	.526
48	85.	99.	92.	1.721	3.029	1.579	2.997	1.093	1.477	1.923	2.499
49	86.	80.	78.	-2.582	.178	-2.819	.100	-.471	.731	-.829	1.237
50	87.	115.	102.	.344	4.454	.364	4.471	.095	1.640	.167	2.776
51	88.	103.	101.	.344	2.673	.350	2.659	.129	1.178	.227	1.993
52	91.	72.	76.	-1.033	-1.425	-1.329	-1.520	.334	.224	.587	.380
53	92.	99.	103.	-.344	-1.247	-.297	-1.239	-.267	-.562	-.469	-.952
54	93.	81.	75.	0	2.316	-.331	2.222	.859	1.718	1.511	2.908
55	96.	80.	76.	-1.721	8.374	-2.003	8.220	.007	4.639	.012	7.852
56	97.	99.	91.	1.033	3.563	.892	3.527	.800	1.732	1.407	2.931
57	98.	86.	90.	1.377	.356	1.215	.299	.998	.622	1.756	1.053
58	99.	92.	99.	2.926	1.247	2.851	1.205	1.426	.853	2.507	1.444
59	100.	100.	92.	1.721	3.029	1.579	3.001	1.093	1.444	1.923	2.445
60	101.	88.	74.	.861	2.494	.498	2.425	1.303	1.577	2.291	2.669
61	103.	87.	88.	-1.033	.178	-1.170	.127	-.079	.505	-.138	.854
62	104.	115.	115.	4.648	6.058	4.748	6.059	1.694	2.405	2.979	4.070
63	106.	102.	91.	-.689	3.029	-.793	3.009	-.018	1.380	-.032	2.336
64	107.	98.	100.	2.582	2.494	2.527	2.464	1.228	1.254	2.159	2.123
65	108.	105.	90.	.689	5.523	.541	5.491	.671	2.473	1.180	4.185
66	111.	82.	72.	-2.754	3.029	-3.067	2.932	-.347	2.026	-.610	3.428
67	113.	85.	86.	-2.066	.713	-2.208	.649	-.501	.824	-.881	1.395
68	114	110	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	2.141	1.448	3.624
69	115.	107.	111.	2.410	2.316	2.545	2.322	.665	.879	1.169	1.487
70	117.	97.	93.	-1.377	.535	-1.411	.518	.414	.352	-.728	.595
71	120.	108.	92.	.516	4.811	.399	[illegible]	.520	2.036	.915	3.446

CASE-NO	DEM2	KABC29	WISCR16	ZDIF2	ZDIF10	ZTDIF2	ZTDIF10	REGRES2	REGRES10	REYREG2	REYREG10
72	121.	100.	106.	2.066	1.604	2.102	1.589	.776	.765	1.364	1.294
73	123.	86.	95.	0	-.356	-.066	-.407	.172	.282	.302	.477
74	124.	96.	99.	.516	1.069	.492	1.044	.280	.639	.492	1.081
75	133.	88.	80.	-2.926	-2.673	-3.130	-2.694	-.704	-.887	-1.238	-1.501
76	135.	123.	107.	.861	4.276	.935	4.325	.169	1.297	.296	2.195
77	138.	104.	105.	1.377	1.425	1.414	1.428	.483	.551	.849	.932
78	139.	95.	105.	.689	-4.454	.740	-4.432	.155	-1.963	.273	-3.322
79	142.	123.	114.	3.443	6.236	3.556	6.266	1.156	2.231	2.032	3.777
80	143.	117.	105.	.344	4.098	.403	4.125	-.008	1.405	-.014	2.379
81	144.	105.	100.	-.344	1.960	-.337	1.961	-.164	.773	-.288	1.309
82	148.	74.	93.	-.344	-.356	-.430	-.453	.077	.670	.135	1.133
83	149.	87.	103.	-.689	-4.098	-.634	-4.110	-.430	-1.535	-.757	-2.597
84	150.	92.	87.	-.689	2.851	-.846	2.794	.119	1.618	.210	2.738
85	151.	81.	95.	.344	-.891	.271	-.956	.336	.189	.590	.319
86	157.	102.	107.	1.205	1.604	1.272	1.596	.332	.700	.584	1.185

CASE-NO	DEM2	KABC29	WISCR16	ZDIF3	ZDIF11	ZTDIF3	ZTDIF11	REGRES3	REGRES11	REYREG3	REYREG11
1	1.	71.	81.	-1.054	-1.205	-1.938	-1.351	-.218	.298	-.959	.533
2	5.	82.	93.	2.214	2.582	1.739	2.447	1.996	1.864	8.783	3.331
3	6.	100.	78.	.316	5.164	-.795	5.096	.974	2.603	4.286	4.652
4	7.	87.	92.	1.476	.861	.994	.777	1.440	.840	6.336	1.501
5	9.	100.	107.	-1.897	-2.926	-1.441	-2.888	-1.744	-1.475	-7.673	-2.636
6	11.	79.	78.	.105	.516	-.994	.392	.806	.916	3.546	1.638
7	12.	106.	96.	.316	3.098	.099	3.091	.384	1.374	1.688	2.456
8	13.	93.	87.	9.000	1.033	9.000	.980	9.000	.739	9.000	1.321
9	15.	97.	103.	.738	.861	.845	.833	.490	.528	2.158	.943
10	16.	86.	84.	-.949	-1.721	-1.689	-1.777	-.232	-.430	-1.022	-.769
11	18.	98.	91.	-.738	2.238	-1.143	2.197	-.294	1.190	-1.292	2.128
12	19.	108.	87.	.527	3.959	-.149	3.952	.847	1.746	3.727	3.120
13	21.	99.	89.	9.000	2.582	9.000	2.542	9.000	1.333	9.000	2.382
14	22.	95.	101.	.422	2.066	.447	2.010	.304	1.197	1.336	2.140
15	27.	119.	116.	3.584	8.090	4.174	8.090	2.336	3.484	10.274	6.228
16	28.	101.	107.	2.424	3.271	2.634	3.233	1.705	1.617	7.502	2.891
17	29.	109.	109.	1.476	3.271	1.839	3.278	.883	1.367	3.883	2.444
18	33.	75.	90.	9.000	4.303	9.000	4.107	9.000	2.950	9.000	5.273
19	35.	106.	89.	9.000	9.639	9.000	9.546	9.000	4.671	9.000	8.350
20	46.	77.	84.	0	0	-.795	-.129	.525	.718	2.310	1.284
21	47.	96.	102.	-.527	-.344	-.398	-.362	-.486	-.049	-2.139	-.087
22	48.	85.	86.	1.476	2.410	.696	2.294	1.637	1.683	7.202	3.009
23	49.	108.	96.	.105	3.443	-.099	3.442	.215	1.485	.947	2.655
24	50.	122.	101.	.316	5.508	.348	5.559	.220	2.089	.966	3.734
25	51.	107.	109.	2.214	3.615	2.534	3.606	1.471	1.603	6.473	2.866
26	52.	92.	97.	9.000	.861	9.000	.805	9.000	.684	9.000	1.222
27	53.	110.	111.	2.108	4.131	2.534	4.133	1.322	1.770	5.815	3.164
28	55.	82.	87.	-.632	-.516	-1.242	-.610	-.078	.302	-.344	.540
29	56.	101.	97.	.949	3.615	.745	3.573	.856	1.791	3.764	3.201
30	58.	100.	96.	.211	-.172	0	-.170	.300	-.087	1.318	-.155
31	60.	100.	105.	9.000	9.000	9.000	9.000	9.000	9.000	9.000	9.000
32	62.	101.	105.	.211	3.271	.447	3.233	.004	1.617	.019	2.891
33	63.	75.	93.	9.000	9.000	9.000	9.000	9.000	9.000	9.000	9.000
34	65.	104.	92.	2.108	6.541	1.590	6.477	1.945	3.172	8.557	5.670
35	66.	97.	84.	-1.054	3.959	-1.789	3.890	-.316	2.089	-1.392	3.734
36	67.	106.	125.	2.635	3.787	3.727	3.771	1.283	1.721	5.644	3.077
37	68.	89.	107.	.316	1.549	.646	1.467	.023	1.124	.100	2.010
38	69.	66.	84.	9.000	9.000	9.000	9.000	9.000	9.000	9.000	9.000
39	72.	88.	81.	-.105	2.238	-1.043	2.141	.539	1.503	2.372	2.686
40	73.	108.	110.	.105	2.582	.596	2.593	-.244	1.052	-1.073	1.880
41	75.	100.	94.	1.054	3.098	.696	3.058	1.038	1.562	4.567	2.791
42	76.	93.	91.	2.741	7.402	2.137	7.265	2.483	3.949	10.922	7.059
43	77.	117.	81.	-1.265	2.066	-2.137	2.134	-.386	.510	-1.699	.912
44	80.	83.	88.	.422	-.861	-.199	-.945	.730	.097	3.213	.174
45	82.	81.	90.	9.000	9.000	9.000	9.000	9.000	9.000	9.000	9.000
46	83.	86.	88.	-1.687	.344	-2.186	.261	-.952	.611	-4.190	1.092
47	84.	93.	105.	1.054	-.689	1.242	-.719	.677	-.128	2.979	-.230
48	85.	99.	92.	0	3.615	-.398	3.562	.262	1.853	1.155	3.313
49	86.	80.	78.	-.422	-1.549	-1.491	-1.641	.385	-.156	1.695	-.279
50	87.	115.	102.	.843	4.475	.894	4.501	.607	1.787	2.672	3.195
51	88.	103.	101.	.527	2.066	.547	2.055	.388	.947	1.706	1.693
52	91.	72.	76.	-2.214	-2.238	-3.280	-2.365	-.979	-.253	-4.308	-.453
53	92.	99.	103.	.422	0	.547	-.006	.238	.031	1.047	.056
54	93.	81.	75.	9.000	9.000	9.000	9.000	9.000	9.000	9.000	9.000
55	96.	80.	76.	-.211	0	-1.391	-.112	.619	.625	2.724	1.117
56	97.	99.	91.	1.054	3.615	.547	3.562	1.137	1.853	5.000	3.313
57	98.	86.	90.	.738	-1.721	.199	-1.777	.917	-.430	4.034	-.769
58	99.	92.	99.	1.054	1.205	.944	1.144	.874	.857	3.846	1.532
59	100.	100.	92.	.316	1.893	-.099	1.869	.515	.954	2.265	1.706
60	101.	88.	74.	.211	1.377	-1.093	1.292	1.021	1.069	4.493	1.911
61	103.	87.	88.	9.000	9.000	9.000	9.000	9.000	9.000	9.000	9.000
62	104.	115.	115.	2.214	4.648	2.832	4.671	1.275	1.874	5.607	3.350
63	106.	102.	91.	-.316	2.582	-.745	2.559	.043	1.239	.189	2.215
64	107.	98.	100.	.949	2.582	.894	2.537	.757	1.364	3.331	2.438
65	108.	105.	90.	.211	3.615	-.298	3.595	.496	1.666	2.184	2.978
66	111.	82.	72.	9.000	4.475	9.000	4.316	9.000	2.818	9.000	5.037
67	113.	85.	86.	-2.003	-.344	-2.584	-.424	-1.139	.295	-5.011	.527
68	114.	110.	107.	9.000	3.615	9.000	3.623	9.000	1.510	9.000	2.698
69	115.	107.	114.	2.214	2.582	2.783	2.587	1.307	1.083	5.752	1.935
70	117.	97.	93.	.316	1.721	-.050	1.682	.482	.961	2.121	1.718
71	120.	108.	92.	.316	5.680	-.099	5.651	.515	2.613	2.265	4.671

CASE-NO	DEM2	KABC29	WISCR16	ZDIF3	ZDIF11	ZTDIF3	ZTDIF11	REGRES3	REGRES11	REYREG3	REYREG11
72	121.	100.	106.	1.687	1.033	1.888	1.019	1.149	.521	5.056	.930
73	123.	86.	95.	.316	-1.549	.050	-1.607	.416	-.344	1.832	-.614
74	124.	96.	99.	.527	.861	.447	.827	.453	.559	1.995	.999
75	133.	88.	80.	-.527	9.000	-1.491	9.000	.236	9.000	1.036	9.000
76	135.	123.	107.	.527	9.000	.845	9.000	.191	9.000	.840	9.000
77	138.	104.	105.	.527	-1.033	.745	-.997	.257	-.646	1.129	-1.154
78	139.	95.	105.	.949	-1.205	1.143	-1.217	.593	-.451	2.609	-.806
79	142.	123.	114.	2.003	5.336	2.584	5.395	1.139	1.971	5.011	3.523
80	143.	117.	105.	1.160	3.959	1.342	4.002	.761	1.465	3.350	2.618
81	144.	105.	100.	-.422	.344	-.398	.368	-.337	.017	-1.480	.031
82	148.	74.	93.	9.000	9.000	9.000	9.000	9.000	9.000	9.000	9.000
83	149.	87.	103.	-.422	-2.926	-.248	-2.961	-.435	-1.069	-1.913	-1.911
84	150.	92.	87.	9.000	9.000	9.000	9.000	9.000	9.000	9.000	9.000
85	151.	81.	95.	.738	-1.549	.447	-1.635	.753	-.187	3.313	-.335
86	157.	102.	107.	.316	1.721	.646	1.710	.023	.805	.100	1.439

CASE-NO	DEM2	KABC29	WISCR16	ZDIF6	ZDIF12	ZTDIF6	ZTDIF12	REGRES6	REGRES12	REYREG6	REYREG12
1	1.	71.	81.	-.516	-2.667	-.757	-2.519	.346	-.442	.637	-.705
2	5.	82.	93.	2.754	-.222	2.603	-.101	1.596	.365	2.938	.582
3	6.	100.	78.	1.033	6.000	.719	6.126	1.214	2.676	2.235	4.266
4	7.	87.	92.	1.721	-.222	1.579	-.136	1.113	.236	2.048	.376
5	9.	100.	107.	-1.549	-3.556	-1.424	-3.630	-.995	-1.586	-1.832	-2.528
6	11.	79.	78.	.689	.444	.382	.601	1.042	.739	1.919	1.179
7	12.	106.	96.	.516	4.222	.452	4.269	.385	1.729	.708	2.756
8	13.	93.	87.	.516	3.333	.333	3.452	.670	1.667	1.234	2.658
9	15.	97.	103.	1.205	.667	1.219	.702	.506	.375	.931	.597
10	16.	86.	84.	1.205	.444	.967	.552	1.109	.559	2.042	.891
11	18.	98.	91.	.344	4.444	.218	4.552	.458	2.034	.842	3.243
12	19.	108.	87.	0	5.333	-.172	5.389	.413	2.173	.760	3.464
13	21.	99.	89.	1.721	4.889	1.539	4.998	1.208	2.206	2.224	3.517
14	22.	95.	101.	2.066	2.222	2.035	2.304	.998	1.120	1.838	1.786
15	27.	119.	116.	5.164	7.111	5.267	7.127	2.067	2.682	3.806	4.276
16	28.	101.	107.	3.271	2.444	3.294	2.489	1.409	1.065	2.594	1.697
17	29.	109.	109.	2.066	2.889	2.141	2.886	.744	1.057	1.370	1.684
18	33.	75.	90.	0	-.889	-.133	-.732	.318	.248	.585	.395
19	35.	106.	89.	-1.205	6.000	-1.325	6.084	-.252	2.522	-.463	4.020
20	46.	77.	84.	.861	.222	.630	.388	.937	.692	1.726	1.103
21	47.	96.	102.	.861	1.556	.869	1.616	.366	.797	.673	1.270
22	48.	85.	86.	2.926	3.111	2.679	3.282	1.904	1.774	3.506	2.829
23	49.	108.	96.	.172	4.667	.115	4.708	.213	1.875	.392	2.990
24	50.	122.	101.	.172	7.111	.182	7.106	.054	2.605	.100	4.153
25	51.	107.	109.	2.754	-1.778	2.815	-1.864	1.088	-.973	2.003	-1.552
26	52.	92.	97.	3.098	2.222	2.993	2.325	1.640	1.197	3.020	1.909
27	53.	110.	111.	1.721	2.889	1.831	2.879	.509	1.031	.937	1.643
28	55.	82.	87.	1.205	0	1.007	.126	1.014	.464	1.867	.740
29	56.	101.	97.	1.033	1.778	.971	1.808	.610	.767	1.124	1.223
30	58.	100.	96.	-.516	-.444	-.559	-.454	-.130	-.198	-.240	-.316
31	60.	100.	105.	1.033	2.222	1.077	2.269	.356	.991	.656	1.580
32	62.	101.	105.	1.377	2.222	1.414	2.262	.528	.965	.972	1.539
33	63.	75.	93.	2.582	6.222	2.435	6.528	1.510	3.420	2.780	5.452
34	65.	104.	92.	2.582	6.222	2.421	6.324	1.542	2.672	2.839	4.260
35	66.	97.	84.	-.689	4.222	-.886	4.332	.165	1.961	.303	3.126
36	67.	106.	125.	4.648	2.889	4.881	2.907	1.524	1.134	2.805	1.808
37	68.	89.	107.	.344	-.889	.430	-.830	-.051	-.113	-.093	-.180
38	69.	66.	84.	.861	-4.444	.630	-4.299	.937	-1.106	1.726	-1.763
39	72.	88.	81.	.172	2.889	-.083	3.034	.689	1.598	1.269	2.547
40	73.	108.	110.	-1.377	.889	-1.215	.851	-1.004	.190	-1.849	.303
41	75.	100.	94.	.861	2.000	.763	2.042	.620	.892	1.141	1.422
42	76.	93.	91.	3.271	4.000	3.082	4.133	1.917	1.964	3.529	3.132
43	77.	117.	81.	-1.549	-.444	-1.768	-.573	-.169	-.636	-.311	-1.014
44	80.	83.	88.	1.033	-.667	.852	-.561	.896	.141	1.650	.224
45	82.	81.	90.	-1.893	-1.333	-1.986	-1.228	-.627	-.105	-1.154	-.167
46	83.	86.	88.	-.689	-.889	-.833	-.809	.038	-.036	.070	-.057
47	84.	93.	105.	1.721	.444	1.751	.503	.700	.379	1.288	.604
48	85.	99.	92.	1.377	4.667	1.242	4.771	.941	2.107	1.732	3.359
49	86.	80.	78.	-1.205	-.889	-1.471	-.767	.098	.119	.180	.190
50	87.	115.	102.	.861	5.111	.869	5.113	.366	1.893	.673	3.018
51	88.	103.	101.	.344	.444	.350	.433	.140	.121	.258	.193
52	91.	72.	76.	-.861	0	-1.161	.196	.333	.722	.613	1.150
53	92.	99.	103.	1.721	-1.333	1.725	-1.354	.763	-.569	1.405	-.907
54	93.	81.	75.	1.033	3.333	.680	3.536	1.309	1.976	2.410	3.151
55	96.	80.	76.	-.344	2.000	-.655	2.182	.591	1.407	1.087	2.244
56	97.	99.	91.	1.205	4.667	1.060	4.771	.887	2.107	1.633	3.359
57	98.	86.	90.	1.893	-.222	1.721	-.129	1.262	.262	2.323	.417
58	99.	92.	99.	2.926	1.333	2.851	1.417	1.491	.801	2.745	1.277
59	100.	100.	92.	1.721	5.333	1.579	5.445	1.113	2.379	2.048	3.792
60	101.	88.	74.	9.000	2.889	9.000	3.034	9.000	1.598	9.000	2.547
61	103.	87.	88.	.172	.222	.009	.318	.467	.434	.860	.692
62	104.	115.	115.	3.615	6.000	3.737	6.020	1.326	2.290	2.442	3.650
63	106.	102.	91.	-.344	4.000	-.456	4.070	.114	1.733	.210	2.762
64	107.	98.	100.	1.377	2.222	1.348	2.283	.687	1.043	1.264	1.662
65	108	105.	90.	2.238	5.556	2.058	5.637	1.434	2.349	2.639	3.745
66	111.	82.	72.	-2.066	-1.778	-2.393	-1.689	-.141	-.329	-.259	-.525
67	113.	85.	86.	1.721	.889	-1.871	1.013	-.414	.783	-.762	1.248
68	111	110	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1.580	1.824	2.910	2.908
69	115	107.	114.	2.410	1.778	2.545	1.766	.757	.613	1.394	.977
70	117.	97.	93.	1.377	1.556	1.441	1.609	-.464	.771	-.855	1.229
71	120	108	[illegible]	9.000	4.889	9.000	4.935	9.000	1.974	9.000	3.148

CASE-NO	DEM2	KABC29	WISCR16	ZDIF6	ZDIF12	ZTDIF6	ZTDIF12	REGRES6	REGRES12	REYREG6	REYREG12
72	121.	100.	106.	1.893	1.778	1.933	1.815	.754	.793	1.388	1.264
73	123.	86.	95.	1.033	-.889	.945	-.809	.674	-.036	1.241	-.057
74	124.	96.	99.	.689	0	.661	.028	.375	.103	.691	.164
75	133.	88.	80.	-2.754	-.444	-2.961	-.370	-.738	.111	-1.359	.177
76	135.	123.	107.	.516	5.778	.598	5.737	.035	1.984	.065	3.163
77	138.	104.	105.	.861	0	.909	-.028	.270	-.103	.498	-.164
78	139.	95.	105.	1.721	-1.556	1.751	-1.553	.700	-.565	1.288	-.901
79	142.	123.	114.	1.721	4.444	1.871	4.376	.414	1.390	.762	2.215
80	143.	117.	105.	1.205	5.333	1.246	5.326	.442	1.941	.814	3.094
81	144.	105.	100.	.172	1.556	.168	1.553	.086	.565	.158	.901
82	148.	74.	93.	9.000	-1.111	9.000	-.952	9.000	.174	9.000	.278
83	149.	87.	103.	9.000	-2.889	9.000	-2.858	9.000	-.954	9.000	-1.520
84	150.	92.	87.	9.000	2.667	9.000	2.779	9.000	1.396	9.000	2.225
85	151.	81.	95.	.172	-2.667	.102	-2.589	.245	-.700	.450	-1.116
86	157.	102.	107.	1.549	1.556	1.609	1.574	.550	.642	1.013	1.024

APPENDIX F

STATISTICAL FORMULAS:

KAPPA
FRIEDMAN TWO-WAY ANOVA BY RANKS
MCNEMAR'S TEST FOR CORRELATED PROPORTIONS

a) Proportions

	Judge A				
	Category	1	2	3	p_{iB}
Judge B	1	.44(.30)*	.07	.09	.60
	2	.05	.20(.09)	.05	.30
	3	.01	.03	.06(.02)	.10
	p_{iA}	.50	.30	.20	$\sum p_i = 1.00$

$$p_o = .44 + .20 + .06 = .70$$

$$p_c = .30 + .09 + .02 = .41$$

$$\kappa = \frac{.70 - .41}{1 - .41} = .492 \quad \text{(Eq. 1)}$$

$$\kappa_M = \frac{(.50 + .30 + .10) - .41}{1 - .41} = .831 \quad \text{(Eq. 6)}$$

$$\sigma_\kappa = \sqrt{\frac{.70\,(1 - .70)}{200\,(1 - .41)^2}} = .055 \quad \text{(Eq. 7)}$$

95% confidence limits = .492 ± 1.96 (.055) = .384 ↔ .600

$$\sigma_{\kappa_0} = \sqrt{\frac{.41}{200\,(1 - .41)}} = .059 \quad \text{(Eq. 10)}$$

$$z = \frac{.492}{.059} = 8.34; \quad \kappa \text{ significant at } P < .001$$

b) Frequencies

	Judge A				
	Category	1	2	3	f_{iB}
Judge B	1	88(60)*	14	18	120
	2	10	40(18)	10	60
	3	2	6	12(4)	20
	f_{iA}	100	60	40	$N = 200$

$$f_o = 88 + 40 + 12 = 140$$

$$f_c = 60 + 18 + 4 = 82$$

$$\kappa = \frac{140 - 82}{200 - 82} = .492 \quad \text{(Eq. 2)}$$

$$\sigma_\kappa = \sqrt{\frac{140\,(200 - 140)}{200\,(200 - 82)^2}} = .055 \quad \text{(Eq. 8)}$$

$$\sigma_{\kappa_0} = \sqrt{\frac{82}{200\,(200 - 82)}} = .059 \quad \text{(Eq. 11)}$$

* Chance expectancy

Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 20, 37-46.

A RANK TEST FOR k CORRELATED SAMPLES

A rank test for k correlated samples is the Friedman two-way analysis of variance by ranks (1937). The data are a set of k observations for a sample of N individuals. Such data arise in many experiments where subjects are tested under a number of different experimental conditions. The corresponding parametric test is an analysis of variance for two-way classification where observations are made on each of a group of individuals under more than two conditions. If there is reason to believe that the assumptions underlying the analysis of variance are not satisfied by the data, the Friedman rank method may be appropriate.

The data are arranged in a table containing N rows and k columns. The rows correspond to individuals, or groups, and the columns to experimental conditions. Table 22.1 shows such an arrangement of data for eight subjects tested under four experimental conditions. The observations in the rows are ordered as shown in Table 22.2. For example, the four observations in the top row are 4, 5, 9, and 3. These are replaced by the ranks 2, 3, 4, and 1. The ranks in each column are summed. If the samples are from the same population, the ranks in each column will be a random arrangement of the numbers 1, 2, 3, and 4. Under these circumstances the sums of ranks for columns will tend to be the same. If these sums differ significantly, the hypothesis that they are from the same population may be rejected.

The null distribution here involves a consideration of the $k!$ arrangements of ranks in any row. These are considered equiprobable. Given N rows, the number of possible equiprobable arrangements of ranks is $(k!)^N$.

Table 22.1
Material recalled after four time intervals for a group of eight subjects

	Time interval			
Subject	I	II	III	IV
1	4	5	9	3
2	8	9	14	7
3	7	13	14	6
4	16	12	14	10
5	2	4	7	6
6	1	4	5	3
7	2	6	7	9
8	5	7	8	9

For each of these arrangements a statistic S may be calculated, where

$$S = \sum^{k} (R_i - \bar{R})^2$$

Here R_i = sum of ranks for the ith column
$\bar{R}$ = mean rank sum
S = sum of squares of rank sums about the mean rank sum

If the samples are for the same population, the expectation is that the R_i's are equal and the expected value of S is 0. At least in theory, for any N and k a frequency distribution may be made of the $(k!)^N$ values of S. This distribution may be used to evaluate particular values of S. If the probability associated with a particular value of S is small, the null hypothesis is rejected. For small values of k and N the exact distributions of S are

Table 22.2
Ranks assigned by rows for the data of Table 22.1

	Time interval			
Subject	I	II	III	IV
1	2	3	4	1
2	2	3	4	1
3	2	3	4	1
4	4	2	3	1
5	1	2	4	3
6	1	3	4	2
7	1	2	3	4
8	1	2	3	4
R_i	14	20	29	17

known. Bradley (1968) provides a table of exact critical values of S for $k = 3$ and N up to 15, and for $k = 4$ and N up to 8.

For values that lie outside the tabled values of S it is customary to use a statistic which is a function of S. This statistic is given by

$$\chi_r^2 = \frac{12S}{Nk(k+1)}$$

This statistic has an approximate chi-square distribution with $k - 1$ degrees of freedom.

For computational purposes a more convenient way of writing χ_r^2 is

$$\chi_r^2 = \frac{12}{Nk(k+1)} \sum^{k} R_i^2 - 3N(k+1)$$

For the data of Table 22.2 we have

$$\chi_r^2 = \frac{12}{8 \times 4(4+1)} (14^2 + 20^2 + 29^2 + 17^2) - 3 \times 8(4+1) = 9.45$$

This result for $df = 4 - 1 = 3$ falls between the .05 and .01 levels of significance. Actually it is a little above the 2 percent level. If this level of confidence is acceptable, we may conclude that the samples are not drawn from the same population and that a difference in the experimental conditions is exerting an effect. In this example $S = 126$. If this is referred to a table of exact critical values of S, as given in Bradley (1968), the associated probability is found to fall between .01 and .05, not far from the .02 level. The chi-square approximation is in close agreement with the more exact test.

The asymptotic relative efficiency of the Friedman test relative to the F test resulting from a two-way analysis of variance with one observation per cell is

$$\left(\frac{3}{\pi}\right)\left(\frac{k}{k+1}\right)$$

The efficiency of the test increases as k increases, and extends from .637 for $k = 2$ to a maximum of .955 for $k = \infty$. For $k = 2$ this test is the same as the sign test for correlated samples.

From: Ferguson, G.A. (1976). Statistical analysis in psychology and education. New York: McGraw-Hill Book Company, pp. 394-396.

17.10 A TEST FOR CORRELATED PROPORTIONS IN A TWO-BY-TWO TABLE

A problem that often arises in psychological research is somewhat different from the problem of association between attributes. As an illustration of this problem, suppose that some N individual subjects are each observed by two independent judges. Each judge places each subject into one of two mutually exclusive and exhaustive categories, such as "high leadership potential" versus "low leadership potential." It is assumed that a judge's ratings of different individuals are independent. Let us call these categories simply "H" and "L" for the moment. We would like to ask if these two judges, given all possible subjects in the population, would show the same true proportion of individuals rated in category "H." In other words, in the population of all subjects to be rated, does $p_1(H) = p_2(H)$, where $p_1(H)$ is the proportion rated in category H by judge 1, and $p_2(H)$ is the proportion rated in that category by judge 2?

This is a problem of **correlated proportions**, since each of the two sample proportions will be based in part on the same individuals. A test due to McNemar (1955) applies to this situation. Suppose that the sample of N individuals were arranged into the following 2 × 2 table:

		JUDGE 1	
		H	L
JUDGE 2	H	a	b
	L	c	d

An *exact* test of the hypothesis that $p_1(H) = p_2(H)$ is possible using binomial distribution. Under this hypothesis, the probability of a given sample result, showing a particular pair of cell frequencies b and c, is just

$$\binom{b+c}{b}(.5)^{b+c}. \qquad [17.10.1^*]$$

To carry out the exact test, let g equal the smaller of the two frequencies, b or c. Then one takes the sum of probabilities

$$2\sum_{h=0}^{g}\binom{b+c}{h}(.5)^{b+c}. \qquad [17.10.2^*]$$

If this number is less than or equal to the value chosen for α, then the null hypothesis may be rejected (two-tailed).

When N is relatively large, the exact probability may be approximated by use of χ^2, where

$$\chi^2 = \frac{(|b-c|-1)^2}{b+c} \qquad [17.10.3\dagger]$$

with one degree of freedom.

For our example, a significant result would let one conclude that the *true* distributions of judgments for the two judges differ. Be sure to notice that this is not an ordinary test of association for a contingency table, but rather a test of the equality of two proportions where each sample proportion involves some of the same observations, making the two sample proportions dependent.

From Hays, W.L. (1973). Statistics for the social sciences. Chicago: Holt, Rinehart, and Winston, pp. 740-741.

LIST OF REFERENCES

LIST OF REFERENCES

Algozzine, B., Forgnone, C., Mercer, C.D., & Trifiletti, J.J. (1979). Toward defining discrepancies for specific learning disabilities: An analysis and alternatives. Learning Disability Quarterly 2, 25-31.

Algozzine, B., & Ysseldyke, J.E. (1981). An analysis of difference score reliabilities on three measures with a sample of low-achieving youngsters. Psychology in the Schools, 18, 133-138.

Algozzine, B., Ysseldyke, J.E., & Shinn, M. (1982). Identifying children with learning disabilities. When is a discrepancy severe? Journal of School Psychology, 20, 299-305.

Bannatyne, A.D. (1968). Diagnosing learning disabilities and prescribing remediaton. Journal of Learning Disabilities, 1, 242-249.

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

Bogen, J.E. (1969). The other side of the brain: Parts I, II, and III. Bulletin of the Los Angeles Neurological Society, 34, 73-105, 135-162, 191-203.

Bogen, J.E., DeZure, R., Tenhouten, N., & Marsh, J. (1972). The other side of the brain: IV. Bulletin of the Los Angeles Neurological Society, 37, 49-61.

Bond, G.L., & Tinker, M.A. (1973). Reading difficulties, Their diagnosis and correction, (3rd ed.). New York: Appleton-Century-Crofts.

Bracken, B.A. (1985). A critical review of the Kaufman Assessment Battery for Children (K-ABC). School Psychology Review, 14, 21-36.

Bureau of Education for the Handicapped (1976). Federal Register. 41, 52105-52107.

Cattell, R.B. (1971). Abilities: Their structure, growth, and action. Boston: Houghton Mifflin.

Childers, J.S., Durham, T.W., & Bolen, L.M. (1985). A predictive validity study of the Kaufman Assessment Battery for Children with the California Achievement Test, Psychology in the Schools, 22, 29-33.

Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 20, 37-46.

Cone, T.E., & Wilson, L.R. (1981). Quantifying a severe discrepancy: A critical analysis. Learning Disability Quarterly, 4, 359-371.

Cooley, E.J., & Ayres, R. (1985). Convergent and discriminant validity of the Mental Processing Scales of the Kaufman Assessment Battery for Children. Psychology in the Schools, 22, 373-377.

Cummins, J.P., & Das, P. (1977). Cognitive processes and reading difficulties: A framework for research. Alberta Journal of Educational Research, 23, 245-256.

Danielson, L.C., & Bauer, N.J. (1978). A formula based classification of learning disabled children: An examination of the issues. Journal of Learning Disabilities, 11, 163-176.

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

Das, J.P. (1984). Simultaneous and successive processes and K-ABC. Journal of Special Education, 18, 229-238.

Das, J.P., Kirby, J., & Jarman, R.F. (1975). Simultaneous and successive synthesis: An alternative model for cognitive abilities. Psychological Bulletin, 82, 87-103.

Das, J.P., Kirby, J., & Jarman, R.F. (1979). Simultaneous and successive cognitive processes. New York: Academic Press.

Das, J.P., & Molloy, G.N. (1975). Varieties of Simultaneous and successive processing in children. Journal of Education Psychology, 67, 213-220.

Denckla, M.B. (1974). Development of motor coordination in normal children. Developmental Medicine and Child Neurology, 16, 729-741.

Dunn, L., & Markwardt, F. (1970). Peabody Individual Achievement Test. Circle Pines, MN: American Guidance Service.

Elliot, M. (1981). Quantitative evaluation procedures for learning disabilities. Journal of Learning Disabilities, 2, 84-87.

Epps, S., Ysseldyke, J.E., & Algozzine, B. (1983). Impact of different definitions of learning disabilities on the number of students identified. Journal of Psychoeducational Assessment, 1, 341-352.

Epps, S., Ysseldyke, J.E., & Algozzine, B. (1985). An analysis of the conceptual framework underlying definitions of learning disabilities. Journal of School Psychology, 23, 133-144.

Erickson, M.T. (1975). The z-score discrepancy method for identifying reading disabled children. Journal of Learning Disabilities, 8, 308-312.

Fields, N.A. (1979). A comparison of techniques for the identification of children with reading disabilities. (Doctoral dissertation, University of North Carolina at Chapel Hill), Dissertation Abstracts International (University Microfilms No. 8005036).

Forness, S.R., Sinclair, E., & Guthrie, D. (1983). Learning disability discrepancy formulas: Their use in actual practice. Learning Disability Quarterly, 6, 107-114.

Gazzaniga, M.S. (1970). The bisected brain. New York: Appleton-Century-Crofts.

Hanna, G.S., Dyck, N.J., & Holen, M.C. (1979). Objective analysis of achievement-aptitude discrepancies in L.D. classification. Learning Disability Quarterly, 6, 107-114.

Harris, A.J., & Sipay, E.R. (1980). How to increase reading ability: A guide to developmental and remedial methods. (7th ed.), New York: Longman.

Hays, W.L. (1973). Statistics for the social sciences. Chicago: Holt, Rinehart, and Winston.

Hooper, S.R., & Hynd, G.W. (1985). The relationship between the K-ABC and the VADS with reading-disabled children: Validity of the K-ABC. Journal of Psychoeducational Assessment, 3, 77-87.

Hooper, S.R., & Hynd, G.W. (1986). Performance of normal and dyslexic readers on the Kaufman Assessment Battery for Children (K-ABC): A discriminant analysis. Journal of Learning disabilities, 19, 206-210.

Hoover, H.D. (1984). The most appropriate scores for measuring educational development in the elementary schools: GEs. Educational Measurement: Issues and Practice, 3, 8-14.

Horn, J.L., & Cattell, R.B. (1966). Refinement and test of the theory of fluid and crystallized intelligence. Journal of Educational Psychology, 57, 253-270.

Inglis, J., & Lawson, J.S. (1986). A principal components analysis of the Kaufman Assessment Battery for Children (K-ABC): Implications for the test results of children with learning disabilities. Journal of Learning Disabilities, 19, 80-85.

Jarman, R.F., & Das, J.P. (1977). Simultaneous and successive synthesis and intelligence. Intelligence, 1, 151-169.

Jensen, A.R. (1984). Black-white differences on the K-ABC, Journal of Special Education, 18, 377-408.

Johnson, D.J., & Myklebust, H.S. (1967). Learning disabilities: Educational principles and practices, New York: Grune & Stratton.

Kaluger, G., & Kolson, C. (1969). Reading and learning disabilities, Columbus, OH: Charles E. Merrill.

Kamphaus, R.W., & Reynolds, C.R. (1984). Development and structure of the Kaufman Assessment Battery for Children. Journal of Special Education, 18, 213-228.

Kaufman, A.S. (1979). Cerebral specialization and intelligence testing. Journal of Research and Development in Education, 12, 96-107.

Kaufman, A.S. (1983). Some questions and answers about the Kaufman Assessment Battery for Children. Journal of Psychoeducational Assessment, 1, 205-218.

Kaufman, A.S. (1984). K-ABC and controversy. Journal of Special Education, 18, 409-444.

Kaufman, A.S., & Kaufman, N.L. (1983a). Kaufman Assessment Battery for Children (K-ABC): Administration and scoring manual. Circle Pines, MN: American Guidance Service.

Kaufman, A.S., & Kaufman, N.L. (1983b). Kaufman Assessment Battery for Children (K-ABC): Interpretive manual. Circle Pines, MN: American Guidance Service.

Kaufman, A.S., & McLean, J.E. (1985). Joint factor analysis for the K-ABC and WISC-R for normal children. Journal of School Psychology. 23.

Kaufman, A.S., & McLean, J.E. (1986). K-ABC/WISC-R factor analysis for a learning disabled population. Journal of Learning disabilities, 19, 145-153.

Keith, T.Z. (1985). Questioning the K-ABC: What does it measure? School Psychology Review, 14, 9-19.

Keith, T.Z. (1986). K-ABC factor structure for referred school children. Psychology in the Schools, 23, 240-246.

Keith, T.Z., & Dunbar, S.B. (1984). Hierarchical factor analysis of the K-ABC: Testing alternative models. Journal of Special Education, 18, 367-376.

Klanderman, J.W., Perney, J., & Kroeschell, Z.B. (1985). Comparisons of the K-ABC and WISC-R for L.D. children. Journal of Learning Disabilities, 18, 524-527.

Lawson, J.S., & Inglis, J. (1985). Learning disabilities and intelligence test results: A model based on a principal components analysis of the WISC-R. British Journal of Psychology.

Lezak, M.D. (1976). Neuropsychological Assessment. New York: Oxford University Press.

Luria, A.R. (1966). Human brain and psychological processes. New York: Harper and Row.

Luria, A.R. (1973). Higher coritical functions in man. New York: Basic Books.

McRae, S.G. (1981). Simultaneous and sequential processing as related to reading and comprehension skills of second and third grade readers. Unpublished Master's Thesis. National College of Education.

Mehrens, W.A., & Lehmann, I. (1975). Measurement and evaluation in education and psychology. (2nd Ed.). New York: Holt, Rinehart & Winston.

Murray, A., & Bracken, B.A. (1984). Eleven month predictive validity of the Kaufman Assessment Battery for Children. Journal of Psychoeducational Assessment, 2, 225-232.

Naglieri, J.A. (1985a). Use of the WISC-R and K-ABC with learning disabled, borderline mentally retarded, and normal children. Psychology in the Schools, 22, 133-141.

Naglieri, J.A. (1985b). Normal children's performance on the McCarthy Scales, Kaufman Assessment Battery, and Peabody Individual Achievement Test. Journal of Psychoeducational Assessment, 3, 123-130.

Naglieri, J.A., & Haddad, F.A. (1985). Learning disabled children's performance on the Kaufman Assessment Battery for Children: A concurrent validity study. Journal of Psychoeducational Assessment, 2, 49-56.

Naglieri, J.A., Kaufman, A.S., Kaufman, N.L., & Kamphaus, R.W. (1981). Cross-validation of Das' simultaneous and successive processes with novel tasks. The Alberta Journal of Educational Research, 27, 264-267.

Nebes, R.D. (1974). Hemispheric specialization in commissurotomized man. Psychological Bulletin, 81, 1-14.

Neisser, U. (1967). Cognitive Psychology. New York: Appleton-Century-Crofts.

Obrzut, A., Obrzut, J., & Shaw, D. (1984). Construct validity of the Kaufman Assessment Battery for Children with learning disabled and mentally retarded. Psychology in the Schools, 21, 417-424.

O'Donnell, L.E. (1980). Intra-individual discrepancy in diagnosing specific learning disabilities. Learning Disability Quarterly, 3, 10-18.

Phillips, S.E., & Clarizio, H. (1986). Some limitations of standard scores in diagnosing learning disabilities. CEPSE Research Series/No. 1, East Lansing, MI.

Reynolds, C.R. (1981). The fallacy of "Two years below grade level for age" as a diagnostic criterion for reading disorders. Journal of School Psychology, 19, 350-358.

Reynolds, C.R. (1984-85). Critical measurement issues in learning disabilities, Journal of Special Education, 18, 451-476.

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

Salvia, J., & Ysseldyke, J.E. (1978). Assessment in special and remedial education. Boston: Houghton-Mifflin.

Sattler, J. (1981). Assessment of children's intelligence and special abilities, (2nd Ed.). Boston: Allyn & Bacon.

Schulte, A., & Borich, G.D. (1984). Considerations in the use of difference scores to identify learning disabled children. Journal of School Psychology, 22, 381-390.

Shepard, L. (1980). An evaluation of the regression discrepancy method for identifying children with learning disabilities. Journal of Special Education, 14, 79-91.

Sperry, R.W. (1970). Cerebral dominance in perception. In F.A. Young & D.B. Lindsley (Eds.), Early experience and visual information processing in perceptual and reading disorders. Washington, D.C.: National Academy of Science.

Stanley, J.C. (1971). Reliability. In R.L. Thorndike (Ed.). Educational Measurement, Washington, D.C. American Council on Education.

Sternberg, R.J. (1984). The Kaufman Assessment Battery for Children: An information-processing analysis and critique. Journal of Special Education, 18, 251-261.

Swerdlick, M.E., & Lewis, M.L. (1983). Concurrent validity of the K-ABC with the WISC-R, Stanford Binet, Cognitive Abilities Test and Iowa Tests of Basic Skills. In Klanderman (Chair), Concurrent/construct validity research with cognitive measures and the K-ABC. Symposium conducted at the meeting of the National Association of School Psychologists, Detroit.

Vernon, P.E., Ryba, K.A., & Lang, R.J. (1978). Simultaneous and successive processing: An attempt at replication. Canadian Journal of Behavioral Science, 10, 1-15.

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

White, T.H. (1979). Correlations among the WISC-R, PIAT, and DAM. Psychology in the Schools, 16, 497-501.

Wikoff, R.L. (1979). The WISC-R as a predictor of achievement. Psychology in the Schools, 16, 364-368.

Wilson, L.R., & Cone, T.E. (1984). The regression equation method of determining academic discrepancy. Journal of School Psychology, 22, 95-110.

Witelson, S.F. (1977). Developmental dyslexia: Two right right hemispheres and none left. Science, 195, 309-311.

Zins, J.E., & Barnett, D.W. (1983). The Kaufman Assessment Battery for Children and school achievement: A validity study. Journal of Psychoeducational Assessment, 1, 235-242.

Zins, J.E., & Barnett, D.W. (1984). A validity study of the K-ABC, the WISC-R, and the Stanford-Binet with non-referred children. Journal of School Psychology, 22, 369-371.

General References

Ferguson, G.A. (1976). Statistical analysis in psychology and education. New York: McGraw-Hill.

Hays, W.L. (1973). Statistics for the social sciences. Chicago: Holt, Rinehart, and Winston.

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

Magnusson, D. (1966). Test theory, Reading, Ma.: Addison-Wesley.

McLoughlin, J.A., & Lewis, R.B. (1981). Assessing special students. Columbus: Charles E. Merrill.

Mehrens, W.A., & Lehman, I. (1975). Measurement and evaluation in education and psychology. (2nd Ed.). New York: Holt, Rinehart & Winston.

Nie, N.H., Hull, C.H., Jenkins, J.G., Steinbrenner, K., & Bent, D.H. (1975). SPSS Statistical package for the social sciences. (2nd. Ed.). New York: McGraw-Hill.

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

Sattler, J. (1981). Assessment of children's intelligence and special abilities. (2nd. ed.). Boston: Allyn & Bacon.