THE USE OF INDIVIDUALLY ADMINISTERED COGNITIVE ABILITIES TESTS WITH ACADEMICALLY "AT-RISK" CHILDREN

by

Dorothy May Hall

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF EDUCATIONAL PSYCHOLOGY

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

1 9 9 4

Order Number 9426334

The use of individually administered cognitive abilities tests with academically "at-risk" children

Hall, Dorothy May, Ph.D.

The University of Arizona, 1994

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

THE USE OF INDIVIDUALLY ADMINISTERED COGNITIVE ABILITIES TESTS WITH ACADEMICALLY "AT-RISK" CHILDREN

by

Dorothy May Hall

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF EDUCATIONAL PSYCHOLOGY

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

1 9 9 4

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Final Examination Committee, we certify that we have read the dissertation prepared by Dorothy May Hall

entitled The Use of Individually Administered Cognitive Abilities Tests with Academically "At-Risk" Children

and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy

Shitala P. Mishra 3-31-94
Date

John R. Bergan 3-31-94
Date

Darrell Sabers 3-31-94
Date

Date

Date

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copy of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Shitala P. Mishra
Dissertation Director

4-11-94
Date

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED Dorothy M. Hall

ACKNOWLEDGEMENTS

I would like to thank Dr. Shitala Mishra for serving as chairman of my committee. His guidance of this project is gratefully appreciated. I would, also, like to thank Dr. Darrell Sabers and Dr. John Bergen for their time, expertise, and valuable contributions.

I would like to thank my family, Bob, Elizabeth, and Melissa, for their belief, support and guidance. I would not have been able to complete this project without them.

TABLE OF CONTENTS

TABLE OF CONTENTS--Continued

LIST OF TABLES

LIST OF TABLES--Continued

ABSTRACT

The purpose of this study was to gather information about criterion-related and construct test validity as applied to a referral or "at-risk" sample. Two individually administered cognitive abilities tests, the Wechsler Intelligence Scales for Children-Third Edition (WISC-III) and the Woodcock-Johnson Psycho-Educational Battery-Revised Tests of Cognitive Ability (WJ-R Cog), frequently used in the special education placement decision process have recently been revised. Before test generalizations obtained from the standardization sample can be presumed to apply to a referral or "at-risk" sample, research is needed to explore the applicability of these assumptions to this population. No previous research has used a random sample of children referred for special education evaluation who are not yet placed into a program to explore the construct and concurrent validity of both complete test batteries.

A random sample of 183 students referred for possible special education placement was used. Ages ranged from 6 to 16 years. Students were classified into three age groups, four ethnic groups, and six possible special education classifications, including a non-placed category. Data were analyzed using a variety of techniques including: exploratory factor analysis, confirmatory factor analysis,

stepwise multiple regression, and multiple discriminant function analysis.

Findings of this study support the four factor WISC-III model and the seven factor WJ-R Cog model. Study results indicate a concurrent verbal component (approximately 50%) comprised of the WISC-III VC factor and the WJ-R Cog Gc, Gs, and Ga factors. This verbal component was the best composite predictor variable of academic achievement. Although "at-risk" students scored lower on these verbal component factors, BWL, or broad written language, was the best single discriminator variable in both a six and two group analyses. Special education classifications were examined for specific category profiles and the ACID/ACIDS profile was investigated.

This study initiated information about the underlying test constructs and how these constructs relate to each other when given to an "at-risk" referral sample. It is hoped that this study will stimulate further research in this area.

CHAPTER 1

INTRODUCTION

The assessment and identification of children with special educational needs has long been a concern for psychologists and educators. Public Law 94-142 (P.L. 94-142), the Education for All Handicapped Children Act of 1975, was designed to protect the rights of all children, but especially children with special needs. This legal regulation was established because the educational and psychological needs of handicapped children were not being met by the schools (Reynolds, Elliot, Gutkin, & Witt, 1984). P.L. 94-142 established safeguards for the evaluation and placement of children, compelled schools to identify those children with special needs who were not being served appropriately by the school, and required nondiscriminatory and valid assessment.

Specific to testing and assessment P.L. 94-142 required: 1) tests and other assessment tools be validated for the specific intent for which they are used; 2) tests be administered by trained professionals with strict compliance to the standardized test instructions; and 3) several tests be used to assess different areas of abilities, including tests of general intelligence and academic achievement.

Also, three year reevaluations were required to evaluate placement, progress, and identify new areas for instruction.

This law defines what needs to be done for the assessment and evaluation of special education placement, but it does not define how it is to be done. Further definition of P.L. 94-142 has been through the court system. Most of the judicial decisions concern the identification, assessment, and placement of minority children in special education classes. A good example of these court cases is Larry P. v. Riles (1979). This case had a major impact on testing practices. Minority students had been labeled and placed in educably mentally retarded (EMR) classes as the result of a low score on a standardized intelligence test. Many questions were raised about intelligence testing during this case, such as: what do intelligence tests measure, do all intelligence tests measure the same ability, what population are intelligence tests standardized on, and what is the basis for valid inferences if the standardization sample did not include members from the group for whom inferences are being made.

Many other court cases have been involved in these issues, but recently three federal courts and two federal appeals courts have labored over placement and assessment questions (Sattler, 1988). In all of these cases the essential assessment tool has been the individually administered standardized intelligence test. Some have even suggested that the courts are not the proper place for the resolution to these complicated questions. In the case of Doe v. Board of Education of Montgomery County (1982), which involved a dispute between a physician and a school psychologist over appropriate assessment and placement of a child, the court concluded that the judicial system was not the appropriate forum to determine the validity of assessment or placement decisions.

At the bottom of these concerns is the question of the validity of intelligence assessment tools in psychological assessment. Educators and school psychologists, using both ethical and legal considerations, have always been concerned with the validity of assessment instruments conventionally used in the evaluation of children for possible special education services. Because cognitive assessments are an essential ingredient of this process, it is imperative that psychologists and educators understand what abilities these tools measure and for whom they measure these abilities. The present study, therefore, was designed to examine the validity of two intelligence measures with children referred for academic difficulties.

Statement of the Problem

Three instruments frequently used in intellectual and achievement assessment procedures are the WISC-R and the WJ

Cog and the WJ Ach (Nuttall, 1987). Both test batteries have recently been revised, the WISC-III was introduced in 1991 and the WJ-R in 1989. At this time no known research study has compared the WISC-III and WJ-R Cog in relation to construct and criterion-related validity using a random referred sample of children not yet placed into any special education cateogry.

Information about the nature of the tests must be known so that appropriate interpretations of test scores can be made. Appropriate score interpretation will allow educators and psychologists to make better informed placement decisions. These placement decision are vitally important to the individual child's future and educational attainment.

Placement decisions should be based on different aspects of the child's abilities; therefore, assessment tools should provide a full range of information (Knoff, 1988). Do these two tests assess many aspects or one global ability? Do the WJ-R Cog seven cognitive factor scores measure a more broad variety of abilities than the traditional intelligence test, such as the WISC-III, as claimed by Woodcock (1990)? Do the two cognitive tests measure the same or different abilities? If the two tests measure the same abilities, children are being subjected to redundant procedures without the benefit of more information. Is enough different information provided by both tests to warrant the inclusion of both in an assessment procedure? Are the cognitive tests indicators of present achievement? Again, if the cognitive tests are too closely correlated with achievement no new information is gained by giving these tests.

The WJ-R Cog authors imply that their test was designed to be used in the special education evaluation process (Woodcock & Johnson, 1989). Does the WJ-R Cog put forth information that is specific to the special education categories (Learning Disabilities, Emotionally Handicapped, Mental Retardation, etc.)? Is this the same information that can be obtained from the WISC-III? Are there unique performance patterns specific to these categories that would help in the identification process? If additional information is present in the WJ-R Cog or the WISC-III, this should improve diagnostic formulations (Reschly, 1990).

The Purpose of the Study and the Research Questions

The purpose of this study was to examine the performances of a referral sample (children referred for academic difficulties, but not yet placed into special education) on two tests of cognitive ability, WJ-R Cog and WISC-III, and one academic achievement test, WJ-R Ach. It was hoped that the findings of this study would provide more information about test validity for each cognitive test and its relationship to each other when used with children who have academic difficulty, so that valid inferences and appropriate educational placement decisions can be made.

Based upon the overall purposes, this study was designed to seek answers to the following research questions:

1. Does the WISC-III subtest performance of a referral sample reflect the same underlying construct pattern as that reported for the WISC-III normative sample? In other words, does the WISC-III four factor model, as stated in the WISC-III manual, fit the referral sample's performance? If the subtests are measuring different underlying constructs, test generalizations and inferences are questionable.

2. Does the WJ-R Cog subtest performance of a referral sample reflect the same underlying construct pattern as that reported for the WJ-R Cog normative sample? Does the WJ-R Cog seven factor model fit the referral sample's performance? If the subtests are measuring the same underlying constructs as specified by the normative sample appropriate behavioral inferences and generalizations are more likely to occur.

3. Do the WJ-R Cog subtests and composite scores and the WISC-III subtests, factor indices, scales, and composite scores measure the same or different cognitive abilities? In other words, do the two cognitive tests contribute unique or redundant information?

4. Using the factor index scores, which cognitive tests, WISC-III or WJ-R Cog, best predicts to current academic achievement?

5. Are there unique performance patterns on these three tests that are representative of the various special education categories?

CHAPTER 2

REVIEW OF THE LITERATURE

The purpose of this study was to examine the performance of a referral sample (children referred for academic difficulties) on two tests of cognitive abilities, WISC-III and WJ-Cog, and one academic achievement test, WJ-R Ach. It was hoped that this study would provide more information about each cognitive test and their relationship to each other when used with children who have academic difficulty, so that valid inferences and placement decision can be made. The questions asked about these tests reflect concerns about validity.

The following review is divided into five major sections. The first section discusses legal and ethical consideration as related to assessment. The second section defines test validity. The third and fourth sections contain validity information obtained from the test manuals. The fifth section examines the accumulated body of research on test validity for each test and summarizes the research comparing the WISC-R/WISC-III and the WJ/WJ-R Cog.

Legal and Ethical Considerations

Assessment plays a vital part in the special education placement decision process. These placement decisions often influence individual educational and occupational futures;

therefore, assessment is an area that has been directed by ethical and legal guidelines (Kehle, Clark, & Jenson, 1993).

Legal direction has come from both the legislative and judicial systems. P.L. 94-142 (1975) represents a comprehensive bill composed from many past laws. It requires that tests be validated for the specific purpose for which they are used, accurately reflect the child's abilities, and that the areas of general intelligence and academic achievement be assessed before placement of a child into special education (Reynolds et al., 1984). Court cases such as: Stell v. Savannah-Chatham County Board of Education (1963) questioned whether tests were being used to support segregation; Diana v. State Board of Education (1970) concluded that invalid testing procedures were used to place Hispanic children into special education; Guadalupe v. Tempe Elementary School District (1972) required that a child's test score be more than two standard deviations below the norm and that more than one test score was needed to place a child into special education; Larry p. v. Riles (1979) stated that the use of standardized intelligence tests were invalid for placing black children into special education programs; and in Georgia State Conference of Branches of NAACP v. State of George (1985) the court examined the validity of standardized tests because a disportionate number of black children where being placed into special education classes. Legal concerns about the use of standardized tests in special education placement decisions have mainly centered around the concept of validity.

Psychologists involved in assessment procedures in schools are bound by two sets of ethical guidelines: the National Association of School Psychologists Principles for Professional Ethics (NASP, 1985) and the American Psychological Association Ethical Principles of Psychologists and Code of Conduct (APA, 1992). Both codes guard against the misuse of tests and require the use of valid assessment tools and appropriate interpretations to support decisions based on the assessed abilities of individuals (APA, Ethical Standard Two, 1992; NASP,Professional Responsibility, 1985).

Validity

According to both ethical and legal guidelines, tests should be validated for the specific purposes for which they are used. What is validity and how is it measured?

Validity, according to Anastasi (1988), is concerned with what a test measures and how test performance relates to other independent measures of performance. The Standards for Educational and Psychological Testing (AERA, APA, NCME, 1985) recognizes three common methods used to investigate validity: content related, criterion-related, and construct-related. This paper is concerned with the criterion-related (concurrent) and construct related validity of two tests of cognitive ability.

Criterion-related validity references test performance to a criterion, either present at the time of testing for concurrent validity or in the future for predictive validity (Kerlinger, 1964). If the objective of testing is for diagnosis, concurrent validity is used. The criterion may be other test scores, classifications, ratings, or behaviors (Sattler, 1988).

Construct validity relates the test to theory (Kerlinger, 1964), in other words, it "...refers to the extent to which a test measures a psychological construct or trait." (Sattler, 1988, p. 31). Construct validity requires the accumulation of a variety information, so that information about what the test is measuring can be related to the individual's performance (Anastasi, 1988). ". . . validity is an integrated evaluative judgement of the degree to which empirical evidence and theoretical rationales support the adequacy and the appropriateness of inferences and actions based on test scores..." (Messick, 1989, p. 5). Information on construct validity may be accumulated by many methods, but two common methods are: 1) correlations with other tests, or differential validity (Campbell & Fiske, 1959), where correlations should be higher with tests measuring similar abilities (convergent validity), and lower with tests measuring different abilities (discriminant validity); and, 2) factor analysis which is based on the assumption that intercorrelations of a battery of tests given to a large number of individuals can be attributed to a set of underlying factors that are fewer in number than the tests themselves (Anastasi, 1988; Sattler, 1988). These two methods were utilized in this study to investigate test validity.

WISC-III Manual

This section presents information obtained from the WISC-III Manual on the background, development, and validity of the WISC-III.

Theory of the WISC-III

The Wechsler scales (Wechsler Adult Intelligence Scale-Revised, WAIS-R; Wechsler Intelligence Scale for Children-III, WISC-III; and Wechsler Preschool and Primary Scale of Intelligence, WPPSI) developed out of the Wechsler-Bellevue, Forms I and II, and adhere to David Wechsler concept of intelligence (Wechsler, 1939).

> Intelligence is the aggregate or global capacity of the individual to act purposefully, to think rationally and to deal effectively with his {sic} environment. It is global because it characterizes the individual's behavior as a whole; it is aggregate

> because it is composed of elements which, though not entirely independent, are qualitatively differentiable (Wechsler, 1939, p. 3).

Wechsler felt that intelligence was a unitary construct or a global capacity that cannot be measured directly; instead, we measure specific multifaceted abilities (Zachary, 1990). The subtests of the WISC-III, as the other Wechsler subtests, are designed to assess multiple cognitive abilities and when combined reflect general intellectual ability. The composition of the tasks used to measure intelligence was somewhat irrelevant, what was important was psychometric criteria and high correlations with other variables thought to reflect intelligent behavior (Zachary, 1990).

Wechsler had no hierarchy of abilities as measured by the subtests, instead they were considered equivalent (Wechsler, 1943). Individuals possess a combination of abilities and it is how these abilities are applied to everyday situations that gives meaning to the term intelligence. According to Wechsler (1939), intelligence cannot be measured in its absolute sense, what we assess is the individual's relative position in relation to a particular reference group. The cognitive subtests did not represent different intelligences, but were different methods of assessing intelligence. Although, Wechsler emphasized a unitary construct which is represented by the Full Scale Score, he organized the tests into two major divisions Verbal and Performance.

Description of the WISC-III

The WISC-III is the third revised edition of the WISC (Wechsler, 1949). The original edition was published in 1949; the second revision, the WISC-R, was published in 1974; and the WISC-III in 1991. The WISC-III is an individually administered, norm-referenced test used to measure cognitive ability with children aged 6 years through 16 years (Wechsler, 1991). The WISC-III is composed of ten standard subtests and two supplemental subtests (Digit Span and Mazes), which were part of the WISC-R, plus a new supplemental subtest called Symbol Search. Five subtests form the Verbal Scale (VS): Information, Similarities, Arithmetic, Vocabulary, and Comprehension. Five Subtests form the Performance Scale (PS): Picture Completion, Coding, Picture Arrangement, Block Designs, and Object Assembly. The Verbal subtests and the Performance subtests are used to compute the Full Scale Score (FS). The description of each test is presented below in the standard order of administration. One supplementary subtest,Test 13 Mazes, is not used in this study because it was not consistently given to all 183 subjects. The WISC-III manual (Wechsler, 1991) states that exploratory and confirmatory factor analysis of the WISC-III was conducted with and without the Mazes subtest scores, and no detectible differences occurred due the exclusion of Mazes.

Test 1, Picture Completion. Common objects and scenes with a missing part are presented to the child. The child's task is to identify the missing part.

Test 2, Information. The child is asked questions about common experiences and learning.

Test 3, Coding. Shapes (Coding A) or numbers (Coding B) are paired with symbols in a key. The child is asked to copy the shape or number that correspondence with the presented symbol as quickly as he/she can. This test is timed, 120 second for each Coding task and bonus points are given for time on Coding A. Coding A is administered to children under the age of eight, and Coding B is administered to children eight years or older.

Test 4, Similarities. The child is orally presented with 19 word pairs in which he/she is asked to explain the similarity between the items.

Test 5, Picture Arrangement. The child is asked to place a series of story pictures in a logical sequence. A time limit is set for the child to arrange each group of stimulus cards, and bonus points are given for time.

Test 6, Arithmetic. The child is asked to mentally solve arithmetic problems. Problems are presented visually and orally and a time limit is set for each problem. Bonus points are given for time.

Test 7, Block Design. The child is asked to physically reproduce a two dimensional abstract design. Children use between four to nine blocks, each with two red, two white, and two red and white sides. A time limit is set for each replication and bonus points are given for time.

Test 8, Vocabulary. The child is asked to define various orally presented vocabulary words.

Test 9, Object Assembly. The child is asked to assembly five puzzles of common objects. A time limit is set for each puzzle and bonus points are given for time.

Test 10, Comprehension. The subtest consists of 18 questions about problem situations, interpersonal relations, and social norms. The child responds orally to these questions and one or two points are given depending upon the quality of the answer.

Test 11, Symbol Search. This subtest is composed of a sequence of paired symbol groups, the target group and the search group. The child is expected to indicate if the search group contains a particular symbol from the target group. Symbol Search A is used with children under the age of eight, while Symbol Search B is used with children eight or above. Both Searches A and B have a 120 second time limit. This subtest can be substituted for Coding in calculating the FS.

Test 12, Digit Span. The child repeats a number sequence that has been present orally by the examiner. In Digit Span Forward the child is asked to repeat the number sequence in the same order as presented; while, in Digit Span Backwards, the child reverse the order of presentation.

The subtest task of the WISC-III are similar to the WISC-R and only (approximately) 27% of the WISC-III items are original (Wechsler, 1991); therefore, much has been retained with only slight modification. The order of presentation of subtests has changed with Picture Completion first, and Information and Coding subsequently. The remaining tests follow in the order of the WISC-R. Color and relevant item content have been used to modernize the test.

The WISC-III Manual (Wechsler, 1991) suggests that four factor based scales can be derived from the WISC-III subtests: Factor 1, Verbal Comprehension (VC) is composed of Information, Similarities, Vocabulary, and Comprehension subtests; Factor 2, Perceptual Organization (PO) is composed of Picture Completion, Picture Arrangement, Block Design, and Object Assembly subtests; Factor 3, Freedom from Distractibility (FD) is composed of Arithmetic and Digit Span subtests; and Factor 4, Processing Speed (PrS) is composed of the Coding and Symbol Search subtests. Normative data on these indexes is not included within the manual.

Normative Sample

The standardization sample consisted of 2,200 children, with 200 children in each age group divided equally according to gender. Race and ethnic representation, coded as Black, Hispanic, White, and Other, were consistent with the 1980 United States Census. Further stratification within ethnic group was by parent education. The United States was divided into four regions and most states were represented within the sample. Approximately 12% of the sample consisted of children with special needs (learning disabled, speech/language impaired, emotionally disturbed, physically impaired or gifted/talented).

Reliability

Average internal consistency reliabilities, or split-half correlations corrected by the Spearman-Brown formula, are .96 for the FS, .95 for the VS, and .91 for the PS (Wechsler, 1991). Subtest split-half reliabilities for all subtests, except Coding and Symbol Search, range from .69 on the Object Assembly to .87 on the Vocabulary and Block Design (Shaw, Swerdlik, & Laurent, 1993). Coding (average reliability of .79) and Symbol Search (average reliability of .76) are timed tested; therefore, test-retest reliabilities were used (Wechsler, 1991).

Stability coefficients, or test-retest reliabilities, on a sample of 353 children (divided into three age group), who were retested after an interval of 12 to 63 days ranged from FS of .94, VS of .93, and PS of .87 (Wechsler, 1991). The subtest reliabilities ranged from .57 on Mazes to a .85 on Information (Wechsler, 1991). Reliabilities varied with age and exhibited an increase with age (Edwards & Edwards, 1993). Average standard errors of measurement ranged from 1.08 on Vocabulary to 1.67 on Object Assembly with 3.53 for VS Score, 4.54 on PS Score, and 3.2 on FS Score (Wechsler, 1991).

Criterion-Related Validity

Evidence of criterion-related validity for the WISC-III has been indicated by comparison with other intelligence test, academic tests, school grades and several multivariate classification studies. Some of the intelligence tests used for comparisons are: WISC-R, Wechsler Preschool and Primary Scale of Intelligence-Revised (WPPSI-R), Wechsler Adult Intelligence Scale-Revised (WAIS-R), Otis-Lennon School Ability Test (OLSAT), and the Differential Ability Scales.

The WISC-III and the WISC-R correlations are the highest of all the intelligence test comparisons, with WISC-III and WISC-R VS scores correlating .90, PS scores .81, and the FS scores .89. When mean scale score differences were compared the FS scores were approximately 5 points less for the WISC-III than the WISC-R, the VS scores were 2 points less, and the PS scores were 7 points less. These correlations were obtained from a sample of children (N=106) ranging in age from 6 to 16 years old and the median interval between testing was 21 days. The sample consisted of 70% Whites, 19% Blacks, 8% Hispanic, and 3% of other ethnic origin. The ratio of female to male participants was 55 female, and 45 male. When a clinical sample (57% LD, 35% ADHD, and 8% anxiety/depression) was used the correlations were similar: .86 VS, .73 PS, and .86 for FS. There was an approximate 7 point score difference in the FS scores (WISC-III lower) and a 5 point difference in the VS scores and PS scores.

The WISC-III correlates highly with the WAIS-R and the WPPSI-R: WISC-III VS scores with WAIS-R VS .90 and WPPSI-R VS .85; WISC-III PS scores with WAIS-R PS .80 and WPPSI-R PS .73; and WISC-III FS scores with WAIS-R FS .86 and WPPSI-R FS .85.

A sample of 65 children (ages ranging from 6 to 16 years) were used in the correlations of the OLSAT and WISC-III. WISC-III FS scores correlated highly with the OLSAT Total School Ability Index (TSAI, .73) and moderately (.69 and .59) with the WISC-III VS scores and the OLSAT Verbal School Ability Index (VSAI) and the WISC-III PS scores and the OLSAT Nonverbal School Ability Index (NVSAI). Cross correlations were, also, low to moderate: VS with NVSAI .44, and .59 PS with VSAI. Moderate correlations (ranging from .51 to .65) were obtained for: VSAI and WISC-III Verbal Comprehension Index (VC), Perceptual Organization Index (PO), and Freedom from Distractibility Index (FD); NVSAI and WISC-III PO; and TSAI and WISC-III PO and Processing Speed Index (PrS).

All WISC-III scores and indices (VS, PS, FS, VC, PO, FD, PrS) correlated high to moderately (.92 to .53) with the DAS General Conceptual Ability (GCA) with the highest correlation between the DAS GCA and the WISC-III FS and the lowest between the DAS GCA and the WISC-III PrS. Evidence of discriminant and convergent validity was exhibited by the DAS Verbal Ability's low correlations with WISC-III PS, PO, and PrS, and high correlations with WISC-III VS, FS, and VC.

Most WISC-III scores and DAS Achievement Tests correlated moderately to low, ranging from a high of .61 for DAS Word Reading and WISC-III FD to a low of .20 DAS Spelling and WISC-III PO. The WISC-III FS scores correlated moderately with the DAS Basic Number SKills (BNS;.54), Spelling (S,.51), and Word Reading (WR,.58). The WISC-III VS correlations with DAS Achievement Tests (.55 BNS, .54 S, and .54 WR) were higher than the WISC-III PS with the DAS Achievement Tests (BNS .35, S .29, and WR .41).

The WISC-III VS scores correlated moderately with the WRAT-R and other group achievement tests on Reading and Arithmetic (median correlations of.66 for Reading and .62 for Arithmetic), while the WISC-III PS scores correlations were generally low with group achievement tests with Arithmetic achievement exhibiting a high of .58. The WISC-III FS scores appeared to be a better predictor for group achievement tests with Total Achievement (.74) and Arithmetic achievement (.68) exhibiting the highest correlations; while, WISC-III FS and WRAT-R subtests correlations exhibiting lower correlations (.53 Reading, .28 Spelling, and .58 Arithmetic).

All WISC-III scores and indices displayed low correlations with school grades. The highest correlation indicated was between WISC-III FS and total grade point average (GPA .47) and WISC-III FS and reading grades (.48).

All multivariate classification studies displayed WISC-III FS, VS, and PS scores in the expected direction. The Gifted study consisted of 38 children ranging in age from 7 to 14 years. The mean FS score was approximately 129 and the amount of variability appeared less than the standardization sample (SD=11). The lowest scores were displayed in FD and PrS indices. The MR sample consisted of 43 children ranging in age from 6 to 16. All FS, VS, and PS scores were depressed (FS=56) and score variability was less than the standardization sample (SD=9). The highest mean score was presented in the PrS index (Coding and Symbol Search subtests), which could indicate that this ability may not contribute greatly to general intellectual ability. The average expected difference between the WISC-R and the WISC-III is approximately 5 points, but this group displayed higher differences: VS 9 points, PS 7 points, and FS 9 points. This suggests that WISC-R/WISC-III discrepancies may be more visible in this disability group.

LD, ADHD, Conduct Disorder (CD) children (LD N=65, ADHD N=68, CD=26) exhibited lower VS, PS, and FS scores when compared to the standardization group (LD FS=93.8, ADHD FS=99.4, and CD FS=78). The PS scores were higher than VS scores and FD and PrS indices were lower in these three groups. The ACID profile, formed by the Arithmetic, Coding, Information, and Digit Span composite score and converted into standard scores (M=100, SD=15), displayed by the LD and ADHD groups was below average. It was over one standard deviation below average with a mean composite of 40; although, the full ACID profile was exhibited by only 5% of the LD group and 12% of the ADHD group. Also, scores from Severe Language Impaired (SLI) children (N=44, ranging in age from 6 to 13) revealed lower mean VS (69), PS (78), and FS (72) scores, with PS scores higher than VS as expected for this group. The PO index was higher than the other indices, also.

Correlations with other tests support the use of the WISC-III as a measure of general intellectual ability; although, small sample sizes were used in the multivariate classification studies. The FD and PrS Indices appear to be less related to general intellectual ability, but may have utility in identifying exceptionality.

Construct Validity

Through the years much evidence has been gathered to support the existence of a global intelligence construct (Brody, 1985; Jensen, 1980) with the WISC-R. Kaufman (1979, 1990) analyzed data from the WISC-R standardization sample using factor analysis with a principal axis extraction method. A large unrotated first factor emerged which he called the g factor and nine of the subtests had factor loadings of .60 or above on this first factor. Because the WISC-III is so similar to the WISC-R, the test manual presents this information to substantiate the use of a large global factor (Wechsler, 1991). Also, the results of a common factor model, or a maximum-likelihood factor analysis using squared multiple correlations (SMCs) as communality estimates, indicated a large first median eigenvalue of 5.61, and a smaller second median eigenvalue of 1.28. This supports use of the FS score as a global intelligence score, as did Kaufman's earlier work on the g factor.

The intercorrelations of the subtests and scales of the WISC-III reveal that all are correlated with the FS Score in differing degrees. According to Shaw, Swerdlik, and Laurent (1993) the mean factor loading of the FS is approximately .66 and is interpreted to mean that 44% of the variance in the FS score is attributable to g.

Two major factors (VC and PO) and two minor factors (FD and PrS) are reported in the WISC-III Manual (Wechsler, 1991). The VC and PO factors were investigated thoroughly by Kaufmann (1975) in his work on the WISC-R, and has been supported by many studies using differing variables of educational placement, ethnic origin, socioeconomic status, and combinations of these variables (Petersen & Hart, 1979; Gutkin & Reynolds, 1980; Reschly, 1978; Carlson, Reynolds, & Gutkin, 1983).

A variety of exploratory factor analysis procedures, including principal components, iterated principal-axis, and maximum-likelihood methods, were used on data from the WISC-III standardization sample (Wechsler, 1991). Analysis was conducted on the total sample (N = 2200) and the sample divided into four age groups: 400, were ages 6-7; 600, were ages 8-10; 600, were ages 11-13; and 600, were ages 14-16. Eigenvalues of one or greater were observed for three factors in three of the age groups, but the younger group only exhibited two. Four or five factors were indicated using the scree test (Cattel, 1966), and the chi-square test suggested four factors for all four age groups. Further rotations using five factors displayed a very weak fifth factor with no loading above .35.

Confirmatory factor analysis using the LISREL-VI program (Joreskog & Sorbom, 1987) was conducted on data from the total sample and the four age group subsamples from the standardization sample (Wechsler, 1991). Five models, ranging from one factor to five factors, were evaluated using a variety of goodness-of-fit indices. The AGFIs for all five models were above .80, which is considered a good model fit (Cole, 1987), except the one factor model at the 14-16 age group. The four factor model's AGFI is the highest in all of the age groups, except the 11-13 group where the five factor model is highest. The RMRs on all factor models were not below .10, which is the suggested level for a good model fit (Cole, 1987). The Tucker-Lewis Index (TLI; Tucker & Lewis, 1973) was used to compare each factor model with the one factor model, and the four factor model exhibited the largest increment although the fifth factor model showed similar increments at two of the age groups. It was not as consistent as the four factor model.

The WISC-III Manual presents further support for the four factors by citing results of three exploratory and confirmatory factor analysis studies using three groups of exceptional children: a clinical sample (N= 167; diagnosed LD and ADHD), a high ability group (N=157; IQ at or above 125), and a low ability group (N=109; IQ at or below 75). The children studied ranged in age from 6 to 16 years. The five factor models, as described earlier, were compared. The four factor model was described as the best fit for the data; although, the Picture Arrangement subtest had loadings on three of the four factors for the clinical sample and appeared as a separate weak fifth factor in the high ability group. As stated earlier, independent researchers (Kamphaus, 1993; Roid, Prifitera & Weiss, 1993; Sattler, 1992, Thorndike, 1992) have, also, presented discrepant information on the two minor factors and the Picture Arrangement and Comprehension subtests.

Summary. Special needs children were included within the standardized sample for the WISC-III. There is much support for use of the FS score as a reflection of global intelligence. Support is evident for the age invariance of two major factors, VC and PO. There has been some disagreement over the presence of the two minor factors, FD and PrS, and the presence of a weak fifth factor.

WJ-R Cog Manual

This section presents information obtained from the WJ-R Cog Manual about the background, development, and validity of the WJ-R Cog.

WJ-R Cog Theory

The WJ-R Cog is based on the Horn-Cattell fluid (Gf) and crystallized (Gc) abilities theory (Cattell, 1963; Horn, 1985; Horn & Cattell, 1966). This theory suggests that there are two major influences affecting the development of cognitive abilities: influences associated with education and cultural experiences which develop Gc abilities, and influences associated with genetic factors and physiological functioning which develop Gf abilities. These two influences may not be entirely distinct or uncorrelated (Horn, 1985).

Cattell (1950) argued that these two abilities show different developmental paths. Fluid ability, or Gf, is more important and rises rapidly in early childhood until adolescence, when it plateaus and declines; while, crystallized ability, or Gc, is more important in later life where it continues to increase through middle adulthood. Gf ability is the foundation upon which Gc abilities are developed (Horn & Cattell, 1967). The Gf-Gc theory recognizes broad versus narrow factors, or it is a "second-order system among the primary factors - a system of factors among factors." (McGrew, Werder, & Woodcock, 1991, p. 213).

Horn (1988) believes that intellectual functioning is comprised of many abilities, and through the use of factor analysis and structural equation modeling he has identified nine broad intellectual abilities. The WJ-R in its entirety measures eight of these nine; although only seven of these abilities are measured in the Standard Battery and Supplementary Battery of the WJ-R Cog.

Description of the WJ-R Cog

The WJ-R Cog is an individually administered, norm-referenced test used to assess intellectual ability of individuals kindergarten through adult (Woodcock & Mather, 1989). The test is composed of three cognitive batteries: the Standard Battery with seven subtests, the Supplementary Battery with seven subtests, and the Extended Battery with seven subtests. This study used the Standard and the Supplementary Batteries with a total of 14 subtests.

The WJ-R Cog Standard Battery is composed of seven tests each measuring a specific ability:

Test 1, Memory for Names measures Long-Term Retrieval (Glr). The subject must recall the names of space creatures, which are unfamiliar auditorially and visually.

Test 2, Memory for Sentences measures Short-Term Memory (Gsm). The subject must repeat words, phrases, or sentences that are auditorially presented.

Test 3, Visual Matching measures Processing Speed (Gs). The subject must match identical numbers that are presented in a row of six numbers.

Test 4, Incomplete Words measures Auditory Processing Ga). The subject identifies complete words after listening to words that are missing one or two phonemes.

Test 5, Visual Closure measures Visual Processing (Gv). The subject identifies pictures of objects that are distorted or have a pattern superimposed upon it.

Test 6, Picture Vocabulary measures Comprehension-Knowledge (Gc). The subject names pictured objects.

Test 7, Analysis-Synthesis measures Fluid Reasoning (Gf). The subject is required to learn a mathematical system to detect the missing element.

The WJ-R Cog Supplemental Battery is composed of seven subtests, each providing further information on the abilities measured in the Standard Battery:

Test 8, Visual-Auditory Learning measures long-term retrieval (Glr). The subject is presented with various symbols that are paired with words. He/she is asked to read these symbols.

Test 9, Memory for Words measures short-term memory and attention (Gsm). The subject is asked to repeat lists of unrelated words.

Test 10, Cross Out measures visual processing speed (Gs). This subtest is timed, so that the subject must quickly scan and cross out matching visual stimuli.

Test 11, Sound Blending measures auditory processing (Ga). The subject is auditorially presented with syllables or phonemes of words and he/she must integrate these sounds into a whole word.

Test 12, Picture Recognition measures visual processing (Gv). The subject is first presented with pictures of common objects. Later, he is asked to identify these objects within a field of distracting stimuli.

Test 13, Oral Vocabulary measures comprehension-knowledge or crystallized intelligence (Gc). This test is composed of two parts: Synonyms, where the subject gives a word similar in meaning to the words presented; and Antonyms, where the subject gives a word opposite in meaning to the word presented.

Test 14, Concept Formation measures reasoning or fluid intelligence (Gf). The subject is presented with a stimulus set and asked to deduct a rule from the set. He/she should not draw upon past learning.

The fourteen subtest scores can be combined to form seven cognitive factor scores. Test 1, Memory for Names, and Test 8, Visual-Auditory Learning, combine to form the Long-Term Retrieval (Glr) Cognitive Cluster. Glr has a median reliability of .90 in the 6 to 18 age range. Short-term Memory (Gsm) is comprised of Test 2, Memory for Sentences, and Test 9, Memory for Words. Gsm has a median reliability of .88 in the 6 to 18 age range. Processing Speed (Gs) consists of Test 3, Visual Matching, and Test 10, Cross Out. It has a median reliability range of .83 for the school age groups. Test 4, Incomplete Words, and Test 11, Sound Blending, compose the Auditory Processing Cluster (Ga). Ga has a median reliability of .88 for ages 6 to 18. Test 5, Visual Closure, and Test 12, Picture Recognition, combine to form the Visual Processing Cluster (Gv). Gv has a median reliability of .86 in the Kindergarten to Grade 12 range. Comprehension-Knowledge (Gc) is formed from Test 6, Picture Vocabulary, and Test 13, Oral Vocabulary. Gc has a median reliability of .91 for the 6 to 18 age range. Fluid Reasoning (Gf) is a combination of Test 7, Analysis-Synthesis, and Test 14, Concept Formation. The Gf median reliability for the ages 6 to 18 is .94.

Normative Sample

The WJ-R norming sample consisted of 6,359 subjects in over 100 different geographical locations (McGrew, Werder, &

Woodcock, 1991). The kindergarten to 12th grade sample (3,245 subjects) comprised more than 50% of the total normative sample. Norming subjects approximated the U.S. population distribution on ten variables: census region, community size, sex, race, origin, funding of college, type of college, education of adults, occupational status of adults, and occupation of adults in the labor force. Special education students, who were enrolled in mainstream classes for a portion of their school day, were also included in the norming sample (McGrew, Werder, & Woodcock, 1991). Norm tables provide grade and age equivalent scores and percentile ranks and standard scores (Woodcock & Johnson, 1987).

Reliability

Reliability, or score consistency, is concerned with that portion of the total variance of test scores which is error variance (Anastasi, 1988). Split-half coefficients using the odd and even test scores were corrected using the Spearman-Brown correction formula on all tests except those tests that are speeded. Visual Matching and Cross Out are the only tests used in this study affected by this. Test-retest correlations were calculated for these tests. The split-half coefficients expresses the measure of consistency in relation to content sampling and is often called the coefficient of internal consistency; while, the test-retest procedure expresses stability (Anastasi, 1988).

Reliability coefficients are reported from ages two to beyond 80 (McGrew, Werder, & Woodcock, 1991). For the purposes of this study only the coefficients between the ages of five and 18 will be reported for each of the cognitive tests. Reliabilities reported for the Standard Battery are: Memory for Names ranged from .940 (age 15) to .878 (age 5); Memory for Sentences ranged from .899 (age 15) to .805 (age 13); Visual Matching ranged from .838 (age 15) to .725 (age 17); Incomplete Words ranged from .889 (age 5) to .603 (age 13); Visual Closure ranged from .822 (age 5) to .621 (age 14); Picture Vocabulary ranged from .923 (age 18) to .700 (age 5); and Analysis-Synthesis ranged from .929 (age 5) to .812 (age 13).

Reliabilities for the Supplementary Battery are: Visual-Auditory Learning ranged from .922 (age 18) to .861 (age 13); Memory for Words ranged from .826 (age 6) to .687 (age 13); Cross Out ranged from .747 (age 18) to .640 (age 6); Sound Blending ranged from .920 (age 6) to .765 (age 13); Picture Recognition ranged from .851 (age 18) to .709 (age 6); Oral Vocabulary ranged from .904 (age 18) to .851 (age 13); and Concept Formation ranged from .953 (age 6) to .927 (age 18).

Standard errors of measurement (SEM), which is an estimate of the amount of error associated with an obtained score (Sattler, 1988), ranged from 2.7 (Memory for Names, age 6) to 10.1 (Memory for Words, age 18) using the W scale. The W scale, using the Raush Model, provides a common scale representing the individual's ability and the difficulty of the tasks that are sampled by the measurement (McGrew, Werder, & Woodcock, 1991).

The Broad Cognitive Ability-Standard Scale (BCA) reliability coefficients ranged from .961 (age 15) to .915 (age 13) and the SEMs, using the W scale, ranged from 1.8 (ages 12 and 8) to 2.2 (age 17). The BCA is a composite score derived from the seven standard battery subtests. It is considered a general measure of cognitive ability.

Criterion-Related Validity

The manual reports concurrent validity correlations of the WJ-R Cog BCA and ability clusters with other intelligence tests for two age groups, 9 and 17 years (McCullough & Wiebe, 1991, data with uncorrected correlations). The manual uses the WJ-R preliminary standard scores calculated for whole year age groups in comparison to the WISC-R, SB-IV, and K-ABC. Correlations range from .570 (WJ-R Cog and K-ABC Composite) to .685 (WJ-R Cog and WISC-R FS and SB-IV Composite) for the 9 year old group and from .640 (WAIS-R FS with WJ-R Cog) to .648 (WJ-R

Cog and SB-IV Composite). WJ-R Cog factors of Gsm, Gc, Gf, and Gq correlated moderately to high (.500 or above) with the: K-ABC Sequential Processing and Achievement; SB-IV Verbal Reasoning, Abstract/Visual Reasoning, Quantitative Reasoning, and Short-Term Memory; and WISC-R Verbal Scale. All correlations with the WJ-R Cog factors of Glr, Gs, Ga, and Gv were low (.468 or lower). These findings would support the interpretation that the WJ-R Cog measures general intellectual ability, but the WJ-R Cog measures unique abilities, also.

The manual briefly cites the Evans and Carsen (1991) study which used data from four groups of children (9 and 17 years): gifted, normal, learning disabled (LD), and mentally retarded (MR). All of the means, BCA and the seven cognitive clusters, were in the expected pattern, with MR the lowest and gifted the highest.

Construct Validity

The WJ-R Cog manual (Woodcock & Mather, 1989) presents intercorrelation information on the seven cognitive tests at nine age levels. Low intercorrelation among the seven standard cognitive subtests were found which provides support that each subtest is measuring different aspects of cognitive ability. High intercorrelations were displayed for subtests purporting to measure the same ability from the standard battery, supplementary battery, and extended battery.

A sample of 2,415 subjects (kindergarten through adult) from the norming sample was used for a confirmatory factor analysis of the fourteen subtests (standard battery and supplementary). Age was controlled for by multiple regression and residual correlations were calculated. The seven factor orthogonal rotation indicated loadings between .7 and .8 on each factor by two subtests each. Each factor contributes 9.3% to 12.5% of the total score variance, or 79.6% total score variance is accounted for by the seven factors.

Validity Studies

WISC-R/WISC-III Studies

Criterion-related validity. There have been many studies comparing the WISC-R Full Scale (FS), Verbal Scale (VS), and Performance Scale (PS) to other intelligence tests, achievement tests, and behaviors (i.e. school grades) and most have supported the validity of the WISC-R (Kaufman, 1979).

Intelligence test correlations with the WISC-R FS are: Stanford-Binet: Fourth Edition .78 (Thorndike, Hagen, & Sattler, 1986); Slosson Intelligence Test .61 (Jeffrey & Jeffrey, 1984); McCarthy Scales of Children's Abilities .72 (Arinoldo, 1982); and Kaufman Assessment Battery for

Children .70 (Naglieri & Haddad, 1984). These correlations indicate that the WISC-R has satisfactory concurrent validity with other intelligence tests (Sattler, 1988).

Wide Range Achievement Test (WRAT) reading achievement correlations with the WISC-R: VS .57, PS .34, and FS .56; while, arithmetic tests correlations were VS .62, PS .46, and FS .52 (Appelbaum & Tuma, 1982; Sattler, 1988). Average reading achievement correlations are approximately .66 for VS, .47 for PS, and .65 for FS; while, average arithmetic achievement correlations are approximately .56 for VS, .48 for PS, and .58 for FS (Sattler, 1988). School grades and WISC-R FS have a median correlation of .39 (Sattler, 1988). The WISC-R VS and FS scores correlate higher with achievement than the WISC-R PS scores.

A current study by Hishinuma and Yamakawa (1993) used: the WISC-III; WISC-R; Detroit Tests of Learning Aptitude, 3rd Edition (DTLA-3); Peabody Picture Vocabulary Test, Revised (PPVT-R); Wide Range Achievement Test-Revised (WRAT-R); and the Piers Harris Children's Self Concept Scale. They found that WISC-III and the WISC-R VSs were highly correlated with Verbal subtests, as were the WISC-III and the WISC-R PSs with Performance subtests; although, as expected the cross correlations (VS with performance subtests and PS with verbal subtests) were not. This suggests good discriminant and convergent validity. FS scores were lower on the WISC-III as compared to the WISC-R and same name subtest correlations ranged from a low of .38 (Picture Arrangement) to a high of .79 (Information). The WISC-III FS correlated .70 with the DTLA-3 General Mental Ability score, .66 between the WISC-III VS and the DTLA-3 verbal composite, and .54 between the WISC-III PS and the DTLA-3 nonverbal composite, also, suggesting evidence of convergent and discriminant validity. The PPVT-R correlations were higher: .75 with WISC-III FS, .81 WISC-III VS, and .82 with WISC-III VC Index. This suggests that the WISC-III FS is highly weighted toward verbal abilities, because the PPVT is usually considered a measure of receptive vocabulary (Dunn & Dunn, 1981). Correlations with the WRAT-R ranged from .51 to the WISC-III FS and .76 with the WISC-III VS (WISC-III VS correlated .65 - WRAT-R Reading; .63 - WRAT-R Spelling; and .76 - WRAT-R Arithmetic), suggesting moderate to high validity when predicting academic ability with the WISC-III VS. Low correlations were found between the WISC-III and the Piers-Harris and this supports discriminant validity, because the Piers-Harris is a personality inventory intended to measure self-concept. Hishinuma and Yamakawa (1993) urge caution in interpreting these correlation because of the small sample size (N=42).

Weiss, Prifitera, and Roid (1993) concluded that the WISC-III is a good predictor of academic achievement across gender and ethnic groups using two large sample groups representative of the U.S. Census Bureau statistics of 1988. Sample one consisted of 700 children, aged 6 to 16 years, who were administered various tests: Comprehensive Test of Basic Skills; Iowa Tests of Basic Skills; California Achievement Test, Sixth Edition; and the Stanford Achievement Test, Edition 7 Plus. Sample two consisted of 1000 children aged 6 to 16 years and classroom grades were used with this group.

The WISC-R and the WISC-III has been shown to differentiate between groups of special education categories, such as Learning Disabled (LD) and Emotionally disturbed (ED), and control groups (Fuller & Goh, 1981; Prifitera & Dersh, 1993). Although use of Kaufman's Freedom from Distractibility Factor (FD) (Arithmetic, Coding, and Digit Span subtests; Kaufman, 1975, 1979) or the ACID profile (lowered scores on Arithmetic, Coding, Information, and Digit Span; Kavale & Forness, 1984) on the WISC-R for measures of attention and distraction for ADHD and LD children has been challenged (Ownby & Matthews, 1985; Stewart & Moely, 1983), several current researchers have suggested using the WISC-III factors FD and Processing Speed (PrS) in this way.

Schwean, Saklofske, Yackulic, and Quinn (1993), using a sample of 45 children between the ages of 8 and 11, found Attention Deficit Hyperactivity Disorder (ADHD) children generally performed lower on the PrS, FD, and the Verbal Comprehension (VC) factors and higher on the Perceptual Organization (PO) factors. Although Schwean et al. concluded that the WISC-III measures similar constructs in the ADHD sample as in the standardization sample, much variability was exhibited in the subtest scores.

Newby, Recht, Caldwell, and Schaefer (1993) using a small sample (N=26) of LD children (ranging in age from 9 to 13) supported the ACID profile analysis when they found FD and PrS factors lower than VC and PO factors. Hishinuma and Yamakawa (1993) supported these results. They found that the WISC-III measured similar constructs in a sample of "at risk" children as in the normative sample; although these children had significantly lower VC, FD, and PrS factor scores.

Prifitera and Dersh (1993) suggest that further research is needed to investigate whether the ACID pattern replicates itself with other LD and ADHD samples, and if the ACID pattern is present with other clinical groups and not unique to LD and ADHD classification. Shaw, Swerdlik and Laurent (1993) after reviewing the research in this area concluded that there is not sufficient research on the WISC-

III at this time to use factor scores or profile analysis for making individual differential diagnosis and that much more research is needed to justify using the WISC-III factor scores in this way.

Construct validity. The WISC-III is purported to measure four underlying constructs: Verbal Comprehension (VC), Perceptual Organization (PO), Freedom from Distractibility (FD), and Processing Speed (PrS) (Wechsler, 1991). Three independent construct validity studies have been done to investigate this.

Sattler (1992), using exploratory factor analyses and the conventional method of eigenvalues greater than 1.0 on eleven different age groups of the standardization data, found that the FD factor was not consistent across age groups; although, the VC, PO, and PrS factors were consistent. Sattler's findings have been disputed on the grounds that his analysis contained more error variance due to the smaller number of subjects in each analysis unit (Roid, Prifitera, & Weiss, 1993).

A WISC-III independent study, conducted by Roid, Prifitera, and Weiss (1993) of the 1,118 children who represented the standardization group of the Wechsler Individual Achievement Test (WIAT; Psychological Corporation, 1992), utilized both exploratory and confirmatory factor analyses. There were approximately 100 children in each year level from 6-16.

The exploratory analysis included principal components, iterated maximum likelihood methods of extraction, and contrasting rotations. Past analysis with the Wechsler scales have traditionally used communality estimates in the diagonals of the correlation matrix (Wechsler, 1991), but this study used squared multiple correlations (SMCs) with subsequent iteration. A large first factor eigenvalue of 5.7 and a smaller second factor eigenvalue (1.2) suggested the presence of Kaufman's g or a global factor. These results are similar to the WISC-III Manual (Wechsler, 1991) results (5.61 for the first eigenvalue and 1.28 for the second).

After a series of extractions, three eigenvalues were larger than or equal to 1.0 (the VC factor, the PO factor, and the PrS factor - Arithmetic and Digit Span composed part of the VC factor). The scree test as proposed by Cattell (1966) suggested four factors, but the maximum likelihood analysis (principal analysis with SMCs in the diagonals) using the chi-square goodness of fit model indicated the presence of five factors. Further rotations exhibited the fifth factor with only one variable, the Comprehension subtest.

The confirmatory factor analysis methods were similar to the confirmatory methods used in WISC-III Manual study except for Model 5A, which consisted of VC, PO, FD, PrS, and Comprehension, and Model 5B, which consisted of VC, PO, PrS, Memory (Digit Span), and Numerical Ability (Arithmetic). All AGFIs were above .80, but the four factor model (.952) and the five factor model 5B (.956) displayed the highest estimates. Likewise, these two models exhibited the lowest RMRs (.034 for model four and .032 for model 5B). Models were compared for greatest contrast improvement with model one vs. model two showing the greatest improvement; and, model two compared to model four showing the second highest improvement. The fifth factor, Comprehension, which emerged in the exploratory analysis was not significant using the compared models chi-squared difference. Also, Roid et al. (1993) felt that the RMRs difference of .002 between the four factor and the 5A factor models was not significant and that conservatively the four factor model was the best fit.

Hishinuma and Yamakawa (1993) using an "at-risk" special education sample (N=78; ages ranged from 6 to 16 years old) conducted an unrotated principal-component factor analysis and a rotated varimax factor analysis of the subtest performances. The unrotated principal-component analysis supported the concept of the 13 subtests measuring a global entity or "g", although the ratio of the first two eigenvalues (2.6 to 1) was not as large as that found in the WISC-III manual (4.5 to 1). The rotated varimax factor analysis suggested that the Picture Arrangement subtest loaded on PrS rather than PO, but the four factor model with two strong factors and two weak factors was supported with this particular sample. The authors suggest that further research is needed on the WISC-III factor structure and its utility with exceptional children.

Thus, when the results of both exploratory and confirmatory factor analysis on independent samples are considered, according to Roid et al. (1993), there is support for two major factors and two minor factors. Roid rationalizes that the three factors put forth by Sattler (1992) are too conservative, the five factors suggested by their own maximum likelihood chi squared criteria are too liberal, and that the four factors supported by the scree test "...provide the most psychologically meaningful solution." (Roid et al., 1993, p.10).

WJ-R Studies

Criterion-related validity. The criterion related validity of the WJ-R has been reported using various intelligence tests, academic tests, and a multivariate classification study.

Scarr (cited in McGrew, Werder, & Woodcock, 1991) studied the relationship of the WJ-R, Tests of Cognitive

Ability (Standard Battery) to the WISC-R with a group of 94 third graders (M=9.4 years) from rural north central Texas. The correlation for the WJ-R Cog and the WISC-R was .732. McCullough and Wiebe (cited in McGrew et al., 1991) using two age groups (72 third/fourth graders and 51 tenth/eleventh graders) administered the WJ-R Cog, WISC-R or the Wechsler Adult Intelligence Scale-Revised (WAIS-R), Kaufman Assessment Battery for Children (K-ABC), and Stanford-Binet Intelligence Scale-Fourth Edition (SB-IV). Correlations with the WJ-R Cog ranged from .703 (SB-IV composite score),.668 (WISC-R), and .614 (K-ABC composite) for the younger group; plus, .654 (SB-IV composite) and .697 (WAIS-R). The WJ-R Cog correlations were similar to the other cognitive tests (K-ABC correlated .663 with WISC-R and .719 with SB-IV in the 3/4 grade sample and SB-IV correlated .698 with the WAIS-R in the 10/11 grade sample). Thus, the correlations using the WJ-R Cog and other intelligence tests display evidence indicative that the WJ-R Cog is measuring general intelligence (McGrew et al., 1991).

Evidence of convergent and discriminant validity are present from correlations of the WJ-R Cog subtests and the subtests of the WISC-R, K-ABC, and the SB-IV. The WJ-R Long-Term Retrieval (Glr), Auditory Processing (Ga), and Visual Processing (Gv) did not exhibit a strong convergent relationships with other battery subtests and only a moderate relationship with one other battery subtest (Glr correlated .430 with SB-IV Short Term Memory, STM; Ga correlated .448 with K-ABC Achievement, Ach; and Gv correlated .387 with K-ABC Simultaneous Processing, Sim). It would appear that these subtests are measuring unique abilities.

WJ-R Short-Term Memory (Gsm) correlated .656 with K-ABC Sequential Processing (Seq), .531 with K-ABC Ach, .714 with SB-IV STM, and .447 with the WISC-R Verbal Scale (VS). WJ-R Processing Speed (Gs) correlated above .583 with only one other score, the WISC-R Freedom from Distractibility (FD). WJ-R Comprehension-Knowledge (Gc) correlated at or above .510 with five other battery subtests (K-ABC Ach, SB-IV Verbal Reasoning (VR), SB-IV STM, WISC-R VS and WISC-R VC). WJ-R Fluid Reasoning (Gf) correlated at or above .514 with ten other battery subtests: all three K-ABC Scales, SB-IV Abstract/Visual Reasoning (AVR), SB-IV Quantitative Reasoning (QR), SB-IV STM, and WISC-R VS, PS, VC, and PO. The WJ-R Quantitative Ability (Gq) correlated at or above .608 with five other battery subtests: K-ABC Ach, SB-IV QR, WISC-R VS, WISC-R VC, and WISC-R FD. McGrew et al. (1991) interpret these correlations with other battery subtests as indications of subtest concurrent ability and as evidence of the WJ-R's ability to measure eight distinct cognitive abilities.

Although the WJ-R manual insists that the WJ-R Scholastic Aptitude Tests (Schol Apt) are to be used for achievement prediction, several studies have been done using the WJ-R Cog. Using the Scarr (cited in McGrew et al.,1991) study sample, as discussed above, the WJ-R Broad Cognitive Ability (BCA) and the WISC-R FS were correlated with the WJ-R Ach Broad Reading and Broad Mathematics Clusters, and the PIAT-R Total Reading and Mathematics scores. All correlations were in the expected direction, but the WJ-R BCA appeared to correlate somewhat higher than the WISC-R FS with all four achievement measures. McCullough and Wiebe (cited in McGrew et al., 1991), using both age samples as mentioned above, correlated the WJ-R BCA, WISC-R FS or the WAIS-R, SB-IV Composite, K-ABC Mental Processing Composite with the following achievement tests: WJ-R Achievement subtests and clusters; Basic Achievement Skills Individual Screener (BASIS) Reading, Mathematics, and Written Language Tests; K-ABC Reading Composite, Arithmetic, Faces and Places, and Riddles; Kaufman Test of Educational Achievement (K-TEA) Reading Composite, Math Composite, and Spelling; PIAT Reading Composite, Mathematics, Spelling, and General Information; and WRAT-R Reading, Mathematics, and Spelling. This study supported the findings of Scarr (McGrew et al., 1991) in that the WJ-R Cog appears to be more highly correlated with academic achievement than the other intelligence tests; although the WJ-R Schol Apt clusters appears to be higher than the WJ-R Cog.

A multivariate classification study presents evidence that the WJ-R Cog differentiates between various special education group memberships. Evans and Carlsen (cited in McGrew et al., 1991) using two age groups found that WJ-R Cog group means for Gifted (120), Normal (102 and 107), Learning Disabilities (83 and 78), and Mentally Retarded (58 and 60) were all in the expected direction. The data from this study was subjected to a series of multiple discriminant function analyses. Three significant functions were revealed through the correlation between the combined WJ-Cog/WJ-Ach Clusters (Broad Reading, Broad Mathematics, Broad Written Language, and Broad Knowledge) and the group membership criterion. Age differences in group membership were found. In the younger group (Grade 3/4), the group membership differed on language arts and fluid reasoning because Broad Reading (.827), Broad Mathematics (.703), and Gf (.516) exhibited the highest loadings on the first function; whereas, the older group (Grade 10/11) exhibited the highest loadings on Broad Reading (.708), Broad Mathematics (.785), Broad Written Language (.526), and Broad Knowledge (.559) on the first function indicating that differences in group membership was based on academic achievement. The highest loadings on the second function for Grades 3/4 were Gc (.667), Broad Mathematics (.594), and Broad Knowledge (.853). Highest loadings on the second function for Grades 10/11 were Glr (.561), Gsm (.529), and Broad Written Language (.694). This could indicate that more memory was involved in the older group membership classification (Evans & Carlsen, 1991). The third function suggested that more processing was involved in membership classification across the two age groups because Auditory Processing (Ga) and Visual Processing (Gv) both exhibited high loadings. Although, the older group showed a high loading in Comprehension-Knowledge (Gc). Classification accuracy was highest for the normal group (at or above 90%), and lowest for the gifted group (at or above 56.5%). Learning Disabilities (LD) accuracy ranged from 75.4% (Grades 3/4) and 78.1% (Grades 10/11) with the most difficulty in identifying LD from normal. Mental Retarded (MR) classification accuracy ranged from 69.2% (Grade 3/4) and 91.2% (Grades 10/11) with the most difficulty in separating MR from LD.

McGrew et al. (1991) concluded that the combined WJ-R Cog and WJ-R Ach Broad Clusters were strongly related to group membership classification across both age groups; although, they suggest that further research needs to be done to refine and support these interpretation. Research involving group membership classification with more age diversity could add information and help to clarify these age differences .

Construct validity. Woodcock (1990) presents confirmatory and exploratory factor analysis information obtained from three data subsets from the norming sample.

1) A sample of 2,261 subjects, kindergarten to adult, were administered the 16 primary measures (seven standard battery cognitive tests, seven supplementary battery cognitive tests, and two mathematics achievement standard battery tests) for the eight cognitive factors of Glr, Gsm, Gs, Ga, Gv, Gc, Gf, and Gq. Confirmatory, with an oblique rotation, and an exploratory factor analysis procedures were carried out on this data set.

2) A sample of 1,425 subjects were administered the 16 tests mentioned above, plus five cognitive supplementary measures (Numbers Reversed, Sound Patterns, Spatial Relations, Listening Comprehension, and Verbal Analogies) and six achievement supplementary measures (Science, Social Science, Humanities, Word Attack, Quantitative Concepts, and Writing Fluency). Confirmatory factor analysis was used on this data set.

3) A sample of 3,063 subjects were administered the 27 tests mentioned above plus two delayed recall tests (Delayed Recall-Memory for Names and Delayed Recall-Visual-Auditory Learning). Confirmatory and a oblique rotation factor

analysis was conducted on this data set by two independent researchers (Woodcock, 1990).

Age was controlled for by a multiple regression technique and the maximum likelihood fitting function in the LISREL-VII program (Joreskog & Sorbom, 1989) for confirmatory factor models was used. The hypothesized model factor structure consisted of two subtests contributing to each of the eight factors (McGrew et al., 1991, p. 164). Three goodness-of-fit indexes were used for evaluation: the Goodness-of-Fit Index (GFI), the Adjusted Goodness-of-Fit Index (AGFI), and the root-mean-square residual (RMR). The GFI was .976, and the AGFI was .957 (1 indicates a perfect fit). GFIs and AGFIs at or above .80 are indicators of good model fit (Cole, 1987). The RMR was .023 and according to Cole (1987) values of .10 are indicators of good fit. Positive and high factor loadings on all sixteen tests for the eight Gc-Gf factors were indicated.

Exploratory factor analysis, using the principal axes procedure where R squared is used in the diagonals and both orthogonal and oblique rotations, extracted the eight Gc-Gf factors (seven measured in the cognitive portion and one measured in the achievement portion). The seven cognitive tests did not show lower secondary loadings in other factors in both the orthogonal and oblique rotations; although, the oblique rotation produced the cleanest results. When the exploratory and the oblique confirmatory analysis are compared much similarity is recognized. Both exploratory and confirmatory analysis reinforce the concept of seven cognitive tests each measuring one of seven cognitive abilities.

Analysis for age invariance of the Gf-Gc model is presented by McGrew et al. (1991). The age invariance study was conducted on sample one data (as mentioned above, using 16 tests) which consisted of six different age groups. The factor loadings and model fit statistics indicated that all eight cognitive abilities are invariant across age levels, but when age invariance was examined for sample three data (3,063 subjects taking 27 tests) the older aged group had weaker fit statistics. Although the model fit this group, this could indicate that more research is needed to explain this phenomenon.

Performance on the complete battery of tests (29 tests) was related to three alternative factor structure models using confirmatory modeling (McGrew et al., 1991). The first alternative structure consisted of one single factor or a large general factor. The second and third alternative structures utilized a verbal/nonverbal dichotomy model and a hierarchical model where the eight first-order factors were assumed to be correlated and a second-order g or general intelligence factor would emerge. Data from sample one were used for the analysis. The fit statistics indicated that the Gf-GC model was the most representative of the data and the hierarchical g model was a weaker second.

No factor analytical study has been done at this time on the WJ-R Cog using a sample of students who exhibit academic difficulties and are referred for possible special education. This study could confirm the existence of the seven WJ-R cognitive factors in a sample with a wide range of abilities.

<u>Studies Comparing WISC-R/WISC-III and the WJ/WJ-R</u>

Although the average correlations between the WISC-R FS and the WJ Cognitive Ability Cluster is .77 across 19 comparisons (McGrew, 1986) thus supporting strong concurrent validity, there has been controversy over the lower WJ Cog scores in comparison to the WISC-R FS scores (Cummings & Moscato, 1984a, 1984b; Thompson & Brassard, 1984; Woodcock, 1984a, 1984b). The median mean score discrepancies, found across 21 studies with diverse samples, are approximately five to six scaled score points (McGrew 1986, 1987; Woodcock, 1984a) and larger discrepancies in LD samples have been reported (McGrew, 1987). Many hypotheses have been suggested to explain these differences (Coleman & Harmer, 1985; Cummings & Moscato, 1984a; Phelps, Rosso, & Falasco, 1984; Thompson & Brassard, 1984). One hypothesis suggests that the differences between the WISC-R and the WJ Cog are due to the proportion of general intellectual ability or g measured in each test (Ysseldyke, Shinn, & Epps, 1981), but McGrew (1984, 1986, 1987) contends that when the first unrotated canonical variate is interpreted as g (Jensen, 1984; Kaufman, 1979) in both the WJ Cog and WISC-R, similar redundancy figures are found (Estabrook, 1984; McGrew, 1987). McGrew (1987) found 23.55% total redundancy for the WJ Cog and 26.89% in the WISC-R for a difference of 3.34%.

Another hypothesis suggests that the WJ Cog is more heavily loaded with verbal content (Shinn, Algozzine, Marston, & Ysseldyke, 1982; Ysseldyke et al, 1980). Estabrook (1984) and McGrew (1987) found that the first rotated variable is verbal ability in both tests, representing the WISC-R subtests of Information, Similarities, Vocabulary, and Comprehension, and the WJ Cog subtests of Picture Vocabulary, Quantitative Concepts, Antonyms-Synonyms, and Analogies. Estabrook (1984) and McGrew (1987) concluded that verbal ability is the greatest area of similarity between the WJ Cog and the WISC-R, and that the approximate 3% difference in this area between the two tests does not support the idea that the WJ Cog is more heavily loaded with verbal content.

A third hypothesis suggests that the WJ Cog and the WISC-R may differ in factor structure or measured underlying constructs (Bracken, Prasse, & Breen, 1984; Estabrook, 1984;

Ipsen, McMillan, & Fallen, 1983; Shinn et al, 1982). It is proposed that the WJ Cog measures more crystallized abilities or academic achievement than the WISC-R, because the WJ Cog exhibits superior academic predictive ability (Ysseldyke et al, 1981). McGrew's (1987) analysis with the WJ Cog, WISC-R, and WJ Achievement Tests does not support this. He found that the total redundancy difference between the WJ Cog and the WISC-R with the WJ Ach was only 3.65% and suggested that the WJ Cog superior predictive ability was due to its measure of unique abilities, such as Fluid Reasoning (Gf) and Auditory Processing (Ga) measures, not assessed by the WISC-R.

Woodcock (1990) examined the factorial congruence of the WISC-R and the WJ-R Cog with several different data sets. He concluded that the WJ-R did measure unique abilities, such as Fluid Reasoning (Gf) and Auditory Processing (Ga); although the WJ-R and the WISC-R have similar measures, such as the WISC-R Verbal Comprehension (VC) factor and the WJ-R Comprehension-Knowledge (Gc) factor, and the WISC-R Perceptual Organization (PO) factor and the WJ-R Visual Processing (Gv). Woodcock did not find the WISC-R FD factor when the WJ-R Cog, SB-IV, and the K-ABC were factorially analyzed with this data set.

A validity study conducted by McCullough and Wiebe (cited in McGrew et al., 1991) on a sample of third and fourth graders revealed moderate to high correlations between the WJ-R Cog Comprehension-Knowledge (Gc) measurement and the WISC-R VS (.779) and VC factor (.768), and a low correlations with the WISC-R PS (.302). WJ-R Cog Fluid Reasoning (Gf) correlated moderately with four the WISC-R measures: VS .645, PS .525, VC .695, and PO .555. WJ-R Cog Quantitative Ability (Gq) correlated moderately with WISC-R VS .681, VC .645, and FD .608. WJ-R Cog Processing Speed (Gs) correlated moderately (.583) with one WISC-R measure, FD. All other WJ-R Cog measures (Long-Term Retrieval - Glr, Short Term Memory - Gsm, Auditory Processing - Ga, and Visual Processing - Gv) did not correlate above .500 with the WISC-R measures (VS, PS, VC, PO, and FD). This suggests that while the WJ-R Cog does have similar measures with the WISC-R, such as Verbal Scale, Verbal Comprehension, Performance Scale, Perceptual Organization, and Freedom from Distractibility, the WJ-R Cog does measure unique abilities.

One current study examines both the WISC-III and the WJ-R Cog Fluid Reasoning (Gf) cluster (Analysis-Synthesis and Concept Formation subtests) with the WJ-R Achievement Broad Measures (Broad Reading, Broad, Mathematics, Broad Written Language, and Broad Knowledge). Teeter and Smith (1993) examined the relationship of the WISC-III VC, PO, FD, and PrS to the WJ-R Cog Fluid Reasoning with emotionally disturbed children (ages 11 to 16 years).

Predictive and discriminant validity of WISC-III scores and WJ-R Gf scores was supported for both the control group and the emotionally disturbed (ED) group. Discrimination accuracy rates were 100% for the ED group and 86.7% for the normal group with an overall accuracy rate of 93%, when nine variables were used for analysis (WISC-III VC, PO, FD,and PrS; WJ-R Cog Gf; WJ-R Ach Broad Reading, Broad Mathematics, Broad Written Language, and Broad Knowledge).

The WJ-R Gf displayed a high correlation with WISC-III PO in both ED (.71) and control (.66) groups; although, Gf correlations in the control group were high to moderate on the other three WISC-III factors (.70 VC, .63 FD, and .51 PS). Three ED group Gf correlations were moderate to low: .57 VC, .56 PS, and.42 FD. This could indicate that Gf may be more strongly related to verbal abilities and attention in the control group than in the ED group and more related to nonverbal ability and processing speed in the ED group as compared to the control group. In the ED group FD appeared more related to PS than in the control, but in the control group it appeared more related to Gf. Teeter and Smith suggested that the FD factor may be measuring motivation or non-intellectual abilities for the ED group. The exact nature of the FD factor is somewhat vague.

Teeter and Smith (1993) cautioned that while it does appear that the WISC-III and the WJ-R Cog Gf tests may be measuring cognitive abilities differently for the ED and control groups, further factor analytical research using a large sample with the full WJ-R Cog Standard Battery administered would help to clarify this issue.

Summary

Are ability inferences based on the performances of the WISC-III and the WJ-Cog standardization samples applicable to students exhibiting academic difficulties? Woodcock (1980) recommends using a sample population taken from students referred for special services but not as yet identified according to disability to examine the applicability of the norms from the standardization sample. According to Woodcock, the referral sample will offer a wide range of abilities that will help to make the analysis more accurate. No research using the WJ-R Cog with a referral group has been reported, and only one study using 78 "exceptional" and "at risk" students for the WISC-III is reported (Hishinuma & Yamakawa, 1993). Although, the WISC-III manual (discussed more thoroughly in the next chapter) does describe three confirmatory factor analysis studies of three different groups of exceptional children: learning disabilities, high-ability gifted, and low-ability.

Does the WISC-III and the WJ-R Cog measure similar or different abilities? It appears that the WISC-R/WISC-III and the WJ-R Cog do measure similar and unique cognitive abilities as indicated by one independent study; although, not much research has been done using both the WISC-III and the complete WJ-R Cog Standard Battery. Research done in the past has used the standardization sample of the WJ-R Cog with the WISC-R, or a specific disability group within a narrow age range and only one cognitive factor of the WJ-R Cog (Gf).

Which test predicts academic achievement better? Neither test, WISC-III nor WJ-R Cog, has been used in conjunction with an academic test on the same sample population so that predictive validity can be evaluated. Does the WJ-R Cog predictive validity surpass the WISC-III, as did its predecessor the WJ and the WISC-R?

Are there unique cognitive patterns related to specific special education classification? Research with the WISC-III suggests that the ACID profile (lowered FD and PrS factors) discriminates LD and ADHD students from controls. Although, all articles presenting this information suggest further research to replicate these results. No specific patterns have been suggested for the WJ-R Cog.

CHAPTER 3

METHODS

The present study was designed to examine the criterion-related and construct validity of the WISC-III and WJ-R Cog with a sample of children who have been referred for possible special education placement due to academic difficulties. One of the primary concerns of this study was whether the same underlying cognitive constructs are measured in a referral group as are measured in the normative sample. Another purpose of this study was to examine the relationship of the two cognitive tests to each other. Lastly, this study examined the relationship of the cognitive measures to measures of achievement.

Subject Sample

A sample of 183 students files were randomly selected from the school year 1992/93 school assessment records from a large urban school district in southern Arizona. The only requirement to selection was that the student file had all three tests (WISC-III, WJ-R Cog, and WJ-R Ach) and that subtests scores be recorded. The students ranged in age from 6 to 16 years and from first to tenth grade. The students were divided into three age groups for descriptive purposes. Table 1 presents further descriptive data on the

Table 1

Descriptive Characteristics of Total Sample (N=183)

Age	n	Sp Ed		Ethnic	
75 to	66	Class	n	Code	n
107 months		#1	12	White	40
(6-03 to		#2	35	Black	3
8-11)		#3	14	Hispanic	22
		#4	1	Other	1
		#5	3		
		#6	1		
108 to	73	Class	n	Code	n
143 months		#1	6	White	37
(8-0 to		#2	36	Black	3
11-11)		#3	22	Hispanic	30
		#4	4	Other	3
		#5	3		
		#6	2		
144 to	44	Class	n	Code	n
192 months		#1	1	White	23
(12-0 to		#2	25	Black	3
16-0)		#3	9	Hispanic	16
		#4	3	Other	2
		#5	6		
		#6	0		

Special Education Class: 1 = Speech/Language Impairment; 2 = Specific Learning Disability; 3 = Referred, but Not Placed in special education; 4 = Mild Mental Retardation; 5 = Emotionally Disabled; 6 = Other Health Impairments.

students in the sample including their ethnicity and placement decision.

Instruments

Woodcock-Johnson Psycho-Educational Battery, Revised Tests of Cognitive Ability

The WJ-R Cog and the fourteen subtests used in this study were described in detail in Chapter 2.

Wechsler Intelligence Scale for Children - Third Edition

The WISC-III and the twelve subtests used in this study were described in detail in Chapter 2.

The Woodcock-Johnson Psycho-Educational Battery, Revised Tests of Achievement

Description of WJ-R Ach. The WJ-R Achievement Standard Battery is an individually administered set of tests designed to assess different aspects of achievement (Woodcock & Mather, 1989). It is composed of nine standard battery tests: Letter-Word Identification (L-W), Passage Comprehension (Comp), Calculation (Cal), Applied Problems (App), Dictation (Dic), Writing Samples (Wri), Science (Sci), Social Studies (SS), and Humanities (Hum). From these nine individual tests five achievement clusters are formed: Broad Reading (L-W and Comp), Broad Mathematics (Cal and App), Broad Written Language (Dic and Wri), Broad Knowledge (Sci, SS, and Hum), and Skills (W-L, App, and Dic). Three of these achievement clusters will be used in this study: Broad Reading (BR), Broad Mathematics (BM), and Broad Written Language (BWL).

<u>Normative sample.</u> The normative sample was the same group described earlier in Chapter 2 under the WJ-R Cog Manual section. The data for the school aged sample was gather throughout the school years from September 1986 to April 1988. Norm tables provide grade and age equivalents, percentile ranks, and standard scores.

<u>Reliability.</u> Split-half procedures corrected by the Spearman-Brown formula were used for all standard battery tests and cluster scores. Median reliabilities for six of the nine standard battery tests (L-W, Comp, Cal, App, Dic, and Writ) were above .90 with the remaining three tests (Sci, SS, and Hum) showing median reliabilities at or above .866. Cluster score median reliabilities for Broad Reading, Broad Mathematics, and Broad Written Language were above .90. The median SEMs for the standard battery tests ranged from a high of .62 (L-W) to a low of 3.6 (Writ). The broad cluster SEMs were generally in the 3.0 to 4.0 range.

<u>Validity.</u> Test and cluster intercorrelations are higher with in same content areas than between differing contents. The results of a study by Evans and Carlsen (University of Dayton; McGrew, Werder, & Woodcock, 1991) indicated that performance of age matched gifted, learning disabled, and mentally retarded subjected at two age levels

(9-year-olds and 17-year-olds) are in the expected direction.

The WJ-R Ach was correlated with the Kaufmann Assessment Battery for Children, Kaufman Test of Educational Achievement, Bracken Basic Concept Scale, Boehmn Test of Basic Concepts-Revised, Peabody Individual Achievement Test, and the Wide Range Achievement Test-Revised at three age levels: 3, 9, and 17 (Harrington, University of Kansas; McCullough & Wiebe, Texas Woman's University; Woodcock & Mather, 1989). Again, correlations were the highest in like content areas for both test and cluster scores.

Summary. According to Costenbader and Perry (1990, p. 183), "...the WJ-R Ach seems very appropriate for determining levels of achievement relative to peers and thus for making decisions about special educational placement." In this study the WJ-R Ach will be used as yardstick to gauge the concurrent validity of the cognitive batteries and as a predictor variable in relation to special education classification.

Procedure

The data for this study were gathered from referral student assessment records. Special education secretaries from the two special education regional centers of a large metropolitan school district were informed by their supervisor to pull and copy files for this study at their convenience. When they were not responsible for other work, these secretaries randomly pulled files. Files were gathered throughout the school year. The following file information was recorded: age, grade, and ethnicity of the student; placement decision; WISC-III subtest scores of Picture Completion, Information, Coding, Similarities, Picture Arrangement, Arithmetic, Block Design, Vocabulary, Object Assembly, Comprehension, Digit Span, and Symbol Search; WISC-III scale scores of VS, PS, FS (index scores were later computed from the subtest scores); WJ-R Cog subtest scores of Memory for Names, Memory for Sentences, Visual Matching, Incomplete Words, Visual Closure, Picture Vocabulary, Analysis-Synthesis, Visual-Auditory Learning, Memory for Words, Cross Out, Sound Blending, Picture Recognition, Oral Vocabulary, and Concept Formation; WJ-R BCA; WJ-R Ach clusters of Broad Reading, Broad Mathematics, and Broad Written Language (these broad clusters are derived from two to four content specific subtests).

Placement decisions were reached by a multidisciplinary team at each of the individual schools. The composition of these teams were usually: the student's parent/parents or guardian, a special education teacher, a school psychologist, and a school social worker. Sometimes a regular classroom teacher, the student, a physician, a physical therapist, a building administrator, or a special education administrator were in attendance. Criteria for placement followed the federal guidelines as outlined in P.L. 94-142.

Certified School Psychologists administered the WISC-III, and the school diagnostician administered the WJ-R Cog and the WJ-R Ach. School diagnosticians are usually special education teachers who have received special assessment training (either college courses or district-wide workshops) on the WJ-R. The sequence or order of test administration was not controlled. It was a matter of the diagnostician's and psychologist's caseload; therefore, it is not known whether the WISC-III was given before or after the WJ-R. All placement decisions were not reached until all of the tests were administered.

Statistical Analysis

Consistent with the nature of the research questions, the following statistical analyses are proposed for each question:

1. Does the WISC-III subtest performance of a referral sample reflect the same underlying construct pattern as that reported for the WISC-III normative sample? In other words, does the WISC-III four factor model (VC composed of Information, Similarities, Vocabulary, and Comprehension subtests; PO composed of Picture Completion, Picture Arrangement, Block Design, and Object Assembly subtests; FD

composed of Arithmetic and Digit Span subtests; and PrS composed of Coding and Symbol Search subtests), as stated in the WISC-III manual, fit the referral sample's performance?

This analysis is based on the confirmatory factor analysis described in the WISC-III Manual (Wechsler, 1991). Confirmatory factor analysis is a statistical technique which allows us to compare the common underlying behaviors assessed in the referral sample to those assessed in the standardization sample. Factor models of one, two, three, four, and five factors are contrasted using the maximum likelihood fitting function in the Linear Structural Relations (LISREL-VI) program (Joreskog & Sorbom, 1989).

Factor model 1 consisted of "g" or global intellectual entity. Factor model 2 was composed of the Verbal/Performance dichotomy, with six verbal subtests (Information, Similarities, Arithmetic, Vocabulary, Comprehension, and Digit Span) and six performance subtests (Picture Completion, Coding, Picture Arrangement, Block Design, Object Assembly, and Symbol Search). Factor model 3 consisted of VC, PO, and PrS and is based on Sattler's (1992) conclusions that these three factors are strong and consistent; whereas, the FD is not. Factor model 4, four factors of VC, PO, PrS, and FD, was based on the research presented in the Wechsler test manual. Factor model 5a, five factors of VC, PO, PrS, FD, and Picture Arrangement was based on the results presented in the WISC-III manual. Finally, factor model 5b with five factors of VC, PO, PrS, FD and Comprehension was based on the results of an exploratory factor analysis presented by Roid et al. (1993).

The correlation matrix obtained from all 12 subtests for the total sample was used in this analysis, as recommended by Joreskog and Sorbom (1987). The six models were tested using a variety of goodness-of-fit statistics as recommended by Marsh et al. (1988). The Goodness-of-Fit Index (GFI) and the Adjusted Goodness-of-Fit Index (AGFI), goodness-of-fit index adjusted for degrees of freedom, from Joreskog and Sorborn (1987) were used. Also, a measure of the degree of reproduction of the correlation matrix from the model estimates or root mean square residual (RMSR) was used.According to Mulaik et al. (1985), these three indices are the most frequently cited in factor analytical reports.

The Tucker-Lewis Index (TLI; 1973) was used to compare the fit of the proposed models with the fit of the null or one factor "g" model. Bentler and Bonett (1980) suggest using a nested models approach to assess goodness of model fit to the data. This procedure compares several factor models to a null model (more restrictive model or baseline model) to assess the goodness of fit. The TLI is an incremental fit index that has been used in the past with the WISC-III (Wechsler, 1991). The TLI assumes that the chi square/degrees of freedom ratio of the null model is larger than the chi square degrees of freedom ratio of the target model. This index is resistent to sample size and degrees of freedom differences (Marsh et al., 1988) and values of less than .9 indicate that the model could possibly be improved for better fit (Loehlin, 1992).

2. Does the WJ-R Cog subtest performance of a referral sample reflect the same underlying construct pattern as that reported for the WJ-R Cog normative sample? In other words, do the seven WJ-R Cog Standard battery subtests and the seven WJ-R Cog Supplementary battery subtests measure seven underlying factors (two measures for each factor) in a referral sample?

This analysis was based on the procedures used by Woodcock in his 1990 study and described in the WJ-R technical manual (McGrew et al.,1991). These procedures are similar to the ones stated above. A confirmatory factor model with the maximum likelihood fitting function in the LISREL program (Joreskog & Sorborn, 1989) and three goodness-of-fit indices (GFI, AGI, and RMRs) was used to estimate model fit. Three models were investigated. Model 1 represented general intelligence or "g" with all subtests loading on one factor. Model 2 was based on past research with the old WJ Cog in which four factors were consistently identified (McGrew et al., 1991) when factor analytic research was synthesized. The four factors are: Gc represented by Incomplete Words, Picture Vocabulary, Sound Blending, and Oral Vocabulary; Gf represented by Analysis-Synthesis and Concept Formation; Gs represented by Visual Matching, Visual Closure, Cross Out, and Picture Recognition; and Gsm represented by Memory for Names, Memory for Sentences, Visual Auditory Learning, and Memory for Words. Model 3 was based on the seven factors of Glr, Gsm, Gs, Ga, Gv, Gc, and Gf as expressed in the WJ-R Cog Examiner's Manual (1989). This model suggests that the 14 cognitive subtests (seven standard battery subtests and seven supplementary battery subtests) measure seven unique constructs.

3. Do the seven WJ-R Cog standard subtests and the 12 WISC-III subtests measure the same or different cognitive abilities? Do the four WISC-III factor indices and seven WJ-R Cog factor indices measure the same or different cognitive abilities? Do the three WISC-III composite scale scores measure the same or different cognitive abilities than the WJ-R Cog BCA? In other words, do the two cognitive tests contribute unique or redundant information?

All subtest scores (7 WJ-R Cog, 12 WISC-III) and composite scores (WJ-R Cog BCA, WISC-III VS, PS, and FS) were converted into standard scores (X=100, SD=15) and Pearson product-moment correlations were obtained.

Correlations of .80 or above were considered highly related (concurrent validity); .60 and above, moderately; and .40 and below, indicative of low relationship (discriminant validity).

An exploratory factor analysis was used to summarize the relationship among the WISC-III and WJ-R Cog index variables (Gorsuch, 1983). Because of the large number of subtests (14 WJ-R Cog and 12 WISC-III) and the limited number of subjects (N=183), the index scores should give more functional results as compared to the individual subtests. A principal components analysis extraction method with varimax rotation and a maximum likelihood extraction method (ML) with no rotation was used to analyze the relationship of the two cognitive tests. The Kaiser-Guttman rule of eigenvalues greater than 1.0 (Loehlin, 1992), the scree test (Cattel, 1966), and the chi square test (ML; West, 1991) were used to determine the number of factors.

4. Using the factor index scores, which cognitive tests, WISC-III or WJ-R Cog, better relates to current academic achievement?

Pearson product moment correlations between the four WISC-III index scores, the seven WJ-R Cog index scores, and the three WJ-R Ach scores were computed and examined for relationships. A stepwise multiple regression technique was used to determine which variable or set of variables (WJ-R

Cog BCA and seven factor indices; WISC-III VS, PS, FS, and four indices of VC, PO, FD, and PrS) explained the largest amount of academic score variability (Broad Reading, Broad Mathematics, and Broad Written Language). Variables were added one at a time, until increments in explained variance were no longer significant at the .05 level. The academic scores of BR, BM, and BWL were the dependent variable and the scores and indices were the independent variable.

5. Are there unique performance patterns on these three tests that are representative of the various special education categories?

Means and standard deviations for the sample on WISC-III scale and index scores and WJ-R Cog BCA and index scores were examined and contrasted across groups. Particular attention was given to the WISC-III ACID or ACIDS profile (depressed Arithmetic, Coding, Information, Digit Span, and Symbol Search subtest scores), as suggested by Roid et al., (1993) for LD, ED, and NP classified children. The WISC-III subtest mean was calculated for each subject in the ED, LD, and NP group. The individual mean was than compared to his/her scores on ACID subtests. A second comparison was made that included these four subtests plus Symbol Search. An individual was attributed to have the ACID or ACIDS profile if the scores in the above subtests were below the total subtest mean. Group percentages were then calculated.

A multiple discriminant function analysis technique was used to analyze the differences between students classified into different special education categories and to examine the utility of using the two cognitive batteries for differentiating students. The correlation matrix for the four WISC-III indices scores, seven WJ-R Cog indices scores, and three broad cluster achievement scores was used. One set of analyses examined discriminant functions for the total six classification group, while another set examined the discriminant functions for the LD and NP (LD/NP) group. Three analyses were carried out for both sets: 1) the four WISC-III index scores and the three WJ-R Ach broad cluster scores of BR, BM, and BWL; 2) the seven WJ-R Cog index scores and the three WJ-R Ach; and 3) the seven WJ-R Cog and the four WISC-III index scores.

CHAPTER 4

RESULTS

The purpose of this study was to accumulate information about two recently revised cognitive ability tests that are frequently used with a referral sample to determine special education eligibility. The literature review revealed that most research with these cognitive tests has used a sample taken from the general population. Yet, a referral sample is where these tests are often used, and referral decisions are based upon the generalization of assumptions gained from the general population sample. Also, the literature review indicated that no previous research had used the two complete revised cognitive batteries on the same sample; although, information about how the tests relate to each other would be helpful in the referral/ decision/ placement process.

Data on the performance of the referral sample on the WISC-III, WJ-R Cog, and WJ-R Ach were collected and analyzed using various statistical procedures: multiple regression correlation, confirmatory factor analysis, exploratory factor analysis, and discriminant analysis. Results from these procedures were used to determine answers to the research questions posed in the previous chapter. This chapter presents the results from the data analysis.

Presentation of Findings

Research Question #1

Does the WISC-III subtest performance of a referral sample reflect the same underlying construct pattern as that reported for the WISC-III normative sample?

Confirmatory factor analysis was used to support or disconfirm conclusions about the hypothesized factor structure measured by the WISC-III. A series of analyses using the LISREL-VII and LISREL-VIII programs (Joreskog & Sorbom, 1990, 1993) was conducted on the total sample deleting those subjects who had missing scores (N=180) on the WISC-III 12 subtests.

The correlation matrix and the covariance matrix for the 12 subtests was computed using the PRELIS program (Joreskog & Sorbom, 1989). A series of analyses using the six models specified in the previous chapter was conducted. Several model to data fit indices were utilized as suggested by Marsh et al. (1988). The Chi Square, Chi Square divided by degrees of freedom, Goodness of Fit Index (GFI), the Adjusted Goodness of Fit Index (AGFI), and the Root Mean Square Residual (RMSR) as produced by the LISREL program were used to analyze model fit. The index fit results from the correlation matrix and the covariance matrix were identical due to the standardization of the WISC-III scores. Table 2 presents the results of the Chi square, Chi square divided by degree of freedom, and the goodness of fit indices.

Bentler and Bonett (1980) have warned that the Chi square statistic, which ranges from zero to infinity with zero indicating a perfect fit, is sensitive to sample size and departures from multivariate normality of the variables; therefore, they suggest that the Chi square statistic not function as a inferential statistical test but as a comparative measure. When the Chi squares of the six factor models are compared it appears that Model 5b has the smallest number (64.54), as well as, the smallest Chi square to degrees of freedom ratio (1.43). Models 4 and 5a, also, appear smaller than the remaining three models with Chi square to degrees of freedom less than 2.00.

The GFI and the AGFI generally range from zero to one although negative indices may occur (Marsh et al, 1988). Joreskog and Sorbom (1989) describe the GFI as a ratio of squared discrepancies to observed variances. In other words, according to Joreskog and Sorbom (1986, p. I.41), "GFI is a measure of the relative amount of variances and covariances jointly accounted for by the model." The AGFI is similar to the GFI, but it uses mean squares instead of total sums of squares in the numerator and one minus the GFI in the denominator (Joreskog & Sorbom, 1989). The AGFI enacts a cost for non-parsimonious models. Cole

Table 2

WISC-III Six Factor Models (N=183)

Model	Chi Sq	df	Chi Sq/df	GFI	AGFI	RMSR
1	921.10	65	14.17	.40	.28	.876
2	171.62	53	3.426	.84	.77	.672
3	118.66	51	2.327	.90	.84	.073
4	87.34	48	1.81	.93	.88	.062
5a	71.72	44	1.63	.94	.89	.048
5b	64.54	45	1.434	.95	.90	.045

Table 3

WISC-III Model Improvement (N=183)

Model Improvement	Chi square	df	TLI	
1 to 2	739.48*	12	(1 to 2)	.82
2 to 3	60.42*	2	(1 to 3)	.90
3 to 4	30.04*	3	(1 to 4)	.94
4 to 5a	15.62	4	(1 to 5a)	.95
5a to 5b	7.18	1	(1 to 5b)	.97
4 to 5b	22.8 *	3		

Note: Model 1= global "g"; Model 2= Verbal/Nonverbal Dichotomy; Model 3= VC, PO, and PrS; Model 4= VC, PO, PrS, and FD; Model 5a= 4 factors and Picture Arrangement; and Model 5b= 4 factors and Comprehension. Ages ranged from 6 years 3 months to 16 years 0 months. * p = less than .001.

(1987) recommends that GFI values equal to or above .90, AGFI values equal to or above .80, and RMSR values below .10 be considered indicative of good model fit.

Factor Model 1, one general intelligence factor, and Factor Model 2, the Verbal/Performance dichotomy, do not appear to fit the data adequately. Low GFIs and AGFIs, and high RMSRs are present. Factor model 3, which consists of VC, PO, and PrS indicates a somewhat better fit, but Factor model 4 with the four factors of VC, PO, PrS, and FD appears to fit the data better than the earlier models. Factor model 5a with the five factors of VC, PO, PrS, FD, and Picture Arrangement offers some improvement, but factor model 5b (VC, PO, PrS, FD, and Comprehension) appears to be the best fit with the largest GFI and AGFI and the smallest RMSR.

To summarize, factor models 3, 4, 5a, and 5b all appear to satisfy Cole's (1987) criteria of good model fit. Confirmatory factor analysis can not confirm a specific model, but does provide information that can disconfirm models and it appears that factor models 1 and 2 do not fit the data for this sample as well as the other models.

Table 3 presents the results of the comparisons of the six models. The largest improvement in Chi squares and degrees of freedom can be seen between factor model 1 and factor model 2 with the second largest step between model 2 and model 3. Improvement between models, after the first two steps, can be seen between factor models 3 to 4 and factor models 4 to 5b.

The Tucker-Lewis Index (TLI) compares the data fit of successive models. The models are nested, or hierarchical related, with each successive model less restrictive than the one before it. Each model is compared to the null model which is the most restrictive. In this study, model 1, the one factor model, is the null model. Unlike the Chi square statistics and the goodness of fit indices, the TLI has been shown to be independent of sample size (Marsh et al., 1988). TLI values equal to or above .90 are considered indicative of good fit, although caution should be used in interpreting absolute values (Marsh et al, 1988). Models 3, 4, 5a, and 5b all have TLI values which fit this criteria, but model 5b (VC, PO, PrS, FD, Comprehension) has the highest TLI.

To summarize, the four factor model, as proposed by the WISC-III manual, does fit the data, but two other models composed of five factors fit the data better.

Research Question #2

Does the WJ-R Cog subtest performance of a referral sample reflect the same underlying construct pattern as that reported for the WJ-R Cog normative sample?

Confirmatory factor analysis was used to investigate this question. The maximum likelihood fitting function in the LISREL VII and VIII (Joreskog & Sorbom, 1992,1993) programs was used on the data from the total sample (N=183). All subjects had recorded scores for all of the 14 subtests; therefore, deletion was not necessary.

The correlation matrix and the covariance matrix for the 14 subtests was computed using the PRELIS program (Joreskog & Sorbom, 1989). A series of analysis using the three models specified in the previous chapter were conducted. The Chi Square, Chi Square divided by degrees of freedom, Goodness of Fit Index (GFI), the Adjusted Goodness of Fit Index (AGFI), and the RMSR as produced by the LISREL program were used to analyze model fit. The fit indices from the correlation matrix and the covariance matrix were identical due to the standardization of the WJ-R Cog scores. Table 4 presents the results of the chi square, chi square divided by degree of freedom, and the goodness of fit indices.

Model 1, which consists of one general intelligence or "g", does not appear to fit the data as evidenced by all indicators. Model 2, which is based on findings of the old WJ Cog of four factors, does not fit the data except for the RMSR which is below .10. All indications appear to favor factor model 3, which is based on the seven factor model proposed by the authors. The Chi Square to degrees of freedom ratio for model 3 (1.64) is within the ideal range of less than 2.00 (Roid et al, 1993). The goodness of fit indices for model 3 are above .90 for the GFI (.94), above .80 for the AGFI (.88), and the RMSR (.04) is below .10. These evaluations are based on Cole's (1987) criteria for evaluation of goodness of fit.

Table 5 presents information based on the comparison of the models. The comparison between model 1 to model 2 exhibits the larger difference between chi square and degrees of freedom, but improvement is also shown between models 2 and 3. The TLI favors model 3 (.94) with an index above .90 as suggested by Marsh et al. (1988).

To summarize, model 3, the seven factor model proposed by Woodcock (1990), fits the data from this referral sample best when compared to the other two models investigated.

Research Question #3

Do the seven WJ-R Cog standard subtests and the 12 WISC-III subtests measure the same or different cognitive abilities? Do the four WISC-III factor indices and seven WJ-R Cog factor indices measure the same or different cognitive abilities? Do the three WISC-III composite scale scores measure the same or different cognitive abilities than the WJ-R Cog BCA? In other words, do the two cognitive tests contribute unique or redundant information?

Table 4

WJ-R Cog Three Factor Models (N=180)

Model	Chi Sq	df	Chi Sq/df	GFI	AGI	RMSR
1	1039.55	90	11.55	.38	.28	.33
2	213.83	71	3.01	.86	.79	.09
3	91.84	56	1.64	.94	.88	.04

Table 5

WJ-R Cog Model Improvement (N=180)

Model Improvement	Chi Sq	df	TLI	
1 to 2	825.72*	19	(1 to 2)	.81
2 to 3	121.99*	15	(1 to 3)	.94

Note: Model 1= global "g"; Model 2= old WJ 4 factor model; and Model 3= WJ-R 7 factor model. Ages ranged for 6 years 3 months to 16 years 0 months of age. * p =less than .001

Correlations. All test sub-scores, WISC-III, obtained from the referral sample with three subjects deleted due to missing scores (N=180) were converted into Pearson product-moment correlations using the PRELIS (Jorkeskog & Sorbom, 1989) program. As stated previously in Chapter 3, correlations of .80 or above were considered highly related (concurrent validity); .60 and above, moderately; and .40 and below, were considered indicative of low relationship (discriminant validity). Table 6 presents the correlation matrix obtained for the subtests scores.

As can be seen no correlations were above .60. The highest correlation obtained was between WJ-R Picture Vocabulary (Gc) and WISC-III Vocabulary subtest (.591; squared correlation coefficient of .348). One other correlation, Picture Vocabulary and WISC-III Information (.509) was above .50. Both of these WISC-III subtests are components of the VC factor index.

The WJ-R Memory for Sentences (Gsm) and Analysis-WISC-III subtests. Three of these four WISC-III subtests (Information, Similarities, and Vocabulary) comprise the VC factor index.

The relationship between the two cognitive batteries is not high or moderate as defined by the stated criteria and the majority of correlations were considered low or indicative of discriminant validity. It appears that there

Table 6

WISC-III and WJ-R Cog Subtest Correlations (N=180)

WISC-III	WJ-R Cog Subtests					Subtests
	MN	MS	VM	IW	VC	PV	AS
I	.42	.46	.22	.31	.25	.51	.48
S	.38	.49	.16	.45	.26	.49	.40
A	.35	.37	.35	.28	.16	.39	.42
V	.44	.49	.24	.41	.24	.59	.44
C	.33	.44	.27	.30	.16	.43	.35
PC	.18	.28	.16	.28	.21	.35	.26
PA	.27	.25	.34	.20	.22	.29	.32
BD	.25	.30	.26	.12	.28	.29	.39
OA	.22	.26	.16	.21	.32	.23	.29
CO	.12	.04	.43	.08	.06	.17	.20
SS	.20	.16	.44	.10	.12	.27	.27
DS	.33	.38	.26	.35	.18	.36	.39

Note: (WJ-R Cog) MN=Memory for Names;MS=Memory for Sentences; VM=Visual Matching; IW=Incomplete Words; VC=Visual Closure; PV=Picture Vocabulary; AS=Analysis-Synthesis; (WISC-III) I=Information; S=Similarities; A=Arithmetic; V=Vocabulary; C=Comprehension; PC=Picture Completion; PA=Picture Arrangement; BD= Block Design; OA=Object Assembly; CO=Coding; SS=Symbol Search; and DS=Digit Span. Ages ranged from 6 years 3 months to 16 years 0 months.

Synthesis (Gf) subtests correlated above .4 with four of the is a modest shared verbal component, but the majority of measured abilities are not related. This suggests that the subtests are not redundant and that each measures unique information.

Table 7 presents the factor index correlations. The WISC-III VC index (composed of Information, Similarities, Vocabulary, and Comprehension) correlated moderately (above .60) with the WJ-R Cog indices of Gc (Comprehension-Knowledge; squared correlation coefficient of .44) and Gf (Fluid Reasoning; squared correlation coefficient of .40). VC correlated above .5 with WJ-R Cog Glr (Long Term Retrieval; squared correlation coefficient of .27), Gsm (Short Term Memory; .28 squared correlation coefficient), and Ga (Auditory Processing; .26 squared correlation coefficient). It appears that VC measures a combination of cognitive abilities or that the WJ-R measures overlap in their measurement of verbal comprehension.

The WISC-III PO index (composed of Picture Completion, Picture Arrangement, Block Design, and Object Assembly) does not correlate above .5 with any of the WJ-R Cog indices; although, the highest correlation is with Gf (.447; .20 squared correlation coefficient) which may indicate that PO measures unique abilities.

Table 7

WISC-III and WJ-R Cog Index Correlations (N=180)

WISC-III Indices	WJ-R Cog Factor Indices						
	GLR	GSM	GS	GA	GV	GC	GF
VC	.52	.53	.33	.51	.29	.67	.64
PO	.40	.30	.39	.32	.36	.43	.45
FD	.43	.49	.36	.38	.30	.45	.49
PRS	.14	.12	.45	.09	.12	.18	.13

Table 8

WISC-III and WJ-R Cog Composite Score Correlations (N=180)

WJ-R	WISC-III Scale Scores		
	VS	PS	FS
BCA	.72	.55	.74

Note: (WJ-R Cog) GLR=Long-Term Retrieval;GSM=Short-Term Memory; GS=Processing Speed; GA=Auditory Processing; GV=Visual Processing; Gc=Comprehension-Knowledge; GF=Fluid Reasoning; (WISC-III) VC=Verbal Comprehension; PO=Perceptual Organization; FD=Freedom from Distractibility; and PRS=Processing Speed. Ages ranged from 6 years 3 months to 16 years 0 months.

The WISC-III FD does not correlate above .5 with any WJ-R index, but does exhibit correlations around .4 with four of the WJ-R indices of Glr, Gsm, Gc, and Gf (respective squared correlation coefficients of .19, .24, .20, and .24). PrS, the WISC-III fourth factor index, did not correlate above .18 with any of the WJ-R Cog indices except for Gs (Processing Speed) with a correlation of .452 (.20 squared correlation coefficient). It would be expected from the factor labels that these two indices of processing speed would have high correlations, but .452 is not indicative of concurrent validity.

The WJ-R Cog index of Gv, Visual Processing, did not correlate above .359 (WISC-III PO index) with any of the WISC-III indices. This suggests that Gv is a unique measurement.

Table 8 provides information about the relationship between the WISC-III scaled scores and the WJ-R Cog BCA. It appears that the BCA is moderately related to the WISC-III VS (.52 squared correlation coefficient) and FS (.54 squared correlation coefficient), but somewhat less related to PS (.30 squared correlation).

To summarize, the correlations between subtests are lowest as compared to the correlations between indices and full scale scores. The highest correlations were between the WISC-III verbal and composite measures and the WJ-R Cog measures of general comprehension, fluid reasoning, and composite measures.

Exploratory factor analysis. An exploratory factor analysis of the four WISC-III index scores and the seven WJ-R Cog index scores was carried out to examine the relationship of the two cognitive batteries to each other. The index correlation matrix was used for both analysis.

First, a principal components extraction with varimax rotation of the 11 index scores using the SPSS/PC statistical package was completed. In accordance with the Kaiser-Guttmann criterion, two factors had eigenvalues greater than 1.0. The large first factor had an eigenvalue of 4.829 and accounted for 43.9 percent of the variance. This suggests the existence of a large global or "g" factor present in both tests. The second smaller factor had an eigenvalue of 1.366 and accounted for 12.4 percent of the variances. Together these two factors accounted for 56.3 percent of the variance. This means that a substantial portion of the variance is not accounted for by these two factors.

Many factor analysis experts (Gorsuch, 1983; Kim & Mueller, 1978; Loehlin, 1992) question the appropriateness of relying solely on the Kaiser-Guttman rule and suggest using additional criteria. The scree test (Cattell, 1966) involves plotting successive eigenvalues of the principal components on a graph. Decisions about meaningful factors are based upon the point of inflection, or the point where decreasing eigenvalues form a flat gradual slope (West, 1991). The scree plot for this analysis indicated one strong factor, one relatively weak second factor, and possibly, a very weak third factor (eigenvalue of .891). The scree plot is presented in the Appendix B (page 162).

The varimax rotation converged in three iterations and the rotated factor matrix is exhibited in Table 9. Indices loading heavily on Factor 1 are VC, Glr, Gsm, Ga, Gc, and Gf. All of these indices involve retrieval of verbal information, either from longer term memory or short term memory. Factor 2 appears to be concerned with perceptual or processing speed with the indices of PrS and Gs loading heavily upon it. The FD and PO indices are factorially complex, because they load moderately on both factors. The Gv index loads more upon Factor 1 than Factor 2, but it is not heavily represented in either factor. When the communalities are considered it appears that all of the variables, but the GV index, can be explained somewhat by some combination of the two factors. The GV index is composed of subtests involving reasoning with visual and spatial patterns.

Next, factors were extracted using the maximum likelihood function of the SPSS/PC statistical program. Two

Table 9

WISC-III Factor Matrix for Principal Components with Varimax Rotation and Final Statistics for Communality (N=180)

Subtests	Factor 1	Factor 2	Communality
VC	.79255	.30310	.72000
PO	.45153	.56200	.51972
FD	.60634	.42297	.54656
PRS	.02712	.85931	.73916
GLR	.69509	.15527	.50726
GSM	.72237	.09816	.53146
GS	.14557	.78401	.63585
GA	.72717	-.01352	.52896
GV	.42682	.21846	.22990
GC	.78891	.12442	.63786
GF	.74601	.20378	.59806

Note: Factor 1 accounts for 43.9% of the variance; Factor 2 accounts for 12.4%; and the total variance accounted for is 56.3%. WJ-R Cog: GLR=Long-Term Retrieval;GSM=Short-Term Memory; GS=Processing Speed; GA=Auditory Processing; GV=Visual Processing; Gc=Comprehension-Knowledge; and GF=Fluid Reasoning; WISC-III: VC=Verbal Comprehension; PO=Perceptual Organization; FD=Freedom from Distractibility; and PRS=Processing Speed. Ages ranged from 6 years 3 months to 16 years 0 months.

factors were extracted within seven interactions. The results of this analysis were similar to the principal components; although, the values were somewhat lower. These discrepancies could be explained by the methods employed. The maximum likelihood method investigates relationships between the variables; whereas, the principal components analysis seeks to account for as much of the original variable variance as possible (West, 1991). The first factor had an eigenvalue of 4.32718 and accounted for 39.3 percent of the variance and the smaller second factor had an eigenvalue of .90466 and accounted for 8.2 percent of the variance. The total variance accounted for by the two factors was 47.6 percent.

The maximum likelihood factor matrix is presented in Table 10. The VC, Glr, Gsm, Ga, Gc, and Gf load on Factor 1 and this is similar to the principal components findings. Several discrepancies are apparent between the two analysis. First, PO, FD, and Gv appeared to be factorially complex in the principal components analysis, but in the maximum likelihood analysis they load more heavily on Factor 1. Another difference between the two analysis is that in principal components Factor 2 was composed of PrS and Gs, and in maximum likelihood Gs is factorially complex with moderate loadings on both factors. Thus, this leaves factor 2 mainly composed of PrS.

In both analysis, as evidenced by the communalities, the Gv index is not explained by the two factors, but the other indices seem to be explained fairly well. This is supported by the Chi square statistic of 40.96 with 34 degrees of freedom for a statistical significance of .1917.

Summary. To summarize, the correlations indicate that the WISC-III subtests and the WJ-R Cog subtests do not correlate highly with each other, but the verbal index scores in both batteries indicate moderate correlation (VC and Gc .665). This is supported in the findings of the composite scores (BCA and VS .72; BCA and FS .74). In both the principal components analysis and the maximum likelihood method, it appears that two factors explain between 47 to 56 percent of the variance in the four WISC-III indices of VC, PO, PrS, and FD and the six WJ-R Cog indices of Glr, Gsm, Gs, Ga, Gc, and Gf; although the Gv index is not well represented in either factor. In both analysis, a large Factor 1 appears to represent a global "g" concerned with verbal comprehension, reasoning, and memory, and a smaller Factor 2 appears to represent perceptual or processing speed. Approximately one half of the variance is not accounted for by these two factors, which indicates that while the two cognitive batteries share some areas of measurement a large portion of these batteries measure different cognitive aspects.

Table 10

WISC-III Factor Matrix for Maximum Likelihood Extraction and Final Statistics for Communality (N=180)

Index	Factor 1	Factor 2	Communality
VC	.84465	-.09273	.72202
PO	.60737	.20145	.40948
FD	.69088	.08122	.48391
PRS	.41254	.69713	.65619
GLR	.63332	-.13075	.41819
GSM	.63773	-.17596	.43766
GS	.44719	.43654	.39054
GA	.59310	-.25248	.41551
GV	.40515	-.00325	.16415
GC	.74296	-.18890	.58767
GF	.72228	-.15754	.54651

Note: Factor 1 accounts for 39.3% of the variance; Factor 2 accounts for 8.2% of the variance; and the total variance accounted for is 47.6%. WJ-R Cog: GLR=Long-Term Retrieval;GSM=Short-Term Memory; GS=Processing Speed; GA=Auditory Processing; GV=Visual Processing; Gc=Comprehension-Knowledge; and GF=Fluid Reasoning; WISC-III: VC=Verbal Comprehension; PO=Perceptual Organization; FD=Freedom from Distractibility; and PRS=Processing Speed. Ages ranged from 6 years 3 months to 16 years 0 months.

Research Question #4

Using the factor index scores, which cognitive tests, WISC-III or WJ-R Cog, best predicts to current academic achievement?

WISC-III indices and WJ-R Ach. The linear relationships between the WISC-III index scores and the WJ-R Ach BR, BM, and BWL were examined using the SPSS/PC statistical package. The predictor variables, index scores, were entered one at time into the regression equation using a forward stepwise procedure. The .05 level of significance was used as the criterion for entering variables into the equation. The minimum tolerance level by default for the SPSS/PC system is .001. The tolerance level "...is a measure of the degree of linear association between the independent variables." (SPSS/PC + Advanced Statistical Guide, 1988, p. B-18). The minimum tolerance level was adjusted to .01. for this study. Variables were not entered into the regression equation if the tolerance level was below .01. Table 11 provides information about the WISC-III indices and WJ-R Ach scores.

The VC index appears to be the best predictor variable in relation to all three areas of achievement. VC accounts for 27 percent of the proportion of variance in BR, 42 percent for BM, and 22 percent for BWL when it is entered alone into the regression equation on the first step. The three indices of VC, FD, and PRS accounted for a substantial proportion of BM variance, 52 percent. From the overall picture, it appears that the WISC-III indices predict math achievement better than reading or written language.

WJ-R Cog indices and WJ-R Ach. The linear relationship of the seven WJ-R Cog indices and the three WJ-R Ach clusters was examined using the same procedures as outlined above. Table 12 presents information on these findings.

It appears that the Gc (Comprehension-Knowledge) and Gs (Processing Speed) indices contribute to the prediction of all three academic measures; although, their contribution is most heavily weighted in BM. The GA index (Auditory Processing), the first variable entered, contributes to the prediction of both BR (adjusted square R = .24) and BWL (adjusted square R = .16). The Gsm index (Short Term Memory) does not appear significant for any achievement cluster other than BR. Also, Gf (fluid reasoning) and Glr (Long Term Retrieval) are exclusive to BM.

The WJ-R Cog indices (total adjusted R squared = .48), as do the WISC-III indices (total adjusted R square = .52), predict the BM cluster better than the other two academic clusters, but the WJ-R Cog indices have a slightly higher adjusted R square for BR and BM than do the WISC-III indices. The WJ-R Cog higher adjusted R squares would be expected due to the same method variance (both Cog and Ach

Table 11

Stepwise Multiple Regression for WISC-III Indices (N=183)

Variable	B	Beta T	Sig	T
WJ-R Ach BR				
VC	.4962	.5198	8.05	.0000
Total Adjusted R Square is .27				
WJ-R Ach BM				
VC	.4748	.4313	6.63	.0000
FD	.3541	.2987	4.50	.0000
PRS	.1976	.1651	2.95	.0036
Total Adjusted R Square is .52				
WJ-R Ach BWL				
VC	.4263	.4307	6.31	.0000
PRS	.1811	.1680	2.46	.0148
Total Adjusted R Square is .25				

Note: WISC-III Indices of VC=Verbal Comprehension; PO=Perceptual Organization; FD=Freedom from Distractibility; and PRS=Processing Speed. Ages ranged from 6 years 3 months to 16 years 0 months.

Table 12

Stepwise Multiple Regression for WJ-R Cog Indices (N=180)

Variable	B	Beta	T	Sig T
WJ-R Ach BR				
GA	.3073	.2804	3.76	.0002
GSM	.1745	.1747	2.31	.0219
GC	.1834	.1855	2.41	.0172
GS	.1280	.1287	1.99	.0478
Total Adjusted R Square = .32				
WJ-R Ach BM				
GF	.3176	.2695	4.00	.0001
GS	.3214	.2778	4.86	.0000
GC	.2585	.2248	3.23	.0015
GLR	.2259	.1822	2.80	.0058
Total Adjusted R Square = .48				
WJ-R Ach BWL				
GA	.3131	.2726	3.53	.0005
GS	.2186	.2096	3.08	.0024
GC	.1704	.1644	2.14	.0340
Total Adjusted R Square = .22				

Note: WJ-R Cog Indices of GLR=Long-Term Retrieval; GSM=Short-Term Memory; GS=Processing Speed; GA=Auditory Processing; GV=Visual Processing; Gc=Comprehension-Knowledge; and GF=Fluid Reasoning. Ages ranged from 6 years 3 months to 16 years 0 months.

portions are part of the complete WJ-R battery); although, the proportion of variance in BWL that is accounted for by indices from either cognitive battery is about equal (WJ-R .22 and WISC-III .25).

Composite cognitive scores and achievement. The composite scores from the WISC-III (Verbal Scale, Performance Scale, and Full Scale) and the WJ-R Cog BCA (Broad Cognitive Ability) were used as predictor variables with the WJ-R Ach cluster as dependent variables, and the stepwise procedure as described above were repeated. The results are presented in Table 13.

The BCA is clearly an explanatory power in all three achievement clusters, but it contributes the most to BR. The proportion of variance accounted for by BCA alone in BR is considerable (squared R = .30). When VS is entered into the regression equation on the second step, adjusted multiple squared R, or the proportion of variance in BR which is accounted for by the combined variables of BCA and VS, increases to .32.

The FS accounts for a large portion of variance in both BM and BWL. FS was entered on the first step for both variables and the adjusted squared R's for FS alone accounted for a considerable amount of the explained variance, .47 and .22 respectively. When BCA was added in the second step of the regression equation for both BM and

Table 13

Stepwise Multiple Regression for Composite Cognitive Scores (N=180)

Variable	B	Beta	T	Sig T
WJ-R Ach BR				
BCA	.4400	.4050	4.57	.0000
VS	.1933	.2037	2.30	.0226
Total Adjusted R Square = .32				
WJ-R Ach BM				
FS	.7666	.6494	4.99	.0000
BCA	.3948	.3128	4.04	.0001
PS	-.2531	-.2193	-2.06	.0413
Total Adjusted R Square = .53				
WJ-R Ach BWL				
FS	.3114	.2926	3.05	.0026
BCA	.2827	.2484	2.59	.0104
Total Adjusted R Square = .25				

Note: BCA=WJ-R Cog Broad Cognitive Ability Score; VS=WISC-III Verbal Scale Score; PS=WISC-III Performance Scale Score; and FS=WISC-III Full Scale Score. Ages ranged from 6 years 3 months to 16 years 0 months.

BWL, the amount of variance accounted for increased to .52 and .25. With the addition of PS into the equation for BM and .25. With the the total adjusted square R increases to .53.

As expected the lowest tolerance levels were indicated during the composite cognitive score analysis, and specifically for the BM variable (VS minimum tolerance level .01); although, at no time during this entire analysis was a variable not entered into a regression equation due to an unacceptable tolerance level.

Summary. When the cognitive batteries are examined for index predictive validity, it appears that both explain about the same proportion of variance from a practical standpoint: 32-27 percent for reading, 48-53 percent for math, and 22-25 percent for written language. The WISC-III VC index emerges as the strongest single academic predictor when compared to the WJ-R Cog Indices and the WISC-III Indices. The WJ-R Cog BCA explained variance in all three academic areas; whereas, the WISC-III FS contributed to the variance on math and written language and the VS contributed to reading.

Research Question #5

Are there unique performance patterns on these three tests that are representative of the various special education categories?

Several difference analyses were performed to investigate performance patterns: descriptive statistics, with specific attention to the ACID or ACIDS profile, and discriminant analysis.

Descriptive statistics. Subjects were grouped according to special education classification, and means and standard deviations were computed using the SPSS/PC statistical package. The descriptive statistics for the six categories are presented in Tables 14 through 19.

The rank order of groups, as arranged from highest to lowest on WISC-III FS and WJ-R Cog BCA group mean scores, were: Non-placed (NP), Emotional Disability (ED), Specific Learning Disability (LD), Speech/Language Impairment (SLI), Other Health Impairment (OHI), and Mild Mental Retardation (MMR). As expected, the NP group mean score was the highest on the WISC-III FS and the WJ-R Cog BCA; although this group was approximately one half of a standard deviation below the mean. The mildly mental retardation group had the lowest FS and BCA with a mean of over two standard deviation below the norm. Two groups, Emotional Disability (ED) and Specific Learning Disability (LD) had WISC-III mean FS scores of 89, but their WJ-R Cog BCA mean score differed slightly (ED, 89; SLD, 84). The difference between the FS mean scores and the BCA mean scores ranged from a -5 (LD) to a +7 (OHI) with one group having the same mean score for both batteries

Table 14

Speech/Language Impaired (N=19)

Variable	Mean	SD	Variable	Mean	SD
Age	102.68	17.13			
WISC-III					
VC	83.63	12.03	VS	82.68	12.59
PO	91.21	11.30	PS	90.95	11.03
PRS	93.95	11.46	FS	85.84	10.91
FD	85.05	13.58			
WJ-R Cog					
GLR	90.53	11.33	BCA	82.11	9.87
GSM	98.32	12.31			
GS	84.84	11.78			
GA	83.58	8.45			
GV	96.26	15.95			
GC	82.00	11.96			
GF	86.00	10.33			
WJ-R ACH					
BR	80.74	7.47			
BM	87.89	12.39			
BWL	84.79	8.76			

Note: WJ-R Cog Indices of GLR=Long-Term Retrieval; GSM=Short-Term Memory; GS=Processing Speed; GA=Auditory Processing; GV=Visual Processing; Gc=Comprehension-Knowledge; GF=Fluid Reasoning; WISC-III Indices of VC=Verbal Comprehension; PO=Perceptual Organization; FD=Freedom from Distractibility; and PRS=Processing Speed; WJ-R Ach Clusters of BR=Broad Reading; BM=Broad Math; and BWL=Broad Written Language.

Table 15

Specific Learning Disabled (N=96)

Variable	Mean	SD	Variable	Mean	SD
Age	124.64	31.59			
WISC-III					
VC	88.40	13.06	VS	87.39	12.95
PO	93.48	13.53	PS	91.88	12.78
PRS	86.38	21.35	FS	88.65	12.01
FD	85.64	16.30			
WJ-R COG					
GLR	91.27	11.34	BCA	84.29	10.91
GSM	88.90	11.54			
GS	85.01	13.74			
GA	85.63	10.46			
GV	97.05	13.98			
GC	85.77	13.05			
GF	88.49	11.53			
WJ-R ACH					
BR	77.72	12.48			
BM	84.70	14.53			
BWL	74.78	13.13			

Note: WJ-R Cog Indices of GLR=Long-Term Retrieval; GSM=Short-Term Memory; GS=Processing Speed; GA=Auditory Processing; GV=Visual Processing; Gc=Comprehension-Knowledge; GF=Fluid Reasoning; WISC-III Indices of VC=Verbal Comprehension; PO=Perceptual Organization; FD=Freedom from Distractibility; and PRS=Processing Speed; WJ-R Ach Clusters of BR=Broad Reading; BM=Broad Math; and BWL=Broad Written Language.

Table 16

Non Placed (N=45)

Variable	Mean	SD	Variable	Mean	SD
Age	122.69	26.76			
WISC-III					
VC	94.76	13.60	VS	94.09	13.39
PO	94.33	14.46	PS	93.16	13.67
PRS	86.38	27.71	FS	92.93	13.25
FD	86.84	23.03			
WJ-R COG					
GLR	99.87	12.17	BCA	93.53	11.62
GSM	96.64	15.24			
GS	91.58	11.24			
GA	92.09	14.43			
GV	98.18	11.30			
GC	93.13	13.00			
GF	95.36	14.66			
WJ-R ACH					
BR	89.02	12.64			
BM	92.56	13.13			
BWL	87.16	11.51			

Note: WJ-R Cog Indices of GLR=Long-Term Retrieval; GSM=Short-Term Memory; GS=Processing Speed; GA=Auditory Processing; GV=Visual Processing; Gc=Comprehension-Knowledge; GF=Fluid Reasoning; WISC-III Indices of VC=Verbal Comprehension; PO=Perceptual Organization; FD=Freedom from Distractibility; and PRS=Processing Speed; WJ-R Ach Clusters of BR=Broad Reading; BM=Broad Math; and BWL=Broad Written Language.

Table 17

Mild Mental Retardation (N=8)

Variable	Mean	SD	Variable	Mean	SD
Age	131.75	29.89			
WISC-III					
VC	69.25	11.38	VS	68.50	10.64
PO	70.13	7.10	PS	69.25	6.44
PRS	74.25	7.24	FS	66.50	6.20
FD	73.38	10.97			
WJ-R COG					
GLR	82.25	8.94	BCA	69.00	10.86
GSM	81.88	14.05			
GS	69.13	10.58			
GA	72.00	10.42			
GV	85.25	18.66			
GC	74.88	8.98			
GF	76.25	11.30			
WJ-R ACH					
BR	70.63	12.28			
BM	61.75	19.42			
BWL	63.63	11.96			

Note: WJ-R Cog Indices of GLR=Long-Term Retrieval; GSM=Short-Term Memory; GS=Processing Speed; GA=Auditory Processing; GV=Visual Processing; Gc=Comprehension-Knowledge; GF=Fluid Reasoning; WISC-III Indices of VC=Verbal Comprehension; PO=Perceptual Organization; FD=Freedom from Distractibility; and PRS=Processing Speed; WJ-R Ach Clusters of BR=Broad Reading; BM=Broad Math; and BWL=Broad Written Language.

Table 18

Emotional Disability (N=12)

Variable	Mean	SD	Variable	Mean	SD
Age	134.75	37.09			
WISC-III					
VC	89.58	11.26	VS	88.75	11.00
PO	95.25	8.63	PS	91.58	8.39
PRS	95.17	12.47	FS	88.92	8.65
FD	87.25	9.24			
WJ-R COG					
GLR	95.25	9.73	BCA	89.83	8.85
GSM	97.00	9.05			
GS	91.42	9.00			
GA	89.33	9.15			
GV	94.67	11.62			
GC	87.25	11.85			
GF	92.33	13.44			
WJ-R ACH					
BR	85.67	15.02			
BM	88.83	12.65			
BWL	82.25	12.87			

Note: WJ-R Cog Indices of GLR=Long-Term Retrieval; GSM=Short-Term Memory; GS=Processing Speed; GA=Auditory Processing; GV=Visual Processing; Gc=Comprehension-Knowledge; GF=Fluid Reasoning; WISC-III Indices of VC=Verbal Comprehension; PO=Perceptual Organization; FD=Freedom from Distractibility; and PRS=Processing Speed; WJ-R Ach Clusters of BR=Broad Reading; BM=Broad Math; and BWL=Broad Written Language.

Table 19

Other Health Impairments (N=3)

Variable	Mean	SD	Variable	Mean	SD
Age	108.33	20.68			
WISC-III					
VC	74.33	10.87	VS	72.33	8.96
PO	82.33	20.95	PS	81.67	16.21
PRS	84.67	4.64	FS	74.33	13.47
FD	77.33	9.43			
WJ-R COG					
GLR	87.33	22.05	BCA	81.33	11.90
GSM	88.00	10.98			
GS	85.00	5.10			
GA	81.33	8.26			
GV	80.33	17.99			
GC	81.67	11.03			
GF	86.67	10.27			
WJ-R ACH					
BR	77.67	11.70			
BM	75.00	10.20			
BWL	69.00	14.45			

Note: WJ-R Cog Indices of GLR=Long-Term Retrieval; GSM=Short-Term Memory; GS=Processing Speed; GA=Auditory Processing; GV=Visual Processing; Gc=Comprehension-Knowledge; GF=Fluid Reasoning; WISC-III Indices of VC=Verbal Comprehension; PO=Perceptual Organization; FD=Freedom from Distractibility; and PRS=Processing Speed; WJ-R Ach Clusters of BR=Broad Reading; BM=Broad Math; and BWL=Broad Written Language.

(ED). These differences were less than one-half a standard deviation in either direction.

All of the groups, except the NP group, had higher PS mean scores than VS or FS mean scores. Two groups, OHI and SLI, had similar WISC-III index mean scores profiles with mean scores increasing from VC (lowest), FD, PO, to PrS (highest). The ED group WISC-III index scores had a similar profile with the PrS and PO (same mean score of 95) as the highest mean scores, but the ED group differed in that the FD was the lowest mean index score. FD, also, was one of the lowest mean index score for the LD and NP groups, but the LD group had PO as the highest mean index score while PrS was next to the lowest mean index score. The NP group had VC and PO as the highest mean index scores (94 for both mean index scores) and FD and PrS as the lowest mean index scores (86 for both mean index scores).

The same three WJ-R Cog mean index scores were the highest for five of the groups: Glr, Gv, and Gsm. The OHI group differed from the other groups in that Glr and Gsm mean scores were the highest, but Gv was the lowest mean index score. The lowest mean index scores for three of the groups (MMR, SLD, and NP) was Gs. Gc was the lowest mean index score for the SLI and ED groups.

Four groups (SLD, NP, SLI, and ED) had BM as the highest mean academic score. Three of these four (ED, SLD, and NP), also, had BR as the second high mean with BWL as the lowest mean academic score. SLI reversed this order with BWL and BR from highest to lowest. MMR and OHI groups exhibited the highest mean academic score in BR, but differed on the lowest mean academic scores with MMR indicating BM and OHI, BWL.

To summarize, it appears that total scores (FS and BCA) for both cognitive batteries are capable of differentiating cognitive abilities in the expected direction in relation to special education classification. Little variability (SD= 6.2) was apparent in the MMR mean score, and this is expected for this category. Both FS and BCA group mean scores were fairly close with the largest discrepancy of a +7 in the OHI group. All of the placed special education categories indicated better perceptual skills than verbal on the WISC-III mean scaled scores.

As expected from previous research (Prifitera & Dersh, 1993; Wechsler, 1991), the FD mean index scores for SLD and ED were the lowest as compared to the other WISC-III mean indices. The PrS mean index score was the lowest WISC-III mean index score for MMR, SLI, and OHI.

The WJ-R Cog mean index scores indicated the Gs, Ga, and Gc were the lowest three index means for all of the groups except the OHI; although these three means scores were low for the OHI group, the Gv mean index was the lowest. The standard deviation for the Gv index indicated a large amount of variability for all of the groups. Gv standard deviations ranged from 18.7 (MMR) to 11.3 (NP).

The BWL mean score was the lowest academic mean score for all of the groups except MMR and SLI. The BR mean scores were the highest academic mean scores for MMR and OHI, while BM mean scores were the highest for all of the other groups.

Acid profiles. The WISC-III individual subtest mean was calculated for each subject in the ED, LD, and NP group. The individual mean was than compared to his/her scores on Arithmetic, Coding, Information, and Digit Span. A second comparison was made that, also, included these four subtests plus Symbol Search. An individual was attributed to have the ACID or ACIDS profile if his/her scores in the above subtests were below the total subtest mean. Group percentages were then calculated.

Only ten subject from the LD group (N=96) or approximately 10% exhibited the complete ACID profile. Also, these same ten subjects had depressed Symbol Search scores for the ACIDS profile. Approximately 11% of the NP group (N=45) displayed the ACID profile and 9% displayed the ACIDS profile. None of the ED group (N=12) exhibited either the complete ACID or ACIDS profile; although 42% of this group had three of the four subtest scores (ACID) depressed and 25% had four of the five subtests scores (ACIDS) depressed. Less than the majority of subjects in both the NP and LD group had three of the four ACID subtests lower (41% and 45% respectively) and an even smaller number in each group had four of the five ACIDS subtests lower than the total subtest means, 27% for NP and 32% for LD. It does not appear that the ACID or ACIDS profile is unique to the LD or ED groups in this study; although, the group subject numbers prevent further generalization.

Discriminant function analysis. The SPSS/PC Statistical Package was used to examine linear combinations of the predictor variables that may serve as the basis for discriminating the subjects into groups. Three separate analyses were carried out for all six classification groups and for two classification groups, LD and NP: 1) the four WISC-III index scores and the three WJ-R Ach were used as predictor variables for the classification groups; 2) the seven WJ-R Cog index scores and the three WJ-R Ach were used as predictor variables; and 3) the four WISC-III and the seven WJ-R Cog indices were used as predictor variables. The variables were entered into the equation one step at a time for all of the analysis. The variable selection rule was based on the smallest Wilks' Lambda and the largest Fs. The Wilks' Lambda is sensitive to differences between groups and the homogeneity within groups. The minimum tolerance level, as set by SPSS/PC default, to exclude variables that are linear combinations of variables already entered was .001. Tables 21 through 24 present information obtained from the six analyses.

The Canonical Discriminant Functions Tables (Tables 20 and 22) give evidence that all of the analyses have at least one significant function that explains a percentage of the variance in the groupings. The WISC-III indices and achievement clusters have two significant functions. The structure matrix for this analysis (WISC-III and Ach; Table 22) indicates that the academic variables of BWL, BM, and BR are highly correlate with function one (.84, .77, and .71 respectively), while two WISC-III indices correlate more with function two (VC .48 and PO .64). When these two functions are used to classify the cases approximately 40% of the cases are correct. When the group centroids or discriminant function group means are compared to the WISC-III FS score group rankings, there is a discrepancy. For function one the SLI group has the highest group mean, but this group is in fourth position on the FS group ranking. Once the SLI group is removed, the FS ranking are concurrent with function one group centroids. This rank discrepancy could be explained by the relationship between function one and BWL. The BWL scores for the SLI group are relatively higher when compared to the other classification groups.

Function two group centroids coincide with the FS rankings except for the placement of OHI and SLI. These two group classifications seem to have traded positions for function two and the only apparent explanations for this would be the amount of variability that is evident in the OHI group for the WISC-III index of PO (SD = 20) and the small group size.

The seven WJ-R Cog indices and three WJ-R Ach clusters exhibit one significant function related to classification. According to the structure matrix (Table 22), the variables correlated with this function are BWL (.79), BM (.67), BR (.65), Gc (.51), Gf (.47), and Gsm (.46). When subjects are classified using this function 36% are classified correctly. There is a rank order discrepancy between the BCA and function one group centroids. The SLI and OHI groups are displaced and this is attributed to the BWL scores, the variability that is evident within the OHI group, and the small group size. When the analysis is conducted using only the eleven index scores (four WISC-III and seven WJ-R Cog), the first function is significant and accounts for 64% of the classification variance. The indices correlated with this function are: VC (.66), Ga (.63), Gs (.60), Glr (.58), Gf (.57), Gc (.54), and FD (.47). The group centroids coincide with the BCA and FS group rankings. Approximately, 43% of the subjects are classified correctly using this function.

Table 20

Canonical Discriminant Functions for Six Classification Groups (N=180)

WISC-III Indices and WJ-R Ach

Fcn	Eigen-value	Pct of Var.	Cum Pct	Can. Corr.	Wilks' Lambda	Chi Sq	Df	Sig
1	.3607	61.4	61.4	.515	.593	89.25	25	.0000
2	.1366	23.2	84.6	.347	.806	36.73	16	.0023

WJ-R Cog Indices and WJ-R Ach

Fcn	Eigen-value	Pct of Var.	Cum Pct	Can. Corr.	Wilks' Lambda	Chi Sq	Df	Sig
1	.4073	71.8	71.8	.538	.609	87.23	30	.0000

WISC-III Indices and WJ-R Cog Indices

Fcn	Eigen-value	Pct of Var.	Cum Pct	Can. Corr.	Wilks' Lambda	Chi Sq	Df	Sig
1	.4146	63.7	63.7	.541	.565	95.53	55	.0006

Note: Ages ranged from 6 years 3 months to 16 years 0 months.

Table 21

Canonical Discriminant Analysis Structure Matrix and Group Centroids for Six Classification Groups (N=180)

WISC-III Indices and WJ-R Ach

Structure Matrix			Group Centroids		
Var	Fcn 1	Fcn 2	Group	Fcn 1	Fcn 2
BWL	.83629	-.39266	1	.65818	-.16102
BM	.77351	.24611	2	-.26739	.23733
BR	.70769	-.31094	3	.44051	.53699
VC	.65792	.47948	4	-.80045	-.23843
PrS	.46460	.31803	5	.18121	.32913
FD	.32385	.15944	6	-.63691	.17855
PO	.53197	.63941			

WJ-R Cog Indices and WJ-R Ach

Structure Matrix		Group Centroids	
Var	Fcn 1	Group	Fcn 1
BWL	.78602	1	.56562
BM	.66684	2	-.30667
BR	.64516	3	.61380
GC	.51498	4	-1.14897
GF	.46585	5	.22700
GSM	.45873	6	-.82003
GV	.25771		
GS	.56542		
GLR	.57502		
GA	.57076		

Table 21 Continued

WISC-III Indices and WJ-R Cog Indices

Structure Matrix			Group Centroids	
Var	Fcn 1		Group	Fcn 1
VC	.65711		1	-.25569
GA	.63361		2	-.18683
GS	.60317		3	.79176
GLR	.57941		4	-2.09040
GF	.57033		5	.60176
GC	.54103		6	-.50629
PRO	.46753			
PO	.51734			
GV	.21403			
FD	.46603			
GSM	.54593			

Note: WJ-R Cog Indices of GLR=Long-Term Retrieval; GSM=Short-Term Memory; GS=Processing Speed; GA=Auditory Processing; GV=Visual Processing; Gc=Comprehension-Knowledge; GF=Fluid Reasoning; WISC-III Indices of VC=Verbal Comprehension; PO=Perceptual Organization; FD=Freedom from Distractibility; and PRS=Processing Speed; WJ-R Ach Clusters of BR=Broad Reading; BM=Broad Math; and BWL=Broad Written Language. Ages ranged from 6 years 3 months to 16 years 0 months.

Table 22

Canonical Discriminant Functions for Learning Disabled (LD) and Non-placed (NP) (N=141)

WISC-III Indices and WJ-R Ach

Fcn	Eigen-value	Pct of Var.	Cum Pct	Can. Corr.	Wilks' Lambda	Chi Sq	Df	Sig
1	.260	100.0	100	.454	.794	30.50	2	.0000

WJ-R Cog Indices and WJ-R Ach

Fcn	Eigen-value	Pct of Var.	Cum Pct	Can. Corr.	Wilks' Lambda	Chi Sq	Df	Sig
1	.296	100.0	100	.478	.772	35.66	3	.0000

WISC-III Indices and WJ-R Cog Indices

Fcn	Eigen-value	Pct of Var.	Cum Pct	Can. Corr.	Wilks' Lambda	Chi Sq	Df	Sig
1	.294	100.0	100	.476	.773	32.81	11	.0006

Note: Ages ranged from 6 years 3 months to 16 years 0 months.

Table 23

Canonical Discriminant Analysis Structure Matrix and Group Centroids for Learning Disabled and Non-Placed Groups (N=141)

WISC-III Indices and WJ-R Ach

Structure Matrix		Group Centroids	
Var	Fcn 1	Group	Fcn 1
BR	.91000	2	-.34007
BWL	.88222	3	.75302
VC	.55699		
BM	.42853		
PRO	.26569		
PO	.19298		
FD	.15049		

WJ-R Cog Indices and WJ-R Ach

Structure Matrix		Group Centroids	
Var	Fcn 1	Group	Fcn 1
BWL	.83887	2	-.36989
BR	.77286	3	.78911
GLR	.63402		
BM	.56261		
GC	.55491		
GA	.49727		
GSM	.44888		
GF	.35972		
GS	.29090		
GV	.15506		

Table 23, Continued

WISC-III Indices and WJ-R Cog Indices

Structure Matrix		Group Centroids	
Var	Fcn 1	Group	Fcn 1
GLR	.65455	2	-.36136
GSM	.61609	3	.80016
GS	.39285		
GA	.29772		
GV	-.22712		
GC	.14860		
GF	.32531		
VC	-.08456		
PO	-.41767		
FD	.20851		
PRO	-.24890		

Note: WJ-R Cog Indices of GLR=Long-Term Retrieval; GSM=Short-Term Memory; GS=Processing Speed; GA=Auditory Processing; GV=Visual Processing; Gc=Comprehension-Knowledge; GF=Fluid Reasoning; WISC-III Indices of VC=Verbal Comprehension; PO=Perceptual Organization; FD=Freedom from Distractibility; and PRS=Processing Speed; WJ-R Ach Clusters of BR=Broad Reading; BM=Broad Math; and BWL=Broad Written Language. Ages ranged from 6 years 3 months to 16 years 0 months.

In those analysis that included the academic clusters, BWL was always the first variable entered, while BR and BM were entered, but later in the analysis. VC was the first variable entered in the last analysis, which included only the indices as variables. Correct classification was approximately 40% using any one of the three analysis.

Results from the two group analysis are shown in Tables 22 and 23. One discriminant function is significant for each of the analysis, and group centroids or functional group means are concurrent with the FS and BCA rankings. (NP scores are consistently higher than LD scores.) In the first analysis, those variables correlating with function one are BR (.91) and BWL (.88). The WISC-III Index of VC is not entered into the equation but does correlate moderately with function one (.56) and this is the highest index-function correlation in this analysis. When the WISC-III Indices and the WJ-R Ach variables are used for function one, the number of subjects correctly classified is approximately 71%.

In the second analysis for the two groups, using the WJ-R Indices and the WJ-R Ach variables, approximately 73% of the subjects are classified correctly. The variables most correlated with function one are: BWL (.84), BR (.77), and GLR (.63).

The third analysis, using the WISC-III and WJ-R indices as variables, indicated that approximately 65% of the subjects are classified correctly. Five of the variables, all from the WJ-R Cog, are moderately correlated with the function: Glr (.65), Gsm (.62), Gc (.55), Gf (.52), and Ga (.50). VC was the highest WISC-III index correlated with function one (.46).

To summarize, it appears that the academic clusters of BWL and BR comprise the first function when these academic scores are used in the canonical discriminant analysis for both sets of analyses. The BWL and BR when combined with the WJ-R Cog index of Glr correctly classify 73% of the LD/NP group, which is the highest percentage of correctly classified in this set. The highest correctly classified (43%) function in the six group classification set utilized the WISC-III VC and FD Indices and the WJ-R Cog Indices of Ga, Gs, Glr, Gf, and Gc.

CHAPTER 5

SUMMARY AND DISCUSSION

This study was conducted to examine the WISC-III, WJ-R Cog, and WJ-R Ach performance of a referral sample in relation to construct and concurrent validity. Several comparisons were made: 1) the referral sample performance to the WISC-III and WJ-R Cog standardization sample performance to investigate the factor construct structure; 2) the referral sample performance between the two cognitive tests to discover if the tests duplicated information; 3) the referral sample cognitive scores were viewed as to prediction of the achievement measures; and 4) the referral sample's performance was scrutinized for specific score profiles in relation to handicapping condition. Previous research had not utilized both cognitive tests with the same sample or used a referral sample to examine test theory generalizability.

Summary of Results

The results of this study indicate that there was support for the four factor WISC-III model; although, factor models 5a and 5b fit the data, also. These three models had statistical values indicative of good fit (Cole, 1987; Marsh et al, 1988; McGrew et al, 1991; Wechsler, 1991): Chi square/df ratios under 2.0; GFI and AGFI readings above .80;

RMSR estimates below .10; and TLI values above .90.

Factor model 3, the seven factor model proposed by the WJ-R authors (Woodcock & Mather, 1989), fit the referral sample data the best. The Chi square, Chi Square divided by df, GFI, AGFI, RMSR, model improvement statistics, and the TLI, all indicated that the seven factor model fits best when compared to the other models investigated.

All of the correlations between the subtests were below .600 (squared correlation .36); therefore, it was concluded that most of the subtests, when examined individually, measured unique cognitive abilities. This is indicative of differential or discriminant validity. Although, four of the seven WJ-R Cog subtests (MN, MS, PV, and AS) correlated with seven of the 12 WISC-III subtests (I, S, A, V, C, CO, and SS) above .4 but below .6. Most of the WISC-III subtests involved in these correlation were from the Verbal subtest portion of the battery. This could indicate that there is a small common verbal component measured in each test battery.

When the factor index correlations are examined, it appears that most of the WJ-R Cog and WISC-III indices were only slightly related. Two WJ-R Cog factor indices, Gc (crystallized intelligence) and Gf (fluid intelligence), correlated moderately (above .60) with the one WISC-III VC (Verbal Comprehension) factor index. Again, this suggests that there is a small common verbal component measured in both batteries.

When the WISC-III VS, PS, and FS scores are correlated with the WJ-R Cog BCA, only two correlations are moderate (above .600), VS (.720) and FS (.737). This suggests that the BCA scores and the VS and FS scores obtained from this sample share a verbal comprehension, reasoning, or memory component, almost 50%. This could be interpreted to mean that the BCA and the FS are measuring a common general or global intelligence factor. Because both batteries purport to be measures of general intellectual ability, this correlation is indicative of concurrent validity.

When the factor index scores from both batteries were examined using an exploratory principal components extraction (PC) with varimax rotation and an exploratory maximum likelihood extraction (ML) with varimax rotation, a large first factor accounted for approximately 43.9 to 39.3 percent of the variance and a weaker second factor accounted for 12.4 to 8.2 percent of the variance. When the two factors are combined a total of 47.6 to 56.3 percent of the variance is accounted for. This first large factor supports the idea of concurrent validity, because both batteries are purported to measure "g" or general intellectual ability.

When the WISC-III and WJ-R Cog Indices are used as predictor variables for WJ-R Ach scores of Broad Reading

(BR), Broad Math (BM), and Broad Written Language (BWL), the WISC-III VC Index is the only variable that is consistently entered as the first variable and is statistically significant in all of the regression equations. Although, the WISC-III composite FS did have a higher regression coefficient and standardized regression coefficient than the VC index when predicting one WJ-R Ach score (BM). No single predictor variable when compared to the WJ-R Cog indices had a higher regression coefficient (B), standardized regression coefficient (Beta), or t-value for all three academic areas; whereas, the WJ-R Cog factors of GC, GS, and GA when combined and sometimes in conjunction with other factors appear to predict academic achievement as well.

The total composite scores (FS and BCA) for both cognitive batteries were capable of differentiating cognitive abilities in the expected direction in relation to special education classification. The rank order group means for both WISC-III FS and WJ-R Cog BCA scores were congruent, with MMR, HI, SLI, SLD, ED, and NP the lowest to highest scores. Both FS and BCA group means scores were fairly close (mean score difference of 3.3 points) with the largest discrepancy of a +7 (approximately .5 S.D. difference) in the OHI group. All of the placed special education categories indicated better perceptual skills than verbal on the WISC-III mean scaled scores. The WJ-R Cog factors of Gc, Gs, and Ga were consistently the lowest scores for all of the six groups. The BWL mean score was the lowest academic mean score for most of the groups.

The WISC-III ACID or ACIDS profile (lower scores on Arithmetic, Coding, Information, Digit Span, and Symbol Search) did not appear in the majority of LD or ED classified subjects.

The two group (LD and NP) canonical discriminant function analyses correctly classified a higher percentage than did the six classification canonical discriminant function analyses. The WJ-R Cog Indices and the WJ-R Ach when used together correctly classified 73% of the LD/NP group, but only 36% were correctly classified for the six classification group. The WISC-III Indices and the WJ-R Cog Indices combined correctly classified 65% of the LD/NP classification group, but only 43% of the six classification group were correctly classified.

In the four discriminant function analyses that included the academic cluster scores, BWL and BR were often the first variables entered. These variables were moderately to highly correlated with the first function, or were the only variables comprising the first function.

Discussion of Results

WISC-III Confirmatory Factor Analysis

When the results of the WISC-III confirmatory factor analysis of this study are compared to the total standardization sample confirmatory analysis results published in the Wechsler manual (1991, pp.194-195), it appears that the goodness of fit indices values vary. The Chi square/df and the AGI values in this study were lower than those published in the manual. Some of these differences could be attributed to the sample size and the number of subtests used (12 versus 13) which in turn affect the degrees of freedom and the Chi square statistics. Both sources do support evidence for the four and five factor models.

If the five factor models are accepted, there is some difficulty with meaningfulness. One factor is represented by one subtest. The very purpose of factor analysis is to reduce the manifest data into common latent variables, which are assumed to be smaller in number (Kim & Mueller, 1978). When one factor is represented by one subtest this would seem to defeat the purpose. The underlying construct can not be separated from the subtest itself (Roid et al, 1993). It would seem that there would be no factor generalizability. Inferences about test performance would only apply to that specific factor as measured by that specific subtest. This is completely contrary to Woodcock's (1990) belief that an analysis using several batteries with multiple subtests (2 or more) for each common factor is necessary in order to obtain factors with a higher level of generality.

Another difficulty with accepting the five factor models is stability. The factor-score reliability would be equal to the subtest reliability (Roid et al, 1993). The age dependent reliabilities of Comprehension vary from .72 to .85 and of Picture Arrangement from .70 to .84. According to Roid et al (1993), neither one of these is consistently high enough for factor index reliability.

Finally, one incurs a statistical warning from the LISREL program (Joreskog & Sorbom, 1993) when trying to support a confirmatory factor with the data from one manifest variable. One other study (Roid et al, 1993) investigated the WISC-III five factor theory, and this study discussed the LISREL statistical warning that occurs. Roid concluded that this warning does not a produce confidence about the generalizability of the five factor theory.

The WISC-III confirmatory factor analysis findings of this study do not coincide with those of Sattler (1992) and Thorndike (1992), who suggest that the four factor model does not fit the data. Instead this study agrees with Roid's findings (Roid et al, 1993) with the WIAT sample. This study lends supports to the appropriateness of the use of the Wechsler four factor model and its ensuing factor generalizations with a referral sample.

WJ-R Cog Confirmatory Factor Analysis

Most of the model fit statistics found in the referral sample in this study fit the data better than the values found in the Kindergarten-Adult standardization sample presented in the WJ-R Technical Manual (1991, pp. 169-173). The Chi square divided by df was 4.79 for the standardization sample and 1.64 for the referral sample; GFI was .92 compared to .94 for the referral sample; and the RMSR values were almost equal (.045 to .04). Only the AGFI in the standardization sample was superior to the referral sample (.902 to .88). A portion of these differences could be attributed to sample size (N = 1,425 in the standardization sample; N = 183 in the referral sample), number of subtests used (29 versus 14), and the wider age range used in the standardization sample. The outcome of both analysis is that the seven factor model, as presented by the WJ-R Cog authors, can not be disconfirmed because it fits the data. These findings could have implications about how test scores should be interpreted.

WISC-III and WJ-R Cog Correlations

The WJ-R Cog subtest of MS correlated at or above .441 but below .500 with the WISC-III subtests of I, S, V, and C.

These WISC-III subtests comprised the Verbal Comprehension factor as identified by the test's authors (Wechsler, 1991). The MS (Memory for Sentences) subtest, according to the examiner's manual (Woodcock & Mather, 1989) measures short-term memory and attention. The subject repeats words, phrases, and sentences that he/she has heard from a tape recorder or the examiner. Some common ground for these subtests is that the stimulus is presented auditorially, replies are produced verbally, and the mental processes involve words or word manipulation. Given this much commonality, correlations would have been expected to be higher.

The WJ-R Cog subtest PV (Picture Vocabulary), which involves the recognition or naming of objects, correlates with the WISC-III Verbal Comprehension factor subtests of I (.509), S (.492, V (.591), and C (.431). Although the stimuli presented in PV is visual and the reply may be either pointing in the earlier portion of the subtest or naming the object, this test, also, involves the word manipulation processes. Correlations were not as high as expected.

The third WJ-R Cog subtest that correlates with four of the WISC-III subtests is AS (Analysis-Synthesis). Here the subject is taught a "miniature system of mathematics" (Woodcock & Mather, 1989, p.21). This task, which utilizes both visual and oral directions, involves symbol manipulation, short term memory, reasoning, and the understanding of basic verbal concepts of equal to, similar to, and others. AS correlated with four WISC-III verbal scale subtests: I (.417), S (.401), A (.422), and V (.444). It would be expected that AS and A would correlate because both involve mathematics; although, to be called a miniature math lesson one would expect somewhat higher correlations. Woodcock and Mather (1989) state that the AS subtest measures fluid reasoning, but the correlations with the four verbal scale WISC-III subtests suggest a small verbal (crystallized intelligence) reasoning component, also.

One other interesting outcome is the relationship of the processing speed subtests from both batteries. The WISC-III subtests of Coding and Symbol Search, which comprise the Processing Speed factor as suggested by Wechsler (1991), correlated at or slightly above .432 with the WJ-R Cog Visual Matching subtest. The Woodcock authors (Woodcock & Mather, 1989) state that VM measures processing speed. Higher correlations were expected because of the authors' claims.

When the factor indices are used instead of the subtests, subtest scores are summarized into one factor score. Two WJ-R Cog subtests are used for each WJ-R Cog factor and four WISC-III subtests constitute the VC factor.

These findings are similar to those found by McCullough and Wiebe (1991) in their study concerning third and fourth grade students Gc correlated with VC .768, Gf correlated with VC .695, and all other correlations were below .600.

In this study the WJ-R Cog BCA, composite score, is moderately related to the WISC-III VS composite score (.72) and FS composite score (.737). This suggests that the BCA scores and the VS and FS scores obtained from this sample contain a substantial verbal comprehension, reasoning, or memory component, almost 50%. It is not possible to determine whether this is a typical outcome because no other study compares all three WISC-III composite scores with the WJ-R Cog BCA score. Only the McCullough and Wiebe study (1991) examine the BCA with the WISC-R FS and this correlation is .668 for their third and fourth grade sample.

To summarize, while it appears that the WISC-III and WJ-R Cog composite scores are moderately related due to a verbal comprehension, reasoning, or memory component measured in both batteries, a large portion of both batteries does not appear to be related. This verbal comprehension, reasoning, or memory component can be traced through composite scores, factor index scores, and subtest scores.

WISC-III and WJ-R Cog Exploratory Factor Analysis

Both principal components and maximum likelihood analyses indicate a large Factor 1 composed of verbal comprehension, reasoning, and memory. Most of the subtests that comprised the index scores that form Factor 1 involve auditory input, symbol manipulation or auditory conceptualization, and verbal output. When the ML extraction method is used more indices are included in Factor 1. The ML method pursues the relationship between the variables; therefore, when the indices are examined together many of them contain a verbal comprehension, reasoning, and memory component or what could possibly called a "g" component.

The weak second factor is consistently composed of those indices that are concerned with processing speed. In the first analysis, an index from both cognitive batteries is present (PrS and Gs), but in the second analysis only the WISC-III PrS index is present. The two subtests that form the Gs Index (Visual Matching and Cross Out) require the subject to quickly compare numbers or geometric symbols and this activity could possibly contain skills of symbol manipulation, verbal mediation, or short term memory. The ML analysis supports this possibility and indicates that Gs is factorially complex with loadings of .447 on Factor 1 and .437 on Factor 2.

The WJ-R Cog Gv Index is not well represented in either factor in both analysis, and this could be interpreted to suggest that the Gv index is unique. This index is constructed from Test 5, Visual Closure, and Test 12, Picture Recognition. According to the test authors, these subtests measure "...capability in perceiving and thinking with visual patterns." (Woodcock & Mather, 1989, p.19). The subtests that form the Gv Index are unique from the other subtests in that the input is visual and the process involves visual closure and visual memory.

To summarize, the exploratory analyses suggest that two factors are present when the indices scores from both batteries are examined. A larger first factor appears to be comprised of verbal comprehension, reasoning, symbol manipulation, and reasoning. Also, a smaller second factor, composed of processing speed, is present.

Approximately one half of the variance is not accounted for by these two factors, which indicates that while the two cognitive batteries share some areas of measurement a large portion of these batteries measure different cognitive aspects.

WISC-III and WJ-R Cog Indices and Achievement Prediction

No single predictor variable when compared to the WJ-R Cog indices had a higher regression coefficient (B), standardized regression coefficient (Beta), or t-value for all three academic areas as did the WISC-III VC index.

The WJ-R Cog did not have one powerful predictor variable that occurred in all three regression equations. When the WJ-R Cog indices are examined, the Gc and Gs indices contribute to the prediction of all three academic measures but were often the last two variables entered into the regression equation; therefore, their regression coefficients, betas, and t-values were not as high as the first entered variable or the WISC-III VC index. The Ga index was the first entered variable for BR and BWL, but the Ga variable did not occur in the BM equation nor were the regression coefficients as high as the VC index.

When the composite scores are examined, the WJ-R Cog BCA appears in all three regression equations; although, the regression coefficients are not as great as the WISC-III VC index. The WISC-III FS appears in two equations, BM and BWL, and for the prediction of BM the FS has higher regression coefficients than the WISC-III VC index. The BCA is comprised of 14 subtests, the WISC-III FS is comprised of 12 subtests, and the WISC-III VC index is comprised of four subtests. It would appear that the VC index is more efficient in the number of subtests (administration time) for academic prediction as assessed by the WJ-R Ach scores.

Descriptive Statistics

When twenty-one WISC-R/WJ studies (McGrew,1986,1987; Woodcock, 1984a) are examined the median mean score discrepancy between the WISC-R FS score and the WJ BCA score was approximately five to six scaled score points with larger discrepancies exhibited for LD subjects. In most of these studies the WJ BCA score is lower than the WISC-R FS.

In this study, the WISC-III FS and WJ-R Cog BCA mean scaled score discrepancy was 3.3 points and 5 points for the LD group. This suggests that the WJ-R Cog and the WISC-III are measuring similar abilities in the referral sample. This is suggestive of concurrent validity, because both batteries purport to measure cognitive abilities and both batteries result in similar scores. The largest discrepancy (7 points) was in the OHI group. Lower WJ-R Cog BCA scores were indicated in only two groups, SLI and LD, while all of the other classification groups the BCA was higher than the WISC-III FS score.

The largest discrepancy occurred in the OHI group. The OHI classification group has much diversity within it; therefore, it is not surprising that the scores reflected this. OHI is comprised of a variety of physical disorders ranging from cerebral palsy to severe asthma. Because of these diverse physical disorders the composite score discrepancies could be attributed to many different things.

An example of one testing component dependent upon physical abilities is the mode of response (pointing, verbal reply, or paper/pencil tasks) required by the particular subtest. Also, the OHI group was the smallest group used in this study. The individual OHI performances in this sample may not be indicative of the OHI group as a whole.

In summary, because the WISC-III FS and WJ-R Cog BCA scores are analogous in this referral sample, it would appear that the two cognitive batteries are measuring similar abilities and exhibit concurrent validity.

All of the placed special education categories indicated better perceptual skills than verbal on the WISC-III mean scaled scores, while the NP group exhibited greater verbal skills. As expected from previous WISC-III research (Prifitera & Dersh, 1993), the FD mean index scores for LD and ED groups were the lowest as compared to the other WISC-III mean indices. The PrS mean index scores were the highest WISC-III index scores for ED, MMR, SLI, and OHI groups. These findings coincide with those found in the multivariate study cited in the Wechsler test manual and support Wechsler's theory (1991) that the PrS ability may not greatly contribute to general intellectual ability.

Three WJ-R Cog mean index scores (Gs, Ga, and Gc) were the lowest index means for all of the groups, except the OHI group where the Gv index was the lowest followed by Gc, Ga and Gs. What do these three abilities measure and what do they have in common?

The WJ-R Cog authors (Woodcock & Johnson, 1989) state that Gs, or processing speed, is dependent upon many things: the ability to remain focused, the speed of comparison and writing, the use of metacognition, the current emotional state of the subject, and the "...features of physiological structures (neural, hormonal)..." (Woodcock & Johnson, 1989, p.15). The Gs factor in the WJ-R Cog is measured by two subtests, Visual Matching and Cross Out. High scores on these subtests require fluent visual comparisons.

Ga, or auditory processing, is the ability to fluently comprehend speech under distracting conditions. Ga scores are, also, dependent upon the ability to remain focused, the ability to fluently compare sounds, the use of metacognition, current emotional state of the subject, and physiological structures (Woodcock & Johnson, 1989).

Gc, or comprehension-knowledge, is a measure of the individual's pool of knowledge. Gc is often referred to as "crystallized knowledge" (Woodcock & Johnson, 1989). It relies on "...communication capabilities, especially verbal abilities..." (Woodcock & Johnson, 1989, p.15), judgment, and cultural and educational experiential background.

According to the Horn-Cattell theory of intellectual processing (Horn, 1985), there is a developmental hierarchy of functions. The Gc ability is a more complex function and it is dependent upon the perceptual organization functions of Ga and Gs. Ga and Gs are the linchpins or foundations upon which Gc develops. Gc has often been equated with academic achievement or learning (Sattler, 1988). It is not surprising then that the Gc, Gs, and Ga measures were the lowest for this group of referral students. All of these students, regardless of their eventual educational classification, were referred for evaluation because of low academic functioning. In the classroom most instruction is done through the visual or auditory modalities. It would stand to reason that if these students are having difficulty with the input modes (Ga and Gs) then they would have difficulty with academic learning (Gc).

ACID or ACIDS Profile

The complete ACID or ACIDS profile did not occur in the majority of LD or ED classified students. Approximately 10% of the LD group, 0% of the ED group, and 11% of the NP exhibited the ACID profile. When the fifth subtest, Symbol Search, was included into the analysis the percentages did not noticeable change.

This is contrary to the earlier findings of Sandoval (1984), who suggested the diagnostic utility of the ACID profile, because approximately two-thirds of his LD group exhibited this profile. One major difference between

Sandoval's work and this study is that Sandoval's sample was comprised of previously identified LD students, while this study utilized a sample of referral students. The LD group was not yet classified and comparisons were made to a non-placed group. Sometimes the WISC intelligence test profile is used as a determining factor in special education classification, and therefore it would not be surprising to find the very deficit that allowed the classification in the first place to be a continuing factor in the placement classification.

Kaufman (1979) questioned the utility of profile analysis when the amount of subtest score discrepancy is unknown for the "normal or average" sample. In this study when the LD versus NP percentages are compared it appears that the ACID profile is not unique to the LD group. It must be remembered that the NP group was referred for academic problems, and therefore the NP group may not be indicative of the "normal" group. Prifitera and Dersh (1993) found the cumulative ACID pattern infrequently in the WISC-III total standardization group (approximately 1.1%) and more common in the 99 children classified as LD in the WISC-III standardization group (approximately 5.1%). Although these numbers are statistically significant at the .01 level, for all practical purposes this could be interpreted as approximately 24 subjects out of 2200 of the standardization group, and approximately 5 out of the 99 LD group would exhibit the full ACID profile. It would hardly seem efficient as a sole means of classifying for there would be many false negatives.

Although the ACID or ACIDS profile did not occur in the majority of LD, NP, or ED classified students, when the mean group index scores are examined for these three classifications the FD index score (composed of Arithmetic and Digit Span subtests) was the lowest of the four WISC-III mean index scores. Also, the mean PrS index score (composed of Coding and Symbol Search) was lower than the VC and PO indices in these groups. Approximately 41% of the NP group and 45% of the LD group displayed three of the four ACID subtests lower. These results are consistent with the findings of Prifitera and Dersh (1993), where 21% of the LD sample displayed lower scores on three of the four ACID subtests. The mean scaled subtest scores of Arithmetic, Digit Span (FD Index), Coding, and Symbol Search (PrS Index) were the lowest mean subtest scores, and that the Information mean scale score was not among the lowest for their LD group.

Could the ACID profile be indicative of other classification groups or of academic problems rather than indicative of an LD classification? It is true that the common element for students used in this study was academic difficulty and approximately 11% of the NP group in this study presented the ACID profile. Also, research by Longman, Inglish, and Lawson (1991) found behavior disordered children exhibited the same pattern of cognitive deficits as LD children. Joschko and Rourke (1985) found 6% of a large reading disabled group (N=3,500) displayed the ACID profile. This would tend to suggest that lower scores on the ACID profile subtests may not be unique to the LD classification group.

The results of this study and other studies (Joschko & Rourke, 1985; Longman, Inglish, & Lawson, 1991;Prifitera & Dersh, 1993) who have investigated the ACID pattern intimate that lower FD and PrS Index scores may be indicative of LD students and other students with different classifications or academic difficulties; therefore, when screening specifically for LD a large number of false negatives are possible. The small sample numbers in this study (LD=96; NP=45) support the use of the FD and PrS indices, but also prevent any firm generalizations.

Discriminant Analyses

The WJ-R Ach BR, BM, and BWL variables discriminated best among the six classification groups and between the two classification groups whenever these variables were included in the analysis. One of the mathematical assumptions of discriminant analysis is a multivariate normal distribution.

Unfortunately, there is some difficulty with the distribution of the achievement scores because one of the study requirements was academic difficulty. All of the test variable scores in this study were subjected to a test of univariate and multivariate normality for continuous variables using the PRELIS 2.01 Program (Joreskog & Sorbom, 1992). All of the test score variables had non-significant p-values, except for the achievement scores of BR (.100), BM (.002), and BWL (.004).

The multivariate normal distribution assumption is used to compare sample results to a theoretical probability distribution for a test of statistical significance. When the assumption is violated the sampling distribution is different from the theoretical probability distribution and tests of significance may not be accurate. Several authors have argued that discriminant analysis is a robust technique which can accept some deviation from the assumptions (Klecka, 1980). Tatsuoka (1971) has argued that discriminant analysis performs reasonably well if the assumption of multivariate normality is violated. Lauhenbruch (1975) has shown that even when the multivariate assumption is violated, if findings are robust and consistent they must be considered with some merit. Thompson (1984) suggests that the issue is not multivariate normality, but attenuation associated with distribution differences. In other words, the coefficients in the correlation matrix must not be attenuated by large deviations in the distribution shapes. According to Klecka (1980, p. 61),the violation of multivariate normality results in "... some reduction in efficiency and accuracy." Thus, the BWL, BR, and BR discriminant analysis results, although consistent and robust, must be cautiously interpreted.

One other difficulty with the six group classification discriminant analysis was different group sizes. When extreme differences are in group size are present, the results must be interpreted with caution. The six groups ranged in size from 96 for the LD to 3 for OHI. This is a large difference and the three OHI subject were randomly selected and may or may not be representative of this group. The six group results were similar to the two group results, where group size was not as grossly discrepant. This does lend some support to the results, but further work in this area is needed.

When the outcomes of the discriminant analyses are examined, it appears to reflect the results of the rest of the study. The WJ-R Cog factors of Gc, Gs, and Ga are related to the WISC-III VC factor. These factors VC, Gc, Gs, and Ga appear to be the best predictor of academic achievement. It is this area, verbal reasoning or comprehension, that all of the referral sample groups were the most deficient. The common denominator, lack of academic achievement and especially lower BR and BWL scores, for the referral sample must relate to verbal skills as measured by the WISC-III and WJ-R Cog.

Conclusions, Implications, and Further Research

Nine conclusions can be drawn from this study: 1) the four factor model, as presented by the WISC-III test author, fits the referral sample; 2) the seven factor model, as presented by the WJ-R Cog test authors, fits the referral sample; 3) according to findings in exploratory factor analysis, subtest correlations, factor correlations, composite score correlations, and discriminant analysis, there is a concurrent verbal component (approximately 50%) measured in both cognitive batteries and a discriminant component (approximately 50%) that measures unique abilities; 4) this verbal component, as measured by the VC factor in the WISC-III and the Gc, Gs, and Ga factors in the WJ-R Cog, appears to be the best predictor variable in relation to all three areas of academic achievement; 5) the students exhibiting academic difficulties in this study scored lower on those verbal component factors of WISC-III VC and WJ-R Cog Gs, Gc, and Ga; 6) BWL, broad written language, is the single best discriminator variable in both the six and two group analyses; 7) the WISC-III VC factor index appears to be the best single discriminator variable when only cognitive batteries are used; 8) there were no large discrepancies between the WISC-III and WJ-R Cog composite scores for the referral sample; and 9) the ACID or ACIDS profile did not occur in the majority of LD or ED classified students.

This study is important because it provides support for the construct validity of the WISC-III four factor model and the WJ-R Cog seven factor model with a referral sample. Because the WISC-III four factor model and the seven WJ-R Cog model fit the referral sample data, this corroborates the proposed underlying theoretical constructs of these tests. These theoretical constructs provide the basis for test generalizations about score interpretation and score meaning.

Further construct validity information was obtained from a variety of analysis, using both convergent and discriminant methods, that examined how the two cognitive tests related to each other. There is a convergent verbal component that comprises almost half of each battery, while at the same time there is a large discriminant portion (approximately 50%) that is unique to each cognitive battery. Therefore, unique information, as well as, common information is provided by the use of both tests.

Because there was little discrepancy between the composite scores, this suggests that both tests measure a concurrent general intellectual ability. Again, this information lends support of the use of the test authors' proposed test generalizations to the "at risk" or referral sample.

These findings may be limited because they are based on the results of one small sample of academically "at-risk" children and further research with other "at risk" samples would reinforce test generalizability. Also, more information may be gained if attention were paid to age differences. Past single battery research was carried out with large groups with diverse age ranges and with groups using a limited age range. At this time, no research using both complete test batteries with a limited age range of at-risks students has been conducted. Further research using this unique sample would ensure that the results of this study were generalizable.

Another area of future research could involve covariance structure analysis, structural equation modeling, or latent variable causal modeling. Further information about the relation between four of the WISC-III factors or seven of the WJ-R Cog factors and achievement in reading, mathematics, and written language could be provided by using this form of confirmatory factor analysis. Both measurement (composition of latent factors) and structural (relations among the latent factors) models should be specified in advance through the use of prior research and theory and analyses as a function of age or special education category would help to clarify the relationship between cognitive abilities and academic achievement for these groups.

APPENDIX A

HUMAN SUBJECTS LETTER

Human Subjects Committee

THE UNIVERSITY OF
ARIZONA
HEALTH SCIENCES CENTER

1690 N. Warren (Bldg. 526B)
Tucson, Arizona 85724
(602) 626-6721 or 626-7575

March 28, 1994

Dorothy M. Hall, M.Ed.
c/o Shitala P. Mishra, Ph.D.
Department of Educational Psychology
Main Campus

RE: **THE USE OF INDIVIDUALLY ADMINISTERED COGNITIVE ABILITIES TESTS WITH ACADEMICALLY "AT-RISK" CHILDREN**

Dr. Ms. Hall:

We received your above-cited research proposal. Regulations published by the U.S. Department of Health and Human Services [45 CFR Part 46.101(b) (4)] exempt this type of research from review by our Committee.

Thank you for informing us of your work. If you have any questions concerning the above, please contact this office.

Sincerely yours,

William F. Denny

William F. Denny, M.D.
Chairman
Human Subjects Committee

WFD:rs

cc: Departmental/College Review Committee

APPENDIX B

SCREE PLOT WISC-III EXPLORATORY ANALYSIS

<u>Scree Plot WISC-III Exploratory Analysis</u>

```
4.829E    *
     X
     X
     X
     X
     X
     X
     X
     X
     X
     X
     X
     X
     X
1.366E        *
     X
 .891E            *
 .510E           *   *   *   *
 .281E                               *   *   *   *
     X
```

REFERENCES

American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME) (1985). Standards for educational and psychological tests. Washington, DC: Author.

American Psychological Association (APA) (1992). Ethical principles of psychologists and code of conduct. Washington, DC: Author.

Anastasi, A. (1988). Psychological testing (6th ed.). New York: Macmillan.

Appelbaum, A. S., & Tuma, J. M. (1982). The relationship of the WISC-R to academic achievement in a clinical population. Journal of Clinical Psychology, 38, 401-405.

Arinoldo, C. G. (1982). Concurrent validity of McCarthy's Scales. Perceptual and Motor Skills, 54, 1343-1346.

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. Psychological Bulletin, 88, 588-606.

Bracken, B., Prasse, D., & Breen, M. (1984). Concurrent validity of the Woodcock-Johnson Psych-Educational Battery with regular and learning disabled students. Journal of School Psychology, 22, 185-192.

Brody, N. (1985). The validity of tests of intelligence. In B. Wolman (Ed.), Handbook of intelligence: Theories, measurements, and applications (pp. 353-389). New York: Wiley.

Carlson, L. C., Reynolds, C. R., & Gutkin, T. B. (1983). Consistency of the factorial validity of the WISC-R for upper and lower SES groups. Journal of School Psychology, 21, 319-326.

Cattell, R. B. (1950). Personality: A systematic theoretical and factoral study. New York: McGraw-Hill.

Cattell, R. B. (1963). Theory for Fluid and crystallized intelligence: A critical experiment. Journal of Educational Psychology , 54, 1-22.

Cattell, R. B. (1966). The scree test for the number of factors. Multivariate Behavioral Research, 1, 245-276.

Cole, D. (1987). Utility of confirmatory factor analysis in test validation research. Journal of Consulting and Clinical Psychology, 55, 584-594.

Coleman, M., & Harmer, W. (1985). The WISC-R and Woodcock-Johnson Tests of Cognitive Ability: A comparative study. Psychology in the Schools, 22, 1277-132.

Costenbader, V. K., & Perry, C. (1990). Test review of the Woodcock-Johnson Psychoeducational Battery-Revised. Journal of Psychoeducational Assessment, 8, 180-184.

Cummings, J., & Moscato, E. (1984a). Research on the Woodcock-Johnson Psycho-educational Battery: Implications for practice and future investigation. School Psychology Review, 13, 33-40.

Cummings, J., & Moscato, E. (1984b). Reply to Thompson and Brassard. School Psychology Review, 13, 45-58.

Diana v. State Board of Education, C-70, 37 Rft (N.D. Cal. 1970).

Doe v. Board of Education of Montgomery County, 696 F. 2d 109 (D.C. Cir. 1982).

Dunn, L. M., & Dunn, L. M. (1981). Peabody Picture Vocabulary Test-Revised. Circle Pines, MN: American Guidance Service.

Education of All Handicapped Children's Act of 1975 (PL 94-142). 20 USC Sec. 401 (Supp. 1975).

Estabrook. G. (1984). A canonical correlation analysis of the WISC-R and the Woodcock-Johnson Tests of Cognitive Ability in a sample referred for suspected learning disabilities. Journal of Educational Psychology, 76, 1170-1177.

Federal Register. Procedures for evaluating specific learning disabilities. December 29, 1977, pp. 65082-65085.

Fuller, G. B., & Goh, D. S. 91981). Intelligence, achievement, and visual motor performance among learning-disabled and emotionally impaired children. Psychology in the Schools, 18, 261-268.

Guadalupe Organization v. Tempe Elementary School District, E.D. AZ., (1972).

Georgia State Conference of Branches of NAACP v. State of GA, 11th Cir Court of Appeals, 84-8771 (1985).

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

Hishinuma, E. S., & Yamakawa, R. (1993). Construct and criterion-related validity of the WISC-III for exceptional students and those who are "at risk". Journal of Psychoeducational Assessment: WISC-III Monograph, 94-104.

Horn, J. L. (1985). Intellectual ability concepts. In R. J. Sternberg (Ed.), Advances in the psychology of human intelligence (Vol.3, pp. 91-96). Norwood, NJ: Ablex.

Horn, J. L. (1988). Thinking about human abilities. In J. R. Nesselroade & R. B. Cattell (Eds.), Handbook of multivariate psychology (2nd ed.) (pp. 645-865). New York: Academic Press.

Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized intelligence. Journal of Educational Psychology , 47, 253-270.

Horn, J. L., & Cattell, R. B. (1967). Second order ability structure revealed in right and wrong scores. Journal of Educational Psychology, 58, 115-122.

Ipsen, S., McMillan, J., & Fallen, N. (1983). An investigation of the reported discrepancy between the Woodcock-Johnson Tests of Cognitive Ability and the Wechsler Intelligence Scale for Children-Revised. Diagnostique, 9, 32-44.

Jeffrey, T.B., & Jeffrey, L.<. (1984). Validity of the Slosson Intelligence Test with young learning disabled children. Psychological Reports, 54, 97-98.

Jensen, A. R. (1980). Bias in mental testing. New York: Free Press.

Joreskog, K. G., & Sorbom, D. (1986). LISREL VI: Analysis of linear structural relationships by maximum likelihood, instrumental variables, and least squares methods (4th ed.). Uppsula, Sweden: University of Uppsula Department of Statistics.

Joreskog, K. G., & Sorbom, D. (1987). LISREL 6.13: User's reference guide. Chicago: Scientific Software.

Joreskog, K. G., & Sorbom, D. (1989). LISREL VII: A guide to the program and applications (2nd ed.). Chicago: SPSS, Inc.

Joschko, M., & Rourke, B.P. (1985). Neuropsychological subtypes of learning-disabled children who exhibit the ACID attern on the WISC. In B. P. Rourke (Ed.), Neuropsychology of learning disabilities: Essentials of subtype analysis (pp.65-88). New York: Guilford Press.

Kamphaus, R. W. (1993). Clinical assessment of children's intelligence. Needham, MA: Allyn & Bacon.

Kaufman, A. S. (1975). Factor analysis of the WISC-R at 11 age levels between 6 1/2 and 16 1/2 years. Journal of Consulting and Clinical Psychology , 43, 135-147.

Kaufman, A. S. (1979). Intelligence testing with the WISC-R. New York: Wiley-Interscience.

Kaufman, A. S. (1990). Assessing adolescents and adult intelligence. Boston: Allyn & Bacon.

Kavale, K. A., & Forness, S. R. (1984). A meta-analysis of the validity of Wecechsler Scale profiles and recategorizations: Patterns or parodies? Learning Disability Quarterly, 7, 136-156.

Klecka, W. R. (1980). Discriminant analysis. Beverly Hills: Sage Publications.

Kehle, T. J., Clar, E., & Jenson, W. R. (1993). The development of testing as applied to school psychology. Journal of School Psychology, 31, 143-161.

Kerlinger, F.N. (1964). Foundations of behavioral research: Educational and psychological inquiry. New York: Holt, Rinehart, & Winston.

Knoff, H.M. (1986). Identifying and classifying children and adolescents referred for personality assessment: Theories, systems, and issues. In H.M. Knoff (Ed.), The assessment of child and adolescent personality (pp. 3-34). New York: Guilford.

Lachenbruch, P. A. (1975). Discriminant analysis. New York: Hafner.

(Larry) P. v. Riles, 343 F. Supp. 1306 (N.D. Cal. 1972) aff'r 502 F.2d 963 (9th Cir. 1974); 495 F. Supp. 926 (N.D. Cal. 1979); appealed docket, No. 80-4027 (9th Cir., Jan. 17, 1980).

Longman, R. S., Inglis, J., & Lawson, J. S. (1991). WISC-R patterns of cognitive abilities in behavior disordered and learning disabled children. Psychological Assessment: A Journal of Consulting and Clinical Psychology, 3, 239-246.

Marsh, H. W., Balla, J. R., & McDonald, R. P. (1988). Goodness-of-fit indexes in confirmatory factor analysis: The effects of sample size. Psychological Bulletin, 103, 391-410.

McGrew, K. S. (1984). Normative based guides for subtest profile interpretation of the Woodcock-JOhnson Tests of Cognitive Ability. Journal of Psychoeducational Assessment, 2, 141-148.

McGrew, K. S. (1986). Clinical interpretation of the Woodcock-Johnson Tests of Cognitive Ability. Orlando, FL: Grune & Stratton.

McGrew, K. S. (1987). A multivariate analysis of the Wechsler/Woodcock-Johnson discrepancy controversy. Journal of Psychoeducational Assessment, 5, 49-60.

McGrew, K. S., Werder, J. K., & Woodcock, R. W. (1991). Woodcock-Johnson Psychoeducational Battery-Revised technical manual. Allen, TX: DLM.

Messick, S. (1989). Meaning and values in test validation: The science and ethics of assessment. Educational Researcher, 6, 5-11.

Naglieri, J. A., & Haddah, F. A. (1984). Learning disabled children's performance on the Kaufman Assessment Battery for Children: A concurrent validity study. Journal of Psychoeducational Assessment, 2, 49-56.

National Association of School Psychologists (NASP) (1985). Principles for Professional Ethics. Washington, DC: Author.

Newby, R. F., Recht, D. R., Caldwell, J., & Schaefer J. (1993). Comparison of WISC-III and WISC-R IQ changes over a 2-year time span in a sample of children with dyslexia. Journal of Psychoeducational Assessment: WISC-III Monograph, 87-93.

Ownby, R. L., & Matthews, C. G. (1985). On the meaning of the WISC-R third factor: Relations to selected neuropsychological measures, Journal of Consulting and Clinical Psychology, 53, 531-534.

Petersen, C. R., & Hart, D. H. (1979). Factor structure of the WISC-R for a clinic-referred population and specific subgroup. Journal of Consulting and Clinical Psychology, 47, 643-645.

Phelps, L., Rosso, M., & Falasco, S. (1984). Correlations between the Woodcock-Johnson and the WISC-R for a behavior disordered population. Psychology in the Schools, 21, 442-446.

Prifitera, A., & Dersh, J. (1993). Base rates of WISC-III diagnostic subtest patterns among normal, learning-disabled, and ADHD samples. Journal of Psychoeducational Asessment: WISC-III Mongraph, 43-55.

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

Reschly, D. J. (1990). Found: Our intelligences: What do they mean? Journal of Psychoeducational Assessment, 8, 259-267.

Reynolds, C. R., Elliott, S. N., Gutkin, T. B., & Witt, J. C. (1984). School Psychology: Essentials of theory and practice. New York: Wiley & Sons.

Roid, G. H., Prifitera, A., & Weiss, L. G. (1993). Replication of the WISC-III factor structure in an independent sample. Journal of Psychoeducational Assessment: WISC-III Monograph, 6-21.

Sandoval, J. (1984, April). Verifying the WISC-R ACID profile. Paper presented at the meeting of the National Association of School Psychologists, Philadelphia.

Sattler, J.M. (1988). Assessment of children (3rd ed.). San Diego: Author.

Sattler, J. M. (1992). Assessment of children: WISC-III and WPPSI-R supplement. San Diego, CA: Author.

Schwean, V. L., Saklofske, D. H., Yackulic, R. A., & Quinn, D. (1993). WISC-III performance of ADHD children. Journal of Psychoeducational Assessment: WISC-III Monograph, 56-70.

Shaw, S. R., Swerdlik, M. E., & Laurent J. (1993). Review of the WISC-III. Journal of Psychoeducational Assessment: WISC-III Monograph, 151-160.

Shinn, M., Algozzine, B., Marston, D., & Ysseldyke, J. (1982). A theoretical analysis of the performance of learning disabled students on the Woodcock-Johnson Psycho-educational Battery. Journal of Learning Disabilities, 15, 221-226.

Stell v. Savannah-Chatham Board of Education, 220 F. Supp. 667 (S.B. Ga.1963), rev'd 333 F. F. 2d 55 (5th cir. 1963), cert. denied, 379 U.S. 933 (1964).

Stewart, K. J., & Moely, B. E. (1983). The WISC-R third factor: What does it mean? Journal of Consulting and Clinical Psychology, 51, 940-941.

Tatsuoka, M. M. (1971). Multivariate analysis. New York: John Wiley.

Teeter, P. A., & Smith, P. L. (1993). WISC-III and WJ-R: Predictive and discrminant validity for students with severe emotional disturbance. Journal of Psychoeducational Assessment: WISC-III Monograph, 114-124.

Thomopson, B. (1984). Canonical correlation analysis. Beverly Hills: Sage Publications.

Thompson, P., & Brassard, M. (1984). Validity of the Woodcock-Johnson Tests of Cognitive Ability: A comparison with the WISC-R in LD and normal elementary students. Journal of School Psychology , 22, 201-208.

Thorndike, R. L. (1992, March). Intelligence tests: What we have and what we should have. Paper presented at the meeting of the National Association of School Psychologists, Nashville, TN.

Thorndike, R. L., Hagen, E. P., & Sattler, J. M. (1986). Technical manual, Stanford-Binet Intelligence Scale: Fourth Edition. Chicago: Riverside Publishing.

Tucker, L. R., & Lewis, C. (1973). The reliability coefficient for maximum likelihood factor analysis. Psychometrika, 38, 1-10.

Wechsler, D. (1939). The measurement of adult intelligence. Baltimore: Williams & Wilkins.

Wechsler, D. (1943). Non-intellective factors in general intelligence. American Psychologist, 5, 78-83.

Wechsler, D. (1949). Manual for the Wechsler Intelligence Scale for Children. New York: The Psychological Corporation.

Wechsler, D. (1991). Manual for the Wechsler Intelligence Scale for Children-Third Edition. San Antonio: The Psychological Corporation.

Weiss, L. G., Prifitera, A., & Roid, G. (1993). The WISC-III and the fairness of predicting achievement across ethnic and gender groups. Journal of Psychoeducational Assessment: WISC-III Monograph, 35-42.

West, R. (1991). Computing for psychologists: Statistical analysis using SPSS and Minitab. London: Harwood Academic Publishers.

Woodcock, R. S. (1978). Development and standardization of the Woodcock-Johnson Psycho-educational Battery. Allen, TX: DLM.

Woodcock, R. S. (1984a). A response to some questions raised about the Woodcock-Johnson: The mean score discrepancy issue. School Psychology Review, 13, 342-354.

Woodcock, R. S. (1984b). A response to some questions raised about the Woodcock-Johnson: Efficacy of the aptitude clusters. School Psychology Review, 13, 355-362.

Woodcock, R.S. (1990). Theoretical foundations of the WJ-R measures of cognitive ability. Journal of Psychoeducational Assessment, 8, 231-258.

Woodcock, R. W., & Johnson, M. B. (1989). Woodcock-Johnson Psycho-Educational Battery-Revised. Allen, TX: DLM.

Woodcock, R. W., & Mather, N. (1989). WJ-R tests of cognitive ability-standard and supplemental batteries: Examiner's manual. In R. W. Woodcock & M. B. Johnson, Woodcock-Johnson Psychoeducational Battery-Revised. Allen, TX: DLM.

Ysseldyke, J., Shinn, M., & Epps, S. (1980). A comparison of the WISC-R and the Woodcock-Johnson Tests of Cognitive Ability (Research Report #36). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities.

Ysseldyke, J., Shinn, M., & Epps, S. (1981). A comparison of the WISC-R and the Woodcock-Johnson Tests of Cognitive Ability. Psychology in the Schools, 18, 15- 19.

Zachary, R. A. (1990). Wechsler's intelligence scales: Theoretical and practical considerations. Journal of Psychoeducational Assessment, 8, 276-289.