INFORMATION TO USERS

This reproduction was made from a copy of a document sent to us for microfilming. While the most advanced technology has been used to photograph and reproduce this document, the quality of the reproduction is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help clarify markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure complete continuity.

2. When an image on the film is obliterated with a round black mark, it is an indication of either blurred copy because of movement during exposure, duplicate copy, or copyrighted materials that should not have been filmed. For blurred pages, a good image of the page can be found in the adjacent frame. If copyrighted materials were deleted, a target note will appear listing the pages in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed, a definite method of "sectioning" the material has been followed. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For illustrations that cannot be satisfactorily reproduced by xerographic means, photographic prints can be purchased at additional cost and inserted into your xerographic copy. These prints are available upon request from the Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases the best available copy has been filmed.

University
Microfilms
International
300 N. Zeeb Road
Ann Arbor, MI 48106

Mather, Nancy

PERFORMANCE OF LEARNING DISABLED SUBJECTS AND GIFTED SUBJECTS ON THE WOODCOCK-JOHNSON PSYCHO-EDUCATIONAL BATTERY AND THE WECHSLER INTELLIGENCE SCALE FOR CHILDREN--REVISED

The University of Arizona PH.D. 1984

PERFORMANCE OF LEARNING DISABLED SUBJECTS AND GIFTED SUBJECTS ON THE WOODCOCK-JOHNSON PSYCHO-EDUCATIONAL BATTERY AND THE WECHSLER INTELLIGENCE SCALE FOR CHILDREN — REVISED

By

Nancy Mather

Dissertation Submitted to the Faculty of the

DEPARTMENT OF SPECIAL EDUCATION

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

1 9 8 4

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Final Examination Committee, we certify that we have read the dissertation prepared by Nancy Mather

entitled Performance of Learning Disabled Subjects and Gifted Subjects on the Woodcock-Johnson Psycho-Educational Battery and the Wechsler Intelligence Scale for Children - Revised

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

	2/23/84 Date
Candace L. Bos	February 23, 1984 Date
	February 23, 1984 Date
	Feb 23, 1984 Date
	Feb 23, 1984 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.

	2/25/84
Dissertation Director	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 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: Nancy Mathes

DEDICATION

This dissertation caused much travail,
In a brief synopsis, I'll tell the tale.
The first draft didn't get very far,
as someone decided to burgle my car
That poor thief must have had a fit,
when all he got was a review of lit.
In retrospect, although confusing,
several moments were quite amusing.
After some seconds of cogitation,
the policeman asked,"How d'ya spell 'dissertation'?"
Quick to respond was my friend, Lynne
"It's phonetic", she replied with a grin.
Then he queried, "How much did it cost?",
"Ten thousand dollars has been lost."
He quickly shoved away his book
and decided he'd better help us look.

Next, all the data had to be re-scored
Punchcards, printouts, being bored
My competent, but weary, statistian,
suggested I hire a clever magician.
But, through it all, we did persevere
and thanks to my committee, I'm still here.

Most sincerely, I thank Samuel Kirk,
who encouraged me to complete all this work.
He spent hours and days of precious time,
discussing aspects of this design,
and Mrs. Kirk brought cookies and tea,
that helped dispel encroaching ennui.
William Healey had to forgive,
an occasional split infinitive.
At times, I feared he'd go plum daft,
when he had to read just one more draft.
For his critical comments and editor's skill,
particular thanks goes to Bill.
Candace Bos gave her concentration
to many aspects of this dissertation.
She was always willing to muster,
the energy to examine one more cluster.
A weekend at her house was spent,
with her Apple computer that she lent.
In spite of the power failures - twice
Erasing all data, which wasn't nice
Candy helped me be forgiving

reminding me - I'd go on living.
Judy Mitchell and John Bradley
also gave of their time quite gladly.
John, I found, was really quite able
at teaching me how to present a table.

I'd be remiss to forget and report,
a few fond friends who provided support.
Lynne Jaffe could not have foreseen,
time she'd spend regressing to the mean.
Anne Udall always proceeded with zest
when she administered one more test.
Dick Woodcock went beyond being nice
to read this study and give me advice.
And Jeanne McCarthy, I must interject
also took her time to inspect.

Hard to believe: no more to defend?
Why I never dreamed, it'd come to an end.
This research progressed through time,
like most of the meter in this rhyme.
But I've learned a lot, I must confess
and you've all played a role in this success...
Acknowledgements are the very best part,
as words can be spoken from the heart.
I would like to express sincere appreciation,
and dedicate to you this dissertation.

TABLE OF CONTENTS

TABLE OF CONTENTS -- Continued

TABLE OF CONTENTS -- Continued

TABLE OF CONTENTS -- Continued

LIST OF TABLES

LIST OF TABLES -- Continued

LIST OF ILLUSTRATIONS

ABSTRACT

The purpose of this study was to compare and analyze test scores on the Woodcock-Johnson Psychoeducational Battery (Battery) and the Wechsler Intelligence Scale for Children-Revised (WISC-R) for a sample of subjects classified as learning disabled and a sample of subjects classified as gifted and talented. Subjects were randomly selected and included 51 school-identified learning disabled students and 46 school-identified gifted and talented students. All subjects were administered the Woodcock-Johnson Tests of Cognitive Ability (WJTCA), the Woodcock-Johnson Tests of Achievement (WJTA) and the WISC-R.

The specific areas of investigation for each sample included: (a) performance on the Achievement-Aptitude Profile of the Battery, (b) comparability of full scale scores, (c) performance on the Cognitive clusters of the Battery, (d) performance on the Achievement clusters of the Battery, and (e) the magnitude of Verbal-Performance Scale Score discrepancies on the WISC-R.

Results indicated that the majority (74.5%) of the subjects classified as learning disabled and a few (10.9%) of the subjects classified as gifted and talented evidenced

a moderate or severe deficit on the Achievement-Aptitude Profile. The subjects classified as learning disabled scored significantly higher on the Full Scale score of the WISC-R than on the full scale score of the WJTCA, while the difference between full scale scores was not significant for the subjects classified as gifted and talented.

An important finding was that scores obtained on the Reasoning cluster were invalid for subjects in both samples. Further analyses indicated that the alternative clusters, Oral Language and Broad Reasoning (Woodcock, 1983) provided a more accurate appraisal of abilities for each sample.

Overall, the study provided insights regarding test performance on the Battery and the WISC-R for these two diverse samples.

CHAPTER 1

INTRODUCTION

One major criterion for the determination of a specific learning disability is a significant discrepancy between aptitude and achievement (U.S. Office of Education, 1977). In practice, this discrepancy has been derived by comparing a student's performance on a test of intelligence such as the Wechsler Intelligence Scale for Children-Revised (Wechsler, 1974) to an achievement test such as the Wide Range Achievement Test (Jastak & Jastak, 1978) or the Peabody Individual Achievement Test (Dunn & Markwardt, 1970).

This procedure has several limitations: (a) the measures have been normed or standardized on different populations, (b) the testing of the student is usually performed by two different individuals, such as a school psychologist and a learning disability or remedial teacher, and (c) the amount of discrepancy needed to qualify for services varies, depending upon the established criterion of the school district. For these reasons, assessment results and placement decisions may often be equivocal.

A recently developed test, the Woodcock-Johnson Psycho-Educational Battery (Woodcock & Johnson, 1977), purports to rectify these limitations and meet many of the assessment needs of professionals involved with the field of learning disabilities. The Woodcock-Johnson Psycho-Educational Battery (Battery) contains both cognitive or intelligence subtests (Part One) and achievement subtests (Part Two). Part One, the Woodcock-Johnson Tests of Cognitive Ability (WJTCA), consists of 12 subtests that are grouped into four clusters (Verbal Ability, Reasoning, Perceptual Speed, and Memory), and four Scholastic Aptitude clusters that are used to predict a student's achievement. Part Two, the Woodcock-Johnson Tests of Achievement (WJTA), consists of 10 subtests that are grouped into four achievement clusters (Reading, Mathematics, Written Language, and Knowledge).

The major advantages of the Battery are (a) the norms were established on the same subjects, (b) both parts may be administered by the same person, and (c) the cognitive and achievement parts may be compared to indicate the amount or level of discrepancy between expected achievement (aptitude) and actual achievement. This comparison, the Achievement-Aptitude Profile, makes the Battery particularly useful for identifying students with learning deficits (Woodcock, 1978).

As a result of these factors, the Battery is becoming widely used. In the spring of 1978, five major institutes for research in learning disabilities suggested this test be used as a common measure for describing subjects (Reeve, Hall, & Zakreski, 1979). The rapid acceptance of this instrument nationally, however, seems to have preceded systematic studies of the test by independent investigators (Ysseldyke, Shinn, & Epps, 1981).

The majority of studies that have been conducted have examined the concurrent validity of the WJTCA by comparing it to the Wechsler Intelligence Scale for Children-Revised (WISC-R). Little agreement exists among the findings of these studies. When compared to full scale WISC-R scores, several of the studies report that learning disabled subjects tend to score lower on the full scale score of the WJTCA (Reeve, Hall, & Zakreski, 1979; Ysseldyke, Shinn, & Epps, 1981). Others show no significant mean differences between the full scale scores of the two tests (Ipsen, McMillan, & Fallen, 1982; McGrew, 1983). Extant research does not confirm whether or not the WJTCA and the WISC-R function comparably as ability measures.

Contradictions also exist regarding performance on the Cognitive clusters. For example, Reeve, Hall and Zakreski (1979) found that learning disabled students scored lowest on the Perceptual Speed cluster; the finding was not

replicated by Ysseldyke, Shinn, & Epps (1981).

Furthermore, few studies have examined the performance of school-aged populations other than low achieving and learning disabled students. In a small pilot study, this investigator tested five students from a Gifted and Talented Education program with the WJTCA and the WISC-R. Of the five gifted and talented children, four scored significantly higher on the Broad Cognitive Ability score (full scale score) of the WJTCA than on the Full Scale IQ of the WISC-R. This finding led the investigator to wonder if gifted and talented students score higher on the WJTCA than on the WISC-R, while learning disabled students demonstrate an opposite pattern.

Many questions about the Battery remain unanswered. Presently, few research studies have analyzed the performance of learning disabled subjects on the Achievement-Aptitude Profile. Research is needed to determine if the Battery is a viable instrument for making educational decisions regarding learning disabled students. Furthermore, no research could be found that investigated the performance of high-achieving or gifted and talented students on this instrument. Investigators agree that individually administered intelligence tests are the most accurate and preferred method for identifying gifted students (Clark, 1979; Kirk, 1972; Marland, 1971; Sattler, 1982) and, conceivably, the WJTCA could demonstrate validity

as a measure for the identification of these students. Professionals involved in gifted education have also been concerned with the identification of the underachieving student (Gallagher, 1975; Whitmore, 1980). The Achievement-Aptitude Profile could prove to be a valid measure for identifying underachieving gifted students.

Statement of Purpose

The purpose of this study was to analyze and compare test scores on the Woodcock-Johnson Psycho-Educational Battery and the Wechsler Intelligence Scale for Children-Revised (WISC-R) for a sample of subjects classified as learning disabled and sample of subjects classified as gifted and talented.

Research Questions

The following research questions were investigated:

1. What percentage of the subjects classified as learning disabled and what percentage of the subjects classified as gifted and talented will display a significant discrepancy (moderate or severe deficit) between aptitude and achievement as determined by the functioning levels on the Achievement-Aptitude Profile of the Woodcock-Johnson Psycho-Educational Battery?

Students are placed in learning disabilities programs because examiners have determined that a discrepancy exists between intelligence and academic achievement. This suggests that the majority of students should exhibit an educationally significant discrepancy on

the Achievement-Aptitude Profile in at least one achievement area.

Students are placed in gifted and talented programs because of high intelligence and accelerated achievement. This suggests that the majority of students in this sample should not exhibit educationally significant discrepancies between aptitude and achievement.

2. Are the full scale scores of the Woodcock-Johnson Tests of Cognitive Ability and the WISC-R comparable for the subjects classified as learning disabled and/or for the subjects classified as gifted and talented?

Previous research suggests that the WJTCA and the WISC-R might not provide comparable global measures of intelligence for learning disabled subjects. No research was found that investigated the performance of gifted and talented subjects on the WJTCA.

3. Do the subjects classified as learning disabled and/or the subjects classified as gifted and talented display a characteristic pattern of performance across the four Cognitive clusters (Verbal Ability, Reasoning, Perceptual Speed, and Memory) of the Woodcock-Johnson Tests of Cognitive Ability?

Some evidence suggests that learning disabled students score lowest on the Perceptual Speed cluster (Ipsen et al., 1982; Reeve et al., 1979). Conceivably, the gifted and talented subjects should score highest on the Verbal Ability, Memory, and Reasoning clusters and lowest on the Perceptual Speed cluster which seems more influenced by physiological development and maturation.

Recently, Woodcock (1983) has released alternative

Cognitive clusters of Oral Language and Broad Reasoning that may be substituted for the Verbal Ability and Reasoning clusters. Although this scoring procedure is not yet published, comparisons were made to the original clusters to determine if use of these alternative clusters would alter the results.

4. Do the subjects classified as learning disabled and/or the subjects classified as gifted and talented display a pattern of performance on the four Achievement clusters (Reading. Mathematics, Written Language, and Knowledge) of the Woodcock-Johnson Tests of Achievement?

Many learning disabled students are referred for reading difficulties, suggesting that the subjects in the learning disabled sample should evidence the greatest deficits on the Reading and Written Language clusters. Conversely, since gifted and talented students are often superior in Language Arts, this sample might excel on the Reading and Written Language clusters.

5. Do the subjects classified as learning disabled and/or the subjects classified as gifted and talented demonstrate significant discrepancies between Verbal and Performance Scale scores of the WISC-R?

Regression effects and research suggest that learning disabled students often exhibit significant discrepancies between Verbal and Performance Scale scores of the WISC-R. A similar analysis needs to be made for gifted and talented students to determine if significant discrepancies also occur in this population.

Significance of the Study

The United States Department of Education (1982) has expressed concern regarding the dramatic growth that has taken place in the field of learning disabilities. Conceivably, the Achievement-Aptitude Profile of the Battery could improve the validity of identification procedures by providing practitioners with a selection criterion or a common standard for determining an educationally significant discrepancy. Also, the Achievement-Aptitude Profile may prove useful for identifying underachieving gifted students.

Additionally, the study provides data regarding the comparability of the WJTCA and the WISC-R. The initial studies by Reeve et al. (1979) and Ysseldyke et al. (1981) reported that learning disabled students scored as much as three-fourths of a standard deviation lower on the WJTCA than on the WISC-R. These findings suggest that the WJTCA provide a lower ability estimate of the learning disabled student than the WISC-R. Since a discrepancy is necessary for placement in a learning disability program, some students may be classified by one method but not by another because their low achievement would appear commensurate with their assessed low ability. To make accurate psychoeducational decisions, examiners must know whether the WJTCA and the WISC-R provide comparable data.

Another outcome of this study will be to provide data on the performance of high achieving and gifted and talented children on the Battery. Marland (1971) reported that individual intelligence tests are the preferred method of identification of gifted students. Since the WJTCA is more of a measure of verbal cognition and more related to academic achievement than the WISC-R (Hessler, 1982), it may be a more sensitive instrument for the identification of gifted students.

Several investigators have questioned the validity of the Verbal Ability and Reasoning clusters for use with special populations (Hessler, 1982; McGue, Shinn, and Ysseldyke, 1982). Woodcock (1983) responded to these concerns by creating the alternative clusters of Oral Language and Broad Reasoning that may be substituted for the Verbal Ability and Reasoning clusters. This study provides an analysis of scores on the original and the alternative Cognitive clusters for the subjects classified as learning disabled and the subjects classified as gifted and talented.

The Woodcock-Johnson Psycho-Educational Battery shows promise for use with learning disabled and gifted and talented students because it is the first individually administered test that purports to examine both aptitude and achievement. A reliable, technically sound instrument is

needed to assist practioners with identification, diagnosis, and placement of exceptional children. Reeve et al. (1979) have discussed the importance of obtaining external verification for diverse populations of students. This study should enhance understanding of these two instruments that are so widely used in educational practice.

Definition of Terms

Cognitive Ability

The construct "cognitive ability" refers to a broad, complex set of abilities that includes verbal and non-verbal functions, ranging from lower mental processes to higher mental processes. These processes have been grouped into broader functions of cognitive ability including discrimination-perception, memory-learning, reasoning-thinking, and knowledge-comprehension (Woodcock, 1978).

Achievement

Achievement, as measured by the Battery, includes assessment of proficiency and skill in the areas of reading, mathematics, written language, and knowledge.

Functioning Levels on the Achievement-Aptitude Profile

The Achievement-Aptitude Profile of the Battery has seven functioning levels that describe the amount of

difference between aptitude and achievement. These verbal labels were based on increments of ten cluster difference score points and include: severe deficit (-26), moderate deficit (-16), below average (-6), average (0), above average (+6), superior (+16), and very superior (+26). For this study, a significant discrepancy between aptitude and achievement was defined as a moderate or severe deficit.

CHAPTER 2

REVIEW OF THE LITERATURE

The purpose of this chapter is to provide a review of literature on the following topics: (a) the use of discrepancies between aptitude and achievement for the identification of learning disabled students and gifted underachievers, (b) the performance of learning disabled students on the Woodcock-Johnson Tests of Cognitive Ability and the Wechsler Intelligence Scale for Children-Revised, (c) the performance of learning disabled students on the Woodcock-Johnson Psycho-Educational Battery, (d) the performance of learning disabled students on the Wechsler Intelligence Scale for Children-Revised, and (e) the performance of gifted students on intelligence tests.

Discrepancies between Aptitude and Achievement

Discrepancies and Learning Disabilities

One salient criterion for identifying a student with a learning disability is the determination of a discrepancy between ability and achievement. The regulations of PL 94-142 mandate that to qualify for learning disability

services, a diagnostic team must determine that a child has a severe discrepancy between intellectual ability and academic achievement (U.S. Office of Education, 1977).

Originally, the United States Office of Education (1976) proposed a formula that would specify the severity level needed for eligibility. The intent of the formula was to provide a common standard for identifying students with extreme difficulty in an academic area. The formula was dropped from the final regulations due to overwhelming opposition. The major concerns regarding classification by formula included: (a) dependency on scores from intelligence tests, (b) failure to account for the number of years a student had been in school, (c) selection of an arbitrary severity level, (d) lack of teacher preparation for use of a formula, and (e) difficulty in determining when special services should be discontinued (Danielson & Bauer, 1978). Others addressed the statistical limitations involved when comparing aptitude and achievement (Anastasi, 1976; Cone & Wilson, 1981; McLeod, 1979).

Presently, although the regulations state that a discrepancy must be identified, it is not specified how large the discrepancy should be or how it should be derived. Consequently, states have been faced with the dilemma of quantifying aptitude-achievement discrepancies.

Several investigators have indicated that varied operational definitions or eligibility criteria result in disparate classifications (Algozzine & Ysseldyke, 1983; Ysseldyke, Algozzine, & Epps, 1983). In some cases, students enrolled in learning disability programs do not seem to possess discrepancies. For example, Shinn, Algozzine, Marston, and Ysseldyke (1982) found that 13 of 50 students identified as learning disabled did not evidence a discrepancy greater than one standard deviation on any of the commonly used measures.

Discrepancies and Gifted Underachievement

The existence of a discrepancy between aptitude and classroom achievement has been the basic criterion for the identification of the gifted underachiever (Dowdall & Colangelo, 1982). Typically, a gifted student is identified as an underachiever when achievement is far below what would be predicted by the results of an intelligence test (Gallagher, 1975). Individual intelligence tests, such as the Wechsler Intelligence Scale for Children-Revised (WISC-R) have been reported to be accurate and useful for identifying underachieving gifted children (Whitmore, 1980).

Despite the interest in underachieving gifted students, definition and identification have remained a problem (Whitmore, 1980). Definitions seem to be either too restrictive or too vague and, consequently, an accepted

prevalence estimate does not exist. The last twenty years of research on the underachieving gifted child have not provided a clear definition of this population (Dowdall & Colangelo, 1982).

Quantifying Discrepancies

One major problem in identifying the learning disabled and underachieving gifted student has been quantifying the existence of severe academic discrepancies. Many investigators have indicated that inadequacies exist in the methods used (Cone & Wilson, 1981; Hanna, Dyck, & Holen, 1979). Anastasi (1976) commented that categorizing students by comparing intelligence with achievement is a fallacious procedure; two tests do not have a perfect correlation and, consequently, measurement error will produce either under-prediction or over-prediction of students.

Cone and Wilson (1981) addressed the issue of the lack of correlation between tests and pointed out that attempts to quantify a discrepancy have failed to take into account the phenomenon of regression to the mean. They explained that students with extreme intelligence scores will not obtain equally extreme achievement scores: "Children with IQs of 115 tend not to obtain equally high achievement standard scores. The distribution of achievement scores has regressed or moved toward the average; the expected achievement value is 109" (p. 366).

McLeod (1979) also explained how comparing intelligence and achievement is invalid because of statistical regression. Students who score 130 on a nonverbal test of intelligence should have an average reading quotient around 115. He commented that "... half of the students with an IQ of 130 are bound mathematically to be 15 or more points behind their IQ in reading. How many times has it been reported that over 50% of gifted children are educationally retarded" (p. 45)? Both measurement and regression error should be considered when examining aptitude-achievement discrepancies of learning disabled and gifted students.

Hanna et al. (1979) proposed new criteria for the development of more adequate procedures for identifying discrepancies. Their discussion included the following principles: (a) all derived scores must be based on a national normative group, (b) standardization samples should be analyzed before assuming the comparability of two tests or tests should be used that have been standardized on the same national sample, (c) aptitude measures should be relevant to the subject field and the kind of educational intervention being considered, (d) comparable derived scores must be used when comparing aptitude and achievement tests, and (e) age and grade scores should not be mixed.

Discrepancies and the Battery

The Woodcock-Johnson Psycho-Educational Battery satisfies many of the Hanna et al. (1979) criteria. The derived scores are based on national norms. The Battery is standardized on the same sample and provides data on aptitude and achievement in four academic areas. The Battery's Achievement-Aptitude Profile provides a scale for determining discrepancies (Hessler, 1982; Woodcock, 1980). Since the Achievement-Aptitude Profile is based on actual data, regression error is not a problem.

One of the major advantages of the Battery is that it can help practitioners identify and quantify discrepancies between aptitude and achievement. For a moderate or severe discrepancy to exist, the cluster difference score must be -16 or lower which is most likely educationally significant, while scores below 20 can definitely be considered educationally significant (Hessler, 1982). Hessler commented that the Battery should certainly be viewed as "an appropriate instrument to use in the process of determining learning disabilities, particularly due to its convenient and accurate means of comparing aptitude and achievement" (p. 390). In summary, the Battery may prove useful to professionals needing to identify and quantify discrepancies between aptitude and achievement.

Performance of Learning Disabled Students on the Woodcock-Johnson Tests of Cognitive Ability and the Wechsler Intelligence Scale for Children-Revised

Since the Woodcock-Johnson Tests of Cognitive Ability (WJTCA) are intended to assess ability or intelligence, the majority of studies that have been conducted on the Battery focused on comparing the performance of learning disabled and normal subjects on the WJTCA and the Wechsler Intelligence Scale for Children-Revised (WISC-R). These comparisons were made primarily to assist in establishing the construct and concurrent validity of the instrument, and to determine if the WJTCA are appropriate for the assessment of learning disabled students.

Correlations between Full Scale Standard Scores

Table 1 provides data from several studies that have reported correlation coefficients between the Broad Cognitive Ability (BCA) of the WJTCA and the Full Scale score of the WISC-R. With the exception of the findings of Ysseldyke, Shinn, and Epps (1981), the correlations indicate a strong relationship between the two tests. Results of the Ysseldyke et al. study (1981) indicated a correlation of .67 between the WJCTA and the WISC-R Full Scale which was somewhat lower than the .79 found by Woodcock (1978) and Reeve et al. (1979). Ysseldyke et al. attributed the lower correlation to the lack of variance obtained from a more

Table 1. Reported Correlations between the Full Scale Standard Scores of the WISC-R and the WJTCA

Study	N	Grade/ C.A.	Correlation
Woodcock (1978)	83	Grade 3	.79
	86	Grade 5	.79
	73	Grade 12*	.83
Reeve, Hall & Zakreski (1979)	51	C.A. 7-2 to 11-5	.79
Ysseldyke, Shinn, & Epps (1981)	50	Grade 4	.67
Nisbet (1981)	28	Grades 1-6	.72
Ipsen, McMillan, & Fallen (1982)	60	C.A. 6-4 to 16-3	.85
McGrew (1983)	52	Grades 1-6	.74

*WAIS

restricted age range. In general, the correlations between the two tests in these samples are above .70, which is comparable to the reported .73 correlation between the WISC-R and the Stanford Binet (Wechsler, 1974).

Comparisons of Full Scale Standard Scores

The results of comparisons between the mean WISC-R Full Scale IQ and the mean WJTCA Broad Cognitive Ability score have been inconsistent. Table 2 illustrates the means, standard deviations, and mean differences reported by the studies. Results have ranged from a significant difference of 12.9 standard score points (Reeve, Hall, & Zakreski, 1979) to non-significant differences (McGrew, 1983; Nisbet, 1981).

In the first reported study, Reeve, Hall & Zakreski (1979) compared the performance of 51 learning disabled students on the WJTCA and the WISC-R. Subjects ranged in age from 7 years 2 months to 11 years 5 months with a mean age of 9 years, 3 months and an average intelligence quotient of 98.61 (SD = 12.91) on the WISC-R. In spite of the high correlation of .79 between the two tests, the sample fell close to the normative mean on the WISC-R, while their performance was significantly lower (12.9 standard score points) on the WJCTA.

Reeve et al. (1979) hypothesized two explanations for this discrepancy. Their first suggestion was that

Table 2. Reported Means, Standard Deviations and Mean Differences for the Full Scale Standard Score of the WISC-R and the WJTCA

	WISC-R			WJTCA		
Study	N	Mean	SD	Mean	SD	Mean Difference
Reeve, et.al. (1979)	51	98.61	12.91	85.69	13.05	12.9*
Ysseldyke, et.al. (1981)	50	100.04	12.48	92.36	11.37	7.68*
Nisbet (1981)	28	94.86	12.19	94.21	11.84	.65
Ipsen, et.al. (1982)	60	97.80	15.75	94.97	14.37	2.83
McGrew (1983)	52	96.80	13.10	97.10	11.60	- .30

*$p < .001$

children with learning problems may have greater difficulty than the standardization sample on the WJTCA because several of the subtests of the WJTCA may involve types of skills not assessed by the WISC-R, such as the subtests of the Perceptual Speed cluster. Also, they suggested that some of the norms or conversion tables of the WJTCA may be in error.

Similarly, Ysseldyke, Shinn, & Epps (1981) compared the performance of learning disabled students on the WJTCA and the WISC-R. Their sample was composed of 50 fourth grade students who had been identified as learning disabled six months prior to the study. The average intelligence quotient was 100.04 (SD = 12.45). In this study the performance of the learning disabled sample was 7.68 standard score points lower on the WJTCA than on the WISC-R Full Scale score.

In discussing the inferior performance of learning disabled students on the WJTCA, Ysseldyke et al. (1981) rejected Reeve et al.'s (1979) hypothesis that the difference might be due to performance on the Perceptual Speed cluster. Ysseldyke et al. concurred that errors in the norm tables may be responsible for the low performance or that learning disabled students may perform poorly in different areas. They postulated, however, that the primary reason for differential performance was due to a deficiency

in achievement rather than in cognitive functioning.

Ysseldyke et al. (1981) suggested that the results may be interpreted within the theoretical perceptions of tested intelligence proposed by Cattell (1963) and Newland (1971). Both Cattell and Newland claim that intelligence is composed of two separate types of ability: the processes necessary for the acquisition of knowledge and knowledge that has been previously acquired. Cattell calls these abilities "fluid" and "crystallized intelligence". Newland labels these abilities "process-dominant" and "product-dominant" intelligence.

Ysseldyke et al. (1981) hypothesized that the WJTCA taps crystallized or product-dominant intelligence, while the WISC-R taps fluid or process-dominant intelligence. They assumed that product-dominant intelligence is equivalent to achievement and suggested that the WJTCA are more a measure of achievement than the WISC-R. Additionally, the authors specified the subtests that seem to be the most achievement-oriented and, consequently, the most difficult for learning disabled subjects: Analogies, Numbers Reversed, Picture Vocabulary, Antonyms/Synonyms, Visual Matching, and Quantitative Concepts.

In a related study, Shinn, Algozzine, Marston, and Ysseldyke (1982) analyzed the subtest performance of the original learning disabled fourth grade population used by

Ysseldyke et al. (1981). They found that the learning disabled sample scored significantly lower than the norm group on Quantitative Concepts, Antonyms/Synonyms, Analogies, and Memory for Sentences or three of the six tests that were identified as achievement-oriented. Additionally, the learning disabled group scored higher than the norming sample on the Blending subtest. The authors concluded that the low performance of learning disabled students on the WJCTA is a result of the achievement-oriented subtests.

In an invited paper, Woodcock (1980) responded to some of the concerns that were being raised about the WJTCA. In discussing the results obtained by Ysseldyke et al. (1981) and by Reeve et al. (1979), Woodcock presented four errors which he felt invalidated their conclusions:

1. The authors used selected WISC-R scores. In both studies the WISC-R scores were obtained from students who were previously placed in programs. Presumably, these students were placed in learning disability programs because of high WISC-R scores and low achievement test scores.

2. They failed to obtain current WISC-R scores for purposes of comparison with the WJTCA.

3. They did not take into account the effects of explicit and/or implicit sample selection. Since the

samples are biased, a marked skewing in the statistics should be expected. Some difference in full scale scores can be attributed to the higher correlation of the WJTCA with school learning (approximately .78) than the WISC-R (.66).

4. **The investigators lacked objectivity in the** discussion of the results. Neither study presented the possibility that the WISC-R scores could be too high or that the observed difference could be valid.

Also, Woodcock reminded users that the appropriate score for identifying learning disabled students is the relevant Scholastic Aptitude score, rather than the Broad Cognitive Ability score.

McGrew (1983) tried to rectify many of the concerns raised by Woodcock. He compared the performance of 52 elementary students on the WISC-R and the WJTCA. The subjects had been referred to a clinic and, consequently, were not part of a preselected sample. McGrew found that the mean differences between the WISC-R and the WJTCA were not significantly different for the total population or for a subsample of thirty-one learning disabled students. McGrew concluded that the two tests generally provide comparable measures of aptitude and that results of previous research may be an artifact of sample preselection. He also noted that despite the lack of significant mean differences

for the total population, a few individual students evidenced marked discrepancies between the two instruments. In conclusion, McGrew recommmended that further research be conducted to discover the diagnostic importance of these individual WISC-R/WJCTA full scale score differences.

Results of another recent study also supported the concurrent and construct validity of the WJTCA. Ipsen, McMillan, and Fallen (1982) analyzed the subtest performance of normal, learning disabled, and emotionally disturbed children on the WJTCA and the WISC-R. The sample consisted of 60 students ranging in age from six years, four months to sixteen years, three months, who had been referred to a clinic to ascertain if a handicapping condition existed. Of the 60 students, 7 children were already receiving special education, 27 were identified as learning disabled, 19 emotionally disturbed, 10 normal, and 4 "other". For all of the groups, their Broad Cognitive Ability scores were about three points lower than their WISC-R Full Scale scores, a much smaller difference than found by Reeve et al. (1979) and Ysseldyke et al. (1981). Additionally, the learning disabled subjects scored lowest on Spatial Relations, Visual Matching, Quantitative Concepts, and Analogies. This finding did not support Ysseldyke et al.'s (1981) hypothesis that learning disabled subjects would score lowest on the achievement-oriented subtests.

Nisbet (1981) also failed to find significant differences between the full scale scores of the two tests. He examined the test scores of 28 elementary students who had been referred to child study teams for assessment and had been administered both the WISC-R and the WJTCA. Furthermore, subtest results did not substantiate the Ysseldyke et al. (1981) claim of a difference between product-dominant and process-dominant intelligence.

In summary, the majority of research conducted, has explored the concurrent and construct validity of the WJTCA and the WISC-R. The results of these studies have been contradictory and inconclusive. Earlier studies (Reeve et al,, 1979; Ysseldyke et al., 1981) found that learning disabled students scored considerably lower on the WJTCA than the WISC-R. Later studies (McGrew, 1983; Nisbet, 1981) attempted to rectify methodological problems and did not find significant differences between the full scale scores.

Performance of Learning Disabled Students on the Woodcock-Johnson Psycho-Educational Battery

The Woodcock-Johnson Psycho-Educational Battery (Battery) contains three sets of clusters: (1) the Cognitive clusters including Verbal Ability, Reasoning, Perceptual Speed, and Memory, (2) the Scholastic Aptitude clusters

including Reading Aptitude, Math Aptitude, Written Language Aptitude, and Knowledge Aptitude, and (3) the Achievement clusters including Reading, Mathematics, Written Language, and Knowledge. Recently, Woodcock (1983) has provided the alternative Cognitive clusters of Oral Language and Broad Reasoning that may be substituted for Verbal Ability and Reasoning.

Cognitive Clusters

Pattern of Performance. Several investigators have found that learning disabled students display a unique pattern of performance on the Cognitive clusters, although findings have been inconsistent. Two studies reported that learning disabled students score lowest on the Perceptual Speed cluster (Ipsen, McMillan, & Fallen, 1982; Reeve, Hall, & Zakreski, 1979). Other studies did not confirm this finding, but reported that learning disabled students scored highest on the Reasoning cluster (McGrew, 1983; Ysseldyke, Shinn, McGue, & Epps, 1980).

Validity of Cognitive Clusters. McGue, Shinn, and Ysseldyke (1982) have questioned the validity of the cluster scores for use with learning disabled students. In factor analyzing the 12 Cognitive Ability Tests using Ysseldyke et. al's (1981) sample of fourth graders, they found that only the Verbal factor was verified and, consequently, suggested that the factor structure for learning disabled

students is different from the structure for normal individuals, but cautioned that their results may be an artifact of sample size and restricted age range.

McGue et al. (1982) conjectured that this absence of validity may be caused by three problems in the WJTCA: 1) the overlap of subtest content, 2) the use of differential weighting of complex variables, and 3) the operationalization of the Aptitude clusters. They further suggested that the use of suppressor weights may be inappropriate for a learning disabled population: "The use of suppression weights to control for the operation of an irrelevant factor in the prediction of some criterion implicitly assumes that factorial invariance holds. Thus, with special populations, which may be characterized by unique factor structures, the use of suppressor weights could prove invalid" (p. 281).

Hessler (1982) specified two situations where the suppressor subtests in the Verbal Ability and Reasoning clusters would significantly alter, and perhaps invalidate, the test scores: 1) when a subject scores high or low on all of the subtests of the cluster, or 2) when the suppressor subtest is much higher or lower than the main subtests of the cluster.

Alternative Cognitive Clusters

Woodcock (1983) responded to these concerns by creating the alternative clusters of Oral Language and Broad Reasoning that may be substituted for Verbal Ability and Reasoning. In these alternative clusters, the suppressor subtests have been eliminated and all of the subtests have been provided equal weights. The Oral Language cluster includes Picture Vocabulary, Antonyms-Synonyms, and Analogies and the Broad Reasoning cluster includes Analysis-Synthesis, Concept Formation, and Analogies.

Aptitude Clusters

The purpose of the Scholastic Aptitude clusters is to provide an estimate of a student's expected performance in the four achievement areas. Each cluster is a weighted combination of four Cognitive subtests that best predicted performance on the Achievement clusters. These predictor measures do not contain the same content as the area of achievement. For example, the Reading Aptitude cluster is highly correlated with reading achievement, but does not require any reading. In theory, the clusters should provide a more accurate measure of potential ability of students with learning problems (Hessler, 1982).

McGue et al. (1982) found little support for the Aptitude clusters when used with learning disabled students. They questioned (a) the operationalization of

the aptitude construct, (b) the use of weighted combinations to predict achievement, (c) the degree of overlap and interdependence among the clusters, and (d) the differential predictive ability of the clusters. They concluded that little evidence supports the validity of these clusters for assessment of the learning disabled.

In reviewing their research, Hessler (1982) noted their concerns, but countered that "the aptitude clusters are the best single cognitive predictors of achievement available at the present time" (p. 78). The most common procedure for estimating ability has been to take a single measure of intelligence, such as the WISC-R, and use it as the level of academic expectancy in all areas. The Scholastic Aptitude clusters appear to introduce more specificity and greater precision into the diagnostic process.

Achievement Clusters

McGue et. al (1982) found that the achievement clusters had high correlations with similar measures and, consequently, demonstrated validity for the assessment of learning disabled students. No overlap in content exists between the clusters, and each subtest has an equal weight.

In assessing handicapped students with the Achievement clusters, Hessler (1982) identified four major profiles: (a) deficient performance on the Reading and Written Language clusters and low average or above on the

Mathematics and Knowledge clusters, (b) deficient performance solely on the Mathematics cluster, (c) deficient performance on all clusters but Knowledge, and (d) deficient performance on all four clusters. He also noted that learning disabled students often perform commensurate with ability on the Knowledge cluster, while evidencing difficulty in reading, writing, and mathematics.

Subtest Scatter on the WJTCA

Finally, little emphasis has been given to examination of subtest scatter on the WJTCA, although this has been a common procedure with the WISC-R. Since psychologists frequently use subtest scatter in interpretation, Marston (1980) analyzed the scatter among the subtests of the WJTCA for the standardization sample. Similar to Kaufman's (1976) findings for the WISC-R, Marston (1980) found that a large proportion of the standardization sample had significant subtest scatter. Since significant scatter is found among normal students, Marston (1980) recommended that the practice of scatter analysis on the WJTCA should be discouraged and not used to validate the diagnosis of learning disabilities.

Summary

Present research has not verified a specific pattern of performance for learning disabled students on the

Cognitive clusters. Several investigators (Hessler, 1982; McGue, Shinn, & Ysseldyke, 1982) have questioned the validity of the Cognitive clusters. Woodcock (1983) has responded to these concerns by creating the alternative Cognitive clusters of Oral Language and Broad Reasoning. Additionally, McGue et al. (1982) questioned the validity of the Aptitude clusters; Hessler (1982), however, suggested that the Aptitude clusters are the best predictors of achievement that are available. On the Achievement clusters, McGue et al. (1982) indicated that the clusters have high correlations with other commonly used achievement measures and appear quite valid for use with learning disabled subjects. Additionally, several profiles have been identified on the Achievement clusters (Hessler, 1982). Finally, the practice of examining subtest scatter on the WJTCA has been discouraged (Marston, 1980).

Performance of Learning Disabled Students on the WISC-R

The WISC-R is commonly used as an assessment instrument for the diagnosis of learning disabilities. Many studies exist that have explored whether or not learning disabled students exhibited (a) Verbal-Performance score discrepancies, (b) more subtest scatter, or (c) a different subtest profile than normal learners. The following examination of the literature is representative of the types of studies that have been conducted and the types of

conclusions that have been made.

Recently, Dudley-Marling, Kaufman, and Tarver (1981) reviewed twenty-four studies that investigated the performance of learning disabled students on the WISC and the WISC-R. They attempted to summarize whether or not learning disabled students exhibited significant Verbal-Performance score discrepancies, subtest scatter, and a characteristic WISC-R profile. They found that the literature was contradictory regarding Verbal-Performance score discrepancies and subtest scatter, but that the majority of studies indicated that learning disabled students (as a group) conformed to a pattern of performance on the subtests. They concluded, however, that subtest profiles may not be useful for the differential diagnosis of learning disabled students, as few individual children actually display this pattern.

Verbal-Performance Score Discrepancies and Subtest Scatter

Researchers have found that patterns of high Performance IQ-low Verbal IQ and subtest scatter are common characteristics of learning disabled students. Many of these studies, however, had methodological limitations. Zingale and Smith (1978) found that Performance IQ was greater than Verbal IQ for a group of 122 elementary-aged learning disabled students. This pattern was common among 40% of their group. Also, these results held across

socio-economic groups. They did not, however, have a normal control group.

Gajar (1979) compared the performance of three groups of students who were identified or placed in special education classes. The subjects included 122 emotionally disturbed, 135 learning disabled, and 121 educable mentally retarded students. After examining variances, she found that the learning disabled students evidenced greater subtest scatter than the other subjects. Unfortunately, Gajar did not report the age of her subjects or compare her populations to a normal control group.

Tabachnick (1979) made comparisons between a sample of normal achieving students and a sample of learning disabled students. The study compared the performance of 105 learning disabled children from two private, non-profit agencies to Kaufman's (1976) normal population. Subjects ranged between the ages of 6-9 to 16-5. By comparing the means for the range of scaled scores on the WISC-R subtests, Tabachnick (1979) found that subtest scatter was significantly greater for the learning disabled group. Tabachnick (1979) also noted that several of the learning disabled students exhibited minimal scatter or very flat profiles, making it difficult to differentiate these students from those in the normal population. She speculated that high scatter may be characteristic of a

particular type of learning disabled student who might respond to a certain type of educational therapy.

Ryckman (1981) found subtest variability in a sample of learning disabled students. His sample included 100 elementary-aged children in self-contained learning disabilities programs who were compared to the WISC-R standardization sample. Like Tabachnick (1979), he also found that some learning disabled children exhibited flat profiles.

Few studies have investigated the performance of students with high IQs who evidence severe learning problems (Faigel, 1983; Schiff & Kaufman, 1981). Schiff and Kaufman (1981) attempted to determine if learning disabled children with superior intelligence would exhibit significant Verbal-Performance score discrepancies and a significant amount of subtest scatter. The subjects included 30 students between the ages of six and thirteen with a mean Full Scale IQ of 123 (SD = 7.9). The means obtained from the sample were compared to the normal standardization sample. They found that 43% of the sample had scaled score ranges of 10 or more points and that the mean Verbal-Performance score discrepancy was 18.6 with 87% having higher Verbal IQs than Performance IQs. The authors suggested that students at higher IQ levels may exhibit more exaggerated discrepancies.

Verbal-Performance Score Discrepancies and Subtest Scatter in Learning Disabled and Normal Samples

Analysis of the literature reveals that, although many studies have found Verbal-Performance score differences and subtest scatter in learning disabled samples, congruence exists with normal samples. For example, Ackerman, Peters, and Dykman (1971) compared 82 learning disabled students between the ages of 8 and 11 with 34 controls. Although they found that Verbal-Performance score discrepancies existed in the learning disabled sample, the discrepancies were as common among the normal controls.

Anderson, Kaufman, & Kaufman (1976) found significant Verbal-Performance score discrepancies for a sample of 41 learning disabled students. However, when the ranges for the LD sample were compared to data on normal children, no significant differences existed. Additionally, no difference existed in the amount of subtest scatter obtained by the LD group and normal children.

Kaufman (1976) demonstrated that subtest scatter is common in normal populations by investigating the amount of test scatter in the standardization sample of 2200 subjects. He analyzed both scaled score ranges and the number of tests deviating from a subject's mean. Kaufman observed that a great deal of scatter exists in the standardization sample and that normal children are not

characterized by flat profiles. Kaufman created new normative tables to assist clinicians in assessing scatter objectively. These tables take into account the amount of discrepancy that can be expected in the normal population.

Similarly, Moore and Wielan (1981) compared the scatter indices of 434 subjects referred for a reading disability to the standardization sample. Subjects ranged in age from 6 to 16. The total sample had a mean discrepancy of 11.2 points. They found that 41% of the reading-referred students evidenced Verbal-Performance score discrepancies of 12 points or greater. In the majority of these cases, Performance IQs exceeded Verbal IQs. Results of the standardization sample were quite similar; this group had a slightly lower mean discrepancy of 9.7 points. Additionally, the reading-referred children evidenced as much subtest scatter as the standardization sample.

Gutkin (1979) also concurred that WISC-R scatter is present in normal children and that an analysis of scatter does not provide useful information for determining the presence of a psychoeducational disability. He analyzed the performance of 101 special education students, including 17 emotionally disturbed, 51 learning disabled, 23 minimally brain injured, and 10 educable mentally retarded. Scatter was calculated by subtracting the lowest from the highest subtest scores for the Full Scale, the Verbal Scale, and the

Performance Scale. Gutkin (1979) reported that "40% of normal children either equalled or exceeded the average of the special education students on all three of these scatter indices" (p. 370).

In summary, several recent studies indicate that learning disabled students do exhibit significant Verbal-Performance score discrepancies and subtest scatter on the WISC-R (Gajar, 1979; Ryckman, 1981; Tabachnick, 1979). Other studies report that learning disabled students do exhibit Verbal-Performance score discrepancies and subtest scatter on the WISC-R, but that these characteristics are also found in the normal population and do not seem to differentiate the learning disabled student from the normal student (Anderson et al., 1976; Gutkin, 1979; Kaufman, 1976; Moore & Wielan, 1981).

Subtest Profile

Attempts have been made to discover if learning disabled students exhibit a characteristic subtest profile on the WISC-R. In examining the "uses and abuses of the WISC-R with the learning disabled", Galvin (1981) discouraged the search for a unique group profile and concluded that a global learning disabilities pattern will not identify all learning disabled individuals. Instead, she proposed that there may be useful subpatterns that are related to a specific type of learning disability.

Ryckman (1981) concurred, stating that "the time is right to discontinue the search for a single profile that discriminates between learning disabled and normal children. The heterogeneity of the LD groups precludes the likelihood of finding common characteristics" (p. 511). Several investigators (Bannatyne, 1968; Ryckman, & Elrod, 1983; Vance, Wallbrown, & Blaha, 1978) have proposed distinct subgroups or syndromes of learning disabled children. These patterns will not be detected in a group of children.

In summary, the literature suggests that a single WISC-R profile or a requisite amount of subtest scatter will not differentiate a heterogenous sample of learning disabled students from other students. Research and regression effects suggest that students classfied as learning disabled are more likely to score higher on the Performance Scale than the Verbal Scale of the WISC-R. Conversely, students classified as gifted and talented might demonstrate an inverse relationship, scoring significantly higher on the Verbal Scale. This study examined the magnitude of Verbal-Performance score discrepancies for both the subjects classified as learning disabled and the subjects classified as gifted and talented to explore this possibility.

Performance of Gifted Students on Intelligence Tests

For more than twenty years investigators have attempted to determine the most effective method for identifying gifted students. For example, Pegnato and Birch (1959) attempted to identify the most efficient and effective procedure for locating gifted junior high school students. They examined seven different methods for identifying mental giftedness (intelligence quotient of 136 or above). Of 781 children selected by the various screening methods, 91 had Stanford-Binet's (Terman & Merrill, 1960) above the criterion. The investigators drew several conclusions from their study: (a) teachers are not accurate in their selection of gifted students, (b) mathematics achievement is a poor predictor of giftedness, and (c) group intelligence tests cannot be relied upon for accurate identification. The investigators concluded further that individual intelligence tests should be used for the final identification of gifted students, even though they are time-consuming and expensive.

Other studies have suggested that commonly used screening instruments do not accurately identify superior students. For example, Rust and Lose (1980) investigated the accuracy of the Slosson Intelligence Test (Slosson, 1975) , a group screening device, for predicting Full Scale WISC-R scores for 132 gifted students in second through

eighth grade. They found a .43 correlation between the Slosson Intelligence Test and the WISC-R. The addition of scores from the Scales for Rating the Behavioral Characteristics of Superior Students (Renzulli, Hartman, Callahan, 1971) did not add to the Slosson's ability to predict WISC-R scores. They concluded that the Scales for Rating the Behavioral Characteristics of Superior Students were not an aid in identifying students with high intelligence and that considerable error can be expected when using these two screening instruments.

In summary, many researchers have found that group intelligence tests are not effective screening devices (Baldwin, 1962; Cornish, 1968; Gallagher & Rogge, 1966; Hunter & Lowe, 1978) and that teacher nomination misses many gifted students (Baldwin, 1962; Cornish, 1968; Jacobs, 1970). Investigators also agree that individually administered intelligence tests, such as the WISC-R, represent the most accurate and preferred method for identifying gifted children (Clark, 1979; French, 1974; Hunter & Lowe, 1978; Kirk, 1972; Marland, 1971; Martinson, 1973; Sattler, 1982).

Patterns of Performance

Many studies involving gifted students were performed during the late fifties and sixties (Whitmore, 1980); however, little research was conducted on the

intellectual patterns of gifted children on the WISC (Littell, 1960; Lucito & Gallagher, 1960). For example, in the standardization sample, 565 students were in the average range (Full Scale 90-110) with an age level between 7 and 11, while only 15 students had intelligence quotients above 130 (Gallagher & Lucito, 1961).

The research that was conducted, suggested that gifted students score higher on the Verbal Scale than the Performance Scale of the WISC. This strength was attributed to a Verbal-Comprehension factor (Gallagher & Lucito, 1961; Thompson & Finley, 1962). For example, Lucito and Gallagher (1960) analyzed intellectual patterns on the subtests of the WISC of 50 superior students who had scored above 150 with a mean intelligence quotient of 160.8 on the Stanford-Binet (Terman & Merrill, 1960). They found that the students scored highest on the subtests of Similarities, Block Design, Information, and Vocabulary and lowest on Picture Arrangement and Picture Completion. They observed that the strengths of these students seemed to be in their abilities to store, associate, and assimilate information.

Using this same sample of superior students, Gallagher and Lucito (1961) compared intellectual patterns on the WISC to the performance of average and mentally retarded samples from previous research. Different patterns of intellect were found between the gifted and retarded

students. The greatest strength of the gifted was their verbal performance or ability to recall past information and associate concepts, while their greatest weakness was ability to use structured concrete visual materials. Opposite findings were observed for the retarded samples. Thompson and Finley (1962) found similar results using protocols on 400 gifted children and 309 mentally retarded children.

More recently, a few investigators have analyzed the performance of gifted underachievers on individual intelligence tests. Bush and Mattson (1973) compared the WISC test patterns of 28 gifted underachievers, 23 gifted-achievers, 36 average underachievers, and 22 normal achievers. They found that the Arithmetic and Digit Span subtests differentiated between the gifted underachievers and the gifted achievers, while the Coding subtest was somewhat low for both groups. In summarizing the research, Whitmore (1980) observed that gifted underachievers may produce widely discrepant results, evidencing very superior ability in abstract reasoning and verbal ability, but deficits in arithmetic and visual-motor skills.

In summary, one common factor among gifted students seems to be their ability to understand, organize, and apply abstract concepts (Gallagher, 1975). Conversely, gifted students seem to score lower on tasks involving arithmetic

computation and visual-motor skills (Whitmore, 1980).

Utility of the WISC-R

The WISC-R is commonly used to measure intellectual giftedness (Gallagher, 1975), and the factor structure appears to be valid with this special group. Karnes and Brown (1979) examined the factor structure of the WISC-R for a group of 946 gifted students, ranging in age from 6 to 16. All subjects had an IQ of 120 or above. They concluded that the factor structure of the WISC-R was similar to other groups studied and that the Verbal and Performance Scales appear to be stable across special groups.

One problem that has been noted in using the WISC-R with superior students is that it does not encompass extremes in intelligence because of a low ceiling (Martinson, 1975; Rubenzer, 1976). One alternative is to employ extrapolated scores. Sattler (1982) suggests that caution be used when employing extrapolated IQs that are derived from linear regression methods. Using a regression equation, Ogdon (1975) has developed extrapolated IQs for students with scaled scores above 160 but warns the user that "relationships that are generally linear may depart from linearity at the extremes of the distribution, and that the further one extrapolates from empirical data, the greater the probability of error" (p. 216).

In spite of this limitation, the WISC-R is an

important instrument and one of the best tests for the identification of gifted students. No research could be found that examined the performance of gifted students on the Woodcock-Johnson Psycho-Educational Battery. Conceivably, the Woodcock-Johnson Tests of Cognitive Ability may also prove to be a sensitive instrument for evaluating the abstract, logical, and integrative abilities of the gifted student.

CHAPTER 3

METHOD

The purpose of this chapter is to describe the (a) selection of subjects, (b) tests used, (c) procedures for data collection and analyses, and (d) methodological assumptions and limitations.

Description of the Subjects

Subjects from this study were selected from the Tucson Unified School District, the largest school district in the state of Arizona. Of the 53,500 students served, approximately 5.5% or 3,000 are enrolled in programs for the learning disabled. Additionally, the district has established both self-contained and resource programs for gifted and talented children that serve approximately 1,240 students.

Subjects Classified as Learning Disabled

Using a random number table, a total of 51 elementary-aged subjects, classified by the school district as learning disabled, were selected from a pool of 200 students. The sample ranged in age from 7-3 to 11-0 with a

mean chronological age of 9-1 and a standard deviation of 1.15. To preserve some semblance of the total learning disability population of the district, the variables of sex and placement (resource or self-contained) were determined by the randomization process. The final sample contained 37 boys and 14 girls with 12 subjects in self-contained classrooms and 39 students in resource programs. Minority representation included 5 Hispanic students and 1 Black student. All subjects had English as their primary language as determined by school district documentation and had been enrolled in a learning disability program for at least one year. Parental permission for testing had been obtained prior to entry into a learning disability program.

Placement in a learning disability program is a district decision that is made by a multidisciplinary team. The criteria used by the Tucson Unified School District (1982) are presented below:

> In order for a student to be considered as Learning Disabled and eligible for adaptive education services, the following criteria will be considered according to the severity, significance, and persistence of the presenting problem(s) as related to:
>
> 1) discrepancy between intellectual ability and actual achievement (performance)
> 2) information processing abilities and disabilities
> 3) elimination of exclusionary factors
> 4) the need for special education services that are required because the student cannot learn through ordinary methods of instruction. (This need will be validated by the Child Study Team who have

> already assisted in designing instructional adaptations for the regular classroom) (p. 54).

Evidence of a severe discrepancy is determined by comparing a student's intelligence and achievement. Intelligence is determined by an individual intelligence test, primarily the WISC-R, and the student's full scale score must not be more than 2 standard deviations below the mean. A student's academic functioning is then compared to the student's intellectual potential. A Table of Severe Discrepancy (p. 59) is used as a guideline to adjust achievement scores to various intelligence quotients and determine severe discrepancies. For example, a 10 year old student with a Full Scale IQ score of 80 would have to obtain a grade score below 1.8 before a discrepancy would be considered severe. A 10 year old student with a Full Scale IQ score of 100 would have to obtain a grade score below 2.5, while a student with an IQ score of 120 would have to fall below a grade score of 3.2.

Identification of a deficit in one or more of the psychological processes is based on a student's patterns of abilities and disabilities in the learning processes of perception, integration, and expression. The most commonly used tests to assess psychological processes are the Woodcock Johnson Tests of Cognitive Ability (1977), the Detroit Tests of Learning Aptitude (Baker & Leland, 1967), and the Illinois Test of Psycholinguistic Abilities (Kirk, McCarthy, & Kirk, 1968. A score of two standard

deviations below the student's mean in one area or 1.5 standard deviations below in three or more areas, is used to determine a severe processing deficit. A deficit may occur in any of the following basic psychological processes: visual, auditory, haptic, receptive or expressive language, or sensory integration.

In addition to the identification of a discrepancy and a processing deficit, the student must display a demonstrated history of school failure and difficulty obtaining expected achievement. The student must present both strengths and weaknesses in academic areas. Termination of services must be considered when a student has progressed to within 80-85% of potential or expected academic achievement. If deficits in psychological processes are remediated but academic achievement remains low, placement in a remedial education program is deemed more appropriate than placement in a learning disability program.

Subjects Classified as Gifted and Talented

All subjects were selected from a district elementary program that has self-contained classrooms for the gifted and talented. Permission slips were sent to the parents of 100 students in second through fifth grade. Eighty-one parents agreed to let their children participate and 46 subjects were selected from this pool using a random number table.

Distribution of age and sex of the subjects was determined by the randomization process. The final sample contained 26 boys and 20 girls, ranging in age from 7-3 to 12-0 with a mean chronological age of 8-8 and a standard deviation of 1.06. Minority representation included 2 Blacks, 1 Asian, and 1 Hispanic student. All subjects had English as their primary language and had been enrolled in the program for at least one year.

Placement in a Gifted and Talented Education program is a district decision that is based upon both objective and subjective information. The classroom teacher completes Scales for Rating the Behavioral Characteristics of Superior Students (Renzulli, Smith, White, Callahan, & Hartman, 1976) which includes the categories of Learning, Motivation, Creativity, and Leadership. Other recommendations are submitted from another school member, such as the librarian, principal, or another classroom teacher and the student's parents. Additionally, psychometric data are obtained on achievement using the California Achievement Test and on cognitive ability using the Otis Lennon School Abilities Test, a group measure of reasoning and predictor of future academic success, or an individual intelligence test, if available. All of these factors are listed on a matrix and assigned a point value. Placement in the program is determined by a composite score above 35. Since

identification is not based solely on intelligence, students enrolled in these programs can have a broad range of both cognitive and academic abilities.

Description of the Instruments

The assessment instruments used in this study included the Woodcock-Johnson Tests of Cognitive Ability (WJTCA), the Woodcock-Johnson Tests of Achievement (WJTA), and the Wechsler Intelligence Scale for Children-Revised (WISC-R). Investigators have suggested that these assessment devices have adequate norms, reliability, and validity (Thurlow & Ysseldyke, 1979).

Woodcock-Johnson Psycho-Educational Battery

The Woodcock-Johnson Psycho-Educational Battery is composed of three parts: Tests of Cognitive Ability (WJTCA), Tests of Achievement, (WJTA) and the Tests of Interest Level. The WJTCA and the WJTA were used in this study.

Cognitive Clusters. The WJTCA contain the following twelve subtests: Picture Vocabulary, Spatial Relations, Memory for Sentences, Visual-Auditory Learning, Blending, Quantitative Concepts, Visual Matching, Antonyms-Synonyms, Analysis-Synthesis, Numbers Reversed, Concept Formation, and Analogies. Raw scores are obtained for each subtest and combined to obtain the following cluster scores: the Broad Cognitive Ability or full scale score, the four Cognitive

Factors - Verbal Ability, Reasoning, Perceptual Speed, and Memory, and the four Aptitude cluster scores - Reading, Mathematics, Written Language, and Knowledge Aptitude.

For the Broad Cognitive Ability score, weights based on principal component analysis were assigned to the subtests. Woodcock (1978) felt that the relative weighting of the subtests would provide a better estimate of cognitive ability than the WISC-R where each subtest has an equal weight. The Cognitive clusters were determined by multiple regression and factor analysis. Two of the Cognitive clusters, Verbal Ability and Reasoning contain subtests with negative weights that appear to help control for an irrelevant factor. In the Verbal Ability cluster, the Analysis-Synthesis subtest is supposed to counterbalance the effect of reasoning in the Antonyms-Synonyms subtest. In the Reasoning cluster, the Antonyms-Synonyms is supposed to counterbalance the effect of verbal ability in the Analogies subtest.

Aptitude Clusters. The Scholastic Aptitude clusters provide a measure of a student's expected achievement in four areas: Reading, Mathematics, Written Language, and Knowledge. The clusters were developed using step-wise multiple regression. Subtests with common content were excluded. The four subtests that best predicted achievement in the relevant area were used.

Achievement Clusters. The WJTA contain ten subtests: Letter-Word Identification, Word Attack, Passage Comprehension, Calculation, Applied Problems, Dictation, Proofing, Science, Social Studies, and Humanities. The derived weights for the subtests are equal. Achievement cluster scores may be determined for Reading, Mathematics, Written Language, and Knowledge.

Scores. Scores from the Battery may be converted into many indices of performance: part scores, grade score equivalents, age score equivalents, cluster scores, cluster difference scores, instructional range scores, expected achievement scores, percentile rank scores and standard scores. Additionally, a Relative Performance Index provides an estimation of a student's performance on tasks similar to the one tested and a functioning level assigns a verbal interpretation to the subject's performance. These functioning levels include: Very Superior, Superior, Above Average, Average, Below Average, Moderate Deficit, and Severe Deficit.

Achievement-Aptitude Profile. Additional information is provided by the Achievement-Aptitude Profile which compares a subject's expected achievement (aptitude scores) and actual achievement (achievement scores). The amount of discrepancy is classified by a functioning level that may range from "Very Superior", indicating a student who is

performing significantly above expectancy level, to "Severe Deficit", indicating a student who is performing significantly below expectancy. This Achievement-Aptitude Profile is the most relevant score for identifying students with learning problems (Woodcock, 1978).

Standardization. Normative data for the Battery were collected for 4,732 subjects. The majority of the sample (N=3900) consisted of students in kindergarten through grade 12. The subjects were selected from 49 communities throughout the United States. Results of the 1970 Census Data were used to distribute subjects on the variables of sex, race, occupational status, geographic region, and type of community to make the norming sample as representative of the United States population as possible. In addition, Woodcock (1978) reported that an individual subject weighting procedure was employed during data analysis to ensure that the subjects were distributed proportionally. Severely handicapped students, or subjects who had less than one year in an English speaking environment, were not included in the sample.

Reliability. In determining subtest reliability, the test-retest technique was used for the two timed or speeded subtests on the WJTCA. Subjects were re-tested one day after the first test had been administered. Visual Matching had a median coefficient of .65 while Spatial Relations had a

median coefficient of .86. The remaining subtest reliability coefficients were calculated by using the split-half procedure and applying the Spearman-Brown formula to correct for length. The median reliability coefficients on the WJTCA ranged from a .80 for Memory for Sentences to a .95 for Visual-Auditory Learning. Median reliability coefficients on the WJTA ranged from .83 on Social Studies to a .95 on Letter-Word Identification.

Median cluster reliabilities on the WJTCA ranged from a .70 on Perceptual Speed to a .90 on Verbal Ability. Median cluster coefficients on the Aptitude clusters ranged from .89 on Mathematics Aptitude to .95 on Reading and Knowledge Aptitude. Median cluster coefficients on the WJTA ranged from .92 on the Mathematics cluster to .96 on the Reading cluster. The Broad Cognitive Ability score (full scale) had a median reliability of .97.

Validity. In establishing the concurrent validity of the Battery, Woodcock (1978) used extremely diverse samples. For example, studies were conducted that examined the performance of preschoolers, third, fifth, and twelfth-grade students, severely learning disabled students, students with severe behavior problems, and trainable mentally retarded students.

In addition, the Battery was correlated with many widely used assessment devices. The Broad Cognitive Ability

score had reported correlations of .79 with the WISC-R Full Scale score, .76 with the WISC-R Verbal scale, and from .58 to .62 with the Performance scale for the elementary school-age population (Woodcock, 1978). Concurrent validity for the WJTA was obtained was obtained by making comparisons with the Peabody Individual Achievement Test (PIAT), the Wide-Range Achievement Test (WRAT), the Key Math Diagnostic Arithmetic Test (Connolly, Nachtman, Pritchett, 1971) and the Iowa Test of Basic Skills (Lindquist & Hieronymus, 1956). Correlations on the reading portions of the tests ranged from .78 to .92, on math from .80 to .82, and on written language from .74 to .84. The Knowledge clusters had correlations between .68 to .83 with the General Information subtest of the PIAT.

Two predictive validity studies were conducted for a kindergarten and first grade sample that correlated cognitive ability scores at the end of kindergarten with achievement scores at the end of first grade. For the kindergarten sample, correlations with the Broad Cognitive Ability score were .54 for Mathematics, .67 for Reading, .68 for Written Language, and .75 for Knowledge. For the first grade sample, correlations were .60 for Mathematics, .63 for Written Language, .67 for Reading, and .75 for Knowledge.

Wechsler Intelligence Scale for Children-Revised

The WISC-R is a widely used test of general intelligence. It consists of twelve subtests that are

categorized as Verbal or Performance. The Verbal Scale includes the subtests of Information, Similarities, Arithmetic, Vocabulary, Comprehension, and Digit Span (optional). The Performance Scale includes the subtests of Picture Completion, Picture Arrangement, Block Design, Object Assembly, Coding, and Mazes (optional).

Scores. Raw scores are obtained for each subtest and then converted to scaled scores that are appropriate for the age of the students. The scaled scores have a normal distribution with a mean of 10 and a standard deviation of 3. The scaled scores of the Verbal and Performance subtests are added to yield Verbal, Performance, and Full Scale scores. These scaled scores are converted to standard scores or intelligence quotients with a mean of 100 and a standard deviation of 15.

Standardization. The WISC-R was standardized on approximately 2,200 students who ranged in age from 6 years through 16 years, 11 months. A stratified sampling plan was used to select subjects in accordance with 1970 United States Census Data. The variables used in selecting the sample included: age, sex, race, geographic region, occupation of head of household, and urban-rural residence. The sample did not include institutionalized children, children with severe emotional problems, or children with limited English.

Reliability. With the exception of the Digit Span and

Coding, reliability of the subtests was obtained by the split-half procedure, corrected for length by the Spearman-Brown formula. Reliability coefficients for Digit Span and Coding were determined by the test-retest procedure. Overall, reliabilities ranged from .77 to .86 on the Verbal tests and from .70 to .85 on the Performance tests. The average coefficients across the entire age range were .94 for the Verbal IQ, .90 for the Performance IQ, and .96 for the Full Scale IQ. The average standard error of measurement was 3.60 for the Verbal IQ, 4.66 for the Performance IQ, and 3.19 for the Full Scale IQ.

The stability of the WISC-R was assessed by re-testing 303 children from six age groups after a one month interval. Obtained coefficients were then corrected for the average variability of the normative group. The corrected stability coefficients were .93 for the Verbal IQ, .90 for the Performance IQ, and .95 for the Full Scale IQ. On the second testing, subjects made average gains of about 3.5 on the Verbal Scale, 9.5 on the Performance Scale, and 7 points on the Full Scale. Wechsler (1974) attributed these differences to practice effects.

Validity. Validity was established by correlating the WISC-R with other measures of intelligence: the Wechsler Preschool and Primary Scale of Intelligence (WPPSI), the Wechsler Adult Intelligence Scale (WAIS), and the Stanford-Binet

Intelligence Scale (Terman & Merrill, 1973). The correlations between the WISC-R and WPPSI were .80 for the Verbal Scale, .80 for the Performance Scale, and .82 for the Full Scale. The correlations between the WISC-R and the WAIS were .95 for the Full Scale IQs, .96 for the Verbal IQs, and .83 for the Performance IQs. The average correlational coefficients for the WISC-R and Stanford-Binet were .71 for the Verbal Scales, .60 for the Performance Scales, and .73 for the Full Scale IQs. Wechsler (1974) reported that a comparison of mean full scale differences revealed that scores on the Stanford-Binet were approximately 2 points higher than the WISC-R at ages 6, 9.6, and 12.6. These small differences suggest that the two tests provide similar intelligence quotients for normal children.

Procedures

Examiners

Administration of the Woodcock-Johnson Psycho-Educational Battery was performed by four examiners. Each examiner possessed the following qualifications: (a) completion of an advanced degree (master's or doctorate) in Special Education or Educational Psychology and, (b) evidence of specific training in the test instruments (coursework and/or inservice training). All examiners were observed by the investigator to ensure that testing was

performed according to the standard procedures outlined in the manuals.

All WISC-R testing was performed by the investigator, who has completed advanced training in this instrument, and two state-certified school psychologists.

Test Administration and Scoring

Administration of the Battery and the WISC-R were counterbalanced to reduce order effects. All testing was performed within a three week period in three sessions. Raw scores on both instruments were verified by the investigator.

WISC-R protocols were scored by the examiner administering the test according to the procedures outlined in the test manual. All Battery protocols were scored by the investigator using a microcomputer program developed by Jay Hauger (1984). This program provides cluster scores, age and grade scores, standard scores, percentiles, Relative Performance Indexes, and functioning levels.

Conversion to Standard Scores

Test scores from the Battery were transformed to standard scores with a mean of 100 and a standard deviation of 15. Table Z in the WJTCA provides a conversion of percentile ranks (1st to 99th) to standard scores. Woodcock (1982) has provided additional standard scores for students

who score in the 1st percentile and in the 99th percentile rank. These scores were necessary for learning disabled subjects who fell in the 1st percentile, or below a standard score of 65, and for gifted and talented subjects who fell in the 99th percentile or above a standard score of 135. Six learning disabled subjects scored below 65 on one or more of their cluster score(s), while 27 gifted and talented subjects scored above 135 on one or more cluster score(s).

Data Analyses

The statistical procedures used to answer the research questions will be presented in the following chapter with the results. For the analyses, all test scores were transferred to data sheets. Keypunching and verification of data cards were performed at the Computer Center at the University of Arizona. All individual data cards were re-checked with the raw data by the investigator. The majority of the statistical analysis was performed using the Statistical Package for the Social Sciences (SPSS), (Nie, Hull, Jenkins, Steinbrenner, & Bent, 1975).

Methodological Assumptions and Limitations

One assumption in the study was that each group contained students with homogenous characteristics. Due to the selection criteria that were used to identify these

students, it is possible that both groups could be quite heterogenous. Although students with varying types of abilities and disabilities exhibited diverse performance, no attempt was made to divide these students into subcategories. For the purpose of this study, the salient and unifying characteristic for the subjects classified as learning disabled was achievement below potential, while the salient characteristics of the subjects classified as gifted and talented group were above average aptitude and achievement. Additionally, since the school district provided the operational definitions for both samples, the results may be generalized only to samples selected by similar criteria.

Another limitation involved the use of descriptive data from the Battery to evaluate student performance. In examining the percentage of students obtaining a certain functioning level, a specific score or cut off was used. For example, to obtain a "Moderate Deficit" a student must have a cluster difference score of -16. Although little difference exists between a score of -15 and -16, data are presented employing these cutoffs and the corresponding descriptive labels because they are used by practitioners to evaluate and assess student performance.

Finally, the subjects classified as learning disabled were placed in the programs because of relatively

high aptitude, usually assessed by the WISC-R, and relatively low achievement scores. Although WISC-Rs were re-administered to all subjects, Woodcock (1980) indicated that if a measure is used for a selection criteria, it should not be compared to another measure and that the preferred method would be to identify a sample of referred students.

CHAPTER 4

RESULTS

The purpose of this chapter is to present the results of the study for the subjects classified as learning disabled and the subjects classified as gifted and talented. The following research topics are included: (a) discrepancies between aptitude and achievement, (b) comparability of full scale scores, (c) performance on the Cognitive clusters, (d) performance on the Achievement clusters, and (e) Verbal-Performance Scale score discrepancies.

Discrepancies between Aptitude and Achievement

The Achievement-Aptitude Profile of the Woodcock-Johnson Psycho-Educational Battery was designed for making aptitude-achievement comparisons and assisting practitioners in determining the presence of a learning deficit (Woodcock, 1978). The appropriate criterion for a significant learning deficit is a moderate or severe deficit. According to Hessler (1982), these functioning levels generally represent a difference of at least one standard deviation, making the probablility of educational significance quite high.

In this study, the Achievement-Aptitude Profile was used to ascertain the number and percent of subjects attaining moderate and severe deficits between aptitude and achievement. This procedure involves subtracting each achievement cluster score from an expected achievement score that is obtained from the relevant Scholastic Aptitude cluster. This difference score is then transformed into one of the functioning levels: Very Superior, Superior, Above Average, Average, Below Average, Moderate Deficit, or Severe Deficit.

The overall performance of each sample on the Achievement-Aptitude Profile was examined. Relative proportions were determined by comparing the number of subjects in each of the seven functioning levels to the total number of subjects in each sample. The percentages of subjects in each level were calculated for the four achievement areas.

Results for Subjects Classified as Learning Disabled

Table 3 provides the number and percentage of subjects who attained a certain functioning level on the Achievement-Aptitude Profile. The results displayed in Table 3 indicate that few of the subjects attained functioning levels of Very Superior or Superior and the majority had functioning levels that were below average. Of the 51 subjects, 65% (N=33) displayed a moderate or severe deficit

Table 3. Number and Percent of Subjects Classified as Learning Disabled Scoring within each Functioning Level of the Battery (N=51)

Achievement-Aptitude Functioning Levels	Achievement-Aptitude Comparisons: Reading N	Reading %	Math N	Math %	Written Language N	Written Language %	Knowledge N	Knowledge %
Very Superior	0	0	1	2.0	0	0	0	0
Superior	1	2.0	0	0	1	2.0	0	0
Above Average	2	3.9	10	19.6	2	3.9	16	31.4
Average	6	11.8	14	27.5	8	15.9	21	41.1
Below Average	9	17.6	13	25.5	24	47.1	13	25.5
Moderate Deficit	13	25.5	10	19.6	11	21.6	0	0
Severe Deficit	20	39.2	3	5.9	5	9.8	1	2.0

in Reading, 25% (N=13) of the sample displayed a moderate or severe deficit in Mathematics, 31% (N=16) displayed a moderate or severe deficit in Written Language, and one subject attained a severe deficit in Knowledge.

Figure 1 depicts the performance of the subjects classified as learning disabled on the functioning levels of the Achievement-Aptitude Profile. These graphs illustrate that the majority of subjects attained below average functioning levels in Reading and Written Language but closer to average functioning levels in Mathematics and Knowledge.

Subjects with Moderate or Severe Deficits. Table 4 illustrates the frequency with which subjects obtained moderate or severe deficits on the Achievement-Aptitude Profile. For the 51 subjects classified as learning disabled, 13 subjects (25.5%) did not evidence a moderate or severe deficit in any of the achievement areas, 38 subjects (74.5%) evidenced a moderate or severe deficit between aptitude and achievement in one to three of the achievement area(s). For example, 17 subjects had 1 deficit in an achievement area such as Reading, Mathematics, Written Language, or Knowledge. None of the subjects evidenced a deficit in all four achievement areas.

Four types of profiles appeared on these clusters: Thirteen students had deficient performance on only the Reading cluster, ten subjects had deficient performance on the Reading and Written Language clusters with adequate performance on the Mathematics and Knowledge clusters two

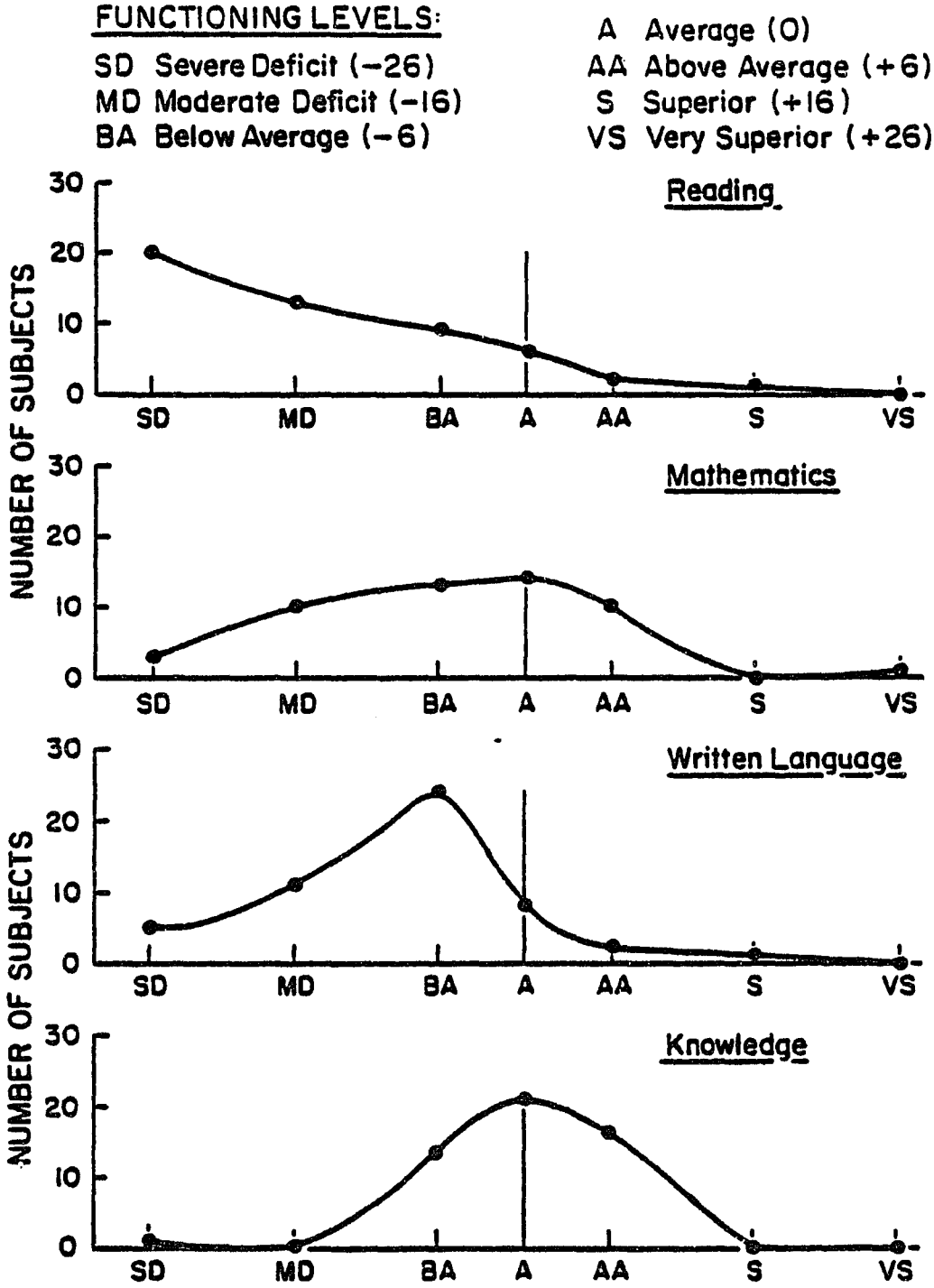

Figure 1. Performance of Subjects Classified as Learning Disabled on the Achievement-Aptitude Profile of the Battery

Table 4. Frequency of Moderate or Severe Deficits(s) on the Achievement-Aptitude Profile of the Battery for Subjects Classified as Learning Disabled (N=51)

Number of Subjects	Number of Deficits	Achievement-Aptitude Comparisons			
		Reading	Math	Written Language	Knowledge
13	0	0	0	0	0
17	1	13	2	1	1
17	2	16	7	11	0
4	3	4	4	4	0
0	4	0	0	0	0

subjects had deficient performance on the Mathematics cluster with adequate performance on the Reading, Written Language, and Knowledge clusters, and four subjects had deficient performance on the Reading, Mathematics, and Written Language clusters with adequate performance on the Knowledge cluster.

Subjects not Evidencing a Moderate or Severe Deficit. Thirteen subjects (25.5%) did not obtain a moderate or severe deficit on this test. To explore the validity of the Battery as a measure for identifying subjects with aptitude-achievement discrepancies, further analyses of the school records were made.

Of the 13 subjects, two appeared to qualify for district learning disability services even though they did not evidence a significant deficit on the Achievement-Aptitude Profile. One subject appeared to have a language disability. The other subject had a signicant aptitude-achievement discrepancy when his Full Scale WISC-R score was used as the aptitude measure.

The other 11 subjects did not appear to have an aptitude-achievement discrepancy and, consequently, failed to meet the school district criteria for placement in a learning disability program (refer to p. 48). Appendix A describes the characteristics of these 11 subjects. This summary includes age, full scale scores on the WISC-R and WJTCA, estimated mental age, achievement scores, reported processing

deficits, emotional factors, relevant observations or characteristics, and a diagnosis by the investigator. This subgroup contained one Hispanic subject and two subjects who had been retained for lack of academic progress.

A summary of the characteristics of the 11 subjects not evidencing a significant deficit on the Achievement-Aptitude Profile demonstrated that they scored 10.9 standard score points lower on the WISC-R ($\bar{X}$ = 83.00, SD = 7.82) than the 38 subjects who had a significant deficit ($\bar{X}$ = 93.9, SD = 10.92). Data in the records suggested the following diagnoses: one subject seemed to be having difficulty because of a bilingual background, five subjects seemed to have primary emotional needs, while five subjects seemed most accurately classified as slow learners. The five slow learners had a mean Full Scale WISC-R score of 76.6 (SD = 3.97), while the five emotionally involved subjects had a mean Full Scale of 87.80 (SD = 6.42). All subjects appeared to be functioning within their ability range as measured by the WISC-R and the WJTCA.

Results for Subjects Classified as Gifted and Talented

For the 46 subjects classified as gifted and talented, 5 subjects (10.9%) evidenced a moderate or severe deficit between aptitude and achievement, while 41 subjects (89.1%) did not.

Table 5 presents the number and percentage of

Table 5. Number and Percent of Subjects Classified as Gifted and Talented Scoring within each Achievement-Aptitude Functioning Level of the Battery (N=46)

Achievement-Aptitude Functioning Levels	Achievement-Aptitude Comparisons							
	Reading		Math		Written Language		Knowledge	
	N	%	N	%	N	%	N	%
Very Superior	0	0	0	0	1	2.2	1	2.2
Superior	4	8.7	5	10.0	4	8.7	0	0
Above Average	9	19.6	13	28.3	16	34.8	10	21.7
Average	26	56.5	17	37.0	19	41.3	22	47.8
Below Average	7	15.2	9	19.6	4	8.7	12	26.0
Moderate Deficit	0	0	1	2.2	2	4.3	1	2.2
Severe Deficit	0	0	1	2.2	0	0	0	0

subjects classified as gifted and talented who attained a certain functioning level on the Achievement-Aptitude Profile. In summary, none of the subjects evidenced a severe or moderate deficit on the Reading cluster, 2 subjects displayed a moderate or severe deficit on the Mathematics cluster, 2 subjects displayed a moderate deficit on the Written Language cluster, and 1 subject displayed a moderate deficit on the Knowledge cluster.

Figure 2 depicts the performance of the subjects classified as gifted and talented on the functioning levels of the Achievement-Aptitude Profile. It can be noted from these graphs that the majority of gifted and talented subjects attained functioning levels within the average range.

To explore the validity of the Battery as a measure for identifying subjects with aptitude-achievement discrepancies, further analyses of the 5 subjects with educationally significant deficits were made. For all 5 subjects, the classroom teachers concurred with the test results; deficits in these specific academic areas had been observed in classroom performance.

Three of the 5 subjects with deviant scores exhibited significant Verbal-Performance Scale score discrepancies on the WISC-R. The subject who attained the severe deficit in Mathematics had a Full Scale of 129 on the

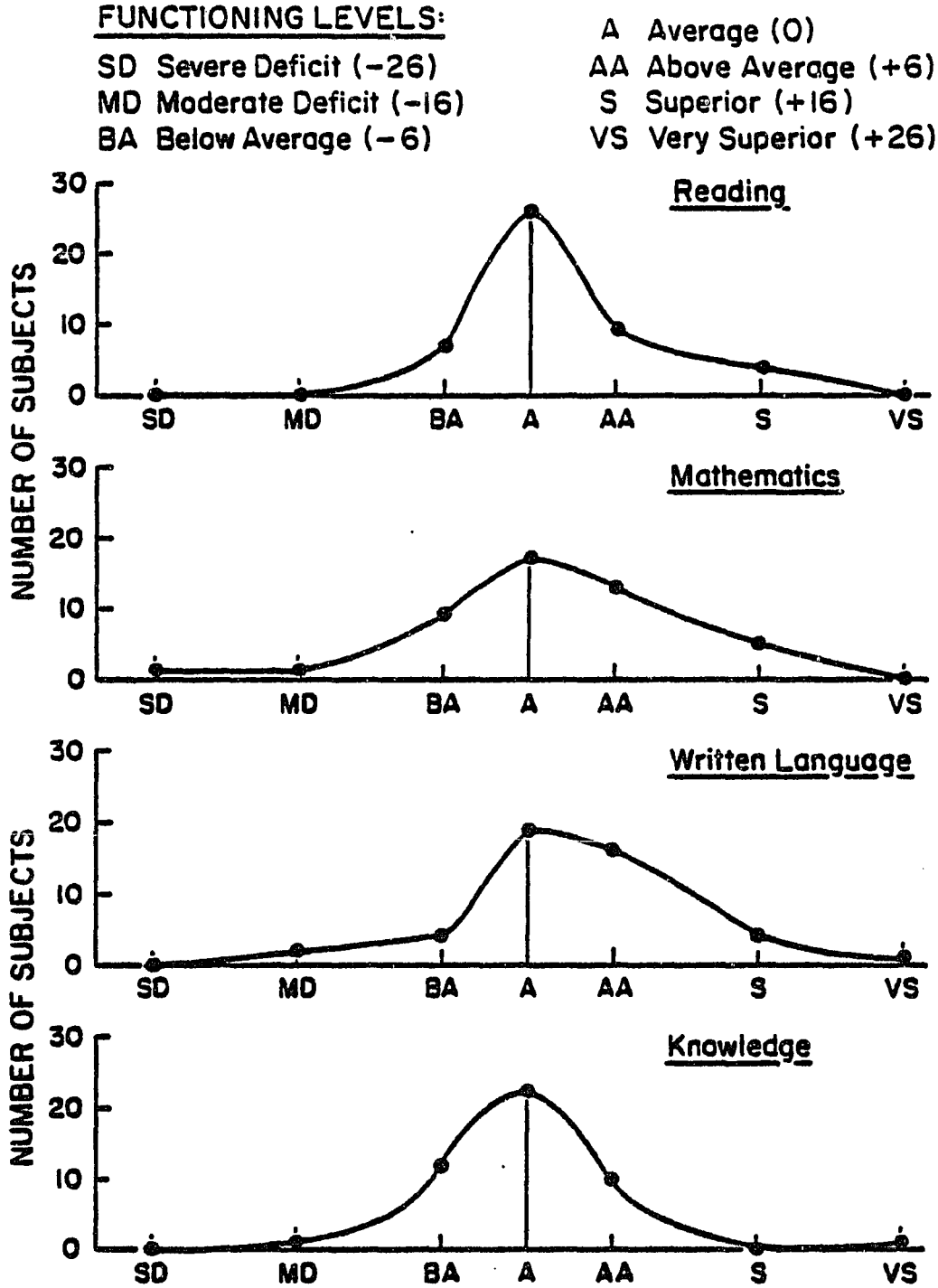

Figure 2. Performance of Subjects Classified as Gifted and Talented on the Achievement-Aptitude Profile of the Battery

WISC-R with a 145 on the Verbal Scale and a 106 on the Performance Scale. One of the subjects who attained a moderate deficit on the Written Language cluster had a Full Scale score of 132 on the WISC-R with a 117 on the Verbal Scale and a 141 on the Performance Scale. The subject who attained the moderate deficit on the Knowledge cluster received a Full Scale score of 101 on the WISC-R with an 87 on the Verbal Scale and 117 on the Performance Scale.

Comparability of Full Scale Scores

This analysis compared mean full scale scores between the WJTCA and the WISC-R for the subjects classified as learning disabled and the subjects classified as gifted and talented. For the subjects classified as learning disabled, the correlation between the WISC-R Full Scale and the WJTCA Broad Cognitive Ability Score (BCA) was .69. For the subjects classified as gifted and talented, the full scale correlation was .62. (Appendix B provides correlations between the major test scores of the Battery and the WISC-R for each sample).

To examine the relationship between the full scale scores of the two tests further, a *t*-test comparison of dependent means was performed. Table 6 illustrates the results for the subjects classified as learning disabled. For this sample, the mean Full Scale on the WISC-R was 5.78 standard score points higher than their BCA score.

Table 6. Mean Difference between Full Scale Standard Scores on the WJTCA and the WISC-R for Subjects Classified as Learning Disabled (N=51)

	Mean	SD	Mean Difference	t-Value	Significance Level
WJTCA	86.08	9.05			
			5.78	5.13	$p < .001$
WISC-R	91.86	10.97			

Individual difference scores ranged from 8 points higher on the BCA score to 28 points lower. Six of the 51 students (12%) had Full Scale WISC-R scores that were one standard deviation or more higher than their BCA scores.

Table 7 illustrates the results for the subjects classified as gifted and talented. For this sample, the difference between mean full scale scores was not significant ($p > .05$). Individual scores ranged from differences of 18 points higher on the BCA score to 17 points lower.

Analysis of individual protocols revealed that 5 of the gifted and talented subjects evidenced a difference of one standard deviation or more between their full scale scores on the two tests. Of the 5 subjects, 4 had higher BCA scores, while one had a higher Full Scale WISC-R score.

Performance on the Cognitive Clusters

To investigate performance on the Cognitive clusters, several analyses were conducted for each sample, including: (a) an analysis of performance using the four original cognitive clusters, (b) an analysis of performance substituting the alternative clusters of Oral Language and Broad Reasoning for Verbal Ability and Reasoning, and (c) a comparison of scores on the Verbal Ability and Oral Language clusters and Reasoning and Broad Reasoning clusters.

A repeated measures ANOVA was conducted for each

Table 7. Mean Difference between Full Scale Standard Scores on the WJTCA and the WISC-R for Subjects Classified as Gifted and Talented (N=46)

	Mean	SD	Mean Difference	t-Value	Significance Level
WJTCA	126.85	8.59			
			1.35	1.00	$p > .05$
WISC-R	128.20	11.51			

sample to determine if a significant difference existed across the four Cognitive clusters. When a significant F ratio was obtained, post hoc comparisons were conducted using the Tukey HSD procedure. The same procedures were followed to determine if use of the proposed alternative Cognitive cluster scores (Woodcock, 1983) would alter the pattern of performance of either group. In this analysis, the Oral Language cluster was substituted for the Verbal Ability cluster and the Broad Reasoning cluster was substituted for the Reasoning cluster. Additionally, t-tests of dependent means were performed to compare scores on Verbal Ability and Oral Language and Reasoning and Broad Reasoning.

Results for Subjects Classified as Learning Disabled

The results of the analysis of variance on the four Cognitive clusters are presented for the subjects classified as learning disabled in Table 8. In this first analysis, differences between cluster scores were not significant ($p > .05$)

The results using the alternative clusters of Oral Language and Broad Reasoning are presented in Table 9. In this analysis a significant difference was detected ($p < .05$). Post hoc comparisons revealed that the Broad Reasoning and Memory clusters were significantly higher than the Oral Language cluster.

Table 8. Repeated Measures ANOVA on Cognitive Clusters for Subjects Classified as Learning Disabled (N=51)

Cognitive Clusters	Mean	SD	F-Value	Significance Level
Verbal Ability	88.00	12.65		
Reasoning	93.55	11.69	1.99	$p > .05$
Perceptual Speed	89.29	13.09		
Memory	91.22	15.36		

Table 9. Repeated Measures ANOVA on Revised Cognitive Clusters for Subjects Classified as Learning Disabled (N=51)

Revised Cognitive Clusters	Mean	SD	F-Value	Significance Level
Oral Language	85.23	12.87		
Broad Reasoning	91.72	11.09	3.68	$p < .05$
Perceptual Speed	89.29	13.09		
Memory	91.22	15.36		

The results of the t-test comparison between the Verbal Ability cluster and the alternative cluster, Oral Language are presented in Table 10. In this comparison, the Verbal Ability cluster was significantly higher than the Oral Language cluster ($p < .01$). Table 11 presents the results of the t-test comparison between the Reasoning cluster and the alternative cluster, Broad Reasoning. In this comparison, the Reasoning cluster was significantly higher than the Broad Reasoning cluster ($p < .05$).

Results for Subjects Classified as Gifted and Talented

The results of the analysis of variance on the four Cognitive clusters are presented for the subjects classified as gifted and talented in Table 12. The differences between the cluster scores were significant ($p < .001$). Post hoc comparisons revealed that the Memory cluster was significantly higher than the Reasoning and Perceptual Speed cluster ($p < .01$) and the Verbal Ability cluster was significantly higher than the Reasoning cluster ($p < .05$).

The results using the alternative clusters of Oral Language and Broad Reasoning are presented in Table 13. Significant differences existed between the cluster scores ($p < .001$). Post hoc comparisons revealed that the Memory cluster was significantly higher than the Perceptual Speed cluster ($p < .01$) and the Broad Reasoning cluster ($p < .05$) and that the Oral Language cluster was significantly higher

Table 10. Mean Difference between Standard Scores on the Verbal Ability and Oral Language Clusters for Subjects Classified as Learning Disabled (N=51)

Cluster	Mean	SD	Mean Difference	t-Value	Significance Level
Verbal Ability	88.00	12.65			
			-2.77	-3.09	$p < .01$
Oral Language	85.23	12.87			

Table 11. Mean Difference between Standard Scores on the Reasoning and Broad Reasoning Clusters for Subjects Classified as Learning Disabled (N=51)

Cluster	Mean	SD	Mean Difference	t-Value	Significance Level
Reasoning	93.55	11.69			
			1.83	2.25	$p < .05$
Broad Reasoning	91.72	11.09			

Table 12. Repeated Measures ANOVA on Cognitive Clusters for Subjects Classified as Gifted and Talented (N=46)

Cognitive Clusters	Mean	SD	F-Value	Significance Level
Verbal Ability	123.22	11.79		
Reasoning	113.72	12.13	11.93	$p < .001$
Perceptual Speed	116.28	13.17		
Memory	128.61	16.06		

Table 13. Repeated Measures ANOVA on Revised Cognitive Clusters for Subjects Classified as Gifted and Talented (N=46)

Revised Cognitive Clusters	Mean	SD	F-Value	Significance Level
Oral Language	124.15	9.12		
Broad Reasoning	121.15	9.70	9.80	$p < .001$
Perceptual Speed	116.28	13.17		
Memory	128.60	16.06		

than the Perceptual Speed cluster ($p < .05$).

The results of the *t*-test comparison between the Verbal Ability cluster and the alternative cluster, Oral Language are presented in Table 14. The difference between these clusters was not significant ($p > .05$). Table 15 presents the results of the *t*-test comparison between the Reasoning cluster and the alternative cluster, Broad Reasoning. In this comparison, the Broad Reasoning cluster was significantly higher than the Reasoning cluster ($p < .001$).

Analysis of individual protocols revealed that in some cases, substitution of the Broad Reasoning for the Reasoning cluster made substantive changes in individual performance. Four subjects had score increases greater than one standard deviation. When contrasted to the main subtests of the Reasoning cluster, these subjects had higher performance on the suppressor subtest, Antonyms-Synonyms.

Performance on the Achievement Clusters

This analysis examined performance on the Achievement clusters. A repeated measures ANOVA was conducted for each sample to determine if a significant difference existed across the four achievement clusters. When an *F* ratio was significant, post hoc comparisons were conducted using the Tukey HSD procedure.

Table 14. Mean Difference between Standard Scores on the Verbal Ability and Oral Language Clusters for Subjects Classified as Gifted and Talented (N=46)

Cluster	Mean	SD	Mean Difference	t-Value	Significance Level
Verbal Ability	123.22	11.79			
			.93	1.02	$p > .05$
Oral Language	124.15	9.12			

Table 15. Mean Difference between Standard Scores on the Reasoning and Broad Reasoning Clusters for Subjects Classified as Gifted and Talented (N=46)

Cluster	Mean	SD	Mean Difference	t-Value	Significance Level
Reasoning	113.72	12.13			
			-7.43	-6.19	$p < .001$
Broad Reasoning	121.15	9.70			

Results for Subjects Classified as Learning Disabled

The results of the analysis of variance on the four Achievement clusters are presented for the subjects classified as learning disabled in Table 16. A significant *F* ratio was obtained ($p < .001$). Post hoc comparisons using the Tukey procedure indicated that the Mathematics and Knowledge clusters were significantly higher than the Reading and Written Language clusters ($p < .05$).

Results for Subjects Classified as Gifted and Talented

The results of the analysis of variance on the four Achievement clusters are presented for the subjects classified as gifted and talented in Table 17. Comparisons between the cluster scores were significantly different ($p < .05$). Results of the Tukey post hoc procedure indicated that the Knowledge cluster was significantly higher than the Reading and Mathematics clusters ($p < .05$) for this sample.

Verbal-Performance Scale Score Discrepancies

In this analysis, two tailed *t*-tests were conducted to determine if either sample illustrated a significant difference in the magnitude of their Verbal and Performance Scale scores. Table 18 illustrates the results for the subjects classified as learning disabled. For this sample the Performance Scale was 9.33 standard score points higher than the Verbal Scale ($p < .001$).

Table 16. Repeated Measures ANOVA on Achievement Clusters for Subjects Classified as Learning Disabled (N=51)

Achievement Clusters	Mean	SD	F-Value	Significance Level
Reading	80.16	10.34		
Mathematics	86.61	11.89	12.45	$p < .001$
Written Language	81.69	8.66		
Knowledge	88.51	11.40		

Table 17. Repeated Measures ANOVA on Achievement Clusters for Subjects Classified as Gifted and Talented (N=46)

Achievement Clusters	Mean	SD	F-Value	Significance Level
Reading	118.99	10.31		
Mathematics	119.11	11.76	3.81	$p < .05$
Written Language	122.13	10.84		
Knowledge	123.93	13.02		

Table 18. Mean Difference between WISC-R Verbal Scale Score and Performance Scale Score for Subjects Classified as Learning Disabled (N=51)

WISC-R	Mean	SD	Mean Difference	t-Value	Significance Level
Verbal Scale	88.25	12.51			
			-9.33	4.88	$p < .001$
Performance Scale	97.59	12.33			

Table 19 illustrates the results for the subjects classified as gifted and talented. The difference between the magnitude of the Verbal and Performance Scale scores was not significant ($p > .05$).

Analysis of individual WISC-R protocols revealed that several of the subjects classified as learning disabled and several of the subjects classified as gifted and talented evidenced significant Verbal-Performance Scale discrepancies. In the normative sample, differences of 22 points were found in 10% of the sample and differences of 26 points were found in 5% of the population (Wechsler, 1974). Eleven of the subjects classified as learning disabled (22%) had discrepancies of 22 points or more: Seven (14%) had discrepancies greater than 26 points, all with the Performance Scale scores greater than the Verbal Scale scores. Nine of the subjects classified as gifted and talented (20%) had discrepancies of 22 points or greater: Four of these subjects had Performance scales that were greater than their Verbal scales. Three subjects classified as gifted and talented (6%) had discrepancies that exceeded 26 points.

Since the Verbal Scale scores of the WISC-R have a higher correlation to the WJTCA (approximately .76) than the Performance Scale scores (between .58 and .67; Woodcock, 1978) and the WJTCA are more a verbal measure of cognition

Table 19. Mean Difference between WISC-R Verbal Scale Score and Performance Scale Score for Subjects Classified as Gifted and Talented (N=46)

WISC-R	Mean	SD	Mean Difference	t-Value	Signficance Level
Verbal Scale	127.09	14.26			
			3.33	-1.47	$p > .05$
Performance Scale	123.76	11.57			

than the WISC-R (Hessler, 1982), one might predict that subjects with marked discrepancies between Verbal and Performance Scale scores would display significant differences between full scale scores of the WJTCA and the WISC-R. Of the 11 subjects classified as learning disabled with Performance Scale scores 22 points greater than Verbal Scale scores, only one student had a difference greater than one standard deviation between full scale scores. For all of the other students, the full scale score differences were less than 8 points with an absolute mean difference of 3.5.

Summary of Results

The results of this study may be summarized as follows:

1. For the 51 subjects classified as learning disabled, 38 subjects (74.5%) evidenced a moderate or severe deficit between aptitude and achievement in one or more of the achievement area(s) on the Battery. Thirteen subjects (25.5%) did not evidence a moderate or severe deficit in any area.

For the 46 subjects classified as gifted and talented, 5 subjects (10.9%) evidenced a moderate or severe deficit between aptitude and achievement in one achievement area, while 41 subjects (89.1%) did not evidence a moderate or severe deficit in any area.

2. The subjects classified as learning disabled scored significantly higher (5.78 standard score points) on the WISC-R than on the WJTCA. The difference between full scale scores was not significant for the subjects classified as gifted and talented.

3. A specific pattern of performance was not identified for the subjects classified as learning disabled on the four Cognitive clusters. Using the alternative clusters, the Broad Reasoning and Memory clusters were significantly higher than the Oral Language cluster. Comparisons between the original Cognitive clusters and the alternative clusters indicated that Verbal Ability was significantly higher than Oral Language and Reasoning was significantly higher than Broad Reasoning.

A specific pattern of performance was identified for the subjects classified as gifted and talented: the Memory cluster was significantly higher than the Reasoning and Perceptual Speed clusters and the Verbal Ability cluster was higher than the Reasoning cluster. This pattern was altered when the alternative clusters were used: The Memory cluster was significantly higher than the Perceptual Speed and Broad Reasoning clusters and the Oral Language cluster was higher than the Perceptual Speed cluster. Comparisons between the original Cognitive clusters and the alternative clusters indicated no difference between Verbal Ability and Oral Language but the Reasoning cluster was significantly lower than the Broad Reasoning cluster.

4. A specific pattern of performance was identified on the Achievement clusters for the subjects classified as learning disabled: The Mathematics and Knowledge clusters were significantly higher than the Reading and Written Language clusters.

A pattern of performance was also identified for the subjects classified as gifted and talented: The Knowledge cluster was significantly higher than the Reading and Mathematics clusters.

5. The Performance Scale score was significantly higher than the Verbal Scale score for the subjects classified as learning disabled. The difference between Verbal and Performance Scale scores was not significant for the subjects classified as gifted and talented.

CHAPTER 5

SUMMARY AND DISCUSSION

The purpose of this chapter is to provide a summary of the study, a discussion of the results including the implications for the field of special education, and directions for future research.

Statement of Purpose

The major purpose of this study was to analyze and compare test scores on the Woodcock-Johnson Psycho-Educational Battery (Battery) and the Wechsler Intelligence Scale for Children - Revised (WISC-R) for a sample of subjects classified as learning disabled and a sample of subjects classified as gifted and talented.

Subjects and Design

Fifty-one subjects classified as learning disabled and 46 subjects classified as gifted and talented were randomly selected to participate in the study. Subjects were administered the Woodcock-Johnson Tests of Cognitive Ability (WJTCA), the Woodcock-Johnson Tests of Achievement (WJTA), and the WISC-R. All testing was performed over a one

year period. Tests were counterbalanced to avoid an order effect.

All WISC-R protocols were scored by the examiner administering the test, while all Battery protocols were computer scored by the investigator.

To facilitate comparisons between measures, all test scores were transformed to standard scores with a mean of 100 and a standard deviation of 15. Data were keypunched and verified at the University Computer Center. Statistical analyses were performed by the investigator and a statistician using the Statistical Package for the Social Sciences (Nie, Hull, Jenkins, Steinbrenner, & Bent, 1975). T-tests of dependent means and repeated measures ANOVAs were used to answer the majority of research questions. All descriptive data were compiled and analyzed by the investigator.

Summary of Results

1. For the 51 subjects classified as learning disabled, 38 subjects (74.5%) evidenced a moderate or severe deficit between aptitude and achievement in one or more of the achievement area(s) on the Battery. Thirteen subjects (25.5%) did not evidence a moderate or severe deficit in any area.

For the 46 subjects classified as gifted and talented, 5 subjects (10.9%) evidenced a moderate or severe deficit between aptitude and achievement in one achievement area, while 41 subjects (89.1%) did not evidence a moderate or severe deficit in any area.

2. The subjects classified as learning disabled scored significantly higher (5.78 standard score points) on the WISC-R than on the WJTCA. The difference between full scale scores was not significant for the subjects classified as gifted and talented.

3. A specific pattern of performance was not identified for the subjects classified as learning disabled on the four Cognitive clusters. Using the alternative clusters, the Broad Reasoning and Memory clusters were significantly higher than the Oral Language cluster. Comparisons between the original Cognitive clusters and the alternative clusters indicated that Verbal Ability was significantly higher than Oral Language and Reasoning was significantly higher than Broad Reasoning.

A specific pattern of performance was identified for the subjects classified as gifted and talented: the Memory cluster was significantly higher than the Reasoning and Perceptual Speed clusters and the Verbal Ability cluster was higher than the Reasoning cluster. This pattern was altered

when the alternative clusters were used: The Memory cluster was significantly higher than the Perceptual Speed and Broad Reasoning clusters and the Oral Language cluster was higher than the Perceptual Speed cluster. Comparisons between the original Cognitive clusters and the alternative clusters indicated no difference between Verbal Ability and Oral Language but the Reasoning cluster was significantly lower than the Broad Reasoning cluster.

4. A specific pattern of performance was identified on the Achievement clusters for the subjects classified as learning disabled: The Mathematics and Knowledge clusters were significantly higher than the Reading and Written Language clusters.

A pattern of performance was also identified for the subjects classified as gifted and talented: The Knowledge cluster was significantly higher than the Reading and Mathematics clusters.

5. The Performance Scale score was significantly higher than the Verbal Scale score for the subjects classified as learning disabled. The difference between Verbal and Performance Scale scores was not significant for the subjects classified as gifted and talented.

Several important implications can be drawn from the

results of this study. Discussion of the findings and implications will be presented topically, following the order of the research questions.

Discrepancies between Aptitude and Achievement

The major advantages of using the Achievement-Aptitude Profile for determining discrepancies are that the errors inherent in comparing individual scores on two different tests are eliminated; regression effect and measurement error are not a problem. Thorndike (1963) observed: "A failure to recognize this regression effect has rendered questionable, if not meaningless, much of the research on underachievement" (p. 13). For example, failure to correct for regression effects has resulted in disproportionate numbers of bright students being labeled as underachievers (Cone & Wilson, 1981; McLeod, 1979). Conceivably, the Achievement-Aptitude Profile could enhance the validity of identifying underachieving students.

Discussion of Subjects Classified as Learning Disabled

In this study, 38 subjects (74.5%) classified as learning disabled by the school district evidenced an educationally significant deficit on the Achievement-Aptitude Profile in one or more achievement area(s). The majority of these subjects evidenced a moderate or severe deficit in Reading. This finding was predictable since a reading problem is

one of the common reasons for referral to a learning disability program and often used as the sole indicator of a learning disability (Lieberman, 1980; Sartain, 1976).

Thirteen of the subjects (25.5%) did not evidence an educationally significant deficit on the Battery. This result was similar to the findings obtained by Shinn, Algozzine, Marston, and Ysseldyke (1982) when they calculated discrepancies for a sample of learning disabled subjects on various standardized measures. This finding was surprising since an aptitude-achievement discrepancy is one criterion for placement in a learning disability program.

Further analysis of the 13 subjects who did not have a discrepancy indicated that 2 of the subjects could be classified by the school district criteria and that 11 of the subjects did not qualify for learning disability services using either the Achievement-Aptitude Profile of the Battery or the school district criteria (refer to p. 48). These 11 students seemed to have school problems that were the result of (a) general low intelligence and/or (b) emotional factors. This finding is in accordance with Chandler and Jones' (1983) observation that emotionally disturbed children are frequently placed in resource rooms for the learning disabled.

This preliminary analysis suggests that the Achievement-Aptitude Profile of the Battery is useful for identifying students who do not have aptitude-achievement discrepancies and, consequently, may not qualify for learning

disability services. A detailed analysis of the 38 students who had a significant deficit on the Achievement-Aptitude Profile was not made.

Discussion of Subjects Classified as Gifted and Talented

Five of the subjects classified as gifted and talented exhibited an educationally significant deficit: 2 on the Mathematics cluster, 2 on the Written Language cluster, and 1 on the Knowledge cluster. The subject who attained the deficit on the Knowledge cluster did not appear to qualify for the gifted program in terms of intelligence (Full Scale WISC-R = 101). The classroom teachers concurred that the other 4 subjects exhibited deficits in a specific academic area and could be categorized as gifted underachievers. The performance of these 4 students is in accordance with Whitmore's (1980) observation that gifted underachievers may produce widely discrepant results, evidencing very superior ability in abstract reasoning and verbal ability, but deficits in arithmetic and visual-motor skills. Schiff, Kaufman, and Kaufman (1981) also observed that learning disabled children with superior intelligence often have difficulties in math, spelling, and handwriting.

None of the subjects classified as gifted and talented evidenced discrepancies in Reading. This finding

was not surprising as accelerated reading ability is most likely one of the initial reasons for referral to the program.

This preliminary analysis suggests that the Achievement-Aptitude Profile of the Woodcock-Johnson Psycho-Educational Battery may prove to be useful measure for identifying underachieving gifted and talented students. Further investigation is needed to validate the absence of aptitude-achievement discrepancies in the other 41 subjects classified as gifted and talented.

Comparability of Full Scale Scores

For both samples, the full scale correlations (.69 and .62) are somewhat lower than the findings reported by other studies (see Table 1). These lower correlations were most likely caused by the restricted ability ranges of the subjects.

For the subjects classified as learning disabled, mean performance on the WJTCA was approximately 5.78 points lower than on the WISC-R. Although significant, the obtained difference is not nearly so pronounced as the 12.91 mean difference reported by Reeve, Hall, and Zakreski (1979). Results are very similar to the findings of Ysseldyke, Shinn, and Epps (1981) who found a mean full scale difference of 7.68, using a sample of 50 fourth grade

students. A significant difference between full scale scores was not obtained for the subjects classified as gifted and talented.

Mean Full Scale Scores

For the subjects classified as learning disabled, the mean full scale score was 91.8 on the WISC-R and 86.08 on the WJTCA, indicating that many students scored significantly below the normative mean of 100. Other studies have also found that the mean full scale IQ score was lower than expected for learning disabled subjects (Anderson, Kaufman, & Kaufman, 1976; Smith, Coleman, Dokecki, & Davis, 1977; Zingale & Smith, 1978). Ryckman (1981) commented that the mean IQ of 84 in the Anderson et al. (1976) study was suspiciously low to be classified as representative of a learning disability. Others have reported that, in general, the average Full Scale WISC-R for students with a learning disability is between 85 to 95 (Faigel, 1981). One may speculate that either the reported low intelligence quotients were caused by interference from a learning disability, or that in some cases, students were placed in learning disability programs because of low intelligence that resulted in poor school performance and failure. As expected, the subjects classified as gifted and talented scored considerably above the normative mean on both

measures with a mean full scale score of 128.19 on the WISC-R and 126.85 on the WJTCA.

Individual Differences

Several subjects in both samples exhibited full scale differences that were greater than one standard deviation. McGrew (1983) suggested the need for determining why some individual students display significant differences. In this study, of the six subjects classified as learning disabled with mean full scale differences greater than one standard deviation, three obtained significantly low scores on the Perceptual Speed cluster. Although the number is too small to make definitive statements, poor performance on this cluster may cause marked deviations between the full scale scores and be indicative of a specific learning disability. This observation has been made by other investigators (Ipsen, McMillan, & Fallen, 1982; Reeve, Hall, & Zakreski, 1979). Conceivably, the large full scale score difference found by Reeve et al. could be attributed to a sample that contained many students who performed poorly on the Perceptual Speed cluster or possessed a specific type of severe learning disability.

Achievement and Intelligence

Ysseldyke et al. (1981) hypothesized that the primary reason for differential performance could be attributed to a greater emphasis on academic content in the WJTCA. One might expect that, if the WJTCA penalize learning disabled students with achievement-oriented subtests, gifted and talented students, who have been partially selected on the basis of high achievement, would score particularly well on these tests. Surprisingly, the difference between full scale scores for the subjects classified as gifted and talented was not significant.

In reality, it becomes impossible to divorce achievement from any intelligence test. In discussing the content of intelligence tests, Sternberg (1982) commented that most tests place heavy academic demands on students, and that intelligence is really a measure of last year's achievement. Jensen (1963) stated that standard IQ tests are fundamentally achievement tests, telling us more about experiential opportunity than a child's learning capacity and that tasks that measure ability to learn provide a better estimate of a child's intellectual capacity. Unlike the WISC-R, the WJTCA do have three subtests that require learning, using instruction and feedback to assess performance.

Examination of the content of the two measures

suggests considerable overlap, although the WJTCA do place a greater emphasis on verbal cognition. Woodcock (1978) suggested that the "heavier emphasis on verbal cognitive ability seems appropriate for most uses in psychoeducational assessment" (p. 132). Additionally, the WJTCA are more highly correlated with school performance than the WISC-R (Hessler, 1982; Woodcock, 1978). Small differences between the tests may be explained by the fact that the WJTCA have a higher correlation with school learning (Woodcock, 1980).

Further discussion of whether or not the WJTCA are more of an achievement measure than the WISC-R seems unnecessary. Intelligence is an artificial construct and, consequently, definitions vary from test to test. In education, the main purpose of an intelligence test is to provide an estimation of potential for school success (Gallagher, 1975). Past achievement contributes to this estimation. The important factor is for educators and clinicians to understand the comparability of instruments that are being used to assess special populations. With this knowledge, intelligent decisions can be made regarding the assessment results.

In summary, the subjects classified as learning disabled exhibited a significant difference between the tests that may be explained by the WJTCA's higher correlation with academic performance. Practitioners should

be aware of this tendency. The two tests seem to provide comparable scores for subjects classified as gifted and talented. This seems to indicate that the WJTCA do have concurrent validity with the WISC-R for identifying students with above average abilities. Several students in both samples, however, did exhibit marked discrepancies between the two measures which indicates that in some cases, further analysis is necessary for an accurate appraisal of student aptitude.

Performance on the Cognitive Clusters

Discussion of Subjects Classified as Learning Disabled

A specific pattern of performance was not detected for the subjects classified as learning disabled on the original clusters. Although significance was obtained using the alternative clusters, the actual differences were minimal. Considering the heterogeneity of a learning disabled sample and the varied dysfunctions of the individual members, these findings were not surprising. Several investigators have noted that individual differences will not be detected using a group design (Doehring, 1978) and that group research cannot be expected to provide a definitive, global pattern for identifying individuals (Galvin, 1981; Mather, 1983; Senf, 1976).

Discussion of Subjects Classified as Gifted and Talented

Although the subjects classified as gifted and talented exhibited a specific pattern of performance using the original Cognitive clusters, common sense dictated that these findings were suspect. Using the original clusters, the subjects scored lowest on the Reasoning and Perceptual Speed clusters. Considering physiological development and findings of previous research (Gallagher & Lucito, 1961; Thompson & Finley, 1962; Whitmore, 1980), one might expect gifted students to score lowest on the Perceptual Speed cluster. Low performance in reasoning, however, seemed inconsistent with expectations regarding high intelligence (Gallagher, 1975). When analysis was performed using the revised clusters, the Broad Reasoning cluster score was more commensurate with the accelerated aptitude of this sample.

In both analyses, the subjects classified as gifted and talented scored highest on the Memory cluster. Woodcock (1978) reported correlations ranging from .69 to .80 between the Memory cluster and the Broad Cognitive Ability score. To explain this finding, one might hypothesize that students with superior short term auditory memory score high on intelligence and achievement tests and, consequently, may be selected for inclusion in gifted and talented programs.

Performance on the Alternative Clusters

For both of the samples in this study, the global scores on the alternative cluster of Broad Reasoning were significantly different from those attained using the Reasoning cluster. Scores on the Reasoning cluster were particularly inaccurate for the subjects classified as gifted and talented; some individuals evidenced marked differences between the scores. Hessler (1982) and McGue, Shinn, and Ysseldyke (1982) alerted investigators to a paradoxical scoring effect that may occur when a student scores (a) consistently high or low on all of the subtests of the Reasoning cluster, or (b) much higher on the suppressor subtest than the main subtests. For these samples, the scoring effect was most pronounced when a subject scored much higher or lower on the suppressor subtest, Antonyms-Synonyms, than the main subtests of the Reasoning cluster.

Several researchers have reported that their learning disabled samples scored highest on the Reasoning cluster (McGrew, 1983; Ysseldyke, Shinn, McGue, & Epps, 1980). Analysis in this study seems to suggest that these earlier results might have been caused by the paradoxical scoring effect, rather than superior reasoning abilities.

Performance on the Cognitive clusters using suppressor weights (Reasoning and Verbal Ability) must be interpreted with caution when assessing special

populations. Elevated and deflated scores on the Verbal Ability and Reasoning cluster may be merely an artifact of the test, rather than a true appraisal of a subject's ability. If results seem inconsistent, practitioners should compare the scores from the alternative clusters before arriving at definitive conclusions.

Although all test interpretation involves a modicum of judgment and discretion, the inaccuracy of the original clusters with these two special populations seems far too subject to misinterpretation. This fallibility in the clusters necessitates diagnostic expertise from the examiner as the scores cannot be trusted. The Standards of Educational and Psychological Tests (APA, 1974) indicate that suppressor weights should not be used if they are invalid for more than one special population. Results of this study for each sample suggest that the alternative clusters of Oral Language and Broad Reasoning have greater validity and seemed to provide a more accurate appraisal of ability. In the future, these alternative clusters should replace the original clusters.

Performance on the Achievement Clusters

In this study, the subjects classified as learning disabled scored significantly lower on the Reading and Written Language clusters than the Mathematics and Knowledge

clusters. This seems to indicate that the majority of subjects exhibited a profile of specific reading and writing disabilities as described by Hessler (1982). Since many students placed in learning disabilities programs evidence reading and writing problems (Kirk & Gallagher, 1983; Lerner, 1981), this finding was expected.

Although significance was obtained for the subjects classified as gifted and talented, the actual differences were minimal. Overall, the subjects classified as gifted and talented evidenced accelerated achievement in all areas.

Verbal-Performance Scale Score Discrepancies on the WISC-R

Similar to findings from other studies assessing learning disabled samples (Anderson, Kaufman, & Kaufman, 1976; Smith, 1978; Smith, Coleman, Dokecki, & Davis, 1977; Zingale & Smith, 1978), the subjects classified as learning disabled exhibited a significantly higher Performance Scale score than Verbal Scale score on the WISC-R (9.33 standard score points). Practically, the amount of discrepancy exhibited by this sample is not highly significant as it does not differ greatly from the magnitude of discrepancy that exists in the normal population. Although a significant difference was not obtained for the subjects classified as gifted and talented, the results were in the

hypothesized direction (mean Verbal Scale score greater than mean Performance Scale score).

Several subjects in both samples exhibited discrepancies that would be classified as educationally significant (Kaufman, 1976; Wechsler, 1974). In both samples, the frequency of severe discrepancies (greater than 22 points) was greater than in the normative population. Although this finding is common for students with learning disabilities, Verbal-Performance Scale score discrepancies may also be more common with gifted students. Kaufman (1976) observed that bright students tended to have greater variability between scaled scores than other samples of children.

An unexpected finding was that subjects with marked discrepancies between Verbal and Performance Scale scores did not exhibit major differences between the two full scale scores. One might expect that students with low performance on the Verbal Scale of the WISC-R would evidence greater difficulty on the WJTCA which are more a measure of verbal cognition (Hessler, 1982). This preliminary finding suggests that a low Verbal intelligence quotient on the WISC-R does not indicate that a student will score lower on the WJTCA or that students with language disabilities will be inordinately penalized by the WJTCA.

Conclusions and Implications

Since measurement error and regression effects are eliminated by using the Achievement-Aptitude Profile, this method of aptitude-achievement comparison is superior to procedures commonly employed by school districts. The results of this study indicated that the Achievement-Aptitude Profile of the Battery is useful for identifying students who do not have aptitude-achievement discrepancies and, consequently, do not qualify for services using the criteria of this district. The Achievement-Aptitude Profile also appears useful for identifying underachieving gifted and talented students. This preliminary analysis suggests that the Achievement-Aptitude Profile can assist diagnosticians in identifying underachieving students.

The subjects classified as learning disabled scored significantly higher on the Full Scale score of the WISC-R than on the Battery. Although this finding may be explained by the WJTCA's higher correlation with achievement, practitioners should be aware of this tendency when evaluating students with learning disabilities.

The two tests provided comparable scores for subjects classified as gifted and talented. This indicates that the WJTCA do have concurrent validity with the WISC-R and may be used to identify gifted students. Several students in both samples, however, did exhibit marked

discrepancies between the full scale scores of the two measures which indicates that in some cases, further assessment is necessary for an accurate appraisal of student performance.

The results of the study suggested that scores obtained on the Reasoning cluster were invalid for subjects in both samples, while scores on the Verbal Ability cluster were invalid for some of the subjects classified as learning disabled. The alternative clusters, Oral Language and Broad Reasoning, provided a more accurate appraisal of abilities for both samples and should be used by practitioners when assessing students who appear to be learning disabled or gifted.

The subjects classified as learning disabled scored lowest on the Reading and Written Language clusters, while subjects classified as gifted and talented evidenced accelerated achievement in all areas. This finding suggests that the majority of students placed in learning disability programs have deficits in reading and written expression.

The subjects classified as learning disabled exhibited a significantly higher Performance Scale score than Verbal Scale score on the WISC-R. Although a significant difference was not obtained for the subjects classified as gifted and talented, the results were in the suspected direction (mean Verbal Scale score greater than

mean Performance Scale score).

Directions for Future Research

More research is needed to ascertain the validity of the Achievement-Aptitude Profile of the Battery as a measure for identifying underachieving and learning disabled students. These studies should compare test results and student placement, assessing the accuracy of each. As Adelman (1979) stated, "Valid classification is a critical step in advancing knowledge related to learning and behavior problems" (p. 13).

An immediate need for research involves further examination of the validity of the Verbal Ability and Reasoning clusters. Studies should be conducted that attempt to discover if the suppressor subtests and weights used in these clusters will affect the test results of other special populations besides subjects classified as learning disabled and subjects classified as gifted and talented. If the alternative clusters are universally more stable and valid, they should be made available to all test users as soon as possible.

Another area for further research involves exploration of whether or not the WJTCA are useful for identifying subtypes or subgroups of learning disabled students. Investigators should determine whether the

Cognitive clusters are successful in identifying and differentiating students with specific deficits in oral language, auditory memory, reasoning, and perception. Furthermore, case studies of individuals who exhibit marked discrepancies between the full scale scores of the WISC-R and the WJTCA are necessary. It is important to determine if atypically low performance on the WJTCA is indicative of a specific type of learning disability, such as a perceptual deficit.

Further controversy regarding full scale score differences seems unnecessary. The WJTCA are more closely related to achievement than the WISC-R (Woodcock, 1980) and minor differences between tests should be expected. Instead, longitudinal studies are needed to help determine the predictive validity of the WJTCA as these subtests may be more useful than the WISC-R as predictors of school success.

Further analyses should be conducted with the data obtained in this study. The data should be re-analyzed after the heterogeneity of the samples has been reduced. This could be accomplished by identifying the subjects in each sample who meet all criteria for placement in district programs. Additionally, an attempt should be made to identify subgroups according to patterns of performance on the Cognitive clusters of the WJTCA. Also, subtest performance on both measures should be examined for each sample. Finally, case studies of the most deviant individuals should be reported.

Concluding Remark

This study examined the performance of a sample of subjects classified as learning disabled and a sample of subjects classified as gifted and talented on the Woodcock-Johnson Psycho-Educational Battery and the WISC-R. The results provided insights regarding test performance and test validity for these two diverse samples. Overall, the Battery appears to provide useful test data that can assist practitioners in making placement decisions involving these types of special students.

APPENDIX A

CHARACTERISTICS OF ELEVEN SUBJECTS CLASSIFIED AS LEARNING DISABLED NOT EVIDENCING A MODERATE OR SEVERE DEFICIT ON THE ACHIEVEMENT-APTITUDE PROFILE OF THE BATTERY

Characteristics of Eleven Subjects Classified as Learning Disabled not Evidencing a Moderate or Severe Deficit on the Achievement-Aptitude Profile of the Battery

Case No.	C.A.	Full Scale WISC-R	Full Scale BCA	Estimated M.A.	WJTA Scores	Processing Deficits	Emotional Factors	Other Observations/ Diagnosis
1	7-3	91	90	6-6	R 6-6 M 6-10 WL 6-2 K 5-5	Auditory and visual memory, visual motor integration	Very shy	Hispanic/bilingual background
2	7-4	86	91	6-6	R 6-11 M 6-10 WL 6-6 K 6-11	Original evaluation: "No pattern of processing deficits are apparent." Re-evaluation: Word discrimination, expressive language, perceptual-motor	Receives regular Adlerian counseling; reluctant to talk; "poor ego strength and poor social skills"; emotional inhibition; quiet, withdrawn behavior	Retained/emotional overlay
10	8-7	76	92	7-8	R 8-11 M 7-6 WL 7-7 K 7-6	Original evaluation: fine and gross motor Re-evaluation: reasoning, verbal ability, visual motor integration, auditory memory, visual and auditory discrimination		Slow learner

Case No.	C.A.	Full Scale WISC-R	Full Scale BCA	Estimated M.A.	WJTA Scores	Processing Deficits	Emotional Factors	Other Observations/ Diagnosis
21	9-3	78	82	7-4	R 7-4 M 7-11 WL 7-7 K 7-6	Auditory and visual association, auditory memory; reasoning, attention, visual motor integration	Fear of failure; negative social interactions and peer relations; suffers from anxiety; turmoil in environment; poor self image	Emotional overlay
25	9-5	72	77	7-3	R 8-4 M 7-2 WL 8-11 K 7-1	Auditory memory verbal ability, difficulty acquiring concepts		Slow learner
33	9-10	76	74	6-7	R 7-3 M 7-4 WL 7-6 K 7-0	Auditory processing verbal skills deficits		Slow learner
34	9-11	83	82	7-5	R 7-6 M 8-2 WL 8-6 K 7-8	Auditory processing, reasoning, visual discrimination, visual memory, visual motor integration		Slow learner

Case No.	C.A.	Full Scale WISC-R	Full Scale BCA	Estimated M.A.	WJTA Scores	Processing Deficits	Emotional Factors	Other Observations/ Diagnosis
39	10-5	82	78	8-2	R 7-2 M 8-10 WL 7-4 K 6-8	Auditory memory, verbal fluency, visual memory fine motor, perception of abstract relationships	Remains stubbornly silent; socially immature, poor personal hygiene; isolationist tendencies; lives in his own world	Emotional
40	10-6	93	96	10-2	R 8-6 M 10-3 WL 8-11 K 10-0	Visual-motor integration	Behaviors including foul language, hitting, fighting, and stealing are interfering with academic performance	Emotional
3	10-7	96	83	8-8	R 7-11 M 10-7 WL 9-8 K 10-2	Auditory memory	Doesn't apply himself; resistant to learning; gives up easily; parents going through divorce	Teacher reports student is capable of grade level work when he applies himself; emotional
48	11-5	76	79	9-1	R 8-5 M 8-10 WL 8-2 K 8-5	Visual and auditory channel deficits; low verbal ability; verbal concept formation; visual motor integration		Retained; slow learner

APPENDIX B

1. PEARSON PRODUCT-MOMENT CORRELATION COEFFICIENTS FOR WISC-R AND BATTERY STANDARD SCORES FOR SUBJECTS CLASSIFIED AS LEARNING DISABLED (N=51)

2. PEARSON PRODUCT-MOMENT CORRELATION COEFFICIENTS FOR WISC-R AND BATTERY STANDARD SCORES FOR SUBJECTS CLASSIFIED AS GIFTED AND TALENTED (N=46)

Pearson Product-Moment
Correlation Coefficients for WISC-R
and Battery Standard Scores for Subjects Classified as
Learning Disabled (N=51)

Battery	WISC-R		
	Full Scale	Verbal	Performance
Broad Cognitive Ability	.69	.60	.56
Verbal Ability	.58	.60	.37
Oral Language	.65	.63	.45
Reasoning	.33	.23	.35
Broad Reasoning	.60	.51	.49
Memory	.40	.25	.41
Perceptual Speed	.09	.05	.12
Reading Aptitude	.67	.69	.41
Written Language Aptitude	.57	.53	.42
Math Aptitude	.63	.59	.44
Knowledge Aptitude	.66	.65	.45
Reading	.32	.40	.12
Mathematics	.58	.56	.40
Written Language	.26	.35	.06
Knowledge	.66	.65	.45

Pearson Product-Moment
Correlation Coefficients for WISC-R
and Battery Standard Scores for Subjects Classified as
Gifted and Talented (N=46)

Battery	WISC-R		
	Full Scale	Verbal	Performance
Broad Cognitive Ability	.62	.59	.43
Verbal Ability	.57	.58	.32
Oral Language	.70	.72	.39
Reasoning	-.15	-.12	-.10
Broad Reasoning	.13	.20	.03
Memory	.32	.32	.24
Perceptual Speed	.10	-.01	.20
Reading Aptitude	.52	.57	.26
Written Language Aptitude	.60	.59	.39
Math Aptitude	.47	.43	.31
Knowledge Aptitude	.71	.71	.41
Reading	.48	.50	.27
Mathematics	.60	.55	.39
Written Language	.37	.43	.13
Knowledge	.62	.70	.25

LIST OF REFERENCES

Ackerman, P., Peters, J., & Dykman, R. Children with specific learning disabilities: WISC profiles. Journal of Learning Disabilities, 1971, 4, 150-166.

Adelman, H. S. Diagnostic classification of LD: Research and ethical perspectives as related to practice. Learning Disability Quarterly, 1979, 2, 5-15.

Algozzine, B. & Ysseldyke, J. Learning disabilities as a subset of school failure: The oversophistication of a concept. Exceptional Children, 1983, 50, 242-246.

American Psychological Association. Standards for educational psychological tests. Washington, D.C.: American Psychological Association, 1974.

Anastasi, A. Psychological testing (4th ed.). New York: MacMillan, 1976.

Anderson, M., Kaufman, A., & Kaufman, N. Use of the WISC-R with a learning disabled population: Some diagnostic implications. Psychology in the Schools, 1976, 13, 381-386.

Baker, H. J., & Leland, B. Detroit Tests of Learning Aptitude (Rev. 1967). Indianapolis: Bobbs-Merrill, 1967.

Baldwin, J. W. The relationship between teacher-judged giftedness, a group intelligence test and a individual intelligence test with possible gifted kindergarten pupils. Gifted Child Quarterly, 1962, 6, 153-156.

Bannatyne, A. Diagnosing learning disabilities and writing remedial prescriptions. Journal of Learning Disabilities, 1968, 1, 242-249.

Bush, W. J. & Mattson, B. D. WISC test patterns and underachievers. Journal of Learning Disabilities, 1973, 6, 251-256.

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

Chandler, H. N. & Jones, K. E. Learning disabled or emotionally disturbed: Does it make any difference? Journal of Learning Disabilities, 1983, 16, 432-434.

Clark, B. Growing up gifted. Columbus, Ohio: Charles E. Merrill, 1979.

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

Connolly, A. J., Nachtman, W. & Pritchett, E. M. Key Math Diagnostic Arithmetic Test. Circle Pines, Mn.: American Guidance Service, Inc., 1971.

Cornish, R. L. Parents', teachers' and pupils' perception of the gifted child's ability. Gifted Child Quarterly, 1968, 12, 14-17.

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

Doehring, D. G. The tangled web of behavioral research on developmental dyslexia. In A. L. Benton & D. Pearl (Eds.), Dyslexia: An appraisal of current knowledge. London: Oxford University Press, 1978.

Dowdall, C. B. & Colangelo, N. Underachieving gifted students: Review and implications. Gifted Child Quarterly, 1982, 26, 179-184.

Dudley-Marling, C. C., Kaufman, N. J. & Tarver, S. G. WISC and WISC-R profiles of learning disabled children: A review. Learning Disability Quarterly, 1981, 4, 307-319.

Dunn, L. W. & Markwardt, F. C. Peabody Individual Achievement Test. Circle Pines, Mn.: American Guidance Service, 1970.

Faigel, H. C. Learning disabilities in adolescents with high IQ scores. Developmental and Behavioral Pediatrics, 1983, 4, 11-15.

Feagans, L. A current view of learning disabilities. The Journal of Pediatrics, 1983, 102, 487-493.

French, J. L. The gifted. In M. V. Wisland (Ed.), Psycho-educational diagnosis of exceptional children. Springfield, Ill.: Charles C. Thomas, 1974.

Gajar, A. Educable mentally retarded, learning disabled, emotionally disturbed: Similarities and differences. Exceptional Children, 1979, 45, 470-472.

Gallagher, J. J. Teaching the gifted child (2nd ed.). Boston: Allyn and Bacon, Inc., 1975.

Gallagher, J. J. & Lucito, L. J. Intellectual patterns of gifted compared with average and retarded. Exceptional Children, 1961, 27, 479-482.

Gallagher, J. J. & Rogge, W. The gifted. Review of Educational Research, 1966, 36, 37-54.

Galvin, G. A. Uses and abuses of the WISC-R with the learning disabled. Journal of Learning Disabilities, 1981, 14, 326-329.

Gutkin, T. B. WISC-R scatter indices: Useful information for differential diagnosis? Journal of School Psychology, 1979, 17, 368-371.

Hanna, G. S., Dyck, N. J. & Holen, M. C. Objective analysis of achievement-aptitude discrepancies in LD classification. Learning Disability Quarterly, 1979, 2, 32-38.

Hauger, J. Compuscore. Allen, Tx.: DLM Teaching Resources, 1984.

Hessler, G. L. Use and interpretation of the Woodcock-Johnson Psycho-Educational Battery. Hingham, Ma.: Teaching Resources Corporation, 1982.

Hunter, J. A. & Lowe, J. D. The use of the WISC-R, Otis, Iowa, and SRBCSS in identifying gifted elementary students. Southern Journal of Educational Research, 1978, 12, 59-65.

Ipsen, S. M., McMillan, J. H., & Fallen, N. H. An investigation of the reported WJ/WISC-R discrepancy. Paper presented at the Virginia State Federation Council for Exceptional Children Annual Conference, March 1982.

Jacobs, J. Are we being misled by fifty years of research on our gifted children? Gifted Child Quarterly, 1970, 14, 120-123.

Jastak, J. F., & Jastak, S. Wide Range Achievement Test (1978 Rev. ed). Wilmington, Del.: Jastak Associates, 1978.

Jensen, A. Learning ability in retarded, average and gifted children. Merrill-Palmer Quarterly, 1963, 9, 123-140.

Karnes, F. A. & Brown, K. E. Factor analysis of the WISC-R for the gifted. Journal of Educational Psychology, 1979, 71, 197-199.

Kaufman, A. A. A new approach to the interpretation of test scatter in the WISC-R. Journal of Learning Disabilities, 1976, 9, 160-168.

Kirk, S. A. Educating exceptional children. Boston: Houghton-Mifflin, 1972.

Kirk, S. A. & Gallagher, J. J. Educating exceptional children (4th ed.). Boston: Houghton-Mifflin, 1983.

Kirk, S. A., McCarthy, J. J., & Kirk, W. D. Illinois Test of Psycholinguistic Abilities (Rev. ed.). Urbana: University of Illinois Press, 1968.

Lerner, J. Learning Disabilities (3rd ed.). Boston: Houghton-Mifflin, 1981.

Lieberman, L. M. The implications of noncategorical special education. Journal of Learning Disabilities, 1980, 13, 65-68.

Lindquist, E. F. & Hieronymus, A. N. Iowa Test of Basic Skills. Boston: Houghton-Mifflin, 1956.

Littell, W. M. The Wechsler Intelligence Scale for Children: Review of a decade of research. Psychological Bulletin, 1960, 57, 132-156.

Lucito, L. & Gallagher, J. J. Intellectual patterns of highly gifted children on the WISC. Peabody Journal of Education, 1960, 38, 131-136.

Marland, S. P. Education of the gifted and talented. Washington: U.S. Government Printing Office, 1971.

Marston, D. An analysis of subtest scatter on the Tests of Cognitive Ability from the Woodcock-Johnson Psycho-Educational Battery (Research Report No. 46). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities, 1980.

Martinson, R. A. Children with superior cognitive abilities. In L. M. Dunn (Ed.), Exceptional children in the schools: Special education in transition. New York: Holt, Rinehart and Winston, 1973.

Martinson, R. A. Identification of the gifted and talented. Reston, Virginia: Council for Exceptional Children, 1975.

Mather, N. Visual perception and reading achievement: The conflict between clinical practice and group research. Paper presented to Arizona CEC Annual State Convention, Phoenix, February 1983.

McGrew, K. S. Comparison of the WISC-R and Woodcock-Johnson Tests of Cognitive Ability. Journal of School Psychology, 1983, 21, 271-276.

McGue, M., Shinn, M. & Ysseldyke, J. Use of cluster scores on the Woodcock-Johnson Psycho-Educational Battery with learning disabled students. Learning Disability Quarterly, 1982, 5, 274-287.

McLeod, J. Educational underachievement: Toward a defensible psychometric definition. Journal of Learning Disabilities, 1979, 12, 42-50.

Moore, D. W. & Wielan, P. O. WISC-R scatter indexes of children referred for reading diagnosis. Journal of Learning Disabilities, 1981, 14, 511-514.

Newland, T. E. Psychological assessment of exceptional children and youth. In W. Cruickshank (Ed.), Psychology of exceptional children and youth. Englewood Cliffs, N.J.: Prentice Hall, 1971.

Nie, H. H., Hull, C. H., Jenkins, J. G., Steinbrenner, K., & Bent, D. H. SPSS: Statistical package for the social sciences. New York: McGraw-Hill Book Company, 1975.

Nisbet, R. E. A comparison of the WISC-R and Woodcock-Johnson with a referral population (Unpublished master's thesis). Moorhead State University, Minnesota, 1981.

Ogdon, D. P. Extrapolated WISC-R IQs for gifted and mentally retarded children. Journal of Consulting and Clinical Psychology, 1975, 43, 216.

Pegnato, C. W. & Birch, J. W. Locating gifted children in junior high schools: A comparison of methods. Exceptional Children, 1959, 25, 300-304.

Reeve, R. E., Hall, R. J., & Zakreski, R. S. The Woodcock-Johnson Tests of Cognitive Ability: Concurrent validity with the WISC-R. Learning Disability Quarterly, 1979, 2, 63-69.

Renzulli, J. S. Hartman, R. K. & Callahan, C. M. Scale for Rating Behavioral Characteristics of Superior Students. Exceptional Children, 1971, 38, 243-248.

Renzulli, J. S., Smith, L. H., White, A. J., Callahan, C. M., & Hartman, R. K. Scales for Rating the Behavioral Characteristics of Superior Students. Mansfield Center, Conn.: Creative Learning Press, Inc., 1976.

Rubenzer, R. L. The gifted or talented student: A neglected educational minority. Guidelines for Pupil Services, 1975-76. Madison: Wisconsin Department of Public Instruction, 14, 63-69.

Rust, J. O. & Lose, B. D. Screening for giftedness with the Slosson and the Scale for Rating Behavioral Characteristics of Superior Students. Psychology in the Schools, 1980, 17, 446-451.

Ryckman, D. B. Searching for a WISC-R profile for learning disabled children: An appropriate task. Journal of Learning Disabilities, 1981, 14, 508-511.

Ryckman, D. B. & Elrod, G. F. Once is not enough. Journal of Learning Disabilities, 1983, 16, 87-89.

Sartain, H. W. Instruction of disabled learners: A reading perspective. Journal of Learning Disabilities, 1976, 9, 489-97.

Sattler, J. M. Assessment of children's intelligence and special abilities (2nd Ed.). Rockleigh, N.J.: Allyn and Bacon, Inc., 1982.

Schiff, M. M., Kaufman, A. S. & Kaufman, N. L. Scatter analysis of WISC-R profiles for learning disabled children with superior intelligence. Part 1: Studies with new or unusual learning disabled populations. Journal of Learning Disabilities, 1981, 14, 400-404.

Senf, G. M. Future research needs in learning disabilities. In R. P. Anderson & C. G. Halcomb (Eds.), Learning disability/minimal brain dysfunction. Springfield, Ill.: Charles C. Thomas, 1976.

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

Slosson, R. L. Slosson Intelligence Test for Children. East Aurora, N.Y.: Slosson Educational Publications, Inc., 1975.

Smith, M. D. Stability of WISC-R subtest profiles for learning disabled children. Psychology in the Schools, 1978, 15, 4-7.

Smith, M. D., Coleman, J. M., Dokecki, P. R. & Davis, E. E. Intellectual characteristics of school labeled learning disabled children. Exceptional Children, 1977, 43, 352-357.

Sternberg, R. J. Lies we live by: Misapplication of tests in identifying the gifted. Gifted Child Quarterly, 1982, 26, 157-161.

Tabachnick, B. Test scatter in the WISC-R. Journal of Learning Disabilities, 1979, 12, 626-628.

Terman, L. M. & Merrill, M. A. Stanford-Binet Intelligence Scale, Form L-M. Boston: Houghton-Mifflin Company, 1960.

Terman, L. M. & Merrill, M. A. Stanford-Binet Intelligence Scale, 1972 norms edition, Form L-M. Boston: Houghton-Mifflin Company, 1973.

Thompson, J. M. & Finley, C. J. A further comparison of the intellectual patterns of gifted and mentally retarded children. Exceptional Children, 1962, 28, 379-381.

Thorndike, R. I. The concepts of over- and underachievement. New York: Bureau of Publications, Teachers College, Columbia University, 1963.

Thurlow, M. L. & Ysseldyke, J. E. Current assessment and decision-making practices in model LD programs. Learning Disability Quarterly, 1979, 2, 15-24.

Tucson Unified School District #1. Adaptive education placement guidelines. Tucson: Tucson Unified School District #1, January 4, 1982.

U. S. Office of Education, Bureau of Education for the Handicapped. State education estimates unserved by type of handicap. Washington, D.C.: Aid to State Branch, May 16, 1975.

U. S. Office of Education. Assistance for states for education of handicapped children: Procedures for evaluating specific learning disabilities. Federal Register, 1977, 42, 65082-65085.

U. S. Department of Education. To assure the free appropriate public education of all handicapped children. Fourth annual report to Congress on the implementation of Public Law 94-142: The education for all handicapped children act, Washington, D.C.: Division of Educational Services, 1982.

Vance, H., Wallbrown, F. H. & Blaha, J. Determining WISC-R profiles for reading disabled children. Journal of Learning Disabilities, 1978, 11, 657-661.

Wechsler, D. Wechsler Intelligence Scale for Children. New York: Psychological Corporation, 1949.

Wechsler, D. Wechsler Adult Intelligence Scale. New York: Psychological Corporation, 1955.

Wechsler, D. Wechsler Preschool and Primary Scale of Intelligence. New York: Psychological Corporation, 1967.

Wechsler, D. Wechsler Intelligence Scale for Children - Revised. New York: Psychological Corporation, 1974.

Whitmore, J. R. Giftedness, conflict, and underachievement Boston: Allyn and Bacon, Inc., 1980.

Woodcock, R. W. Development of standardization of the Woodcock-Johnson Psycho-Educational Battery. Hingham, Ma.: Teacher Resources Corporation, 1978.

Woodcock, R. W. A response to some concerns raised about the Woodcock-Johnson Psycho-Educational Battery. Leadership Series in Special Education, Minneapolis Public Schools, X, 107-117.

Woodcock, R. W. Standard scores. In G. L. Hessler, Use and interpretation of the Woodcock-Johnson Psycho-Educational Battery. Hingham, Ma.: Teaching Resources Corporation, 1982.

Woodcock, R. W. Oral Language and Broad Reasoning Clusters. Unpublished rough draft, March 1983.

Woodcock, R. W. & Johnson, M. B. Woodcock-Johnson Psycho-Educational Battery. Hingham, Ma.: Teaching Resources Corporation, 1977.

Ysseldyke, J., Algozzine, B. & Epps, S. A logical and empirical analysis of current practice in classifying students as handicapped. Exceptional Children, 1983, 50, 160-166.

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

Ysseldyke, J., Shinn, M., McGue, M. & Epps, S. Performance of learning disabled and low achieving students on the Wechsler Intelligence Scale for Children -- Revised and the Tests of Cognitive Ability from the Woodcock-Johnson Psycho-Educational Battery. In W. M. Cruickshank & A. A. Silver (Eds.), Bridges to tomorrow: Vol. 2, The best of ACLD. Syracuse, N.Y.: Syracuse University Press, 1981.

Zingale, S. A. & Smith, M. D. WISC-R patterns for learning disabled children at three SES levels. Psychology in the Schools, 1978, 15, 199-204.