1955

A Study in the Validation of Intelligence Tests for Cerebral Palsied Children

Elizabeth Richardson
Loyola University Chicago

A STUDY IN THE VALIDATION OF INTELLIGENCE TESTS FOR CEREBRAL PALSIED CHILDREN

by

Elizabeth Richardson

A Thesis Submitted to the Faculty of the Graduate School of Loyola University in Partial Fulfillment of the Requirements for the Degree of Master of Arts

June

1955

TABLE OF CONTENTS

Chapter		Page
I.	STATEMENT OF THE PROBLEM	1
	The need for a valid estimate of the intelligence of cerebral palsied children--Inadequacy of standard tests--Two promising tests--The purpose of this study.	
II.	REVIEW OF THE LITERATURE	4
	Common clinical practice in testing cerebral palsied children--Studies using the Binet with cerebral palsied children--Description, development, and standardization of the Raven and Ammons--Reliability and validity findings of the Raven and Ammons with normal children--Use of the Raven and Ammons with cerebral palsied children--Summary of the literature.	
III.	DESIGN OF THE RESEARCH	25
	Description of the subjects--Administration of the tests--Scoring--Method of statistical analysis.	
IV.	FINDINGS AND INTERPRETATIONS	29
	Descriptive analysis of scores--Statistical treatment of validation data--Discussion of findings--Interpretation of findings.	
V.	SUMMARY AND CONCLUSIONS	38
	Summary of the problem, literature, procedure, and statistical findings--Conclusions.	
BIBLIOGRAPHY		42
APPENDIX I		46

APPENDIX II 48

APPENDIX III 50

LIST OF TABLES

Table		Page
I.	RANGE, MEAN, AND STANDARD DEVIATION OF MENTAL AGES AND INTELLIGENCE QUOTIENTS	30
II.	BI-VARIATE, PARTIAL, AND MULTIPLE CORRELATIONS USED IN THE VALIDATION OF THE AMMONS AND RAVEN TESTS	31

CHAPTER I

STATEMENT OF THE PROBLEM

Measuring the intelligence of the cerebral palsied child provides a stimulating challenge to psychologists today. There is a growing interest in the welfare of these children with neuro-muscular disabilities and much money and effort is being spent in their education and habilitation. In planning treatment it is important to distinguish those children who are mentally retarded from those who are mentally average or superior. Intelligence not only limits a child's ability to profit academically, but also profoundly affects his ability to improve through occupational, physical, and speech therapy. A key factor, therefore, in the education and habilitation of the cerebral palsied child is a valid estimate of his intelligence.

Intelligence tests as standardized on normal children require either speech or fine motor manipulation on the part of the subject, or both. In many cases the cerebral palsied child lacks the speech or motor coordination necessary to respond to such tests. Consequently there is a need for an intelligence test or tests which are within the response capacity of these handicapped subjects.

Two such tests which appear to meet this need are the Raven Progressive Matrices Test and the Ammons' Full-Range Picture Vocabulary Test. Neither test requires speech or fine motor manipulation; the subject responds simply by pointing his finger or nodding his head. The fact that they have

no time limit and can be administered within a period of minutes is also an advantage for these children who are slower to respond and who fatigue easily.

The purpose of this thesis is to explore the validity of these two promising tests when used with cerebral palsied children. The Revised Stanford-Binet, Form L, was selected as the criterion with which to validate these tests because it is recognized as the most dependable of available standard measures for the age group used in this study.[1] A comparison will be made between the performances of mildly handicapped cerebral palsied children who have adequately completed the Stanford-Binet and their performances on the two new tests. If the results show a favorable correlation, one can assume that the two new tests are valid instruments for measuring the intelligence of mildly handicapped cerebral palsied children. In that the present group of subjects can be measured by a standard test, this is not the group for whom we are attempting to find adequate measures. However, there seems to be no criterion with which one could empirically validate these two tests with more severely handicapped children who cannot complete a standard test. It would seem, therefore, that if one could validate these new tests on children slightly handicapped, one would come as close as possible to inferring validity with the more severely handicapped group. There would be no reason why this is not acceptable on the assumption that a more disabling

---

1 Florence L. Goodenough, Mental Testing, New York, 1950, 315.

neuro-muscular disturbance would not necessarily distort the mental functioning which these tests measure.

CHAPTER II

REVIEW OF THE LITERATURE

The review of the literature will include studies describing common clinical practices in testing the intelligence of cerebral palsied children. The use of standard procedures with this group, and more specifically, the use of the Stanford-Binet, the criterion test in this thesis, will be investigated. Pertinent quantitative data from the use of the Binet with the cerebral palsied will be given. Studies employing the Progressive Matrices and Full-Range Picture Vocabulary tests will then be reviewed. These will include the description, development, and standardization of the tests, reliability and validity findings on normal children, and their use with cerebral palsied children. Findings relating to the present thesis problem will then be summarized.

The most common clinical practice in the testing of cerebral palsied children is the use of a battery of standard intelligence tests individually selected according to the particular handicap of each child.[1] One of the earliest investigations of methods for psychological examination of the cerebral palsied child was made by Doll, Phelps, and Melcher[2] in 1932. Twelve children were tested with the Stanford-Binet, Myers Mental Measure, Goodenough

---

1 Harry V. Bice, "Psychological Services for the Cerebral Palsied," The Nervous Child, VIII, April, 1949, 183-191; B. B. Burgemeister and L. H. Blum, "Intellectual Evaluation of a Group of Cerebral Palsied Children," The

Drawing Test, Healy Pictorial Completion Test, Porteus Maze, Witmer Formboard, and Ohio Literacy Test. Of these tests the Binet was found the most satisfactory because it had a comparatively wider range of application and was apparently less influenced by handicaps of speech, motor control, experience, training, and effects of practice. Since this pioneer investigation other standard tests such as the Cornell-Coxe Performance Ability Scale, Vineland Social Maturity Scale, Kohs Block Design, Bender-Gestalt, Wechsler Bellevue Intelligence Scale, Thematic Apperception and Rorschach tests have been adopted into general clinical procedure for testing cerebral palsied children. However, most investigators have been in agreement with Doll in finding the Binet the most satisfactory of the standard tests.[3]

The Binet has thus become the test most widely utilized in obtaining a measure of intelligence with these children.[4] Some writers have urged

---

Nervous Child, VIII, April, 1949, 177-180; Eric Denhoff, "Needs in the Field of Psychologic Appraisal of Children with Cerebral Palsy," New England J. of Medicine, CCILIII, October, 1950, 524-527; Raymond H. Holden, "Improved Methods in Testing Cerebral Palsied Children," American J. of Mental Deficiency, LVI, October, 1951, 349-353; Raymond H. Holden, "A Review of Psychological Studies in Cerebral Palsy: 1947 to 1952," American J. of Mental Deficiency, LVII, July, 1952, 92-99.

2 E. A. Doll, W. M. Phelps, and R. T. Melcher, Mental Deficiency Due to Birth Injuries, New York, 1932.

3 L. G. Portenier, "Psychological Factors in Testing and Training the Cerebral Palsied," Physiotherapy Review, XXII, 1942, 1-3; Bice, "Psychological Services for the Cerebral Palsied," The Nervous Child, VIII, 183-191; Patria Asher and F. E. Schonell," "A Survey of 400 Cases in Cerebral Palsy in Childhood," Archives of Disease in Childhood, XXV, 1950, 360-379.

4 Holden, "Improved Methods in Testing Cerebral Palsied Children," American J. of Mental Deficiency, LVI, 350.

a flexible use of this test, modifying time, speech, and motor requirements.[5] Other writers feel that there are serious difficulties involved in the evaluation of abbreviated or selected test items.[6] Blum, Burgemeister, and Lorge[7] feel that adaptation of standardized items to the limitations of the handicapped child affects the difficulty and placement of items. Omission of motor items gives undue weight to verbal items, enlarging of materials makes differences easier to see, and naming of multiple choices creates a task of recognition rather than one of recall. Burgemeister and Blum[8] further believe that errors are probably more frequently made in excusing failure on the grounds of incapacity than in penalizing a child because of that incapacity. Most investigators from both schools of thought emphasize range of abilities rather than a specific IQ.

---

5 Katherine M. Maurer, "Mental Evaluation of Cerebral Palsied Children," American Psychologist, I, 1946, 288-289; E. A. Doll, "Mental Evaluation of Children with Cerebral Palsy," Crippled Child, June, 1952, 6; B. T. Jewell and Helmut Wursten, "Observations on the Psychological Testing of Cerebral Palsied Children," American J. of Mental Deficiency, LVI, January, 1952, 630-637.

6 C. R. Strother, "Evaluating Intelligence in Children Handicapped by Cerebral Palsy," Crippled Child, XXIII, 1945, 182-183.

7 L. H. Blum, B. B. Burgemeister, and I. D. Lorge, "The Mental-Maturity Scale for the Motor Handicapped," School and Society, LXXIII, April, 1951, 232-233.

8 Burgemeister and Blum, "Intellectal Evaluation of a Group of Cerebral Palsied Children," The Nervous Child, VIII, 177-180.

Most studies involving the intelligence testing of cerebral palsied children have been of the descriptive or commentary type. Those offering quantitative data have dealt for the most part with distribution of IQ for the cerebral palsied group as a whole, differences in IQ levels of the various diagnostic sub-groups, correlation between motor impairment and IQ, and increase of IQ after retest.

For the purposes of this thesis investigations of the reliability of the most common standard procedure are pertinent. Maurer[9] administered the 1937 Stanford-Binet to eighty-five cerebral palsied children with IQ's ranging from twenty-seven to 136, with a mean of seventy-nine for both sexes. Half of the subjects were retested at least once after an interval of from one to six years. The test-retest correlation for the total group was .90. Special schooling did not raise IQ's. The Binet was judged to be suitable for testing cerebral palsied children as early as five years of age.

Asher and Schonell[10] retested twenty children between the ages of three and fifteen years with the Binet after a twelve month period. The mean IQ on the original test was 87.7 with a standard deviation of 12.69; the mean IQ on the retest was 89.9 with a standard deviation of 17.09. In the group as a whole there was no significant difference between test and

---

9 Maurer, "Mental Evaluation of Cerebral Palsied Children," American Psychologist, I, 288-289.

10 Asher and Schonell, "A Survey of 400 Cases of Cerebral Palsy in Childhood," Archives of Disease in Childhood, XXV, 360-379.

retest IQ's, but in six of the twenty cases, all of them severely handicapped, the difference exceeded seven IQ points. The authors concluded that while the Binet gave fairly accurate indications of intellectual level with the moderately handicapped, with the severely handicapped it could only be used as a rough estimate.

Jewell and Wursten[11] retested one hundred cerebral palsied children with a median time of nine months between testing. Both the Binet and Wechsler test were used in this study and while the majority of children were given the former, the authors do not specify the exact number. Twenty-seven children showed an IQ increase of five or more points. Of the cases with initial intelligence quotients of fifty to eighty-nine, 39 per cent showed IQ increases of five or more points, and 25 per cent showed increases of ten or more points. Such increases are considered substantial.

The authors report in the same study the heterogeneity of results when testing 166 children with the Binet. Heterogeneity was defined as any record in which the results extended through three or more units above basal age. Of the 166 cases, 118 were found heterogeneous, indicating inconsistent and unpredictable performances on the Binet.

Thus most authors agree that though the Binet is the most useful of the standard tests, it has its limitations, particularly when applied to the more severely handicapped. They emphasize the need for new techniques

---

11 Jewell and Wursten, "Observations on the Psychological Testing of Cerebral Palsied Children," American J. of Mental Deficiency. LVI, 630-637.

better suited to the limited responsiveness of the physically handicapped.

Penrose and Raven[12] expressed the need for a "true" intelligence test in which differences in educational background would be eliminated and tests for special abilities which depend on physical factors would be avoided. They therefore felt that both verbal and technical items were inadmissable. Also, tasks which involve repetition were felt not to be pure intelligence tests. The authors describe a series of fifty new perceptual tests based on Spearman's theory of educative rather than reproductive, nature of intelligence.

In a later report Raven[13] goes on to describe some experimental work done with these new perceptual tests. The series was given to more than six hundred children between the ages of four and fourteen in a small urban area in England. Subsidiary groups of children were tested outside the area in order to determine the reliability of the test in various forms. Norms and standard deviations were computed and a revised series of sixty tests was prepared on the basis of this study.

Raven with Miller[14] next investigated the influence of position in the choice of the fixed number of alternatives in the perceptual tests.

---

12 L. S. Penrose and J. C. Raven, "A New Series of Perceptual Tests: Preliminary Communication," British J. of Medical Psychology, XVI, 1936, 97-104.

13 J. C. Raven, "The R. E. C. I. Series of Perceptual Tests: An Experimental Survey," British J. of Medical Psychology, XVIII, 1939, 16-34.

14 F. M. Miller and J. C. Raven, "The Influence of Positional Factors on the Choice of Answers to Perceptual Intelligence Tests," British J. of Medical Psychology, XVIII, 1939, 35-39.

The subjects were school children of slightly above normal intelligence. The number of children was not specified. The investigation showed that the frequency with which an alternative is chosen depends partly on its own position and partly on the relative positions of the alternative figures shown with it.

Raven and Waite[15] sought to determine of these tests of education, termed the R. E. C. I. Perceptual Tests, could be used to differentiate between physically handicapped children who are backward because of loss of schooling and those who are backward because of mental impairment. The subjects consisted of ninety-one physically defective and ninety-seven mentally defective children, evenly distributed in age from six to sixteen. The physically defective children were classified according to the nature of their defect and according to teachers' estimates of their mental ability. These classifications were compared to their standard scores on the R. E. C. I. The mentally defective children were classified according to the nature of their defects and to their progress in school. These classifications were compared to their scores on the R. E. C. I. as well as the Stanford-Binet. It was found that the physically defective children on the average differed little from normals inspite of the fact that teachers rated many of them as backward. The Mentally defective children, on the other hand, had mean scores which were consistently and significantly lower than normal. The authors concluded

---

15 J. C. Raven and A. Waite, "Experiments on Physically and Mentally Defective Children with Perceptual Tests," British J. of Medical Psychology, XVIII, 1939, 40-43.

that this test can be used to differentiate backwardness because of loss of schooling and backwardness because of mental deficiency.

On the basis of his surveys with the preliminary R. E. C. I. Perceptual Tests, Raven prepared a revised series of sixty tasks which he called Progressive Matrices (1938), which hereafter, when abbreviated, will be called the Raven. Each task consists of one pattern or matrix in which one part has been removed. Six or eight cut outs are pictured with each matrix, all of the proper geometric shape to fit into the blank space of the matrix, but each with a different design. It is the task of the subject to select among these alternative cut outs, by pointing or nodding, the one which would correctly complete the pattern. The sixty patterns are divided into five sets, each representing a different abstract or logical process. They are arranged in order of increasing difficulty. Instructions are brief and are given orally by the examiner. There is no time limit. Norms are presented in the form of percentile points at relatively coarse intervals. The original edition of the test was designed to cover the whole range of intellectual development from infancy to adulthood.[16]

The individual form of the test was standardised on 660 children from Ipswich, an English port, manufacturing center, and market town in an agricultural district.[17] To obtain a random sample, three in every forty entries from the Maternity and Child Welfare Register between May, 1924, and

---

16 J. C. Raven, Guide to Using Progressive Matrices (1938), 5th ed., London, 1952, 1-13.

17 J. C. Raven, "Standardization of Progressive Matrices, 1938," British J. of Medical Psychology, XIX, 1941, 137-150.

December, 1932, were selected. For each age group the individuals were divided into four classes according to the occupations of their fathers. The children tested individually were then used as a basis for selecting a random sample of 1407 children to be tested in groups. Age, school, and occupational class of fathers were proportional in the two groups. Percentile points were calculated at half-yearly intervals from the natural score distribution of the children tested. Separate norms were calculated for the two groups, also for 3665 male adults. Raven found that the rate of development of eductive ability in childhood does not seem to be constant and that Progressive Matrices mental ages therefore should not be used like Binet mental ages for the calculation of IQ's.

Results of a clinic group of 301 children given both the Binet and the Progressive Matrices individually are compared in table form, but no correlations are stated. In 170 cases the Progressive Matrices agreed with the Binet IQ classes. In 109 cases the Matrices scores deviated one grade, and in twenty-two cases the Matrices scores deviated two grades from the corresponding Binet IQ class. Raven concludes that if the Binet scale can be regarded as correctly standardized for English children, norms for his test are approximately correct. He found the following differences between the two tests: verbal fluency effected a higher score on the Binet; emotional disturbance interferred less with performance on the Matrices; the latter test more clearly differentiates between genuine intellectual superiority and verbal fluency, and backwardness due to specific defects in education and genuine intellectual defect.

One hundred and fifty normal school children, ages six to thirteen, were given the Progressive Matrices, Mill Hill Vocabulary, and Stanford-Binet tests by Raven.[18] Correlation between the Matrices scores and Binet mental ages was found to be .855, between Mill Hill Vocabulary and Binet mental ages, .926, and between the combined scores of the Matrices and vocabulary test, and Binet mental ages, .918. Retest reliability on the Raven test for an unstated number of normal Colchester school children, age thirteen plus or minus one year, was reported as .88. In comparing performances on the vocabulary and perceptual tests Raven found that these two processes do not develop at the same rate, do not mature at the same age, and do not remain constant for the same length of time. He felt, therefore, that the assumptions in calculating intelligence quotients are untenable, and reaffirmed his belief that the most satisfactory method of comparing people is to evaluate ability in terms of percentage frequency with which a similar degree of ability is found to occur amongst people of the same age.

During the war the Progressive Matrices was widely used with adults. Little was known about its value as an intelligence test for children. Keir[19] presents a preliminary report using the Raven as a group test with 296 London children ages ten to fourteen. Using the split-half method, she found the

---

18 J. C. Raven, "The Comparative Assessment of Intellectual Ability," British J. of Psychology, XXXIX, 1948, 12-19.

19 Gertrude Keir, "The Progressive Matrices as Applied to School Children," British J. of Psychology, Statist. Sect., II, 1949, 140-150.

correlation between odd and even items to be .76. Data is also presented for forty-one eleven year old children in primary school who had been tested twice at an interval of two years. The average correlation for both sexes was .64. Keir also reports an incomplete and unpublished investigation started in 1939 by Miss Horton. On retesting after three weeks a sample of 123 boys and girls, aged eleven to thirteen and a half years, she obtained a reliability coefficient of .71. Both investigators obtained higher correlations on other tests when applied to the same children. Their figures are considerably below Raven's (.88), and they express serious doubt as to the reliability of the 1938 Progressive Matrices.

As judged by its correlation with other tests, Keir found the validity of the Raven also lower than had been claimed. Its correlation with the Simplex Junior Intelligence Test was only .56 and with educational tests below .50 with one exception, much lower than the corresponding correlations furnished by the Simplex test. For the present groups correlations between the Raven and the Mill Hill Vocabulary Test, .413 for boys, .36 for girls, are much lower than for Raven's group (.57). Keir again mentions the work of Miss Horton who found a correlation of only .62 between the Progressive Matrices and Burt's revision of the Terman-Binet Scale. In regard to the relative difficulty it was found that the test items are less evenly spaced than those of the Binet Scale, that the test contains too many items of medium difficulty. A factor analysis of the intercorrelations between items shows that the test as a whole is more homogeneous than the Binet, although its own general factor contributes less to its total variance. Keir felt that more

extensive investigations are needed before the 1938 Progressive Matrices can be accepted as a serviceable school test.

Banks and Sinha[20] applied the Progressive Matrices (1938) to 310 London children ranging in age from eight to thirteen years. Children over eleven years were also tested for intelligence with Richardson's Simples test; those under eleven were tested individually with the Binet scale. Reliability of the Raven with the split-half method (corrected) was .86, with Burt's analysis of variance, .90, and with the Kuder-Richardson formula, .91. The average of these figures is .88 which the investigators felt barely reaches the minimum usually thought necessary for a satisfactory intelligence test. The average validity, whether judged by internal or external criteria, was .54. An item-analysis, based on bi-serial correlations, indicated that the validity of the separate items differed widely, ranging from .20 to .80, averaging about .45. It was suggested that if the poorest items were eliminated, the coefficient could probably be raised to .65. The items were scaled by the method originally suggested by Burt for grading the Binet test. It was found that there were too many items of medium difficulty. The authors conclude that the 1938 Progressive Matrices is unquestionably promising, but that there is much room for improvement.

Because of the wide applicability of the 1938 Progressive Matrices, Raven felt it necessary to revise the test for younger children, defectives,

---

20 Uma Sinha, "A Study of the Reliability and Validity of the Progressive Matrices Test," British J. of Educational Psychology, XXI, November, 1951, 238-239; Charlotte Banks and Uma Sinha, "An Item-analysis of the Progressive Matrices Test," British J. of Psychology, Statist. Sect., IV, 1951, 91-94.

and people of more than average ability in order to provide a wider dispersion of scores for these groups. Progressive Matrices (1947), Sets A, Ab, B, were constructed to measure the first two groups; Progressive Matrices (1947), Sets I and II to measure the latter group. Progressive Matrices (1947), Sets A, Ab, B, differs in form from the original edition in that there are less items, thirty-six patterns divided into three sets, which are in color, a feature intended to attract and hold the attention of younger children. It is this booklet form of the Progressive Matrices which was used in the present study. Norms for Sets A, Ab, B are based on scores obtained by 608 Dumfries School children aged five to eleven and a half years.[21]

Foulds and Raven[22] assessed the consistency, retest reliability, and validity of the 1947 revision, Sets I and II. The test was administered to 471 children, ten and a half plus or minus one-half years of age, and to a larger number of older children and adults. An item-analysis revealed that with a few exceptions the total number of problems solved appeared to provide a consistent estimate of intellectual efficiency. The retest reliability after an interval of six weeks for 109 of the children was .76, considerably lower than the coefficients obtained for the older groups in the same study. The authors conclude that it is clear that this form of the test cannot be

---

21 J. C. Raven, Guide to Using Progressive Matrices (1947), Sets A, Ab, B, 2nd ed., London, 1951, 1-16.

22 G. A. Foulds and J. C. Raven, "An Experimental Survey with Progressive Matrices (1947)," British J. of Educational Psychology, XX, June, 1950, 104-110.

used satisfactorily before the age of eleven years. Validity is illustrated in terms of differences in mean scores between two groups of children, 351 in number, from two different school areas. The test results show a striking agreement with the actual number of children taking up scholarships from these two areas. The total sample of 471 children at ten and a half years and 596 children at twelve and a half years were used to establish working norms for this form of the test.

The only validity study using the same form of the Raven as used in this thesis was done by Martin and Weichers.[23] They administered the Progressive Matrices (1947), Sets A, Ab, B and the Wechsler Intelligence Test for Children to one hundred Indiana school children between the ages of nine and ten. Each test was administered individually. Correlations of .91, .84, and .83 were obtained between the Raven scores and WISC Full Scale, Verbal, and Performance IQ's, respectively. Martin and Weichers predict that in view of the high correlations and ease and speed of administration, this form of the Raven will be used more extensively in the clinical testing of children.

Four studies in the literature deal with the clinical evaluation of the Raven test with cerebral palsied subjects; there are no quantitative studies. Tracht[24] tested seventeen cerebral palsied persons of widely varying

---

23 A. W. Martin and J. E. Wiechers, "Raven's Colored Progressive Matrices and the Wechsler Intelligence Scale for Children," J. of Consulting Psychology, XVIII, April, 1954, 143-145.

24 Vernon S. Tracht, "Preliminary Findings on Testing the Cerebral Palsied with Raven's 'Progressive Matrices'," J. of Exceptional Children, XV, December, 1948, 77-79.

degrees of disability, both physical and meantal, within the age range of ten to thirty-five. The 1938 Progressive Matrices was used. In spite of the great variability of physical disability, the scores on the test ranged from the mentally defective to the mentally superior level. Tracht pointed out the test's important advantages of minimal stress and motor response, absence of speech response and time limits, and brief and straightforward directions.

Taibl[25] sought to discover if the 1938 Progressive Matrices is a significantly effective medium in evaluating the intelligence of the cerebral palsied. He administered the test individually to 115 subjects ranging in age from six to adulthood, and in mental ability from mentally defective to superior. He found that the performance of spastics in this study was significantly different from that of the athetotic group. He concluded that the 1938 Raven test may be an effective tool for cerebral palsied subjects who cannot approximate performances on the Stanford-Binet.

Sievers[26] tested thirty-three cerebral palsied subjects ranging in age from nine months to fifty-three years and eight months. The effectiveness of the Binet, Vineland, Gesell, Wechsler, and Raven scales were investigated. Three children were given the 1947 Raven, Sets A, Ab, B. Sievers concluded that the standard tests have limitations and that the Raven may be used with subjects with cerebral palsy who have severe speech and physical defects.

---

25 R. M. Taibl, An Investigation of Raven's 'Progressive Matrices' as a Tool for the Psychological Evaluation of Cerebral Palsied Children, Unpublished Doctoral Dissertation, University of Nebraska, Lincoln, Nebraska, 1951, Chapter V.

26 D. J. Sievers, Psychometric Problems Related to Cerebral Palsy, Unpublished Master's Thesis, University of New Mexico, Albuquerque, New Mexico, Chapters II and IV.

However, in that it measures perceptual intelligence, the Raven should be interpreted with caution.

Holden[27] also voices an objection to testing the cerebral palsied with the Progressive Matrices Test since it is based on visual form perception which is often impaired in brain injured individuals. However, he feels that this test may prove useful and recommends further evaluation.

A rapid and valid intelligence test which would be available to children with verbal retardation or physical handicaps was sought by Ammons and Huth.[28] They point out that vocabulary tests have been recognized as giving excellent estimates of intellectual ability. The vocabulary sub-test has been found to correlate most highly with the whole scale on both the Binet and Wechsler tests. In reviewing various vocabulary tests, the authors found the Van Alstyne multiple-choice, recognition-principle picture vocabulary idea best suited to the verbally retarded and physically handicapped. Feeling that the existing Van Alstyne test was inadequately standardized and dated, Ammons and Huth devised a series of sixty-four ink-line drawing sketches and a list of 253 words pertinent to the pictures.

For the preliminary standardization fifty-two white children, two boys and two girls from each grade level, kindergarten through twelfth grade, were given the present form of the picture vocabulary test and the vocabulary

---

27 Holden, "Improved Methods in Testing Cerebral Palsied Children," American J. of Mental Deficiency, LVI, 352.

28 R. B. Ammons and R. W. Huth, "The Full-Range Picture Vocabulary Test: I. Preliminary Scale," J. of Psychology, XXVIII, 1949, 51-64.

test from the 1937 Stanford-Binet. Mental age levels of difficulty for each item were derived from 50 per cent success performances at each grade level. Selection of approximately four of the best items at each age level, five through sixteen, was then made. A scale of forty-eight items was selected using the following criteria: irregularity of scatter, relative failure to discriminate, ambiguity, cultural bias, sex differences, number of words to be retained. Norms are given for the various Binet vocabulary mental ages. The odd-even reliability coefficient was found to be .95 ± .01, and when corrected for the full length, .97 ± .01. A correlation of .96 ± .01 was obtained between total vocabulary scores and Binet vocabulary scores.

Ammons with Rachiele[29] next sought to set up two final forms of the picture vocabulary test. Using the same plates as the above study, 243 words chosen from dictionaries and free associations were added to the forty-eight words previously selected. Of these, forty-three were eliminated on the basis of the above criteria, leaving 248 items for pretesting. Pretesting consisted of administering these words to a small sample of children and adults of widely varying age and ability. Words were again eliminated by the same criteria, and the resulting 226 words were listed by plates and by difficulty level. These were then administered to six hundred white American born subjects ranging in age from two to thirty-four years. Between the ages of two and seventeen, the subjects were selected by age or grade levels with respect to the fathers' occupations in proportion to the 1940 Census. Subjects were

---

29 R. B. Ammons and L. D. Rachiele, "The Full-Range Picture Vocabulary Test: II. Selection of Items for Final Scales," Educational and Psychological Measurement, X, 1950, 307-319.

drawn from urban Denver area and rural districts in Colorado and Nebraska. All subjects were given either the Binet or Wechsler and the vocabulary items as developed thus far. Using the same procedure as above, items were again evaluated and a final selection of 170 items was divided into two groups of equal length and difficulty.

The final form consists of 16 eight and one-half inch by eleven inch plates, each with four cartoon-like line drawings. The subject is asked to indicate by pointing or nodding which picture on a card best illustrates the meaning of a word given orally by the examiner. There are approximately eighty words, divided into two parallel forms, covering the range of verbal ability from early infancy to superior adult. Norms are given in the form of mental ages for children and percentiles for adults. The test can be administered in fifteen minutes or less.[30]

Standardization of the Full-Range Picture Vocabulary Test, which hereafter, when abbreviated, will be called the Ammons, on a representative preschool-age population was carried out by Ammons and Holmes.[31] A sample of 120 American-born white children in public and private nurseries in the urban and rural Denver area, ranging in age from two through five, was used. The children were equally divided as to sex and proportionally divided among

---

30 Robert B. Ammons, Full-Range Picture Vocabulary Test, Unpublished Summary, University of Louisville, Louisville, Kentucky, n.d., 1-2.

31 R. B. Ammons and J. C. Holmes, "The Full-Range Picture Vocabulary Tests: III. Results for a Pre-school-age Population," Child Development, XX, 1949, 5-14.

the occupational status of their fathers according to the 1940 Census. All children were given the preliminary scale of 226 words and the 1937 Binet Scale. The results of the standardization include the evolvement of two equivalent forms of the Full-Range Picture Vocabulary Test with a reliability coefficient, based on the intercorrelation of the two forms, of .93, and separate norms for the two forms by age and sex and also for the combined sexes at each age level. Validity as measured by the correlations of the Binet and Full-Range Picture Vocabulary Test was found to be .85 for Form A and .83 for Form B.

Ammons, Arnold, and Herrmann[32] standardized the Full-Range Picture Vocabulary Test on a white school population of 360 children, fifteen boys and fifteen girls, at each grade level from one through twelve. The sample was controlled for the occupational status of the parents. The preliminary vocabulary scale of 226 words and the 1937 Binet scale were administered. An item-analysis was made and eighty-eight items were selected and divided into two forms, combined with the items from the pre-school and adult levels to make up the final scale. The intercorrelations of the two forms yielded a reliability coefficent of .987. The median correlations of the Binet vocabulary scores and the Full-Range Picture Vocabulary scores as computed separately by forms and age groups were found to be .67 for Form A and .69 for Form B.

---

32 R. B. Ammons, P. R. Arnold, and R. S. Herrmann, "The Full-Range Picture Vocabulary Test: IV. Results for a White School Population," J. of Clinical Psychology, VI, 1950, 164-169.

Steady progression of performance from age to age was given as another indication of the validity of the test. Norms based on the standardisation group were calculated for ages six to sixteen.

The literature includes no further quantitative studies of the Full-Range Picture Vocabulary Test which would relate to this thesis. Holden[33] is the only author commenting on its use with the cerebral palsied. He found it valuable because no speech and only minimal motor performance is necessary on the part of the subject. He found its usefulness limited, however, by lack of opportunity to obtain qualitative clinical data due to its brief administration time as well as information about other important intellectual functions such as abstract thinking, practical reasoning, concentration, and memory. Holden states that the test offers definite possibilities for a more accurate intellectual measurement of the cerebral palsied, but that it needs further evaluation.

In summary, then, it has been found that standard intelligence tests are most commonly used in testing cerebral palsied children. Of these standard tests the Binet has been found the most satisfactory, although its limitations for this group are universally recognised. Studies on the Raven have for the most part used the original 1938 edition. While its author found the reliability and validity of the test adequate, three other investigators found lower coefficients, and questioned its use in its original form.

---

33 Holden, "Improved Methods in Testing Cerebral Palsied Children," American J. of Mental Deficiency, LVI, 352-353.

However, the one validity study using the revised 1947 edition, Sets A, Ab, B, the same form as in this thesis, was very favorable. Although there are no quantitative studies on the use of the Raven with the cerebral palsied, clinical evaluations have been encouraging. A possible limitation is that the Raven, being based on visual form perception which is often impaired in the brain injured, might underestimate the intelligence of this group. However, further data of the quantitative kind will be necessary before adequate evaluation can be made. Fewer studies have been made of the Arnmons test, but all indicate good reliability and validity. One author comments favorably on its possibilities as a test for the cerebral palsied, but feels it may be limited because it leaves unmeasured such psychological functions as conceptual thinking, practical reasoning, concentration, memory, and furnishes little opportunity for clinical observation. Again as with the Raven, further quantitative data is needed before an adequate evaluation can be made.

CHAPTER III

DESIGN OF THE RESEARCH

The subjects in the present study consisted of thirty-two cerebral palsied children who were tested through the co-operation of the Cerebral Palsy Clinic of Mercy Hospital, Chicago; the Parents Association for Cerebral Palsied Children of Chicago; and the Bureau of Child Study, Chicago Public Schools. There were twenty-two boys and ten girls in the sample. The group ranged in chronological age from sixty-one to 138 months; the mean chronological age was 107.45 months with a standard deviation of 22.60. The age range was determined by the age limits of the Raven test. The diagnosis of cerebral palsy for each child was made or confirmed by medical specialists in the field. The children were classified in the following sub-diagnostic categories: spastic, twenty-eight, or 87 per cent of the cases; athetoid, four or 13 per cent of the cases. The group included no ataxias. With respect to physical status, approximately one-third of the children used prosthesis for ambulation. The remainder of the group were able to walk with varying degrees of efficiency. Coordination was judged to be good in fifteen or 47 per cent of the cases; fair in thirteen or 40 per cent of the cases; and poor in four or 13 per cent of the cases. Socialization information about the children was gathered from a composite of the Vineland Social Maturity Scale, case history data, and personal observation. Socialization was judged to be good in 56 per cent of the cases, fair in 13 per cent, and

poor in 31 per cent of the group. In order to make the group as homogeneous as possible, children with complications additional to cerebral palsy, and children other than white, were ruled out of the sample.

Each child was given a battery of four tests: the Revised Stanford-Binet, Form L; Raven's Progressive Matrices (1947), Sets A, Ab, B; and Ammons' Full-Range Picture Vocabulary Test, Form A. and Form B. Because the Stanford-Binet was adopted as the criterion for validation, it was the first test administered to each child. In that it was necessary for each child to finish a complete criterion test, this procedure eliminated the possibility of needless testing. All four tests were for the most part administered at the same testing session. Approximately one and one-half hours was generally required to complete the four tests. The two forms of the Ammons were given in rotated order to counterbalance the effects of practice.

The four tests were administered individually to each child in a private room. A conversation period preceded the testing in an attempt to establish rapport with the child. Encouragement and praise were used frequently throughout the testing in order to keep the child interested and at ease. Subjects were allowed as much time as was necessary. The Stanford-Binet was administered and scored according to familiar standard procedure.[1] A brief description of the administration of the lesser known Raven and Ammons will be given.

---

1 Lewis M. Terman and Maud A. Merrill, Measuring Intelligence, Boston, 1937, 75-131.

In the administration of the Raven,[2] the first problem was used as illustration. Explanation was continued until the child grasped the nature of the problem. If the child failed on the second problem, Problem I was re-demonstrated. When the child demonstrated that he understood the nature of the task, no further instructions were necessary, and he was allowed to proceed at his own rate. However, care was taken throughout the test to ensure that the child looked carefully at each pattern and was satisfied that the alternative he chose was the only one he thought would complete the pattern. The entire set of thirty-six problems was given to each child, regardless of the number of failures. Each correct answer represented one point. Correct answers were totaled and converted into equivalent percentile points from the table of norms. The scores on the Raven were converted to T scores to facilitate subsequent computation.

In the administration of the Ammons,[3] a few simple words appropriate to Plate I, but not used in the test, were given as practice. The subject was instructed not to guess and asked to define words or point again later in the test on words thought guessed. Words on a card were given until three point-levels were consecutively passed and failed. In order that motivation not be lost, order of presentation of words was sometimes varied, and several easy words much below the child's verbal level given. The printed

---

2 Raven, Guide to Using Progressive Matrices (1947), Sets A, Ab, B, 7-16.

3 Robert B. Ammons, Use of the Full-Range Picture Vocabulary Test, Unpublished Guide, n.d.

words were not visible to the child at any time. Each item answered correctly represented one point. All items below the three passed consecutively for each card were assumed to have been passed. Correct answers were totaled and converted into equivalent mental ages from the sheet of norms. Interpolation was sometimes necessary. Corresponding IQ's were then found by means of the Revised Stanford-Binet tables.

The mean scores and standard deviations for all tests were then computed in order to get a descriptive picture of the sample. The statistical analysis of validity was correlational. Product-moment correlations were run between the criterion Binet and the Ammons, both forms, between the Binet and the Raven, and between the two forms of the Ammons. In that these correlations might be spurious because of the common factor of chronological age, partial correlations, holding chronological age constant, were run between the same tests. Correlations between chronological age and IQ were also computed as a check on the above. Multiple correlations between the Binet, on the one hand, and the Raven and one and both forms of the Ammons, on the other, were run to see if a combination of Raven and Ammons would yield a higher coefficient than if taken separately. All correlations were analyzed for significance.

CHAPTER IV

FINDINGS AND INTERPRETATIONS

A brief descriptive analysis of the scores on the four tests will be presented first. This will be followed by the factual presentation of the validation data, its discussion and interpretation.

The descriptive data consists of the range, mean, and standard deviation of mental ages and IQ's for each test. The mental age ranges in months for three of the tests were found to be thirty-one to 162 for the Binet, twenty-nine to 173 for the Ammons Form A, and thirty-one to 183 for the Ammons Form B. The mean mental ages in months with their respective standard deviations were 84.59 and 37.40 for the Binet, 96.69 and 36.50 for the Ammons Form A, and 99.50 and 37.41 for the Ammons Form B. Raven scores were not given in mental ages. The range in IQ was thirty to 127 for the Binet, thirty to 135 for the Ammons Form A, and thirty-one to 153 for the Ammons Form B. The number of correct responses on the Raven ranged from five to thirty-five. The mean intelligence quotients and their respective standard deviations were 78.88 and 25.90 for the Binet, 90.12 and 29.10 for the Ammons Form A, and 92.94 and 28.10 for the Ammons Form B. The mean of the Raven scores, converted to T scores, was 51.75 with a standard deviation of 9.33. The above descriptive data is summarized in Table I.

TABLE I

RANGE, MEAN, AND STANDARD DEVIATION OF MENTAL AGES AND INTELLIGENCE QUOTIENTS

Test	Mental ages in months		
	Range	Mean	Standard deviation
Stanford-Binet	31-162	84.59	37.40
Ammons Form A	29-173	96.69	36.50
Ammons Form B	31-162	99.50	37.41
	Intelligence quotients		
	Range	Mean	Standard deviation
Stanford-Binet	30-127	78.88	25.90
Ammons Form A	30-135	90.12	20.10
Ammons Form B	31-153	92.94	28.10
Raven	5-35*	51.75 ≠	9.33 ≠

* Raw scores.

≠ T scores

Validation data consists of nineteen correlations which are presented in composite form in Table II. The Ammons Form A, Ammons Form B,

TABLE II

BI-VARIATE, PARTIAL, AND MULTIPLE CORRELATIONS USED IN THE VALIDATION OF THE AMMONS AND RAVEN TESTS

Variables	Correlation coefficient
IQ . IQ	
Binet . Ammons Form A	.88
Binet . Ammons Form B	.90
Binet . Raven′	.74
Ammons Form A . Form B	.94
Ammons Form A . Raven′	.64
Ammons Form B . Raven′	.63
IQ . CA	
Binet . CA	.002*
Ammons Form A . CA	-.032*
Ammons Form B . CA	-.073*
Raven′ . CA	.372**
IQ & IQ . CA	
Binet & Ammons Form A . CA	.88
Binet & Ammons Form B . CA	.90
Binet & Raven′ . CA	.80
Ammons Form A & Form B . CA	.94
Ammons Form A & Raven′ . CA	.70
Ammons Form B & Raven′ . CA	.71
IQ . IQ & IQ (& IQ)	
Binet . Ammons Form A & Raven′	.92
Binet . Ammons Form B & Raven′	.93
Binet . Ammons Form A & Form B & Raven′	.93

Note.--All coefficients are significant at the .01 level of confidence unless otherwise indicated.

′ T score. * Not significant. ** Significant at the .05 level of confidence.

and Raven were first correlated with the criterion Binet by means of Pearson's product-moment coefficient of correlation as computed from a scatter diagram.[1] The product-moment correlations were .88, .90, and .74 respectively. When N is thirty-two with thirty degrees of freedom the correlation must be .449 to be significant at the .01 level of confidence. The above correlations are clearly beyond this as well as substantial in size. Such high correlations with the Binet indicate that the Ammons Form A, Ammons Form B, and Raven are valid instruments for measuring the intelligence of this population.

Because all three tests correlated well with the Binet, one would expect them to correlate to some extent with each other. The correlation between the two forms of the Ammons was .94, between the Ammons Form A and the Raven, .64, and between the Ammons Form B and the Raven, .63. These correlations, again, are clearly beyond the .01 level of confidence. As would be expected, the two forms of the Ammons correlated the highest, even higher than with the Binet, in that they were constructed to measure the same verbal function and to be of near equal difficulty. Again, because the Raven and the Ammons are designed to measure different aspects of intelligence, it is not surprising to find somewhat lower correlations between them. However, in that all correlations are significant, it is indicated that these three tests are measuring to a great extent the same functions, which confirms the first set of correlations and suggests the feasibility of using

---

1 J. P. Guilford, Fundamental Statistics in Psychology and Education, New York, 1942, 205-208.

the tests interchangeably.

It was possible that these correlations might be spurious because of the common factor contributed by chronological age, and because the relatively few cases showed a wide range of both chronological and mental age. Therefore the partial correlation technique[2] with chronological age held constant was applied to the above variables. The partial correlations between the Binet and the Ammons Form A was .88, between the Binet and the Ammons Form B, .90, and between the Binet and the Raven, .80. Only the correlation between the Binet and the Raven was changed, this being raised considerably. The partial correlation between the Raven and the Ammons Form A was .70, between the former and the Ammons Form B, .71. Here again the correlation for the Raven was increased. The two forms of the Ammons with chronological age held constant correlated .94, indicating no change. When N is thirty-two with twenty-nine degrees of freedom the correlation must be .522 to be significant at the .01 level of confidence. All correlations again are well beyond this.

From the above partial correlations it can be seen that chronological age, although a common factor, did not effect the first set of correlations between the Binet and the Ammons Form A, the Binet and Ammons Form B, and the Ammons Form A and Form B. Chronological age did effect the correlations in which the Raven was a variable, reducing them considerably.

---

2 Ibid., 269

Correlations between IQ and chronological age were also calculated as a check on the above. The correlation for the Binet was .002, for the Ammons Form A, -.032, for the Ammons Form B, -.073, and for the Raven, .372. Only the correlation for the Raven was significant, this at the .05 level of confidence. These results confirm the above findings that chronological age effected only the Raven.

Although the above bi-variate correlations are significantly high, the possibility of increasing the correlations through multiple correlation was investigated.[3] The Binet was the criterion or dependent variable in the multiple correlations. A correlation of .92 was obtained between the Binet on the one hand, and the Ammons Form A and the Raven on the other. The multiple correlation of the Binet with the Ammons Form B and the Raven was found to be .93. Both correlations are well beyond the .01 level of confidence. A correlation by the Doolittle method[4] of the Binet with both forms of the Ammons as well as the Raven gave a coefficient of .93. When N is thirty-two with twenty-eight degrees of freedom the correlation must be .573 to be significant at the .01 level of confidence. This correlation is therefore highly significant.

The above multiple correlations are all higher than the bi-variate correlations with the Binet, indicating that a combination of the Raven and Ammons gives a slightly more valid estimate of intelligence than does the

---

3 Ibid., 258.

4 Ibid., 263-265.

Ammons separately, and a considerably more valid estimate than does the Raven alone. It appears to be of little significance which form of the Ammons is combined with the Raven, or if one or both forms are employed.

In summary, then, all correlations with the criterion are substantial. This indicates that both the Raven and the Ammons are valid instruments for measuring the intelligence of cerebral palsied children. However, it is to be noted that the Ammons correlates higher with the Binet than does the Raven. It would seem that this difference could in part be explained in terms of the psychological functions which are emphasized in each test. The Ammons, being a vocabulary test, is completely verbal and would therefore be expected to correlate highly with the Binet, which is primarily verbal. The Raven, being a perceptual test, emphasizes discrimination, seeing logical relations, and visual form perception. These psychological functions are not as heavily weighted in the Binet as are the verbal functions. Hence one would anticipate a lower correlation between the Raven and the Binet.

Another explanation of the lower correlation of the Raven might be the previously cited argument by Holden that a perceptual test would underestimate the intelligence of cerebral palsied subjects since their form perception is often impaired. Other investigators would seemingly substantiate this argument. Strauss and Lehtinen[5] together with Lewis[6] consider

---

5 Alfred A. Strauss and Laura E. Lehtinen, Psychopathology and Education of the Brain-Injured Child, New York, 1950, 25-50.

6 Richard S. Lewis, Alfred A. Strauss, and Laura E. Lehtinen, The Other Child, New York, 1951, 21-28.

perception as the activity of the mind between sensation and thought, and therefore a preliminary to thinking. Thus an impairment of perception does not mean an impairment of intellect. However, if a measurement of intelligence were to be dependent upon perception, and this intermediary process were disturbed, the estimate of intelligence would be distorted. These authors have found in many brain injured children perceptual disturbances of figure-ground relationships, inability to see wholes, fixation to details, and perseveration. Dolphin and Cruickshank[7] in a series of three studies came to substantially the same conclusions. Although this argument may contribute to the explanation of the Raven correlation being relatively lower than the Ammons, it must not be ignored that this correlation was significant, and that the argument therefore may be overly emphasized.

To meet this objection to the Raven as well as that to the Ammons, that it measures only one psychological function and is limited by lack of qualitative clinical data due to its brief administration time, theoretically it would seem preferable to combine the two tests. This suggestion is supported by the multiple correlations which are the highest in the study.

A word of caution should be interjected concerning the interpretation of the validation data. The wide range in both chronological age and

---

7 Jane E. Dolphin and William M. Cruickshank, "Visuo-Motor Perception in Children with Cerebral Palsy, "Child Behavior, III, April, 1951, 198-209; "The Figure-Background Relationship in Children with Cerebral Palsy, J. of Clinical Psychology, VII, 1951, 228-251; "Pathology of Concept Formation in Children with Cerebral Palsy," American J. of Mental Deficiency, LVI, 1951, 386-392.

IQ and the sizeable standard deviations in the present sample must be considered in view of the relatively small population. Since the magnitude of the correlation coefficient varies with the degree of heterogeneity of a sample, and since the present sample is quite heterogeneous, the very high correlations found for the relatively small sample should be interpreted with some restraint until further evidence accrues. Also it must be remembered that the present sample is restricted in that it represents milding handicapped cerebral palsied children. That these tests are valid for more seriously handicapped children can merely be inferred, for there is no way by which to validate the tests for this group. Findings derived from more disabled groups should therefore be interpreted with prudence.

CHAPTER V

SUMMARY AND CONCLUSIONS

A review of the literature indicates that at present there is no adequate instrument for measuring the intelligence of severely handicapped children such as those found in the cerebral palsied group. Standard intelligence tests, which are most frequently used when testing cerebral palsied children, require either speech or fine motor manipulation on the part of the subject, or both. There are many cerebral palsied, however, who cannot speak and perform fine motor activities. It is with such a group that an estimate of intelligence is important, for upon this estimate depends the decision of spending time and money for education and habilitation.

There are two recently developed intelligence tests, Raven's Progressive Matrices and Ammons' Full-Range Picture Vocabulary Test, which require neither verbal nor fine motor response, and which take only a brief period of time to administer. With these tests one could estimate the intelligence of even the most severly handicapped child. The purpose of this thesis was to investigate the validity of these two tests when used with cerebral palsied children.

There are at present apparently no quantitative studies on the use of these two tests with the cerebral palsied. Studies on the reliability and validity of the original form of the Raven with non-handicapped children

are contradictory; the one study using the revised 1947 edition, as is used in this thesis, was very favorable. The fewer studies on the Ammons when used with non-handicapped children all denote good reliability and validity. Clinical evaluation of these two tests with the cerebral palsied has been encouraging, and several authors have felt that research in this area would prove promising.

The Revised Stanford-Binet, Form L, the most satisfactory of standard tests, was chosen as the criterion of validation. A comparison was made between the performances of thirty-two mildly handicapped cerebral palsied children who adequately completed the criterion test and their performances on the Raven and the two forms of the Ammons. This sample ranged in chronological age from sixty-one to 138 months; in IQ, as measured by the criterion Binet, from thirty to 127. The tests were administered individually the Binet being given first, the others in rotated order. The mean scores and standard deviations for each test were computed in order to get a descriptive picture of the group. The statistical analysis of validity was correlational.

Bi-variate correlations were run between the criterion Binet and the Ammons, Forms A and B, between the Binet and the Raven, and between the two forms of the Ammons, to estimate the validity of each test separately. Because of the possibility of a spurious result due to the common factor of chronological age, partial correlations, holding chronological age constant, were run between the same tests. Correlations between chronological age and IQ were also computed as a check on the above. Multiple correlations between

the Binet, on the one hand, and the Raven and one and both forms of the Ammons, on the other hand, were computed to see if a combination of the Raven and the Ammons would yield a higher correlation than if employed separately. These correlations were analyzed for significance.

All correlations, with the exception of those between chronological age and IQ, were found to be highly significant. However, a reasonable degree of restraint should be used when interpreting these high correlations in view of the heterogeneity and relatively small size of the sample. It also must be noted in interpreting the results that the present group could be measured by a standard test and is therefore not the kind of group for which we are attempting to find adequate measures. However, there appears to be no criterion by which to validate these tests with a more severely handicapped sample. Therefore, it seems that to establish the validity of these new tests on children slightly handicapped would be to come as close as possible to inferring validity with the more severly handicapped group. Such an inference would be acceptable on the grounds that a more severe neuro-muscular disturbance would not necessarily affect the mental functioning which these tests measure.

With due caution as indicated above, the following conclusions can be drawn from the present statistical data:

a. The Raven, Ammons Form A, and Ammons Form B are all valid instruments for measuring the intelligence of cerebral palsied children.

b. Although both tests are valid, if used separately, the Ammons is to be preferred to the Raven.

c. However, arguments that a perceptual test such as the Raven would underestimate the intelligence of brain injured children have perhaps been overly emphasized.

d. Either form of the Ammons can be utilized. There is no preference between them.

e. The most favorable procedure would be to administer both the Raven and the Ammons, either form. Here the statistical data reinforces the qualitative evaluation that a combination of tests would be preferable because of additional clinical data and insight into more than one psychological function.

f. Raven's Progressive Matrices and Ammons' Full-Range Picture Vocabulary Test appear to fulfill the need for adequate instruments by which to measure the intelligence of cerebral palsied children with varying degrees of handicap.

BIBLIOGRAPHY

I. PRIMARY SOURCES

A. BOOKS

Raven, J. C., Guide to Using Progressive Matrices (1938), 5th ed., London, 1952.

Raven, J. C., Guide to Using Progressive Matrices (1947), Sets A, Ab, B, 2nd ed., London, 1951.

Terman, Lewis M., and Merrill, Maud A., Measuring Intelligence, Boston, 1937.

B. ARTICLES

Ammons, R. B., Arnold, P. R., and Herrmann, R. S., "The Full-Range Picture Vocabulary Test: IV. Results for a White School Population," J. of Clinical Psychology, VI, 1950, 164-169.

Ammons, R. B., and Holmes, J. C., "The Full-Range Picture Vocabulary Test: III. Results for a Preschoolage Population," Child Development, XX, 1949, 5-14.

Ammons, R. B., and Huth, R. W., "The Full-Range Picture Vocabulary Test: I. Preliminary Scale," J. of Psychology, XXVIII, 1949, 51-64.

Ammons, R. B., and Rachiele, L. D., "The Full-Range Picture Vocabulary Test: II. Selection of Items for Final Scales," Educational and Psychological Measurement, X, 1950, 307-319.

Miller, F. M., and Raven, J. C., "The Influence of Positional Factors on the Choice of Answers to Perceptual Intelligence Tests," British J. of Medical Psychology, XVIII, 1939, 35-39.

Penrose, L. S., and Raven, J. C., "A New Series of Perceptual Tests: Preliminary Communication," British J. of Medical Psychology, XVI, 1936, 97-104.

Raven, J. C., "The R. E. C. I. Series of Perceptual Tests: An Experimental Survey," British J. of Medical Psychology, XVIII, 1939, 16-34.

Raven, J. C., "Standardisation of Progressive Matrices, 1938," British J. of Medical Psychology, XIX, 1941, 137-150.

C. UNPUBLISHED MATERIALS

Ammons, Robert B., Full-Range Picture Vocabulary Test, Unpublished Summary, University of Louisville, Louisville, Kentucky, n.d.

Ammons, Robert B., Use of the Full-Range Picture Vocabulary Test, Unpublished Guide, n.d.

II. SECONDARY SOURCES

A. BOOKS

Doll, E. A., Phelps, W. M., and Melcher, R. T., Mental Deficiency Due to Birth Injuries, New York, 1932.

Goodenough, Florence L., Mental Testing, New York, 1950.

Guilford, J. P., Fundamental Statistics in Psychology and Education, New York, 1942.

Lewis, Richard S., Strauss, Alfred A., and Lehtinen, Laura E., The Other Child, New York, 1951.

Strauss, Alfred A., and Lehtinen, Laura E., Psychopathology and Education of the Brain-Injured Child, New York, 1950.

B. ARTICLES

Asher, Patria, and Schonell, F. E., "A Survey of 400 Cases in Cerebral Palsy in Childhood," Archives of Disease in Childhood, XXV, 1950, 360-379.

Banks, Charlotte, and Sinha, Uma, "An Item-analysis of the Progressive Matrices Test," British J. of Psychology, Statistical Section, IV, 1951, 91-94.

Bice, Harry V., "Psychological Services for the Cerebral Palsied," The Nervous

Child, VIII, April, 1949, 183-191.

Blum, L. H., Burgemeister, B. B., and Lorge, I. D., "The Mental-Maturity Scale for the Motor Handicapped," School and Society, LXXIII, April, 1951, 232-233.

Burgemeister, B. B., and Blum, L. H., "Intellectual Evaluation of a Group of Cerebral Palsied Children," The Nervous Child, VIII, April, 1949, 177-180.

Denhoff, Eric, "Needs in the Field of Psychologic Appraisal of Children with Cerebral Palsy," New England J. of Medicine, CCILIII, October, 1950, 524-527.

Doll, E. A., "Mental Evaluation of Children with Cerebral Palsy," Crippled Child, June, 1952, 6.

Dolphin, Jane E., and Cruickshank, William M. "The Figure-Background Relationship in Children with Cerebral Palsy," J. of Clinical Psychology, VII, 1951, 228-251.

Dolphin, Jane E., and Cruickshank, William M. "Pathology of Concept Formation in Children with Cerebral Palsy," American J. of Mental Deficiency, LVI, 1951, 386-392.

Dolphin, Jane E., and Cruickshank, William M., "Visuo-Motor Perception in Children with Cerebral Palsy," Child Behavior, III, April, 1951, 198-209.

Foulds, G. A., and Raven, J. C., "An Experimental Survey with Progressive Matrices (1947)," British J. of Educational Psychology, XX, June, 1950, 104-110.

Jewell, B. T., and Wursten, Helmut, "Observations on the Psychological Testing of Cerebral Palsied Children," American J. of Mental Deficiency, LVI, January, 1952, 630-637.

Holden, Raymond H., "Improved Methods in Testing Cerebral Palsied Children," American J. of Mental Deficiency, LVI, October, 1951, 349-353.

Holden, Raymond H., "A Review of Psychological Studies in Cerebral Palsy: 1947 to 1952," American J. of Mental Deficiency, LVII, July, 1952, 92-99.

Keir, Gertrude, "The Progressive Matrices as Applied to School Children," British J. of Psychology, Statist. Sect., II, 1949, 140-150.

Martin, A. W., and Wiechers, J. E., "Raven's Colored Progressive Matrices and the Wechsler Intelligence Scale for Children," J. of Consulting Psychology XVIII, April, 1954, 143-145.

Maurer, Katherin M., "Mental Evaluation of Cerebral Palsied Children," American Psychologist, I, 1946, 288-289.

Portenier, L. G., "Psychological Factors in Testing and Training the Cerebral Palsied," Physiotherapy Review, XXII, 1942, 1-3.

Raven, J. C., "The Comparative Assessment of Intellectual Ability," British J. of Psychology, XXXIX, 1948, 12-19.

Raven, J. C., and Waite, A., "Experiments on Physically and Mentally Defective Children with Perceptual Tests," British J. of Medical Psychology, XVIII, 1939, 40-43.

Sinha, Uma, "A Study of the Reliability and Validity of the Progressive Matrices Test," British J. of Educational Psychology, XXI, November, 1951, 238-239.

Strother, C. R., "Evaluating Intelligence in Children Handicapped by Cerebral Palsy," Crippled Child, XXIII, 1945, 182-183.

Tracht, Vernon S., "Preliminary Findings on Testing the Cerebral Palsied with Raven's 'Progressive Matrices'," J. of Exceptional Children, XV, December, 1948, 77-79.

C. UNPUBLISHED MATERIALS

Sievers, D. J., Psychometric Problems Related to Cerebral Palsy, Unpublished Master's Thesis, University of New Mexico, Albuquerque, New Mexico, 1950.

Taibl, R. M., An Investigation of Raven's 'Progressive Matrices' as a Tool for the Psychological Evaluation of Cerebral Palsied Children, Unpublished Doctoral Dissertation, University of Nebraska, Lincoln, Nebraska, 1951.

APPENDIX I

FULL-RANGE PICTURE VOCABULARY TEST
ANSWER SHEET—FORM B

Plate 1
vegetable (3.8)
human (4.4)
dessert (4.5)
agriculture (10.7)
anti-socialness (13.2)
segment (15.0)
intimidation (16.6)
translucent (A2.5)
depredation (A4.0)

Plate 2
phonograph (3.3)
transport (8.4)
terpsichorean (A6.0)

Plate 3
car (1.6)
fight (2.8)
paying (6.0)
customer (6.3)
fuel (7.5)
sale (7.9)
purchase (10.4)
transaction (14.6)
aggressiveness (A3.6)

Plate 4
panels (13.9)
domicile (A4.0)

Plate 5
island (5.3)
munificence (A5.7)

Plate 6
bird (1.6)
fly (2.5)
conveyance (14.5)

Plate 7
passion (12.5)
impact (13.5)
dialogue (13.6)
discourse (A4.5)

Plate 8
music (3.0)
laundry (4.7)
sudden (9.1)
garment (9.8)

Plate 9
manufacturing (7.2)
skyscraper (7.8)
industrial (10.0)
pecuniary (A4.9)

Plate 10
spoon (1.8)
razor (3.0)
thermometer (4.1)
mercury (10.7)
beverage (10.9)
tonsorial (A4.4)

Plate 11
circle (2.7)
sentiment (13.9)
lobe (15.5)
chronometer (15.7)
pendant (17.7)

Plate 12
meal (3.9)
perspiration (9.6)
humid (14.7)
felony (16.7)
gourmand (A4.6)
repast (A5.2)
mendicant (A6.3)

APPENDIX I (continued)

FULL-RANGE PICTURE VOCABULARY TEST
ANSWER SHEET—FORM B

Plate 13
cheerful (6.8)
collision (7.4)
sympathy (9.6)
mishap (11.1)
propulsion (13.3)
condolence (16.2)
lacrimation (A6.3)

Plate 14
policeman (2.5)
listening (5.3)
broadcast (5.9)
uniform (6.2)
safe (6.5)
protection (6.7)
authority (10.4)
gravitation (11.8)
catastrophe (12.0)
constabulary (A3.2)
fortuitous (A6.4)

Plate 15
bathtub (1.6)
operation (3.1)
cleanliness (8.7)
crisis (12.5)
somnolent (16.2)
supine (A5.5)

Plate 16
train (1.5)
airplane (1.8)
intersection (8.5)

APPENDIX II

FULL-RANGE PICTURE VOCABULARY TEST
ANSWER SHEET--FORM A

Plate 1
pie (1.7)
window (1.7)
seed (6.5)
sill (6.7)
transparent (13.3)
rectangular (14.7)
sector (16.0)
illumination (16.0)
culinary (17.2)
egress (A6.3)

Plate 2
athletes (8.6)
competition (15.0)
revelry (A4.0)
ebullience (A6.4)

Plate 3
counter (4.0)
pump (4.4)
clerk (6.4)
sport (7.6)
recreation (10.8)
pugnacity (16.9)
replenishment (A3.1)
retaliation (A4.1)

Plate 4
shrubbery (9.8)
dwelling (11.7)

Plate 5
surf (12.5)
isolation (12.9)

Plate 6
horse (1.5)
wagon (2.3)
insect (6.7)
transportation (8.6)
antiquated (A3.8)

Plate 7
discussion (7.7)
skill (10.9)
amour (13.8)

Plate 8
firecracker (2.7)
clothes (3.0)
explosion (4.9)
clean (5.5)
dehydration (A4.3)

Plate 9
farm (4.1)
currency (12.2)
tranquility (16.5)
agrarian (A6.2)

Plate 10
furniture (4.4)
steel (6.0)
refreshment (6.2)
liquid (7.3)
container (9.5)
centigrade (14.5)

Plate 11
clock (1.6)
locket (3.0)
numbers (3.4)
engraving (9.8)

Plate 12
hot (5.2)
fear (7.4)
nutrition (10.4)
gorging (12.8)
poverty (13.9)
mastication (A2.6)
itinerant (A4.5)
coercion (A4.6)
corpulence (A5.5)
insatiable (A5.6)

APPENDIX II (continued)

FULL-RANGE PICTURE VOCABULARY TEST
ANSWER SHEET--FORM A

Plate 13
telephone (2.1)
crying (2.9)
accident (3.0)
vehicles (9.5)
destruction (10.0)
portrait (10.2)
communication (10.6)
consolation (13.4)
negligence (14.3)
bereaved (15.4)
deleterious (A6.2)

Plate 14
danger (5.6)

Plate 15
bed (1.6)
newspaper (2.5)
anaesthesia (11.7)
immersion (14.6)
displacement (A5.0)
perusing (A5.0)

Plate 16
propellers (3.7)
harbor (8.1)
locomotive (8.2)
nautical (16.5)

APPENDIX III

FULL-RANGE PICTURE VOCABULARY TEST
NORMS BY FORMS

FORM A

Score	Mental Age (Years)
6	2.5
12	3.5
17	4.5
21	5.5
25	6.5
29	7.5
33	8.5
37	9.5
42	10.5
46	11.5
50	12.5
54	13.5
58	14.5
62	15.5
69	16.5

FORM B

Score	Mental Age (Years)
6	2.5
12	3.5
17	4.5
20	5.5
25	6.5
29	7.5
32	8.5
37	9.5
41	10.5
45	11.5
49	12.5
53	13.5
57	14.5
62	15.5
69	16.5

Adult Percentiles

Score	Percentile
84	99
82	95
80	90
78	80
75	70
73	60
70	50
67	40
65	30
60	20
57	10
55	5
52	1

Score	Percentile
84	99
82	95
81	90
78	80
75	70
73	60
71	50
68	40
65	30
62	20
58	10
55	5
52	1

APPROVAL SHEET

The thesis submitted by Elizabeth Richardson has been read and approved by three members of the Department of Psychology.

The final copies have been examined by the director of the thesis and the signature which appears below verifies the fact that any necessary changes have been incorporated, and that the thesis is now given final approval with reference to content, form, and mechanical accuracy.

The thesis is therefore accepted in partial fulfillment of the requirements for the Degree of Master of Arts.

April 27, 1945
Date

Frank X. Kohler
Signature of Adviser FK