

1980

## The validity of the Leiter international performance scale in measuring the intelligence of intellectually superior children

Susan Braxton Paltrow  
*Portland State University*

Let us know how access to this document benefits you.

Follow this and additional works at: [http://pdxscholar.library.pdx.edu/open\\_access\\_etds](http://pdxscholar.library.pdx.edu/open_access_etds)

 Part of the [Child Psychology Commons](#), and the [School Psychology Commons](#)

---

### Recommended Citation

Paltrow, Susan Braxton, "The validity of the Leiter international performance scale in measuring the intelligence of intellectually superior children" (1980). *Dissertations and Theses*. Paper 2979.

[10.15760/etd.2972](https://doi.org/10.15760/etd.2972)

This Thesis is brought to you for free and open access. It has been accepted for inclusion in Dissertations and Theses by an authorized administrator of PDXScholar. For more information, please contact [pdxscholar@pdx.edu](mailto:pdxscholar@pdx.edu).

AN ABSTRACT OF THE THESIS OF Susan Braxton Paltrow for the Master of Science in Psychology presented June 3, 1980.

Title: The Validity of the Leiter International Performance Scale  
in Measuring the Intelligence of Intellectually Superior Children

APPROVED BY MEMBERS OF THESIS COMMITTEE:

[REDACTED]

Walter G. Klopfer, Chairperson

[REDACTED]

Ronald E. Smith

[REDACTED]

Richard Lazere

The thrust of this project was to compare the obtained IQ scores between the Wechsler Intelligence Scales for Children-Revised Full Scale score at 120 and above of public school students who were in classes for "gifted and talented" with the scores obtained when the Leiter International Performance Scale was administered.

Fifty such subjects were tested with their verbal consent and prior written permission from their parents for voluntary participation in this research project. Parent conferences to provide feedback were held when so requested and forty parents took this option.

Pearson Product-Moment Correlation Coefficients were computed for

the Wechsler Intelligence Scale for Children-Revised (WISC-R) Verbal versus the Leiter International Performance Scale (LIPS) IQ scores ( $\underline{r} = .28$ ), WISC-R Performance versus LIPS ( $\underline{r} = .16$ ) and WISC-R Full Scale versus the LIPS ( $\underline{r} = .43$ ). Only the latter was significant (at the .01 level). The first conclusion drawn was that the LIPS could not substitute for the Verbal or Performance sections of the WISC-R. The second conclusion was that the WISC-R Full Scale and the LIPS were not significantly different and suggest that the LIPS might well serve as a satisfactory substitute for the full WISC-R.

The  $\underline{t}$  test for related mean IQ scores obtained between the WISC-R and the LIPS was not statistically significant ( $\underline{t} = 1.12$ ). Therefore, one could conclude that the two intelligence tests are measuring essentially the rate of intellectual development of those students who participated in this study in a similar manner.

It appears from the results obtained that the LIPS and WISC-R are not interchangeable but that together they would provide a more thorough assessment of the superior individual and could better aid in his/her educational programming.

It is hoped that this study will encourage others to do further research with these instruments. It is anticipated that the superior functioning students within the public school system would be the beneficiaries of programs designed to more effectively meet their individual needs.

THE VALIDITY OF THE LEITER INTERNATIONAL PERFORMANCE SCALE IN MEASURING  
THE INTELLIGENCE OF INTELLECTUALLY SUPERIOR CHILDREN

by

SUSAN BRAXTON PALTROW

A thesis submitted in partial fulfillment of the  
requirements for the degree of

MASTER OF SCIENCE  
in  
PSYCHOLOGY

Portland State University

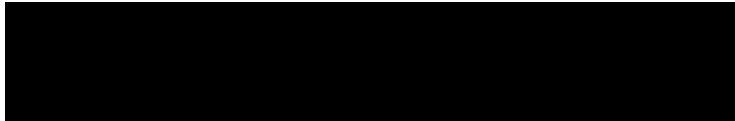
1980

TO THE OFFICE OF GRADUATE STUDIES AND RESEARCH:

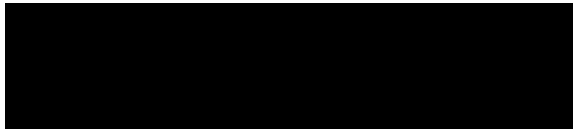
The members of the Committee approve the thesis of  
Susan Braxton Paltrow presented June 3, 1980.



Walter G. Klopfer, *W*Chairman

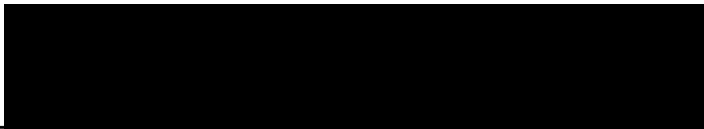


Ronald E. Smith

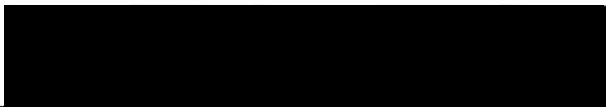


Richard Lazere

APPROVED:



Robert E. Jones, Jr., *R*Head, Department of Psychology



Stanley E. Rauch, *S*Dean of Graduate Studies

## ACKNOWLEDGMENTS

The author wishes to express her gratitude and appreciation to all the students and their parents for their time and their efforts. Special thanks to the students, because they spent several hours of their time taking intelligence tests.

My sincere appreciation goes to my chairperson, Walter Klopfer, who has also been my teacher and at times, my advisor. To my colleague, Marlin Pezner, for the loan of the Leiter Scales and her moral support and encouragement. To Georgene Inaba, who served as the Certified School Psychologist, supervising and teaching the use of the Leiter and even more of her time and effort when things looked so black. Kudos are also in order to the members of my committee, Ronald E. Smith and Richard Lazere, for their excellent advice. Although James Paulson was not a member of my committee, I wish to thank him for his faith in my ability to carry through the statistical analyses.

To Eugene Bates, Director of Special Services, Oregon City, Oregon, without whose support and cooperation this project would still be on the drawing board, goes my gratitude and thanks.

My appreciation to Linda McEwen for the many hours spent typing and for the excellence of the finished product.

Especially to my husband, Kenneth G. Paltrow, who provided emotional support and encouragement to bring this project to fruition. And, to the rest of the family for putting up with my mood swings and unavailability, as well as my intrusion into their space.

## TABLE OF CONTENTS

	PAGE
ACKNOWLEDGEMENTS . . . . .	iii
LIST OF TABLES . . . . .	vi
CHAPTER	
I INTRODUCTION . . . . .	1
The Leiter International Performance Scale . . . . .	2
Purpose of the Study . . . . .	9
Definitions . . . . .	10
Scope of the Study . . . . .	11
Method of Study . . . . .	12
Limitations of the Study . . . . .	13
II SURVEY OF THE LITERATURE . . . . .	15
Historical Overview . . . . .	15
Adaptation of the Leiter with Cerebral Palsy Victims . . . . .	15
Psychological Evaluation of the Deaf with the Leiter . . . . .	15
Stability and Validity with Language Delayed . . . . .	16
Comparison of the Leiter with Mental Defectives . . . . .	17
Comparison of Binet Revision with Normal Children . . . . .	17
Comparison of 1937 Revised Binet and Leiter . . . . .	18
Comparison of the 1937 Revised Binet, Leiter, Wechsler Intelligence Scale for Children (1949) and California Achievement Test and the Leiter with Superior Children . . . . .	18
III FINDINGS . . . . .	21

CHAPTER	PAGE
IV SUMMARY, CONCLUSIONS AND RECOMMENDATIONS . . . . .	25
BIBLIOGRAPHY . . . . .	30
APPENDIX A . . . . .	32
APPENDIX B . . . . .	33
APPENDIX C . . . . .	34



LIST OF TABLES

TABLE		PAGE
I	Comparison of Average Chronological Age Average	
	Mental Age . . . . .	5
II	Pearson Product-Moment Correlation Coefficients WISC-R	
	Verbal vs LIPS, WISC-R Performance vs LIPS, WISC-R	
	Full Scale vs LIPS . . . . .	22
III	Preferences . . . . .	23
IV	Male versus Female Pearson Product-Moment Correlation	
	Coefficients . . . . .	24

## CHAPTER I

### INTRODUCTION

Since 1905, when Binet and Simon constructed their intelligence scale to aid in the identification of feebleminded (currently referred to as developmentally delayed) school children, there have occurred numerous attempts to expand and refine measuring instruments which yield information about intelligence.

Public Law 94:142, the Education for All Handicapped Children of 1975 Law, provides powerful impetus for change in the educational system for handicapped children. It also articulates the right to education for all persons and spells out the financial obligations at the federal, state and local levels. Built into PL 94:142 are guidelines for the identification of these handicapped individuals, the requirement that handicapped pupils be educated in the least restrictive available environment commensurate with their abilities and needs, i.e., within the regular classroom. Although PL 94:142 does not address itself to the superior functioning students, it may be that this population ought to be included under the law as well.

Many of the administrators of programs for the superior functioning students would push for their inclusion under PL 94:142 as well, because the regular classroom might be a restrictive environment for these individuals. For example, in some public schools an enrichment program has been provided for these superior functioning students outside the regular classroom, while others have not addressed

this issue at present.

The present trend towards meeting the educational requirements of each individual child has warranted the use of more refined instruments with which to make assessments of intellectual abilities. Performance scales have been especially useful because of the opportunity for clinical observation of the subject. Freeman (1955) addresses this aspect:

Clinical psychologists are agreed that, where indicated, the use of performance scales can provide more information than just a rating in the form of a numerical index. These tests provide an opportunity to observe qualitative aspects of behavior under standardized conditions in a variety of situations. A subject's approach to a problem might reveal, for example, a state of depression or agitation; hesitation or impetuosity; thoughtful deliberateness, bull-headed persistence, or easy discouragement; an insightful approach or one of haphazard trial and error (p.609).

Correlations between the Stanford-Binet and performance scales are low positive and suggest that performance scales are not interchangeable with verbal-type intelligence tests. (Cronbach 1949). Nonetheless, attempts have been made periodically to devise performance scales which measure in a manner comparable to verbal-type measures.

#### The Leiter International Performance Scale

The Leiter International Performance Scale is a non-verbal test of intellectual functioning designed to measure functions of intelligence comparable to those verbal-type tests in situations where the verbal scale may be inappropriate. Werner (1965) states:

(a) It requires no verbalizations on the part of the examiner or the child. This makes it especially useful for the testing of children with speech and hearing difficulties, mental retardates, foreign-born or bilingual children, and shy or withdrawn children. (b) It has no time limits. (c) It reaches down to lower chronological age levels than other performance scales. (d) The tests lowest in the scale are tests of ability to learn rather than tests of acquired skills or materials already

learned. (e) The materials for the test are interesting to children (pp. 814-5).

The Leiter was first published in 1940. The 1940 form of the test was based upon a 1938 version which Dr. Arthur Leiter developed in Hawaii, using Japanese and Hawaiian children as his subjects. He did not use any Caucasian children, because he felt that the Caucasian population in Hawaii was almost entirely composed of professional people and their families. Thus, the population was not representative of the general white population of the United States.

Dr. Arthur Leiter's work in Hawaii, and his previous experiments in non-verbal test construction (his short 1936 scale which he claimed correlated .79 with the 1916 Binet), led him to assume that:

if the language factor in mental tests could be eliminated, the difference in native intelligence between children of various races, which had been found in previous investigators using tests which required the use of language, would no longer be found to exist (Leiter, 1952, p.10).

(The 1938 scale was composed of 56 tests in groups of four per year from year 3 to year 10, and even years from 12 to 22. Each test was worth three months mental age credit up to year 10, six months from years 12 to 16, nine months at year 18 and twelve months at year 20 and 22. This scale was the one administered to 764 Japanese and Chinese children.)

In the Fall of 1938, Dr. Leiter returned to the mainland and administered his test to 280 middle-class white children. The difference between the performance of the two groups was great enough, he felt, to demonstrate that the test was unsuitable for use with a Caucasian group. From Leiter's study (Goulard 1940) it is clear that he found that the chronological ages and mental ages of his Caucasian group were so disparate as to invalidate the test for this group. Consequently, he

abandoned his plans for developing an international test of performance and he undertook to develop a performance scale suitable for use with Caucasians. The revised 1936 scale became the basis for the 1938 scale. The 1938 scale, using 107 white children, had a correlation coefficient of .81 with the 1916 Binet. (McNemar 1949).

The major changes in the 1940 scale were (1) the relocation of tests already in the scale, and (2) the addition of new tests suitable for use about the ten year level. The 1940 scale was applied to 280 middle-class American white children in California who "were equally distributed between ages 5-0 and 12-11" (Leiter 1940). There were thirty-five subjects tested at each age. The reliability of the test was determined by finding the split-half method and comparing the standard deviations of scores at each age with the S.D. of the revised Stanford-Binet at the same age levels. The coefficient of reliability equalled .89 (the Spearman-Brown correction raised this to .94). The standard deviations are "well below those reported by Terman and Merrill at the same age levels". (Goulard 1940).

Further use of the 1940 revision indicated that the test would be easier to administer if tests appeared only at the even year levels beyond year ten.

Dr. Grace Arthur (1949) had also prepared a revision of the Leiter up to year twelve, and to simplify administration made some changes in the form of the test. The XI-4 was substituted for IX-2, XI-1 substituted for X-4, a single frame was used instead of six of varying lengths, a lighter carrying case was developed and no tests were given at the eleven year level. Children having a mental age up to 7.99 can be tested on either the Arthur adaptation or the 1948 Revision. Above that

TABLE I  
 COMPARISON OF AVERAGE CHRONOLOGICAL AGE  
 AND AVERAGE MENTAL AGE\*  
 (N = 35, all levels)

Age level (in years)	Average C.A. (in months)	Average M.A.
5	66.09	72.34
6	77.05	95.05
7	89.20	109.62
8	101.60	127.11
9	113.42	142.65
10	126.82	154.34
11	138.65	164.54
12	149.68	169.28

\*Comparison at each age level from 5 to 12 on 1938 scale when that scale was applied to 280 middle-class American white children. The table quoted from Table IX, page 13 of Leiter (1940).

age, it is necessary to use the 1948 version of the Leiter.

Dr. Arthur's changes made the Leiter scale for the younger age group a point scale, rather than a mental age scale.

Comparisons of the 1940 and 1948 revision led the author to state that the two tests were measuring the same type of "general intelligence" ( $r = .92$ ). (Goulard 1940).

Leiter (1940) states, however, that:

it must not be assumed that the intelligence quotients of the two scales are exactly comparable because whereas the Binet scale follows the theory that the amount of information children pick up through incidental learning is a good index of brightness, the LIPS is built on the principle that children's ability to cope with entirely new situations is a truer indication of their general intelligence (p.10).

Leiter (1969) states that the norm for the IQ on the LIPS is 95 for children in the continental U.S. Since the mean IQ obtained from other intelligence tests is 100:

the only practical thing to do was to add a constant, namely 5 points of adjusted IQ to the first obtained IQ. This, or the adjusted IQ, is the one that is always reported, but it is never labeled the adjusted IQ in a psychological report; it is given as the IQ obtained from the application of the Leiter International Performance Scale (p.4).

Below are listed the tests from year five to eighteen, since that was the lowest level reached in establishing a basal age for the subjects in this study. Complete instructions for administration and scoring can be found in the Leiter manual.

Year V

(4 tests, 3 months each)

1. Genus
2. Two color circles
3. Clothing
4. Block Design

Year VI

(4 tests, 3 months each)

1. Analogous progression
2. Pattern completion test
3. Matching on basis of use
4. Block Design

Year VII

(4 tests, 3 months each)

1. Reconstruction
2. Circle series
3. Circumference series
4. Recognition of age differences

Year VIII

(4 tests, 3 months each)

1. Matching shades of gray
2. Form discrimination
3. Judging mass
4. Series of radii

Year IX

(4 tests, 3 months each)

1. Dot estimation
2. Analogous designs
3. Block Design
4. Line completion

Year X

(4 tests, 3 months each)

1. Foot print recognition
- \*2. Block Design
3. Concealed cubes
- \*4. Block Design



## Year XII

(4 tests, 6 months each)

- \*1. Block Design
2. Similarities of two things
3. Recognition of facial expressions
4. Classification of animals

## Year XIV

(4 tests, 6 months each)

1. Concealed cubes
2. Analogous designs
3. Memory for a series
4. Form completion

## Year XVI

(4 tests, 6 months each)

1. Code for a number series
2. Reversed clocks
3. Dot estimation
- \*4. Block Design

## Year XVIII

(6 tests, 6 months each)

1. Position analogy
2. Dot estimation
3. Form completion
4. Concealed cubes
5. Spatial orientation
6. Concealed cubes

The figures in Appendix C (see pages 36-41) represent the frame and some of the test materials used in the 1948 Revision of the Leiter International Performance Scale.

\*Timed tests.

### Purpose of the Study

This study was undertaken to determine the validity of the Leiter International Performance Scale in measuring the intelligence of public school children from the age of six to age sixteen, who had been identified as functioning in the superior and above range of intellectual ability. The attempt to determine validity was to be accomplished by means of comparing the resulting scores on the Leiter International Performance Scale to those of the Wechsler Intelligence Scale for Children-Revised. The LIPS is a non-verbal test of intelligence often used with hearing and/or speech impaired, shy or withdrawn individuals. It was hoped that this study might confirm the general validity of the LIPS with superior functioning students, so that it could be used with more assurance when indicated, for the measurement of intelligence in this population.

Since 1974, when the Wechsler Intelligence Scale for Children was revised, no research had been conducted to determine the validity of the LIPS as compared to the WISC-R.

As previously stated, the current trend is to identify early any handicapping condition that might require an Individual Educational Plan for the student in the least restrictive environment (with additional requirements that the assessments include an assessment of adaptive behavior and observations of the student under diverse conditions).

Anastasi (1954) addressed the issue of the promise of the performance-type intelligence scale yielding additional information as to provide a more thorough intellectual assessment, versus the verbal-type instrument alone:

On the other hand, the "verbalist" type of individual may obtain a deceptively high score on certain verbal tests, although his understanding of most problems may be very superficial and his practical judgment may be seriously deficient. It is now generally recognized that performance or non-language tests are not simply a substitute for verbal tests. Each type of test predicts somewhat different criteria. Together, they provide a more complete picture of the individual and serve as mutual corrections in the evaluation of his test performance (1954,p.236).

Paul Witty (1951) has commented on the uniqueness and novelty of the test materials presented in the LIPS. He noted that it attempts to present items that would minimize previous learning and appear to require more individual management, control and organization of these new materials than do most verbal-type intelligence scales.

It would appear that a combination of a verbal and non-verbal type intelligence test might have value when assessing the intellectual capabilities of students who give evidence of superior intellectual achievement. The combination of verbal and non-verbal intelligence tests might provide for a more thorough and differentiated assessment of the intellectually superior group and allow for school placement decisions which will best serve the needs of the intellectually gifted individual.

### Definitions

"Gifted and talented children are those identified by professionally qualified persons who, by virtue of outstanding abilities, are capable of high performance. These are children who require differentiated educational programs and/or services beyond those normally provided by the regular school program in order to realize their contribution to self and society." U.S. Office of Education as a guideline.

Stoddard (1943): ".....ability to undertake actions that are characterized by (1) difficulty, (2) complexity, (3) abstractions, (4) economy, (5) adaptiveness, (6) social values, (7) the emergence of originals, and to maintain such actions under conditions that demand a concentration of energy and a resistance to emotional forces."

Wechsler (1944) states that "intelligence is the aggregate or global capacity to act purposefully, to think rationally, and to deal effectively with his environment....."

Leiter (1969) states that "general intelligence is the ability to solve problems with which an individual has had no previous experience."

Terman (1937) defines intelligence as "the ability to carry on abstract thinking."

Verbal test: A verbal test is one which involves the use of language, either written or spoken. Most pencil and paper tests require the use of written language but a few, such as the Porteus Maze Test, are non-verbal in character. The Stanford-Binet and Wechsler Intelligence Scales for Children-Revised are examples of individual verbal tests.

Non-Verbal tests: A performance test in which the subject is required to manipulate materials rather than to say or write something, is commonly spoken of as a non-verbal test.

Examples of this type of test are the Kohs Block Design Test, the Army Beta and the Leiter International Performance Scale.

### Scope of the Study

The subjects for this study included fifty-three public school students from the greater Portland area. The subjects ranged in age from six years, eleven months to sixteen years, eleven months. All of the subjects were Caucasian. No minority students volunteered for the study nor did it appear that any had been identified in the schools from which the population was drawn as "gifted and talented". Fifty subjects were selected (twenty-one female subjects and twenty-nine male subjects) according to the criterion set forth by Wechsler in his definition of superior functioning individuals. Those individuals whose Full Scale IQ's were 120 and above and who had been identified by the public school system as "gifted and talented" were the subjects of this study. Most of the subjects were drawn from the Clackamas County area. The director of Special Services in Oregon City lent his cooperation and support to

the project and through his efforts the bulk of the subjects were obtained.

### Method of Study

Each subject was informed as to the nature of the study. All students tested gave verbal consent for their participation. No students were tested without prior written consent from their respective parents or guardians. Individual parent conferences were made available to any participant in this research project who requested such a conference. Forty such conferences took place at the parents request and lasted anywhere from one half hour to one and one half hours.

All subjects were selected and were appropriately assigned after the intellectual testing. The Wechsler Intelligence Scale for Children-Revised (WISC-R) protocols were independently scored. Those subjects who achieved an IQ score at 120 or above were included in the study. The WISC-R and the Leiter International Performance Scale (LIPS) were administered to all subjects. The WISC-R was administered first to all subjects and the LIPS was administered on the same day, allowing for one subject to be tested per day. Three of the fifty-three subjects did not score an IQ of 120 and were, therefore, eliminated from this study since they did not meet the previously agreed upon criterion of intellectual ability.

The testing was conducted in well-lighted rooms with a minimum of outside distractions. In the case of some of the younger children, it was necessary to break for their recess periods and for all subjects a lunch break was included. Some of the elementary and high schools provided the examiner with access to a room where the testing could be

accomplished in somewhat familiar surroundings. In some instances, the testing was conducted in a testing room provided by Oregon City Special Services. In all cases, the conditions for testing were comparable to the usual atmosphere available for testing of students by the district's certified school psychologists.

The data resulting from the administration of these two tests were then compared by means of the Pearson Product Moment Correlation Coefficient to determine the validity of the LIPS in measuring the intelligence of superior functioning students.

The correlation coefficients were computed between the WISC-R Verbal versus the LIPS, the WISC-R Performance versus the LIPS, and the WISC-R Full Scale versus the LIPS.

#### Limitations of the Study

The most obvious and serious limitation is that the validity of the LIPS was determined by comparison with another instrument (WISC-R), so that errors inherent in the criterion instrument influenced the validation data.

The second limitation is the restricted range of the sample (Full Scale IQ Scores between 120 and 145) of twenty-five IQ points, and third, all were Caucasian students. However, it is interesting to note that the subjects involved in this study were from a broad range of socio-economic status; heads of household who are custodians, well-drillers, elementary school teachers, college professors, lawyers, architects and physicians.

The fourth, Sattler (1974) in discussing the limited floor and ceiling of the WISC-R states:

Another difficulty with the WISC-R is that the range of Full Scale IQ's (40 to 160) is insufficient, so that children who

have a mental age below six or who are gifted may not be properly assessed. However, the range of IQ's is greater than possible on the WISC, which yields a range of 46 to 154..... The highest IQ that can be obtained by children aged 16 years, 8 months and older is 158. As in the case of the WISC, it is likely that the ceiling on the WISC-R is too low to make the test appropriate for use with gifted children (p.157).

Fifth, Paul Witty (1940) writing in "School and Society" states:

If by gifted children we mean those youngsters who give promise of creativity of a higher order, it is doubtful if the typical intelligence test is suitable for use in identifying them. For creativity posits originality, and originality implies successful management, control and organization of new materials.....The content of the intelligence test is patently lacking in situations which disclose originality or creativity (p.504).

In some ways, the LIPS addresses the remarks by Paul Witty (1940) but, it too, does not allow for creativity in responses. The practicality of administering two individual intelligence tests which require a minimum of two hours and fifteen minutes of the examiner's and the subject's time would probably rule out using both tests except in very unusual circumstances.

## CHAPTER II

### SURVEY OF THE LITERATURE

The preponderance of the studies on the LIPS have been comparisons with the Stanford-Binet Intelligence Scale and its revisions (1916, 1937, Form L). Most of these studies have been conducted on populations with various physical and mental handicaps such as: cerebral palsy, deafness, language delay, brain-injury and mental deficiency.

Arnold (1951) studied twenty-five children with no apparent physical defects and with IQ's between 90 and 110 on the Revised Stanford-Binet Scale, Form L, and twenty-five subjects with no apparent physical defects and an IQ between 50 and 70 (mentally retarded). Arnold reports that the three groups were matched for chronological age and sex but does not report the range of IQ scores for the experimental group with the physical defect diagnosed as cerebral palsy.

The study reports that the examiner placed the blocks as directed by the child for the experimental group (cerebral palsied) rather than the child doing this for himself.

Arnold's conclusions are quoted below:

The results of this study showed that: 1. According to these data the adaptation techniques of administration did not invalidate the Leiter Scale and Maze Test. 2. The Leiter adaptation, the Maze adaptation, and the Binet score tended to rank subjects in the same general order. 3. Both of the adapted tests had a high index of reliability.....(p.177).

Birch and Birch (1951) were interested in studying the psychological evaluation of the deaf child's learning capacity and its direct



relationship with his educational program.

They studied fifty-three subjects to whom the LIPS and two or more intelligence tests had been administered. Their conclusions are below:

It appears that the Leiter scale gives IQ's which are considerably lower than those of the other tests commonly used with deaf.....(p.506). It may be that when the Leiter score is considerably below the scores of the Arthur and the Hiskey, one can predict that the child will be a serious teaching problem, oral speech and reading. Also, it may be that when both the Leiter and the Goodenough scores are low and the Hiskey and Arthur scores are high, the learning problems of the child will be even more serious (p.507).

Weiner (1971) reports on his study of the stability and validity of the Arthur Adaptation of the LIPS (AALIPS) and the Peabody Picture Vocabulary test with children whose language development was delayed.

His sample at the onset of the study consisted of thirty children who had been diagnosed as language delayed by a speech pathologist. The mean age in months when the AALIPS was administered for the first time was 54.70. At the second testing there were still thirty subjects and the mean age in months was 61.07. By the third test administration, the number of subjects dropped to twenty-two and the mean age in months was 79.18.

The AALIPS IQ did not change significantly on either retesting and the product-moment coefficient correlation between the results of test session one and two was 0.64 and between one and three was 0.63 ( $p < 0.01$  in each instance).

Weiner (1971), in his summary, states:

.....the AALIPS seems to be a highly useful test for determining the adequacy of nonverbal intellectual functioning of preschool, language delayed children who are similar to those included in the present (p.260).

He further concluded that the AALIPS showed reasonable stability

over time and could serve to predict later functioning on the IQ tests which assess a broad range of nonverbal behaviors. It appears to be applicable without restriction in group studies, but in the clinical setting is best limited to discriminating between normal and subnormal functioning.

Beverly and Bensberg (1952) did a comparison of the LIPS, the Cornell and Coxe Performance Ability Scale (1934), and the Revised Stanford-Binet, Form L (1937) with mental defectives.

Their study examined fifty students ranging in age from six years, eleven months to sixteen years, two months. The mean IQ on the Stanford-Binet was 58.16, Cornell-Coxe was 63.18 and the LIPS was 54.16.

The Pearson Product-Moment Correlation Coefficients were .67 between the Stanford-Binet and the Cornell-Coxe, .62 for Leiter versus the Stanford-Binet, and .82 for LIPS versus Cornell-Coxe.

They found significant differences between the three tests, with the LIPS scoring below the Stanford-Binet and Cornell-Coxe scoring above the Stanford-Binet. The suggestion was made that the LIPS norms for mental defectives lack some adequacy.

Glenn (1951) chose the Binet 1937 Revision, because of its reliability and also because he had access to children who had been given the Binet previously, as the instrument with which to compare the LIPS. His study was conducted with fifty-three, six-year-old children (19 girls and 34 boys).

He concluded:

The Leiter International Performance Scale does not differentiate between normal and above normal intelligence but may more safely be used to differentiate those children who are mentally deficient from those who have normal intelligence (p.26).

It appears that he tested three subjects whose IQ's fell in the superior range of intellectual ability and no conclusions could be drawn from such a small sample.

The study by Williams (1941) made use of fifty children in a Los Angeles school. Williams randomly picked every fourth child until she selected ten subjects at age five, six, seven, eight, nine and ten years. These fifty subjects were given the Stanford-Binet (1937 Revision) and the LIPS. She took into account past measures of intelligence when available and labeled these "Previous IQ's". She included all group tests as well as the former Binet scores.

A Pearson Product-Moment Correlation Coefficient was calculated to determine the degree of relationships between the LIPS and Binet IQ's. A comparatively high correlations of  $.67 \pm .078$  was found. Williams states (1941):

A  $r$  of this size gives a 26% reduction in error of prediction of one variable from another. -Hull says most are less than .70.

Williams found that the LIPS almost consistently underestimated the child's intelligence as measured by the Binet.

The only study that was directed specifically towards the superior functioning public school student was conducted by Earl F. Peisner (1956). His study compared the 1937 Revision Stanford-Binet with the LIPS, the 1949 version of the Wechsler Intelligence Scale for Children (WISC) and the California Achievement Test (CAT) on thirty-five "selected" superior sixth grade pupils.

Peisner's criterion instrument was the 1937 Revision of the Stanford-Binet and obtained IQ scores of 120 or above.

He did not delineate the breakdown between male and female

subjects included in his study. The subjects ranged in age from eleven years, one month to twelve years, four months, and were enrolled in a special class for gifted sixth grade pupils in a public school in Corvallis, Oregon.

Twenty-eight of his subjects earned an IQ score on the Stanford-Binet at 120 or above before this study was undertaken. An additional eight subjects were included after their achievement scores suggested that they could obtain IQ scores at 120 and above on the Stanford-Binet. The Stanford-Binet was then administered to these remaining eight subjects bringing the total to thirty-five subjects for this particular study.

Peisner states that the LIPS was administered first, since the subjects in this study had previously had the Stanford-Binet administered. He then administered the WISC, thereby separating the two verbal intelligence tests with the non-verbal test. He states that the CAT's had been administered five months prior to the LIPS administration.

Peisner reports the following Pearson Product-Moment Correlation Coefficients between the LIPS and the WISC: WISC Full Scale Score  $\underline{r} = .60$ , WISC Verbal  $\underline{r} = .64$ , and WISC Performance  $\underline{r} = .55$ , all values significant at the .01 level.

Peisner concluded:

The primary conclusion to be drawn from the results of this study is that, with samples like the one employed and with criteria comparable to the standardized scales employed, the validity of the Leiter International Performance Scale in measuring the intelligence of superior children would likely be low. This conclusion would still be warranted if, as Leiter suggests, five I.Q. points are added to an individual's I.Q. score earned on the Leiter Scale.....(p.39).

.....Another conclusion is offered on the basis of the writer's observations during the administration of the intelligence scales. Some of the Leiter tests, such as the Form Completion Test, Year XIV, and the Concealed Cubes Test, Year XVIII,

because of their difficulty and novelty at the preadolescent level, demand a degree of adaptiveness and persistence which does not appear to be required on the performance items of the Wechsler (p.40).

## CHAPTER III

### FINDINGS

The questions asked at the onset of this research project were:

- 1) Is the LIPS interchangeable with the WISC-R in identifying superior functioning children in the public schools?
- 2) Can the LIPS substitute for either the verbal and/or performance sections of the WISC-R?
- 3) Does the LIPS give information in addition to the WISC-R that is valuable?

The findings of this study were reported via the results of the computations of the correlation coefficients between the WISC-R Verbal scores versus the LIPS, the WISC-R Performance scores versus the LIPS, and the WISC-R Full Scale scores versus the LIPS.

The Pearson Product-Moment Correlation Coefficients are set forth in Table II.

TABLE II  
 PEARSON PRODUCT-MOMENT CORRELATION COEFFICIENTS  
 (N = 50)

LIPS	WISC-R	RESULTS
Range: 110-161	115-152 Verbal	$\underline{r} = .28^*$
110-161	108-139 Performance	$\underline{r} = .16^*$
110-161	120-145 Full Scale	$\underline{r} = .43^{**}$

\* not significant

\*\* significant at .01 level

With the limited range on the WISC-R at 25 points and the LIPS range at 51 points,  $\underline{r} = .43$  is a high correlation and would suggest that the two intelligence tests are measuring the same intellectual abilities.

Even though there is a significant correlation between the Full Scale Score on the WISC-R and the LIPS IQ scores, one could not conclude that the LIPS is interchangeable with the WISC-R unless the subjects in the study were similar to those in this study who had already scored above 120 IQ points.

The insignificant correlations between the WISC-R Verbal and Performance scores when compared with the LIPS indicate that the LIPS could not be substituted for either with any assurance of accuracy.

A t test for related mean IQ scores was obtained on the WISC-R Full Scale Scores and the LIPS IQ score with the following results:

The mean scores obtained on the WISC-R Full Scale and the LIPS are not significantly different ( $t = 1.12$ ); therefore, one could conclude that the two intelligence tests are measuring essentially the rate of intellectual development in a similar manner.

When the testing was completed, the examiner asked each subject the following question: "If, at some time in the future, you were given the opportunity to choose between the two tests you have just taken, which would you choose?" Their preferences are listed below.

TABLE III  
PREFERENCES

	Prefer LIPS	Prefer WISC-R
Females:	5	16
Males:	<u>3</u>	<u>26</u>
Total:	8	42

It is also interesting to note that two sixteen-year-old males stated that they were having a difficult time making this decision and that they thought they would prefer a combination of the two tests rather than take either alone. They mentioned that they thought the two tests were measuring the same "thing", but approaching it from a different angle.

Another interesting finding presented itself when the subjects were divided between females (N = 21) and males (N = 29) and the Pearson



Product-Moment Correlation Coefficients were computed on the WISC-R Full Scale IQ scores versus the LIPS IQ scores obtained.

TABLE IV  
PEARSON PRODUCT-MOMENT CORRELATION COEFFICIENTS  
WHEN DIVIDED BY SEX

	LIPS	WISC-R	RESULTS
Range	110-165	120-145	$\underline{r} = .52^{**}$
Females			
(N = 21)			
Range	111-152	120-144	$\underline{r} = .30^*$
Males			
(N = 29)			

\* not significant

\*\* significant at .02 level

One could conclude that the more limited range for the male subjects greatly influenced the results of the correlations obtained.

## CHAPTER IV

### SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

The most significant question addressed in this study is: Is the LIPS a valid measure of the intelligence of superior functioning children, the criterion of validity being the WISC-R?

The method employed when investigating the question stated above was the administration of the WISC-R and the LIPS to fifty subjects enrolled in public schools in Clackamas County, Oregon. Forty-three of the subjects were selected on the basis of prior placement in classes for "gifted and talented" students. Seven of the subjects were candidates for admission to the "gifted and talented" programs. These students had been tested with the Slosson Intelligence Test (1963) and had obtained IQ scores at 120 and above.

Admission to such classes were said to be dependent upon receiving high scores on one or more of the following tests: Otis-Lennon Mental Ability Test (1959), (a group test, high score not defined), high achievement test scores (not defined and a group type test), or by teacher recommendations, which led to individual testing with the Slosson Intelligence Test (1963).

Two specific methods were employed to address the validity of the LIPS as compared to the WISC-R; (a) the Pearson Product-Moment Correlation Coefficient was computed between the WISC-R Verbal, Performance and Full Scale IQ scores obtained and the LIPS IQ scores obtained, (b) a t test for related means was computed in order to determine whether or not

there was a significant difference between the mean IQ scores obtained on the WISC-R and LIPS IQ tests.

These correlation coefficients revealed that there was no significant relationship between the WISC-R Verbal and Performance IQ scores obtained compared to the LIPS IQ scores obtained. There was a high correlation, considering the limited range of IQ scores obtained, between the WISC-R Full Scale IQ score and the LIPS ( $r = .43$ ) significant at the .01 level.

The  $t$  test for related measures resulted in a  $t = 1.12$ , which is not a significant difference between the mean Full Scale IQ scores on the WISC-R and the LIPS for this particular population.

The above reported results suggest that the LIPS cannot be substituted for either the Verbal or Performance sections of the WISC-R alone when testing superior functioning students, but might well serve as a satisfactory substitute for the full WISC-R.

The results lead this author to conclude that for a more thorough and accurate picture of the intellectual ability of the superior functioning student, the administration of both the WISC-R and the LIPS would be worthwhile in planning the superior functioning student's individual educational program. The author recognizes the impracticality of routinely administering both a verbal and non-verbal type intelligence test except under unusual circumstances. The study does suggest that the LIPS might serve as a viable test of intellectual ability for superior functioning students who for various reasons cannot respond to the WISC-R.

This investigation of the validity of the LIPS in assessing the intellectual ability of superior students has yielded some interesting observations which may be of value to persons who employ the LIPS in

their work:

1) The "Similarities; two things" test Year XII-2 appears to be easier for the subjects below the age of fourteen. They appear to approach the problem in a more simplistic manner, therefore, passing the test. The older subjects tended to look for a more complex solution and sometimes became quite frustrated although most were able to complete the task correctly.

2) The subjects below age fourteen had more difficulty attacking the problem presented at Year XII-4, "Classification of animals", often misplacing the seal and the frog.

Scoring. The test is passed if the blocks are arranged in the following order: owl, squirrel, seal, dog, frog, bee. This is the only arrangement for which credit is allowed. The following are representative interpretations of the items in this test which were made by subjects who passed the test: (1) the owl goes with the bird because both are birds; (2) the squirrel goes with the rat because both are rodents, both live in holes, both are animals; (3) the seal goes with the fish because both are sea animals; (4) the dog goes with the tiger because both are larger animals; (5) the frog goes with the alligator because both are amphibians, both live in swamps and marshy places; (6) the butterfly goes with the bee because both are insects. (Leiter 1965 Manual, p.48).

3) Some of the subjects expressed their opinions regarding how the Year XIV-2, "Analogous designs" ought to be solved. They frequently stated that the blocks (coded by the examiner, reading from left to right) one and three ought to be reversed.

4) It was found that when testing subjects whose chronological age was fourteen years or more that it was necessary to start the testing at Year X, in order to be sure that the subject understood how to go about solving the "Block design" subtests and the "Concealed cubes" test (as the manual states that XIV-1 "Concealed cubes" must be preceded by X-3, "Concealed cubes"). This usually meant that superior functioning

students could be confronted with "Dot estimation" three times during the test administration, (at IX-1, XVI-3 and SVIII-2), and not only did the novelty wear off, but the subjects tended to become negative and hostile towards these subtests, sometimes becoming haphazard in their approach or not even attempting the task at all.

5) Not one subject of the fifty tested was able to successfully complete the Year XVIII-5, "Spatial orientation" test even if the subject figured out that the solution lay in matching right and left. Most of the subjects looked for a more complex solution and, thereby, failed that subtest.

6) The scoring system does not allow for the examinee to get partial credit on any subtest even if he/she gets seven out of eight correct responses, and this seems to be an unwarranted penalty.

This author would recommend that if a subject at or above fourteen years of age was able to successfully complete the first two presentations of "Dot estimation", to eliminate the negativism and hostility that appears in most cases with the third presentation, that they automatically receive credit for the third, as does a perfect execution of the "Block Design" Year V-4 automatically receive credit at Year IX-3.

Further, the penalties for placing one block out of order are too great. A method of scoring which would enable the examinee to earn at least partial credit would seem more appropriate and would not adversely effect the objective scoring system now employed.

This investigation with superior functioning students who are usually highly verbal, and in this study all were, revealed that the conditions under which the LIPS is administered at times created some unnecessary anxiety and negativism. In the future, it would be to the

examiner's benefit to be prepared to spend an extended period of time establishing rapport with the student when they consider using the LIPS with the superior functioning population.

It appears that the LIPS is a test of verbal and non-verbal conceptual abilities and further research could be designed to investigate this hypothesis.

## BIBLIOGRAPHY

- Anastasi, Anne. *Psychological Testing*. New York, MacMillan, 1954, 682p.
- Arnold, Gwen F. "A Technique for Measuring the Mental Ability of the Cerebral Palsied", The Psychological Service Center Journal, Volume III No.4, September, 1951.
- Arthur, Grace. "Arthur Adaptation", J. Clinical Psychology, (October, 1949), 5:345-349.
- Beverly, Louise and Gerald J. Bensberg. A Comparison of the Leiter, the Cornell-Coxe and Stanford-Binet with Mental Defectives. The American Journal of Mental Deficiency. 57:89-91. 1952.
- Birch, Jane R. and Birch, Jack W. "The Leiter International Performance Scale as an Aid in the Psychological Study of Deaf Children." American Annals of the Deaf 96:502-11; November, 1951.
- Buros, O.K. (Ed.), The Sixth Mental Measurements Yearbook. Highland Park, New Jersey, The Gryphon Press (1965).
- Cronbach, Lee J. *Essentials of Psychological Testing*. New York, Harper and Brothers, 1949. 475p.
- Freeman, Frank S. *Theory and Practice of Psychological Testing*. New York, Henry Holt, 1955. 609p.
- Glenn, Robert T. A Comparison of Intelligence Quotients Derived by the Leiter International Performance Scale and the 1937 Stanford Revision of the Binet. Master's thesis. Pittsburg. University of Pittsburgh, 1951. 42 numb. leaves.
- Goulard, L.J., A Study of the Intelligence of Eleven and Twelve-Year-Old Mexicans by Means of the Leiter International Performance Scales, unpublished Master's thesis, The University of Southern California, 1940.
- Leiter, R.G., A Comparative Study of the General Intelligence of Caucasian, Chinese and Japanese Children as Measured by the Leiter International Performance Scale, unpublished Doctoral dissertation, The University of Southern California, 1938.
- Leiter, R.G., *The Leiter International Performance Scale, Volume I* (Santa Barbara, California, Santa Barbara State College Press, 1940).
- Leiter, R.G., "Caucasian Norms for the Leiter International Performance Scale", The Psychological Service Center Journal, Volume I No. 4 December, 1949.

- Leiter, R.G., Manual, 1948 Revision, Volume II, Washington, D.C. Psychological Service Center Press, 1952.
- Leiter, R.G., General Instructions for the Leiter International Performance Scale, California, Western Psychological Services, 1969. 4p.
- McNeMar, Quinn, Psychological Statistics, New York: John Wiley and Sons, Inc., 1949.
- Peisner, Earl F., The Validity of the Leiter International Performance Scale in Measuring the Intelligence of Selected Superior Children, Doctoral thesis, Oregon State College, June 1956.
- Sattler, J.M., Assessment of Children's Intelligence, Saunders, Philadelphia, London, Toronto, 1974. 519p.
- Stoddard, G.D., The Meaning of Intelligence. New York, MacMillan, 1943. 504p.
- Terman, L., Intelligence and Its Measurement. Journal of Educational Psychology 12:123-147. 1921.
- Terman, Lewis and Maud A. Merrill. Measuring Intelligence. New York: Houghton Mifflin Co., 1937.
- Wechsler, David, Measurement of Adult Intelligence, Third Edition. Baltimore, Williams and Wilkins, 1944, 519p.
- Weiner, Paul S., "Stability and Validity of Two Measures of Intelligence Used with Children Whose Language Development is Delayed." Journal of Speech and Hearing Research 14:254-261; 1971.
- Werner, Emmy E. "Review of the Arthur Adaptation". The Sixth Mental Measurements Yearbook. Highland Park, New Jersey: The Gryphon Press (1965).
- Williams, Ruth M., A Clinical Examination of the Leiter International Performance Scale, Master's thesis, The University of Southern California, June 1941.
- Witty, Paul. Contributions to the I.Q. Controversy from the Study of Superior Deviates. School and Society 51:504; 1940.
- Witty, Paul (ed.) The Gifted Child. Boston, D.C. Heath, 1951, 338p.



APPENDIX A

SUBJECT #	AGE	SEX	VERBAL	PERFORMANCE	FULL SCALE	LEITER	OCCUPATION OF HEAD OF HOUSEHOLD
801	10-2	M	119	135	130	115	
802	10-10	F	125	126	128	130	Law clerk - Artist
803	11-8	F	118	120	121	116	Well driller
804	10-2	F	127	121	127	132	
805	8-7	M	139	123	134	130	Teacher
806	8-4	M	145	124	139	152	Pipe-fitter - Housewife
807	11-11	M	147	126	141	122	Coach
808	12-11	F	120	121	123	115	Computer
809	11-8	M	140	139	144	142	Custodian
8010	14-6	M	125	108	120	113	Insurance
8011	11-2	M	120	115	121	116	
8012	8-3	M	135	129	135	129	
8013	9-11	M	130	130	133	161	Airline ticket agent
8014	10-3	F	149	129	143	125	Ph.D. Forest Service
8015	7-11	F	141	141	145	156	Ph.D. Forest Service
8016	12-11	M	115	121	121	113	Teacher
8017	14-4	M	135	121	131	136	Electrician
8018	7-10	M	131	108	123	129	
8019	7-10	F	136	128	135	129	Artist - R.N.
8020	11-6	F	122	128	135	129	Lawyer
8021	13-9	M	131	112	125	128	Secretary
8022	11-11	F	135	121	131	110	Insurance
8023	7-7	M	127	121	127	143	Architect

APPENDIX B

INFORMED CONSENT

I (We) as the parent(s)/guardian of \_\_\_\_\_, hereby agree to allow \_\_\_\_\_ to be tested with the Leiter International Performance Scale and the Wechsler Intelligence Scale for Children-Revised, as a participant in the research project entitled: "The Validity of the Leiter International Performance Scale in Measuring the Intelligence of Superior Children".

I understand that the study involves approximately three hours of the student's time in the taking of a non-verbal and a verbal intelligence test.

It has been explained to us that the purpose of the study is to learn whether or not the Leiter is a useful tool in identifying superior functioning children in the public school system.

We may not receive any direct benefit from the participation of our child in this study, but his/her participation may help to increase knowledge which may benefit others in the future.

Susan Braxton Paltrow has offered to answer any questions we may have about the study and what will be required of our son/daughter in the study.

We have read and understand the foregoing information.

Date: \_\_\_\_\_

Signature: \_\_\_\_\_

Date of birth: \_\_\_\_\_

## APPENDIX C

SUBJECT #	AGE	SEX	VERBAL	PERFORMANCE	FULL SCALE	LEITER	OCCUPATION OF HEAD OF HOUSEHOLD
8024	16-7	M	143	117	134	136	Lumber salesman
8025	7-4	M	152	129	144	131	Physicist
8026	10-7	M	145	111	132	142	Sheetmetal worker
8027	12-1	M	131	112	125	125	Longshoreman
8028	9-11	F	133	138	140	148	Teacher
8029	11-5	M	130	136	138	128	Biologist
8030	16-9	M	139	124	135	136	College professor
8031	16-4	M	140	131	140	140	
8032	10-7	F	127	131	133	137	
8033	6-11	M	135	112	126	117	Teacher
8034	9-4	F	133	121	130	140	Physician
8035	7-3	M	147	121	139	118	Physician
8036	13-11	M	142	124	138	131	City employee
8037	9-5	M	131	108	123	140	Physician
8038	11-8	M	135	132	138	111	Physician
8039	14-10	F	149	129	143	136	College professor
8040	11-10	F	124	112	121	119	Newspaper publisher
8041	7-1	F	142	132	142	165	Engineer
8042	7-11	F	133	128	133	138	Sales
8043	15-11	M	135	112	126	124	Teacher
8044	16-10	F	139	130	139	140	College professor Ph.D.
8045	11-4	F	139	111	128	128	Insurance agent
8046	14-6	F	136	115	128	113	
8047	16-11	M	133	129	131	140	Car dealer
8048	8-3	F	145	128	141	117	
8049	15-4	F	142	133	142	132	Biologist
8050	11-2	M	139	121	133	126	Sales

