

Student Work

5-1-1972

The Validation of the Omaha Public School System Special Education Test

John Albert Hippe Jr.
University of Nebraska at Omaha

Follow this and additional works at: <https://digitalcommons.unomaha.edu/studentwork>
Please take our feedback survey at: https://unomaha.az1.qualtrics.com/jfe/form/SV_8cchtFmpDyGfBLE

Recommended Citation

Hippe Jr., John Albert, "The Validation of the Omaha Public School System Special Education Test" (1972). *Student Work*. 2403.
<https://digitalcommons.unomaha.edu/studentwork/2403>

This Thesis is brought to you for free and open access by DigitalCommons@UNO. It has been accepted for inclusion in Student Work by an authorized administrator of DigitalCommons@UNO. For more information, please contact unodigitalcommons@unomaha.edu.

THE VALIDATION OF THE OMAHA PUBLIC SCHOOL SYSTEM
SPECIAL EDUCATION TEST

A Field Study

Presented to

The Faculty of the Graduate College
University of Nebraska at Omaha

In Partial Fulfillment

of the Requirements for the Degree
Specialist in Educational Administration

by

John Albert Hippe, Jr.

May 1972

UMI Number: EP73948

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI EP73948

Published by ProQuest LLC (2015). Copyright in the Dissertation held by the Author.

Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against unauthorized copying under Title 17, United States Code



ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

Accepted for the faculty of The Graduate College of the University of Nebraska at Omaha, in partial fulfillment of the requirements for the degree Specialist in Education.

Graduate Committee Kenneth Burkholder Ed. Admin
Name Department
Robert Ackerman - Chem. & Earth Sci.

Darrell Kellam Ed Ad
Chairman

PREFACE

The writer wants to offer his appreciation for the cooperation received from the principals of the schools where students were tested. Thanks is also offered to Mr. Dale Samuelson, Mrs. Richard Galusha, Mrs. Geraldine Nesvan, and Mrs. Annie Green for their aid in performing this study. A very special thanks to my building principal, Mrs. Eleanor Snellenberg for her help and understanding during the efforts of this study. Also, gratitude should be expressed to the Omaha Public Schools for the opportunity to conduct this study. Last, a special appreciation to my major adviser, Dr. Darrell Kellams, and to Dr. Robert Ackerman and Dr. Kenneth Burkholder for their guidance and counsel during this study.

TABLE OF CONTENTS

Chapter	Page
1. INTRODUCTION	1
STATEMENT OF THE PROBLEM	2
HYPOTHESIS	3
ASSUMPTIONS	3
LIMITATIONS	3
DEFINITIONS	4
ORGANIZATION OF STUDY	4
2. RELATED LITERATURE	5
3. THE STUDY	20
The Procedure	20
The Findings	21
4. CONCLUSIONS AND RECOMMENDATIONS	23
Conclusions	23
Recommendations	24
BIBLIOGRAPHY	25
APPENDIX	28

Chapter 1

INTRODUCTION

Teachers are usually first to detect the inadequate performance of some children in reading, spelling, and writing. Many parents also recognize that something is wrong and desire to have the situation corrected. It is therefore the obligation of public schools to provide education for all children, and so it becomes the responsibility of school administrators and teachers to find and administer the necessary techniques to realistically meet a student's individual needs.

With failure in the language skills, the gap between grade level and academic achievement tends to widen. By the time the failures of the ten, eleven, and twelve-year olds are fully recognized, all concerned are aware of the need for help. Where assistance is available, children may be sent to medical or psychiatric clinics for diagnosis. Too often, however, they are left to "sink or swim," to waste formative years, become dropouts, or never attain their full potential. They often become discouraged, resentful, and embittered people resigned to mediocrity and feelings of inferiority. Emphasis must therefore be placed upon early identification and placement in the proper program, before an individual's problems become too complex. If a child's problem cannot be identified and remediated on the elementary level, what chance does he have in the complex secondary programs of our modern high schools? No longer can teachers wait for the child to suffer continuous academic failure before being referred

for special help. Early identification of problems and preventive practices are the only solution.¹

Screening tests should be administered to groups of children to detect reading disability--severe or borderline. These tests are for screening, rather than testing in depth. They are not meant to take the place of clinical testing when it appears needed, or available.²

The Wide Range Achievement Test and the Omaha Public School System Special Education Test are designed to determine the instructional reading level of primary and intermediate educable mentally handicapped children. The results of the tests provide guidelines as to what reading level materials should be used by the child. Therefore, individualization of the child's curriculum can be implemented.

STATEMENT OF THE PROBLEM

The purpose of this study was to verify the validity of the Omaha Public School System Special Education Test for the identification and placement of primary and intermediate educable mentally handicapped children on a certain reading level. These students have had intelligence tests administered by school psychologists to qualify them for placement in special education classes. The students must have an IQ which ranges between 60 and 85 to qualify for placement in

¹Bernie Straub, and Jerome Hellmuth, Educational Therapy (Seattle: Special Child Publications, 1966).

²Bernie Straub, and Jerome Hellmuth, Disadvantaged Child (Seattle: Special Child Publications, 1967).

an educable mentally handicapped classroom.

HYPOTHESIS

The Omaha Public School System Special Education Test is a valid instrument for determining the instructional reading level for primary and intermediate educable mentally handicapped children enrolled in Omaha Public School System Elementary Schools.

ASSUMPTIONS

1. All educable mentally handicapped children who attend the Omaha Public Schools are required to take the Omaha Public School System Special Education Test.

2. The sample included all the primary and intermediate educable mentally handicapped children who were attending Omaha Public Schools during April and May, 1971.

3. Every child who took the Omaha Public School System Special Education Test had the test administered with identical procedures and given by the same person.

4. The Omaha Public School System Special Education Test and the Wide Range Achievement Test are designed to show instructional reading levels.

LIMITATIONS

1. This study was limited to primary and intermediate educable mentally handicapped school children attending Omaha Public Schools.

2. This study was limited to the area of reading and did not concern any other subject.

DEFINITIONS

SPED--the title of the Omaha Public School System Special Education Test.

WRAT--the title of the Wide Range Achievement Test.

ORGANIZATION OF STUDY

The main purpose of the opening chapter is to introduce the study. It is very important to analyze the literature of a given field of study, therefore, the second chapter attends to this particular task. Chapter three simply presents the study with regard to procedure and findings. Finally, the last chapter indicates the conclusions made by the writer and recommendations to be made to the Omaha Public School District.

Chapter 2

RELATED LITERATURE

A coefficient of reliability is the most common fact reported concerning a test and certainly it is much easier to obtain than is a coefficient of validity or factor loading. Reliability is the minimum information one should have concerning a test, but it is certainly not the most useful information. If it is quite high, it can sometimes give the test user a false sense of security. For example, should one feel very comfortable about a test whose r_{tt} is .90, but whose proportion of variance in the factor it wishes to measure is only .30? What scores mean and what we can predict from them is entirely restricted to the relevant factors.

What, then, is the importance of knowing a coefficient of reliability? It is safe to say that in developing the vast majority of tests constructed today the makers strive toward internal consistency. The coefficient r_{tt} indicated how well the test constructor has approached that goal. A spread of scores is no indicator, since the spread may be due entirely to error variance. This last statement is reasonable when we remember that each individual could decide the answer to every item (in a multiple-choice test) by the throw of a die and there could still be a dispersion of scores.

There are certain tests and measures concerning which reliability is the most important thing to be known. Achievement

examinations are usually regarded as self-validating in the sense that they were tailored to measure the incomes of well-identified objectives in training or education. They serve their own criteria. To know their accuracy of measurement in terms of index of reliability, then, is of first importance. Accuracy is the only question raised when assuming the selection of relevant material. Tests used in clinical diagnosis should be known to have a very substantial degree of reliability. Tests used in personnel classification and in vocational guidance should also be highly reliable. In the latter, it is differences between scores that count, as in the ups and downs on a profile. We shall see that different scores are almost always lower in reliability than the test scores from which they come, sometimes very much lower. Much depends upon the ratio of the reliabilities of the separate scores to their own correlations.

As to how high reliability coefficients should be, no hard-and-fast rules can be stated. For research purposes, one can tolerate much lower reliabilities than one can for practical purposes of diagnosis and prediction. For some purposes, even a test of low reliability adds enough to the prediction to justify its use, particularly when it is given with a battery of other tests.

It is sometimes said that reliability is important because it contributes to validity and that validity is the important goal. The relationships between reliability and validity are by no means simple and not always direct. If a test is found to be valid because it predicts some practical criterion, we then add to it more of the same kind of true variance, thus making it more reliable, and also adding to its validity.

Reliability should be the minimum information provided concerning a test and is not to be considered the most important or most useful information. Reliability coefficients do not, as a rule, represent the psychological stability of a test. The overriding factors determining the size of a reliability coefficient, regardless of the method used, are primarily statistical in nature. We must, therefore, distinguish between clinical reliability and statistical reliability.

The clinical reliability of a psychological or educational test will not be known until the validity variances of an ability are accurately determined. Statistical reliability on the other hand is a function of the following conditions: the number of items in a test, the time it takes to complete the test, the population sampling used, the range of ability tapped, the specific technique of comparing or splitting two forms, the manner of scaling of the items within the test, the method of answering the test items, the unit of measurement, and the dispersion of item difficulties within a scale.

The Wide Range Achievement Test satisfied the statistical conditions of reliability most adequately. Numerous population groups of different degrees of homogeneity have been studied by the authors of the above-mentioned test during the past twenty years. The correlation coefficients (when in such homogeneous groups as policemen and nurses) ranged from .92 to .98 for the reading and spelling tests.³

³J. P. Guilford, Psychometric Methods (New York: McGraw-Hill Book Co., Inc., 1954), pp. 388-389.

The WRAT was studied rather intensively by DeLong, especially concerning:

The changes in test scores from administration to administration of educable mentally retarded children to determine (a) the extent of such variations for individuals on various tests and (b) if such variations can be ascribed, at least in part, to individuals rather than entirely to the tests.

A group of seventy-seven retarded persons ranging in age from fifteen to seventeen years was given five successive administrations of five standardized tests within a three week period. Among these five tests was the WRAT. DeLong tested the repeat findings for homogeneity of variance of raw scores using Bartlett's test and found the results significantly heterogeneous. He concludes, therefore, that significant changes in scores tend to occur from one administration to another.

The WRAT showed the smallest variations of all tests, when analyzing the average differences between the high and low scores of the individuals who took all five administrations. These differences for the total group, as well as for the subgroups, were found to be statistically significant for all scales except the WRAT. Mid-scores and low scores were compared for each individual. Again, the WRAT differences were the only one not significant. Seventy-three of the seventy-seven subjects were found to vary less than ten per cent from one WRAT administration to another.

On the basis of clinical experience, and some validity calculations made in the past, the most reasonable guess concerning the clinical reliability of the WRAT is that the coefficients vary from .90 to .95 for each subtest with an average of .93.⁴

⁴A. R. DeLong, The Limits of Accuracy of the Test Scores of Educable Mentally Retarded Individuals (Journal of the Association for Research in Growth Relationships, 1962), pp. 3, 26-44.

The WRAT is distinctive in that it measures reading, spelling, and arithmetic skills from kindergarten to college, and may thus be said to employ the principle of the power scale. It is designed primarily as an individual clinical test. A valuable feature of this instrument is the provision made for analyzing personality adjustment from uneven scores in the three parts of the test. This feature has apparently been made possible by checking test results with clinical case studies in the Delaware State Mental Hygiene Clinic and elsewhere.

The author makes the defensible claim that a word-pronunciation test is superior to both paragraph and picture-reading tests as a measure of how well a child can read.

Item analysis was made by the method of computing the percentage of successes at each grade level. Reliability coefficients, which were determined by repeated testing, ran from .88 (120 cases) for spelling to .95 (110 cases) for reading. One might question the use of clinical cases (280) in the determination of internal consistency because of the uneven scores made by many emotionally maladjusted individuals.

Standardization of test administration regards the reading test as being more crucial than the others. Word recognition test correlates to a greater extent than other tests with future achievement.

The manual of directions of the WRAT has provided a very helpful analysis of (a) strephic, (b) phonic, (c) ideational, and (d) expressional reading errors, especially as these relate to slow introverted and fluctuating extroverted types of individuals. It also shows how varying achievement patterns may be suggestive of specific personality anomalies.⁵

⁵Louis P. Thorpe, The Third Mental Measurement Yearbook (New Brunswick, N. J., Rutgers University Press, 1949), p. 21.

The WRAT, often administered in clinics to obtain an estimate of reading ability, was included in the battery because its validity could be demonstrated; the Columbia Mental Maturity Scale (CMMS) and Full-Range Picture Vocabulary Test (FRPV) might prove to be valuable predictors of educability even if they did not correlate well with the Wechler Intelligence Scale for Children (WISC). The WRAT might also contribute to indicating its relative worth in comparison with the WISC, since the WISC might have a spuriously high correlation with the Full Scale Test of which it is a part.

One hundred students ranging in age from six to eleven and eight to ten, fifty-one male and forty-nine female, in four second grades of the white elementary schools of Newport News, Virginia, were individually tested at the rate of one child per day between September, 1959, and April, 1960. The WISC, CMMS, FRPV, and WRAT were given in one session, averaging about an hour and a quarter. Half the children were given the WISC first to obtain a possible criterion of their reading ability. School authorities furnished information regarding the grade level of the twelve readers on the list.

The means, standard deviations, and product-moment intercorrelations of all variables were obtained. Partial correlation was used to determine the effect of age on correlations between IQs. Multiple correlation coefficients were calculated to demonstrate that an average was the best weighing of two-variable combinations of certain tests.

All tests correlated significantly with the WRAT, ranging from the CMMS r of $.30 \pm .06$ to the WISC r of $.61 \pm .04$. The average of CMMS

and FRPV IQs correlates $.56+$ to $.046$ with the WRAT.⁶

Wagner of Richmond and McCloy of the Mobil Psychiatric Clinic have compared the reading test of the WRAT with the Woody-Sangren Silent Reading Test and the New Stanford Reading Test.

Wagner's coefficient for twenty-nine children was $+.78$. McCloy's coefficient for fifty-seven children was $+.74$ with the Woody-Sangren and $+.80$ for forty-seven children with the New Stanford Reading Test.

All three coefficients were significantly beyond the $.01$ level of confidence.⁷

The WRAT appears to have adequate face validity as coarse screening measure in three restricted areas: spelling, word pronunciation, and arithmetic computation. With a range from kindergarten to college, it is obvious that the number of pupils per grade on which it has been standardized is, likewise, relatively small. Its unique "individual" feature is limited to the word-pronunciation section and the test is apparently adequate.⁸

The WRAT is designed for the use of clinical workers and purports to measure the school achievement of any individual (from kindergarten to college) in a total of not more than forty-three minutes

⁶Bessie S. Smith, The Relative Merits of Certain Verbal and Non-Verbal Tests At The Second-Grade Level (Journal of Clinical Psychology, 1961), pp. 17, 53-54.

⁷R. F. Wagner, and F. McCloy, Two Validity Studies of the Wide Range Achievement Test (Virginia Academy of Science, Richmond, Virginia, 1962).

⁸Paul Douglas Courtney, The Third Mental Measurement Yearbook (New Brunswick, N. J.: Rutgers University Press, 1949), pp. 46-47.

of testing time. Success in such an undertaking obviously involves an extremely difficult job of sampling from the total of school learnings. In this test, the sampling is done, first, by limiting the measurement to certain areas.

The aspects of the three subjects tested are: (a) skill in pronouncing printed words (ranging from letters of the alphabet to such words as "remiges," "heinous," and "conduit;"; (b) skill in mathematical computation (ranging from counting dots to simple algebra and logs); (c) skill in spelling words dictated (ranging from letters of the alphabet to such unfamiliar words as "pusillanimous," "iridescence," and "eleemosynary").

The test does have value in remedial and vocational studies. The manual accompanying the test contains a considerable amount of "evidence," such as analyses of errors, and a study of patterns of scores on the three tests.⁹

The chief goal in the development of reading skills for the educable mentally handicapped child is to meet his probable adulthood needs. He will need to be able to locate and recognize street signs, bus numbers and names, house numbers, names of public buildings, dates, names and numbers in telephone directories, public notices and warnings, the names of grocery, clothing, and household items. He will also need to be able to make out bills, to read simple instructions, and to sign various documents.

⁹Verner M. Sims, The Third Mental Measurement Yearbook (New Brunswick, N. J.: Rutgers University Press, 1949), pp. 47-48.

Though a mentally handicapped child may enter school at the chronological age of six, his mental age is such that he may not be ready to read for three or four years. Unfortunately, he may be forced to remain in the regular classroom for several years, continually meeting failure in the reading situation. Such experiences may well result in frustration and in avoidance reactions. The special teacher must supplement feelings of self worth and confidence before actual reading instruction.¹⁰

The reading ability of young retarded children appears often to be affected by the nature of their environmental backgrounds. Adverse effects can be overcome, at least partially, through intensive language stimulation programs. The reading patterns of the retarded may be further characterized by the nature of their cognitive styles. On the other hand, it has been hypothesized that the learning styles developed through reading may well influence the formation of cognitive styles.¹¹

No differences have been found in the overall reading achievement of various etiological groupings. Nonetheless, reading subtests did tend to favor one group over another. One variable, the sex of the child, has differentiated the mentally handicapped in their achievement. Investigations were consistent in their findings of superior achievement

¹⁰S. A. Kirk, and G. C. Johnson, Educating the Retarded Child (Boston: Houghton-Mifflin, 1951).

¹¹M. Dunn, P. Ponchaart, and P. Pfohl, The Efficacy of the Initial Teaching Alphabet and the Peabody Language Development Kit With Disadvantaged Children in the Primary Grades: An Interim Report After Two Years (Institute on Mental Retardation and Intellectual Development, 1967), pp. 4, 1-15.

in favor of girls. Non-intellectual variables have been hypothesized as accounting for this differentiation.

Boys tend to attain lower reading achievement levels than do girls, regardless of placement. A significantly larger number of retarded males tend to be placed in special situations than do females. Furthermore, it has been hypothesized that the learning of boys may be adversely affected by negativism. The special class curriculum is not likely to have as strong an academic orientation as does the regular class curriculum.¹²

When reviewing the literature dealing with the importance of reading to educable mentally handicapped children, one finds that there is no one best method for teaching reading to classes consisting of these students.¹³

Token reinforcement programs in special education classrooms indicate that positive results are almost invariably obtained, even with different types of target behaviors and various kinds of populations. Reading achievement definitely improves.¹⁴

For the best results in remedial reading instruction, the educable mentally handicapped child should meet the following requirements: (a) have an individual IQ test of at least 60; (b) be less than

¹²R. J. Capobianco, and D. Y. Miller, Quantitative and Qualitative Analysis of Exogenous and Endogenous Children in Some Reading Processes (Syracuse: Syracuse University Research Institute, 1958).

¹³Patricia A. Cegelka, and Walter J. Cegelka, A Review of Research: Reading and the Educable Mentally Handicapped (The Council for Exceptional Children, November, 1970), pp. 187-199.

¹⁴Saul Axelrod, Token Reinforcement Programs in Special Classes (The Council for Exceptional Children, January, 1971), pp. 341-379.

fourteen or fifteen years old; (c) have a reading ability below the fourth grade; (d) have a reading age of at least one year below his arithmetical computation. While such strict criteria are not necessarily requisite for remedial reading instruction, smaller average improvement should be expected when children are selected on a less rigorous basis.¹⁵

Parent involvement is crucial to the success of a reading achievement program in a school. The community around the school also has an important role in formulating a reading achievement program.

Children should be treated as citizens and individuals instead of as an anonymous mass.¹⁶ Instruction should be planned and carried out on the basis of a knowledge of the aptitude of the learner and must be guided by frequent appraisals of pupil progress. The present day classroom teacher has the opportunity and the responsibility for selecting and using a variety of procedures to determine aptitude and to evaluate achievement. Preparation for teaching must include these aspects of instruction.

Information on pupil aptitude will be obtained from a number of sources, some as simple as the teacher's day-by-day observation of pupil behavior. The truly professional teacher, however, will make extensive use of such instruments as special tests for measuring general intelligence, aptitude, and readiness.

¹⁵T. G. Hegge, Special Reading Disability With Particular Reference to the Mentally Deficient (American Journal of Mental Deficiency, 1934 - Houghton-Mifflin, 1940), pp. 39, 324 and 169.

¹⁶Seymour Fliegel, Practices That Improved Academic Performance in an Inner-City School (Kappan, Bloomington, Indiana, February, 1971), pp. 341-343.

Teachers may help pupils improve their reading by using three basic steps in the educational process: establishing goals; teaching; evaluation. These three steps should be combined into an effective program.

The one ability that is given the greatest emphasis in the elementary school is reading. As the student masters certain fundamental reading skills, there will be minor changes in the objectives of instruction and resultant changes in the types of abilities evaluated. On the other hand, there are certain basic objectives dealing with the understanding of word meanings and the comprehension of sentences, or larger passages, which are important at all levels. To the extent that this is true, suggestions concerning the evaluation of reading objectives at one grade level will be equally applicable at all levels.

Since reading is a communication skill, practically all of the objectives of instruction in this subject are concerned basically with comprehension. Some deal with comprehension of the meaning of words. The student may be expected to know the meaning of a word and to associate it with the picture of the object that it describes.¹⁷

The basic reason that the school staff administers tests is to promote the education of a child according to his or her unique abilities. The school in its attempt to meet the needs of the child plans its testing program with some of the following questions in mind: What types of information does the school staff need to know about its pupils? How will tests help supply this information? How can the staff

¹⁷C. M. Lindwall, Testing and Evaluation: An Introduction (New York: Harcourt, Brace, and World, Inc., 1961), pp. 1-46, 81-89.

use the information once they have it? Tests help the teacher to deal with individual differences in students by indicating which children have the same level of skill in a particular subject. These tests can also assist the teacher in helping to present to the student a realistic picture of his or her strengths and weaknesses. Aside from the above mentioned knowledge, tests provide information in helping the student decide on educational and vocational goals.

These can be of great value in discovering educationally and socially maladjusted children--children who present severe problems and are to be found in virtually every school. Some of these students are introverted, neurotic, or malcontent, and thus are simply not adjusting to normal patterns of behavior. Tests, therefore, render assistance to teachers in their attempts at understanding and helping these types of children.¹⁸

The traditional manner of determining the reliability of a test is to administer that test at least twice to a group of students, with the expectation that the students have not changed during the time interval between the two administrations with regard to the characteristics being measured. If the test is perfectly reliable, the two scores for any student will be identical. In reality this does not happen very often. Instead, highly reliable tests will yield results which place the student in the same relative position, with regard to his classmates, at one administration as at any other administration, provided, of course, that the students have not changed with regard to

¹⁸Louis J. Karmel, Testing in Our Schools (New York: The MacMillan Company), pp. 1-37.

the characteristic being measured.

The degree to which a test is reliable can be represented by a coefficient of reliability. There are four types of such coefficients: coefficients of stability; coefficients of equivalence; coefficients of stability and equivalence; coefficients of internal consistency.

Determining the degree of validity of a test is a more complex activity than determining the degree of the test's reliability. Ideally there should be as many determinations of the degree of validity as there are important uses of the test. These uses can be classified into four major categories:

1. The test scores can be interpreted in terms of the student's present ability to perform with respect to a well-defined type of situation or subject matter.
2. The test scores can be used to predict future student behavior.
3. The test scores represent student characteristics not directly measured by the test.
4. The test scores can be interpreted in terms of psychological theory, that is the test scores have meaning, psychologically speaking.¹⁹

No test is valid for all purposes, situations, or individuals. There are three aspects of validity: content validity, criterion-related validity, and construct validity.

1. Content validity is of primary importance in achievement

¹⁹ Stanley J. Ahmann, Testing Student Achievement and Aptitudes (Washington, D. C.: The Center For Applied Research in Education, Inc., 1962), pp. 42-62.

tests. Subject-matter experts devise and arrange test items that they feel cover the topics pertinent to the area represented by the test. Criterion-related validity is also necessary to check against a later criterion of performance. An achievement test may be used for a selection program.

2. Criterion-related validity is of primary importance in intelligence or scholastic aptitude tests to reveal the ability to predict school success. The kind of aptitudes measured is evaluated, many times, by the content of the test items and correlations with other tests.

3. Construct validity is of primary importance in personality tests, especially where projective techniques are used.²⁰

²⁰Louis J. Karmel, Measurement and Evaluation In The Schools (Toronto, Ontario: The MacMillan Company, 1970).

Chapter 3

THE STUDY

The Procedure

A list of all primary and intermediate educable mentally handicapped children was supplied by the Omaha Public Schools Special Education Department. A table of random numbers was applied to this list to secure the desired sample size of 12%. Both the SPED and WRAT were administered by the writer. Each child was given both tests. The tests were administered in a varied order. One child took the SPED first and the WRAT second while another child took the WRAT first and the SPED second.

The tests were administered on a one to one basis. Each pupil's classroom teacher, or principal, introduced the writer to the children who were to be tested. The children were tested either in the principal's office or the office of the nurse, and none of the students had any prior knowledge that he or she would be given any kind of test. (A copy of the WRAT and the SPED is in the appendix.)

Level 1 was administered to children ages eleven and under; level 2 was administered to all children over age twelve for the WRAT. A child was stopped when he had accumulated twelve consecutive failures, as recommended in the manual.

The Wide Range Achievement Test by J. F. Jastak and S. R. Jastak is published by Guidance Associates, 1526 Gilpin Avenue, Wilmington, Delaware and was copyrighted in 1965.

The Omaha Public School System Special Education Test was prepared by Mrs. Geraldine Nesvan of the Psychological Services of Omaha Public Schools in 1969 and revised in 1970.

The Findings

There were 439 primary and intermediate educable mentally handicapped children attending Omaha Public Schools during the time when the tests were administered. The tests were administered during the months of April and May of 1971. A sample of fifty-three children was used in this study ($439 \times 12\% = 53$).

The data collected from both tests were correlated to determine the validity of the SPED test. The SPED test's coefficient of correlation was calculated by a comparison to the WRAT test and its coefficient of correlation. The WRAT test is a widely and extensively used test which has a proven validity. On the basis of clinical experience and validity calculations made in the past, the clinical reliability of the WRAT test is that the coefficients vary from .90 to .95 for each subtest with an average reliability of .93. The results of both tests were correlated to establish the validity of the SPED test.

Dr. William A. Smith, of University Testing at the University of Nebraska at Omaha, recommended the WRAT test for a comparison to the SPED test. Dr. Rene Hlavac, Assistant Superintendent of Schools at Omaha, Nebraska for the Department of Pupil Personnel Services, also recommended the WRAT test for a comparison.

The Computer Center at the University of Nebraska at Omaha programmed a Linear, Regression, Correlation, Coefficient on the sample of fifty-three children. This program computes linear regression

coefficients and standard error of estimate for regression lines of X on Y and Y on X. The product-moment correlation is then computed.

Formulas used:

$$\bar{X} = \frac{\sum x}{n} \quad \bar{Y} = \frac{\sum y}{n} \quad \sum x^2 = \sum X^2 - \frac{(\sum x)^2}{n} \quad \sum y^2 = \sum Y^2 - \frac{(\sum y)^2}{n}$$

$$\sum xy = \sum XY - \frac{\sum x \sum y}{n} \quad b_{yx} = \frac{\sum xy}{\sum x^2} \quad b_{xy} = \frac{\sum xy}{\sum y^2}$$

$$a_{yx} = \bar{Y} - b_{yx} \bar{X} \quad a_{xy} = \bar{Y} - b_{xy} \bar{X}$$

$$S_{yx} = \sqrt{\frac{\sum y^2 - \frac{(\sum xy)^2}{\sum x^2}}{n - 2}}$$

$$S_{xy} = \sqrt{\frac{\sum x^2 - \frac{(\sum xy)^2}{\sum y^2}}{n - 2}}$$

$$r = \frac{n \sum XY - \sum X \sum Y}{\sqrt{(n \sum x^2 - (\sum x)^2) (n \sum Y^2 - (\sum Y)^2)}}$$

The Computer Center calculated the following results:

The X mean (WRAT test) = 2.15; The Y mean (SPED test) = 1.50.

The equation for the Y dependent: $X = 1.5000 + 0.000 Y$

Standard Error = 0.68.

The equation for the X dependent: $X = 0.1955 + 0.6075 Y$

Standard Error = 0.65.

Correlation coefficient = 0.64.

Chapter 4

CONCLUSIONS AND RECOMMENDATIONS

Conclusions

The Omaha Public School System Special Education Test is a valid instrument for determining the instructional reading level for primary and intermediate educable mentally handicapped children enrolled in the Omaha Public Elementary Schools.

The SPED test would also be an excellent instrument for determining the instructional reading level for regular primary and intermediate elementary children.

The administering of the test is done in a very simple manner which is easily understood by the children. Also, the administering of the test is not a heavy time burden on the classroom teacher. Approximately five minutes is all that is needed to administer and correct the test.

The results of the test aid the classroom teacher in structuring a more individualized curriculum for the child.

The results obtained from the Computer Center at the University of Nebraska at Omaha's program of a Linear, Regression, Correlation, Coefficients, on the sample of fifty-three children determined a correlation coefficient of 0.64. A correlation coefficient of 0.64 is not a significant difference as compared with the coefficient of the Wide Range Achievement Test which has an average reliability of 0.93.

Recommendations

Omaha Public Schools should continue to use the Special Education Test as an expedient method for determining the reading level for primary and intermediate educable mentally handicapped children.

The SPED test is also easy to administer, and calculate, without being time consuming for the teacher of the educable mentally handicapped children. The results certainly suggest consideration for additional testing, such as diagnostic testing, which would secure additional information so that specific recommendations could be made for the child who has reading problems.

The test clearly and definitely displays areas in which the individual child has reading difficulties and deficiencies.

The SPED test would be an excellent instrument used in the regular class of primary and intermediate elementary grades to determine the instructional reading level of these school children.

BIBLIOGRAPHY

BIBLIOGRAPHY

- Ahmann, Stanley J. Testing Student Achievements and Aptitudes. Washington, D. C.: The Center For Applied Research In Education, Inc., 1962.
- Axelrod, Saul. Token Reinforcement Programs in Special Classes. The Council for Exceptional Children, January, 1971.
- Capobianco, R. J., and D. Y. Miller. Quantitative and Qualitative Analysis of Exogenous and Endogenous Children in Some Reading Processes. Syracuse: Syracuse University Research Institute, 1958.
- Cegelka, Patricia A., and Walter J. Cegelka. A Review of Research: Reading and the Educable Mentally Handicapped. The Council for Exceptional Children, January, 1971.
- Courtney, Paul Douglas. The Third Mental Measurement Yearbook. New Brunswick, New Jersey: Rutgers University Press, 1949.
- DeLong, A. R. The Limits of Accuracy of the Test Scores of Educable Mentally Retarded Individuals. Journal of the Association for Research in Growth Relationships, 1962.
- Dunn, M., P. Ponchaart, and P. Pfof. The Efficacy of the Initial Teaching Alphabet and the Peabody Language Development Kit With Disadvantaged Children in the Primary Grades: An Interim Report After Two Years. Institute on Mental Retardation and Intellectual Development, 1967.
- Fliegel, Seymour. Practices That Improved Academic Performance in an Inner-City School. Bloomington, Indiana: Kappan, February, 1971.
- Guilford, J. P. Psychometric Methods. New York: McGraw-Hill Book Company, Inc., 1954.
- Hegge, T. G. Special Reading Disability With Particular Reference to the Mentally Deficient. American Journal of Mental Deficiency, 1934. Boston: Houghton-Mifflin, 1940.
- Karmel, Louis J. Testing in Our Schools. New York: The MacMillan Company.
- Kirk, S. A., and G. O. Johnson. Educating the Retarded Child. Boston: Houghton-Mifflin, 1951.

- Lindwall, G. M. Testing and Evaluation: An Introduction. New York: Harcourt, Brace, and World, Inc., 1961.
- Sims, Verner M. The Third Mental Measurement Yearbook. New Brunswick, New Jersey: Rutgers University Press, 1949.
- Smith, Bessie S. The Relative Merits of Certain Verbal and Non-Verbal Tests At The Second-Grade Level. Journal of Clinical Psychology, 1961.
- Straub, Bernie, and Jerome Hellmuth. Educational Therapy. Seattle: Special Child Publications, 1966.
- Thorpe, Louis P. The Third Mental Measurement Yearbook. New Brunswick, New Jersey: Rutgers University Press, 1949.
- Wagner, R. F. and F. McCloy. Two Validity Studies of the Wide Range Achievement Test. Richmond, Virginia: Virginia Academy of Science, 1962.

Other Sources

- Cegelka, Patricia A., and Walter J. Cegelka. A Review of Research Reading and the Educable Mentally Handicapped. The Council for Exceptional Children, Vol. 37, No. 3, November, 1970.
- Kluckhorn, Clyde. Culture and Behavior. New York: The Free Press, 1962.
- Lilly, Stephen M. Improving Social Acceptance of Low Sociometric Status, Low Achievement Students. Arlington, Virginia: Official Journal of the Council for Exceptional Children, 1960.
- Shuey, Audrey M. The Testing of Negro Intelligence. Lynchburg, Virginia: J. P. Bell Company, Inc., 1958.
- Tyler, Leona E. Tests and Measurements. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1963.

APPENDIX

327- OMAHA PUBLIC SCHOOLS SPED WORD RECOGNITION LIST

Reading Level
School
Grade

Name _____ Age _____ Date _____

Last		First		Middle	School		
List P (20)	List 1 (40)	List 2 (60)	Examiner _____				
1. see	1. with	1. game	List 3 (80)	List 4 (100)	List 5 (120)		
2. look	2. friends	2. hide	1. safe	1. harness	1. cushion		
3. mother	3. came	3. grass	2. against	2. price	2. generally		
4. little	4. horse	4. across	3. smash	3. flakes	3. extended		
5. here	5. ride	5. around	4. reward	4. silence	4. custom		
6. can	6. under	6. breakfast	5. evening	5. develop	5. tailor		
7. want	7. was	7. field	6. stream	6. promptly	6. haze		
8. come	8. what	8. large	7. empty	7. serious	7. gracious		
9. one	9. bump	9. better	8. stone	8. courage	8. dignity		
10. baby	10. live	10. suddenly	9. grove	9. forehead	9. terrace		
11. three	11. very	11. happen	10. desire	10. distant	10. applause		
12. run	12. puppy	12. farmer	11. ocean	11. anger	11. jungle		
13. jump	13. dark	13. river	12. bench	12. vacant	12. fragrant		
14. down	14. first	14. lunch	13. damp	13. appearance	13. interfere		
15. is	15. wish	15. sheep	14. timid	14. speechless	14. marriage		
16. up	16. basket	16. hope	15. perform	15. region	15. profitable		
17. make	17. food	17. forest	16. destroy	16. slumber	16. define		
18. ball	18. road	18. stars	17. delicious	17. future	17. obedient		
19. help	19. hill	19. heavy	18. hunger	18. claimed	18. ambition		
20. play	20. along	20. station	19. excuse	19. common	19. presence		
			20. understood	20. dainty	20. merchant		

List 6 (140)	List 7 (160)	List 8 (180)	High School (200)	SCORE
1. installed	1. administer	1. prairies	1. traverse	List P _____
2. importance	2. tremor	2. evident	2. affable	List 1 _____
3. medicine	3. environment	3. nucleus	3. compressible	List 2 _____
4. rebellion	4. counterfeit	4. antique	4. excruciating	List 3 _____
5. infected	5. crisis	5. twilight	5. pandemonium	List 4 _____
6. responsible	6. industrious	6. memorandum	6. scrupulous	List 5 _____
7. liquid	7. approximate	7. whimsical	7. primordial	List 6 _____
8. tremendous	8. society	8. proportional	8. chastisement	List 7 _____
9. customary	9. architecture	9. intangible	9. sojourn	List 8 _____
10. malicious	10. malignant	10. formulated	10. panorama	List H.S. _____
11. spectacular	11. pensive	11. articulate	11. facsimile	Raw _____
12. inventory	12. standardize	12. deprecate	12. auspicious	Score _____
13. yearning	13. exhausted	13. remarkably	13. contraband	(Total number of correct words including the words below starting level.)
14. imaginary	14. reminiscence	14. contrasting	14. envisage	
15. consequently	15. intricate	15. irrelevance	15. futility	
16. excellence	16. contemporary	16. supplement	16. enamoured	
17. dungeon	17. attentively	17. inducement	17. gustatory	
18. detained	18. compassionate	18. nonchalant	18. decipher	
19. abundant	19. complexion	19. exuberant	19. inadequacy	
20. compliments	20. continuously	10. grotesque	20. simultaneous	

DIRECTIONS FOR READING TEST

OMAHA PUBLIC SCHOOLS SPED TEST

1. Allow the child to read from one sheet while you keep score on another. At the start, say the following: "I want to see how many of these words you can read. Please begin here and read each word aloud as carefully as you can." (Indicate at what list to start.) "When you come to a difficult word, do the best you can and if you can't read it, say 'blank' and go on to the next one."
2. Start a child with a list where you think he can pronounce all 20 words in that one list correctly. Note that each list of words is graded. List P (primer) is for the first few months of first grade. List 1 is for the balance of first grade, List 2 is for second grade, etc. If the starting list is too difficult and the child makes even one mistake, go back until you reach an easier list where he can pronounce all 20 words correctly.
3. After you have found the starting list, go on into more advanced lists until you find the stopping list, where he mispronounces or is unable to read all 20 words. When you reach a point where the words become very difficult, say: "Look quickly down this list and read the words you think you know."
4. When a child reads very slowly and takes more 5 seconds on each and every word, move him along by saying the "blank" for him. Or call out the number of the word at a rate of about 5 seconds each.

Still another plan is to use a small card or piece of paper, covering up a word after a 5 second exposure forcing him on to the next word.

5. Count as an error each mispronounced or omitted word as well as a word which takes more than about 5 seconds to pronounce. (If a child has a speech defect such as a stutter, disregard the 5 second interval and allow as much time as necessary.) Count it an error when a child is uncertain about a word and gives more than one pronunciation, even though one of them may have been correct. Be particularly careful about scoring the word endings as they must be absolutely correct. Keep score by putting a check mark (✓) after each error or a plus sign (+) after each correct word. Enter the number of correct words at the bottom of each list as you go along. An analysis of the types of errors made, will indicate areas of weakness.
6. To find a child's raw score for reading, count the total number of words he was able to pronounce correctly in all lists and add the words below the starting list for which he automatically receives credit. To obtain the Reading Level, look up the value of this raw score in Table 1 below. A simple way to determine the Reading Level is to take half the raw score. For example, if the raw score were 46, half of this number would be 23 and the Reading Level would be 2.3 or the 3rd month of 2nd grade.

TABLE I
CHANGING THE RAW SCORE TO READING LEVEL

(Reading Grade Level is given in years and months. For example, 5.2 means the 2nd month of 5th grade.)

SCORE	GRADE	SCORE	GRADE	SCORE	GRADE	SCORE	GRADE	SCORE	GRADE	SCORE	GRADE	SCORE	GRADE
0-1	0.0	26-27	1.3	52-53	2.6	78-79	3.9	104-105	5.2	130-131	6.5	156-157	7.8
2-3	0.1	28-29	1.4	54-55	2.7	80-81	4.0	106-107	5.3	132-133	6.6	158-159	7.9
4-5	0.2	30-31	1.5	56-57	2.8	82-83	4.1	108-109	5.4	134-135	6.7	160-161	8.0
6-7	0.3	32-33	1.6	58-59	2.9	84-85	4.2	110-111	5.5	136-137	6.8	162-163	8.1
8-9	0.4	34-35	1.7	60-61	3.0	86-87	4.3	112-113	5.6	138-139	6.9	164-165	8.2
10-11	0.5	36-37	1.8	62-63	3.1	88-89	4.4	114-115	5.7	140-141	7.0	166-167	8.3
12-13	0.6	38-39	1.9	64-65	3.2	90-91	4.5	116-117	5.8	142-143	7.1	168-169	8.4
14-15	0.7	40-41	2.0	66-67	3.3	92-93	4.6	118-119	5.9	144-145	7.2	170-171	8.5
16-17	0.8	42-43	2.1	68-69	3.4	94-95	4.7	120-121	6.0	146-147	7.3	172-173	8.6
18-19	0.9	44-45	2.2	70-71	3.5	96-97	4.8	122-123	6.1	148-149	7.4	174-175	8.7
20-21	1.0	46-47	2.3	72-73	3.6	98-99	4.9	124-125	6.2	150-151	7.5	176-177	8.8
22-23	1.1	48-49	2.4	74-75	3.7	100-101	5.0	126-127	6.3	152-153	7.6	178-179	8.9
24-25	1.2	50-51	2.5	76-77	3.8	102-103	5.1	128-129	6.4	154-155	7.7	180-200	11.5

Prepared by Geraldine Nesvan
9-18-69 Psychological Services
Approved:

WIDE RANGE ACHIEVEMENT TEST WORD RECOGNITION AND ALPHABET LIST

Percentiles and Standard Scores corresponding to grade rating and age may be found in the Manual.

Level I—Reading—Grade Norms										Level II—Reading—Grade Norms																	
Score	Grade	Score	Grade	Score	Grade	Score	Grade	Score	Grade	Score	Grade	Score	Grade	Score	Grade	Score	Grade	Score	Grade								
1	N.5	16-17	Kg.6	36-37	1.0	53	3.3	60	5.3	79	8.1	92	12.9	0	Pk.5	16	1.3	29	4.4	42	6.8	55	9.3	68	13.0	81	16.8
2	N.8	18	Kg.7	38	2.0	54	3.5	67	5.5	80	8.4	93	13.3	1	Pk.8	17	1.5	30	4.6	43	6.9	56	9.0	69	13.2	82	17.1
3	Pk.1	19-20	Kg.8	39-40	2.1	55	3.6	68	5.7	81	8.7	94	13.7	2	Kg.1	18	1.7	31	4.8	44	7.1	57	9.9	70	13.5	83	17.4
4	Pk.2	21	Kg.9	41	2.2	56	3.8	69	5.9	82	9.0	95	14.1	3-4	Kg.2	19	1.8	32	5.0	45	7.3	58	10.2	71	13.8	84	17.7
5	Pk.3	22	Gr.1.0	42-43	2.3	57	3.9	70	6.1	83	9.3	96	14.5	5-6	Kg.3	20	2.0	33	5.2	46	7.5	59	10.5	72	14.1	85	18.0
6	Pk.5	23	1.1	44	2.4	58	4.1	71	6.3	84	9.7	97	14.9	7	Kg.4	21	2.2	34	5.4	47	7.7	60	10.8	73	14.4	86	18.3
7	Pk.7	24-25	1.2	45-46	2.5	59	4.2	72	6.5	85	10.1	98	15.4	8	Kg.5	22	2.4	35	5.6	48	7.9	61	11.0	74	14.7	87	18.6
8	Pk.9	26-27	1.3	47	2.6	60	4.4	73	6.7	86	10.5	99	15.8	9	Kg.6	23	2.6	36	5.8	49	8.1	62	11.6	75	15.0	88	19.0
9	Kg.1	28-29	1.4	48	2.7	61	4.5	74	6.8	87	10.9	100	16.2	10-11	Kg.7	24	2.8	37	6.0	50	8.3	63	11.9	76	15.3	89	19.3
10-11	Kg.2	30-31	1.5	49	2.8	62	4.7	75	7.0	88	11.3			12	Kg.8	25	3.2	38	6.2	51	8.5	64	12.2	77	15.6		
12	Kg.3	32-33	1.6	50	2.9	63	4.8	76	7.2	89	11.7			13	Kg.9	26	3.5	39	6.3	52	8.7	65	12.4	78	15.9		
13-14	Kg.4	34	1.7	51	3.0	64	5.0	77	7.5	90	12.1			14	Gr.1.0	27	3.9	40	6.5	53	8.9	66	12.6	79	16.2		
15	Kg.5	35	1.8	52	3.1	65	5.1	78	7.8	91	12.5			15	1.1	28	4.2	41	6.6	54	9.1	67	12.8	80	16.5		

LEVEL 2

Two letters in name (2) A B O S E R T H P I U Z Q (13) 15

milk	city	in	tree	animal	himself	between	chin	split	form	25
grunt	stretch	theory	contagious	grieve	toughen	aboard	triumph			33
contemporary	escape	eliminate	tranquillity	conspiracy	image	ethics				40
deny	rancid	humiliate	bibliography	unanimous	predatory	alcove				47
scald	mosaic	municipal	decisive	contemptuous	deteriorate	stratagem				54
benign	desolate	protuberance	prevalence	regime	irascible	peculiarity				61
pugilist	enigmatic	predilection	covetousness	soliloquize	longevity	abysmal				68
ingratiating	oligarchy	coercion	vehemence	sepulcher	emaciated	evanescence				75
centrifugal	subtlety	beatify	succinct	regicidal	schism	ebullience				82
misogyny	beneficent	desuetude	egregious	heinous	internecine	synecdoche				89

LEVEL I

cat	see	red	to	big	work	book	eat	was	him	how	36
then	open	letter	jar	deep	even	spell	awake	block	size		46
weather	should	lip	finger	tray	felt	stalk	cliff	lame	struck		56
approve	plot	huge	quality	sour	imply	humidity	urge				64
bulk	exhaust		abuse		collapse	glutton	clarify				70
recession	threshold		horizon		residence	participate	quarantine				76
luxurious	rescinded		emphasis		aeronautic	intrigue	repugnant				82
putative	endeavor		heresy		discretionary	persevere	anomaly				88
rudimentary	miscreant		usurp		novice	audacious	mitosis				94
seismograph	spurious		idiosyncrasy		itinerary	pseudonym	aborigines				100

A R Z H I Q S E B O 10

Two letters in name (2) A B O S E R T H P I U Z Q 25