Comparative Study of Three Intelligence Tests Administered to Deaf Children

Angela Lucille Chemazar

Loyola University Chicago

Follow this and additional works at: https://ecommons.luc.edu/luc_theses

Part of the Psychology Commons

Recommended Citation

Chemazar, Angela Lucille, "Comparative Study of Three Intelligence Tests Administered to Deaf Children" (1959). Master's Theses. 1555.
https://ecommons.luc.edu/luc_theses/1555

This Thesis is brought to you for free and open access by the Theses and Dissertations at Loyola eCommons. It has been accepted for inclusion in Master's Theses by an authorized administrator of Loyola eCommons. For more information, please contact ecommons@luc.edu.

This work is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 License.
Copyright © 1959 Angela Lucille Chemazar
COMPARATIVE STUDY OF THREE INTELLIGENCE TESTS
ADMINISTERED TO DEAF CHILDREN

by

Angela Lucille Chemazar

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

February
1959
LIFE OF AUTHOR

Angela Lucille Chemazar was born in Chicago, Illinois on May 9, 1932.

She was graduated from Saint Mary High School, Chicago, Illinois in June, 1950, and received the degree of Bachelor of Arts in Psychology from Mundelein College, Chicago, in June, 1954.

She taught sixth grade in a Chicago parochial school during the 1954-1955 school year. In 1955, she began graduate training in psychology at Loyola University. She is presently employed as a psychologist by the Catholic Charities Guidance Center.
TABLE OF CONTENTS

Chapter | Page
-------|------
I. INTRODUCTION | 1
-Problems in the selection of a suitable instrument for testing deaf children--Difficulties encountered in administering a test to the deaf--Purpose of present research--Hypotheses formulated.

II. REVIEW OF RELATED LITERATURE | 6
-Studies conducted with the Ontario--Studies conducted with the Leiter--Studies conducted with the Raven--Criticisms of former studies.

III. PROCEDURE | 17
-Description of tests used--Analyses of tests used--Selection of subjects--Contacting the subjects--Administration of tests--Discussion of the sample--Statistical analysis employed.

IV. RESULTS | 31
-Presentation of statistical data--Comparison of data presently obtained with that from previous studies--Correlations between Ontario, Leiter, and Raven--Analysis of I.Q. point changes.

V. SUMMARY AND CONCLUSIONS | 34
-Re-statement of the problem--Purpose of research--Hypotheses formed--Summary of literature--Methods used--Conclusions drawn

BIBLIOGRAPHY | 41

APPENDIX I | 44

APPENDIX II | 45
I. SUMMARY OF RELATED LITERATURE ........................................... 15
II. ANALYSES OF ONTARIO, LEITER, AND RAVEN ............................ 27
III. FREQUENCY OF DISTRIBUTION OF RAVEN SCORES ....................... 44
IV. CHANGES BETWEEN LEITER AND ONTARIO I.Q. POINTS .................... 33
V. SUMMARY OF STATISTICAL RESULTS ........................................... 37
CHAPTER I

INTRODUCTION

One of the problems encountered by a psychologist in the clinical situation is the testing of deaf children who can neither hear nor speak. Conventional tests, such as the Stanford-Binet and the Wechsler Intelligence Scale for Children, cannot be used because they involve verbalization on the part of the examiner and the examinee. Although there is a performance part in the WISC, the directions are rather long and not readily adaptable to unvarying and simple pantomime directions. An adaptation in directions to enable the testing of the deaf would present another problem of standardization. That is, the Performance Scale would theoretically have to be re-standardized since the norms could not be used.

The situation of testing deaf children can be approached from two aspects. First, the selection of a suitable instrument; and second, the difficulties commonly encountered in administering a test to the deaf.

When a test is being selected to be used in testing the deaf, both a technical and a practical criteria must be considered. On the technical

---

1 Throughout the remainder of this paper the Stanford-Binet Scale will hereafter be referred to as the Binet.

2 Throughout the remainder of this paper the Wechsler Intelligence Scale for Children will hereafter be referred to as WISC.
side, it must have reliability, validity, standardized scoring, and appropriate norms. The practical criterion would include cost of materials and scoring, ease of and time required for administration and scoring, availability of comparable forms, test acceptability and lastly, usefulness of results.

Certain difficulties peculiar to the testing of deaf subjects commonly arise. In the literature the following points are generally stressed.

(1) It is difficult to convey to a deaf child the need to work with speed. At times he may misinterpret the tester, think he is wrong, and stop working. More time is lost and results are lowered. Thus, a deaf child is frequently at a disadvantage on time items.

(2) The test should be non-verbal in both directions and responses. A deaf child is usually older than most children before he realizes the existence of language. Years of schooling are required before he obtains a working knowledge of the language. Acquisition of speech is dependent on many factors, such as degree of hearing, use made of sounds heard, opportunity of obtaining a hearing aid, educational opportunities and methods, etc. These facts would indicate that tests which employ language, either in directions or test proper, tend to penalize deaf subjects.

(3) Test items should be interesting to the testee since the examiner can do little to "sell" the test. This would be especially important at the lower levels.
Numerous measures are available to evaluate intellectual ability. After more than a quarter century of research in intelligence testing, an adequate test of intelligence for the deaf remains to be developed. Clinicians have made use of several available examinations to obtain some estimate of the intelligence of the deaf. Three of these are: the Ontario School Ability Examination, the Leiter International Performance Scale, and the Raven's Coloured Progressive Matrices.

At the Catholic Charities Guidance Center, much use has been made of both the Leiter and the Ontario in the examination of deaf children. On occasion, re-examinations with the scale not previously utilized is requested. That is, if the Ontario was initially used, then the Leiter would be given during the second visit. A preliminary comparison of results obtained by the two scales frequently indicated many differences. A search was made to discover studies which indicated to what extent the two scales were related. Since none could be found the need for a controlled study to determine to which extent the two instruments measure the same functions became evident. Thus, the present study began to be organized.

3 Throughout the remainder of this paper the Ontario School Ability Examination will hereafter be referred to as the Ontario.

4 Throughout the remainder of this paper the Leiter International Performance Scale will hereafter be referred to as the Leiter.

5 Throughout the remainder of this paper the Raven's Coloured Progressive Matrices will hereafter be referred to as the Raven.
A search was made for a suitable criterion for the Ontario and Leiter. Tests of intelligence are frequently compared with results obtained by the Binet and WISC. As mentioned above, these scales would not be suitable for use with deaf subjects. A preliminary survey of literature indicated that the Raven which involves minimal pantomime directions and no verbal responses may be an adequate criterion. A study by Martin and Wiechers demonstrated that the Raven measures to a considerable extent the same factors as the WISC (9, 4). In studies with cerebral palsied and normal subjects, Richardson (34) and Tracht (36) found highly significant correlations between the Raven and Ammons Full-Range Picture Vocabulary Test and between the Raven and the Verbal Scale of the Wechsler-Bellvue Scale. Because of the high correlations found by Martin and Wiechers and by Richardson and Tracht, and because of the ease with which Raven could be administered to the deaf, the Raven was considered to be an adequate criterion for comparison with the Leiter and Ontario I.Q. 's.

The following hypotheses were formulated which would be tested in this study.

(1) It is hypothesized that there will be a significant relationship between Ontario and Raven results when the two tests are administered to a group of deaf children from seven to nine years of age.

(2) It is hypothesized that there will be a significant relationship between Leiter and Raven results when the two tests are administered to a group of deaf children from seven to nine years of age.
(3) It is hypothesized that there will be a significant relationship between Ontario and Leiter results when the two tests are administered to a group of deaf children from seven to nine years of age.
CHAPTER II

REVIEW OF RELATED LITERATURE

A survey of the literature reveals that the following studies are available on the use of the Ontario, Leiter, and Raven to determine the intelligence of deaf children. In these studies the tests were compared to other tests of intelligence or to teachers' ratings of intelligence and achievement. However, in none of these studies were the Ontario and Leiter compared to each other or to the Raven.

In 1948 Kirk and Perry (17) published a study in which they compared the ratings from the Ontario with those from the Nebraska Test of Learning Aptitude. Their method was two-fold. First, they compared the Ontario I.Q.'s with the Nebraska I.Q.'s when both were administered to forty-nine deaf children. Secondly, they compared the Ontario I.Q.'s and the Nebraska I.Q.'s with Binet I.Q.'s when all three were administered to forty-nine hearing children. The age range of the children was from five to eleven years. The results obtained from the former group were an Ontario mean I.Q. of 102.9 and a Nebraska mean I.Q. of 95.8. This difference was reported to be statistically significant but the level of confidence was not stated.

---

Throughout the remainder of this paper the Nebraska Test of Learning Aptitude will hereafter be referred to as the Nebraska.
The mean I.Q.'s from the hearing group were: Binet, 107.3; Ontario, 102.9; and Nebraska L.Q., 95.7. The differences between the Ontario and Binet were not found statistically significant, while the differences between the Binet and Nebraska were significant. Again the level of confidence was not stated.

The investigators reached following conclusions: 1. The Ontario consistently tested higher than the Nebraska for both groups; 2. With hearing children the results obtained with the Ontario were closer to the results obtained from the Binet than those obtained from the Nebraska; 3. The Ontario is superior to the Nebraska, if the Binet is considered to have any relation to learning ability; and 4. The Ontario is more practical in terms of time, cost and ease of administering and scoring.

The Nebraska was standardized entirely on deaf and hard-of-hearing children. Norms were based on 166 children who ranged from four to ten years of age and who attended state schools for the deaf in six states. Now, in the study by Kirk and Perry just reviewed the subjects ages ranged from five to eleven years. The test results of the eleven-year-old children would have been compared with normative data based on children not older than ten. The oldest children in Kirk and Perry's study would thus tend to obtain lower Learning Quotients. This may account to some extent for the lower Nebraska mean L.Q.'s obtained by both the deaf and hearing children in the study. For a more accurate analysis of data it is recommended that the test results for eleven-year-old subjects be omitted and the statistics re-calculated. It is probable that results obtained and conclusions drawn would then vary from those originally stated.
Grace Ross, who was a research assistant at Oakland Public Schools in California, also made use of the Ontario in the intelligence testing of deaf children. (35) She studied the relative effectiveness of the Ontario, the Nebraska, and the Vineland Social Maturity Scale. Sixty-one deaf children from three years to ten years six months of age were initially examined. Of these, the results for thirty-six whose scores were recorded within time limits, were chosen for comparison. No mention was made of their geographical locations. Ross reported an Ontario mean I.Q. of 104.6, a Nebraska mean L.Q. of 104.8, and a Vineland mean S.Q. of 94.7. The correlation between the Ontario and Vineland was .91; between the Nebraska and Ontario was .84; and between the Nebraska and the Vineland was .85.

Ross found that the differences between the mental age and chronological age of deaf children increased as they advanced in school. She recommended further investigation in order to determine whether this difference was by chance or limited to the geographic area where these tests were administered. Because of the high correlation obtained, Ross considered the Vineland as a satisfactory substitute for the Ontario or Nebraska when neither of these could be administered. She particularly stressed its usability with young deaf children who often had emotional and personality disturbances to overcome. It would seem that such disturbances may also hinder social development and thus a low social quotient would result. An indication of the present level of social maturity could be obtained but it would seem unfair

7 Throughout the remainder of this paper the Vineland Social Maturity Scale will hereafter be referred to as the Vineland.
to the child to predict social and mental capacity from Vineland results obtained of a disturbed child. Another point that should not be overlooked is that the age spread of the subjects used ranged from two years to ten years, six months. The high correlation received may be due possibly to the range in age and not as much to the similarity of test content. A further criticism could be levied against the criterion used, the Vineland. It is a scale of social maturity and would thus be an inadequate criterion for an intelligence test.

Ross does not mention whether or not the age range of the final group of thirty-six subjects was identical to that of the group of sixty-one children initially tested, which was from three to ten years. As was mentioned above, Nebraska norms are based on children from four to ten years of age. Unless other suitable norms were available, the three-year-olds should not have been included in the study.

Several studies were also found in which the Leiter Scale is the instrument used to determine intelligence. Birch and Birch administered the Leiter to fifty-three deaf children who were experiencing learning difficulties. The obtained I.Q.'s were compared with those obtained on the Arthur Point Scale of Performance Tests, Form I, the Hiskey's Nebraska Test of Learning Aptitude for Young Deaf Children, the Performance Scale of the Wechsler Adult Intelligence Test, Form I, and the Goodenough Drawing Test. The investigators omitted the age range of the group. Although they presented the scores for twenty-six out of fifty-three cases, they gave no indication of group results on the five tests administered. All five tests were not administered to all subjects. The writers failed to state how many
subjects were administered each test. It would seem to this writer that if sufficient data were not available for some of the subjects that they should have been excluded.

The main conclusion drawn by Birch and Birch was that the Leiter I.Q.'s were consistently lower than those of other tests. They suggested that the Leiter predicted learning ability more accurately than other tests used. That is, if the Leiter score was considerably below the four other test scores one could predict that the child would find it difficult to learn. The subjects in this study consisted only of those who were experiencing learning difficulties. It is possible that children who were successful in school may have manifested similar clusters of test scores. It would have been better for Birch and Birch to have studied test results of both achievers and non-achievers before they concluded that Leiter I.Q.'s consistently lower than those of the other tests were indicators of academic inaptitude.

Birch and Birch (6) in a second study administered the Leiter to thirty-five enrollees at the Pennsylvania School for the Deaf. Their ages ranged from three years, eleven months to eight years, six months. The enrollees were examined approximately six months after entrance. At the same time they were rated on general intelligence by their supervising teachers who did not know the Leiter results. Approximately a year after the testing had been completed, these same teachers rated the thirty-five children again on their educational achievement. At the time the subjects had been in the school from one to three years. The Pearson product moment correlation between the Leiter I.Q.'s and the teachers' ratings of general intelligence was .86; the correlation between the Leiter I.Q.'s and the teachers' achievement ratings
was .71; and the correlation between the teachers' intelligence ratings and
teachers' achievement ratings was .82. No mean scores were presented. On
the basis of these results Birch and Birch suggested that the Leiter Scale
should prove useful in predicting the future success of young deaf children.

The design of this second study by Birch and Birch appears better than
that of their first one. Instead of comparing Leiter results with other
intelligence tests they have chosen teacher ratings as their criteria. Both
achievers and non-achievers were included. Additional information would have
been provided if they had compared Leiter results and teachers' ratings with
standardized measurements of academic achievement, such as the Stanford
Achievement Test.

Investigations have also been conducted in which the Raven Progressive
Matrices were used as measures of intellectual ability. Levine and Iscoe
(23) investigated the feasibility of using the Raven Progressive Matrices
(1938) to evaluate the intelligence of the adolescent deaf. The test was
administered individually to seventy-three subjects, thirty-six males and
thirty-seven females, who were residents in a school for the deaf. Some had
previously taken the Chicago Non-Verbal Examination while others had taken
the Performance Scale of the Wechsler-Bellevue, Form I. The correlation

---

8 Throughout the remainder of this paper the Raven Progressive Matrices
will hereafter be referred to as Raven (1938).

9 Throughout the remainder of this paper the Chicago Non-Verbal Examination
will hereafter be referred to as C.N.V.E.

10 Throughout the remainder of this paper the Performance Scale of the
Wechsler-Bellevue, Form I, will hereafter be referred to as W.B., I.
between the W.B.,I and Raven ('38) was .55, significant at the one per cent level of confidence. Although this correlation was not high enough for individual prediction, Levine and Iscoe considered it sufficiently high for the Raven ('38) to be included in a test battery for screening purposes.

In the study, Levine and Iscoe stated they had made use of "recent scores on C.N.V.E. and W.B.,I. They did not state the exact amount of time between administration of the different tests. It is possible that the difference found between test scores could be attributed to changes in the skills of an individual in the period of time between tests rather than to real differences between the tests themselves.

The next study reviewed is one in which the Raven was administered to one hundred, hearing children, by Martin and Wiechers. (27) It was the high correlations obtained by them that prompted the writer to employ the Raven as a criterion for the Leiter and Ontario. All the children were nine years of age, the same as a third of the subjects used by the writer. Martin and Wiechers individually administered the Raven and WISC to the one hundred children. The correlation obtained with the Verbal Scale I.Q. was .84, with the Performance Scale I.Q. was .83, and with the Full Scale I.Q. was .91. The intercorrelations with the subtests ranged from .74 (Block Design) to .47 (Information). The mean Full Scale I.Q. was 107, the Raven mean was 25.3. The investigators stated, "It can be inferred that the Raven score and WISC full scale I.Q. are measuring to a considerable extent the same factors." (27, 144)

The final inquiry summarized is that of Oleron (29). He studied what he considered to be the abstract intelligence of the deaf by means of the
Raven (*38). There were eighty-five subjects who ranged from nine to thirteen years of age. The medians rather than the means and standard deviations were presented. Cleron compared the medians of each deaf group with the corresponding age of the standardization group. On the basis of the medians obtained, he found that the deaf did poorer in the area of abstract reasoning. He also studied the influence of the age of onset of deafness. He found that children who became deaf after their sixth year were superior to those congenitally deaf and to those who became deaf at an early age. In examining the influence of the amount of hearing loss, he found that residual hearing did not have a marked influence.

Available literature indicates that few studies have been conducted on the applicability of the Ontario, Leiter and Raven to deaf subjects. Of the studies made, samples were usually small in number and limited in geographic location.

Another criticism that can be made is the frequent failure of the investigators to report sufficient descriptive information regarding procedure and data obtained. In any experimental study the variables and procedures must be adequately described so that another experimenter can repeat it under similar conditions. A further criticism of some of the studies is that results should be reported with suitable statistics. If these analyses are available, results of later investigations can be compared with those previously obtained. In reviewing the studies summarized above, three failed to state mean results and few report standard deviations. Thus, it is difficult to contrast results obtained in one study with those of another study.
Table I is a tabular summary of the studies reviewed in the foregoing paragraphs. The following facts are presented, if they were included in the original reports.

1. Those who conducted the study
2. Number of subjects used. Unless otherwise stated, all subjects were deaf.
3. The subjects' age range
4. Instruments employed
5. Mean test results
6. Correlation between two instruments
### TABLE I

**SUMMARY OF RELATED LITERATURE**

<table>
<thead>
<tr>
<th>Investigator</th>
<th>N</th>
<th>Age</th>
<th>Test</th>
<th>Mean</th>
<th>Re-test</th>
<th>Mean</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kirk and Ferry</td>
<td>49</td>
<td>5-0 to 11-0</td>
<td>Ontario</td>
<td>I.Q.</td>
<td>Nebraska</td>
<td>L.Q.</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>49</td>
<td>5-0 to 11-0</td>
<td>Ontario</td>
<td>I.Q.</td>
<td>Nebraska</td>
<td>L.Q.</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>49</td>
<td>5-0 to 11-0</td>
<td>Ontario</td>
<td>I.Q.</td>
<td>Binet</td>
<td>I.Q.</td>
<td>*</td>
</tr>
<tr>
<td>Ross</td>
<td>36</td>
<td>3-0 to 10-6</td>
<td>Ontario</td>
<td>I.Q.</td>
<td>Vineland</td>
<td>S.Q.</td>
<td>.91</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(possibly)</td>
<td></td>
<td>104.6</td>
<td></td>
<td>94.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>3-0 to 10-6</td>
<td>Nebraska</td>
<td>L.Q.</td>
<td>Vineland</td>
<td>S.Q.</td>
<td>.85</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(possibly)</td>
<td></td>
<td>104.8</td>
<td></td>
<td>94.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Ontario</td>
<td>I.Q.</td>
<td>Nebraska</td>
<td>L.Q.</td>
<td>.84</td>
</tr>
<tr>
<td>Birch and Birch</td>
<td>53</td>
<td>*</td>
<td>Leiter</td>
<td>*</td>
<td>Arthur</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>(1951)</td>
<td></td>
<td></td>
<td>Miskey</td>
<td>*</td>
<td>Wechsler</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Goodenough</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Birch and Birch</td>
<td>35</td>
<td>3-11 to 8-6</td>
<td>Leiter</td>
<td>*</td>
<td>teachers' intelligence ratings</td>
<td>.86</td>
<td></td>
</tr>
<tr>
<td>(1956)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Leiter</td>
<td>teachers' achievement ratings</td>
<td>.71</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>teachers' achievement ratings</td>
<td>.82</td>
</tr>
</tbody>
</table>
### TABLE I (Continued)

**SUMMARY OF RELATED LITERATURE**

<table>
<thead>
<tr>
<th>Investigator</th>
<th>N</th>
<th>Age</th>
<th>Test</th>
<th>Mean</th>
<th>Re-test</th>
<th>Mean</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levine and</td>
<td>73</td>
<td>15-0 to</td>
<td>Raven</td>
<td>(m)</td>
<td>Chgo. Non- Verbal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iscoe</td>
<td></td>
<td>19-0</td>
<td></td>
<td>34.1</td>
<td>31.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Raven</td>
<td>(m)</td>
<td>Perf. of Wechsler-</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>34.1</td>
<td>103</td>
<td>(m)</td>
<td>.55</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Raven</td>
<td>(f)</td>
<td>Bellevue (f)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>31.5</td>
<td>98</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Martin and</td>
<td>100*</td>
<td>9-0</td>
<td>WISC: Verbal</td>
<td>Raven</td>
<td>25.3</td>
<td>.84</td>
<td></td>
</tr>
<tr>
<td>Wechsers</td>
<td></td>
<td></td>
<td>Performance</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Full</td>
<td>107</td>
<td>Raven</td>
<td>25.3</td>
<td>.91</td>
</tr>
<tr>
<td>Clerion</td>
<td>88</td>
<td>9-0 to</td>
<td>Raven</td>
<td>*</td>
<td>Raven</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>13-0</td>
<td>(study group)</td>
<td></td>
<td>(normative group)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(results given as median score and percentile for each age group)

* Not reported

+ Hearing children
CHAPTER III

PROCEDURE

One function of the Catholic Charities Guidance Center is to administer psychometric examinations to deaf children in order that some estimate of their intelligence can be obtained. This information is an aid in placing a child in a suitable school group. More uniform school groups make teaching more effective and more pleasurable to both pupils and teacher. A teacher can also make provisions for individual differences that are revealed through tests, observations, and interviews.

The intelligence tests administered to deaf children at the Guidance Center are usually the Ontario School Ability Examination or the Leiter International Performance Scale. On occasion, it is necessary or requested that a child be re-tested at a later date. Frequently, the test used initially is not administered a second time. The Ontario or the Leiter, depending on what was utilized initially, is administered. When the test results are compared, they are frequently found to differ considerably. The present study is an attempt to gain more information on the extent of the differences between these two tests.

Description of Tests Used

Prior to a description of the procedure employed in the study, a discussion of the tests utilized is in order. The Ontario School Ability Examination (1) is a performance test constructed by Harry Amoss. It was
standardized on Canadian children of the Ontario Province. It was published specifically to measure the intelligence of children who are lacking in language facility. It was originally used in an attempt to decide the eligibility of candidates who applied for admission to the Ontario School for the Deaf, Ontario, Canada. The Ontario is divided into six examinations:
(a) Manipulation (paper folding, block building, Healy Fermald puzzle, judging weights, etc.)
(b) Colour Patterns (block designs)
(c) Knox Blocks
(d) Dominoes (adapted from Drever and Collins)
(e) Drawings (drawing geometric figures and designs)
(f) Tapping (rhythmic taps on blocks in prescribed order to be imitated by examinee)

Amoss stated that the Ontario is a composite of items from other tests such as the Gesell Block Building and the Drever Collins Block Design, etc.

Standardization of the initial form of the Ontario was begun by administering the test to thirty deaf pupils in the Day Classes for the Deaf in Toronto, Canada, and to approximately fifty hearing children who had been previously tested by the Revised Stanford-Binet. Certain portions were modified on the basis of the results of this testing and the succeeding form was administered to a group of hearing subjects. The number of subjects in this group was not stated. A mental age table was compiled on the basis of the per cent passing each item. Amoss states:
An item is placed under the mental age at which seventy-five per cent of the hearing candidates of that mental age as ascertained by a Stanford Revision Examination succeed in passing the test, save when in a sequence examination, such as the Knox Cube, fifty per cent of the candidates succeed in passing the second test beyond, in which case the first item beyond will appear at that mental age (1, 8).

The Ontario was then administered to 288 students at the Ontario School for the Deaf whose ages ranged from five to twenty-two years. Norms were determined from the results of this administration. It was found that the Ontario and Binet correlated more closely with children below twelve years of age. Amoss did not cite correlational data obtained nor did he state the reliability or validity of the test. He indicated the need for further experimentation. He stated a belief that the Ontario would probably be more valid in the Province of Ontario for determining school ability up to and including the mental age of twelve years.

The testing manual includes a complete description of materials used and administering procedures. In each of the six subtests, testing continues until a certain consecutive number of errors are made. Each subtest item is credited with a designated number of points. The mental age of the examinee equals the total number of points obtained. The intelligence quotients are obtained by the usual formula: M.A./C.A. x 100.

The second test utilized in the present study is the Leiter International Performance Scale. The Leiter is a cross-cultural test which originated in 1927. The individual test items were constructed and developed by Russell Graydon Leiter (21) through use with different ethnic groups in Hawaii. Later, Stanley D. Porteus (21) collaborated in applying the test to
groups in Hawaii and Africa. Several revisions were made. The latest full
scale revision was made in 1948 and this revision was used in this study.
The 1948 revision was based on the examination of children, high school
students and World War II Army recruits who were reared in the United States.

The coefficient of correlation between the 1940 and 1948 revisions was
.92. In regard to the intelligence quotient obtained Leiter states:

This study seemed to verify what was already
known about the difficulty of the 1940 Scale, namely,
that the 1940 Scale, and consequently the 1948
Revision, are scaled about six months or five points
of I.Q., too low. However, time did not permit a
full scale revision in 1948 and it may be 1960 or
later before a complete revision can be made and the
tests relocated so that an I.Q. of 100 will be the
norm. In the meantime whenever psychological
examiners wish to compare the results of the 1948
Scale with the results of the other tests, they may
do so very conveniently by adding five points to
the I.Q. obtained from the application of the 1948
Revision. (21, 58)

Leiter cites numerous studies on the reliability and validity of the
various revisions of the Leiter Scale (21). Those for the 1948 revision, the
one used in the present study, are:

(1) Miriam E. Tate found a reliability coefficient of .75 when she
applied the Leiter, Binet, and Arthur Point Scale, Form I, to
108 children who were five years old. (21, 58-60)

(2) Fifty mental defectives from six years, eleven months to sixteen
years, two months were administered the Leiter, Cornell-Coxe
Performance Ability Scale and the Binet by Louise Beverly and
Gerard Bensberg. The intercorrelations of the three scales were:
Leiter and Cornell-Coxe, .82; Leiter and Binet, .62; Cornell-Coxe
and Binet, .67. (21, 61)
(3) The study by Birch and Birch (5) the sole investigators who utilized deaf subjects, was presented in the previous chapter, "Review of Related Literature", pages 8 and 9.

(4) Robert Glenn compared the results of the Leiter and Binet Scales when they were administered to fifty-three unselected American white children: from six years to six years, eleven months of age. A correlation of .77 + .05₄ was obtained. (21, 63-64)

(5) Atwood Hudson compared the test responses of athetoid and spastic children who ranged from four years, eleven months, to twelve years, four months. He applied the Leiter, Ammons Full-Range Picture Vocabulary Test and the Raven. The intercorrelation of the three measures were: Leiter and Ammons, .78; Leiter and Raven, .84, Ammons and Raven, .83.

Arthur (3) also devised an adaptation of the Leiter Scale with the items in year levels II to XII which are applicable for testing children from four to eight years of age. She gave additional credit at the basal age established in order to make the scale more comparable with other tests. The Revised Form II of the Point Scale of Performance Tests is used in conjunction with the Arthur Adaptation of the Leiter Scale in order to increase the reliability of the rating obtained.

The Leiter Scale necessitates minimal directions either spoken or pantomime. The materials consist of a wooden frame with an adjustable card holder, card materials, and blocks. A card containing printed pictures is placed in the frame. Wooden blocks of the same shape and size on which are pasted proper response pictures are placed before the subject in the order
directed by the manual. The examinee then inserts the blocks in the appropriate stalls. In the 1938 Revision there are time limits only for four block design items which appear at the tenth, twelfth, and sixteenth year levels. In Arthur's revision, these time limits are eliminated.

Comprehension of the task presented is considered a part of the scale. At the lower levels of the test, each type of task is generally presented in a simple form which is readily comprehended by most subjects. The difficulties of each type increase as the test progresses. The scale measures a number of mental functions by means of block designs, analogies, matching colors or forms, counting concealed blocks, completing forms, memorizing a series, estimating number of dots, etc. The Leiter ranges from year level II to XVIII, all of which contain four tests except for the last level which contains six tests. Three months credit is given for every success from Year II to Year X. From Year XII, upwards, each test receives six months credit. When administering the Leiter a basal age is established and testing proceeds through two consecutive levels of failure. Scoring is similar to that of the Stanford-Binet. The number of years and months credited are summed in order to obtain the mental age. The I.Q. is calculated by means of the usual formula. When comparing results with those obtained by other tests, five points are added to the I.Q. as suggested by Leiter. (21, 58)

The third test employed in this study is the Raven's Coloured Progressive Matrices. (30 & 31) Both a board and a book form of the test are available. The former is utilized with pre-school children and the latter is suitable for children over six years of age. Occasionally a child succeeds in solving most of the problems in Sets A, Ab, and B of the Coloured Matrices. In such a
case Raven suggests that the subject proceed to Sets C, D, and E of the Matrices 1938 Scale in order to obtain a more complete assessment of intellectual capacity. (33, 10)

The Coloured Matrices Test is used with children and any person who cannot speak or understand the English language. Raven considers it "a test of observation and clear thinking". (33, 1) It consists of three sets, each with twelve problems which develop a consistent theme of thought. Each problem is presented as an abstract design from which a part is missing. The subject is to select the missing insert from one of six to eight alternatives. Simple directions are given in English or in pantomime. The Raven can be administered as an individual or group test. The number of correct responses is summed and computed with the normative group of the same chronological age. A percentile score is thus obtained.

The percentile norms were obtained on a representative sample of six hundred children living in the burgh of Dunfris, England. Their ages ranged from five to eleven and a half years. Subsequently every third, nine year old child of that group was administered the Stanford-Binet and Chrichton Vocabulary Scale. The coefficient of correlations between the Binet and Raven was .66; between the Binet and Chrichton, .83; and between the Raven and Chrichton, .65.

**Analyses of Tests Used**

As was discussed earlier in the paper, both a technical and a practical criterion must be considered in the selection of a suitable psychometric instrument. Table II summarizes the analyses made of the Ontario, Leiter, and
Raven. Table II is located on page 27.

There appears to be a deficiency of information regarding the reliability and validity of the Ontario and Raven. Amoss (1) discusses studies made to determine the reliability and validity of his scale but does not state coefficients of correlation obtained. He concludes: "The Ontario School Ability Examination is set forth as valid in the Province of Ontario and probably valid in other comparable areas for the purpose of determining the school ability up to and including mental age twelve years of deaf children, retarded children and children whose home is other than English." (1, 12) No further studies were available.

Raven quoted a test, re-test reliability of .90 for children nine years of age, and a reliability of .65 for those under seven years. No statistics were presented for test validity.

Leiter (21) had recently published numerous evidences of the reliability and validity of the various forms of the Leiter tests. Of the six studies conducted with the 1948 revision, one had utilised deaf subjects. From a review of the literature, one can readily see the necessity for many more studies which would obtain indications of the reliability and validity of the three tests used in this study.

In the table the norms for the Ontario, Leiter, and Raven are labeled "questionable". On the Ontario, norms are based on Canadian deaf children. A search of literature did not disclose a normative study with this examination, based on deaf children residing in the United States. Therefore, there is some doubt as to the applicability of available norms to Chicago residents.
The 1948 revision of the Leiter Scale has its norms based on Americans. In this respect the obtained norms are acceptable. However, there are no norms available for the deaf. Whether or not deaf subjects should be compared with a representative sample of the deaf or with a representative sample of the normal population is a much disputed question. Many opinions have been presented for both sides. It is the writer's opinion that norms on both types of samples should be available for tests extensively used with deaf subjects. In this way the testee can be compared both with others who have the same handicap and also with a normal sample.

The norms for the Raven Matrices were obtained on a sample of English children. Whether these norms are applicable to a group of deaf children residing in the United States is questionable.

In the experience of the writer and of other examiners consulted, a major drawback of the Ontario is the difficulty with which several of its portions are administered. For example, on the items which involve judgment of weights, a majority of deaf subjects appear unable to grasp what is expected of them. Also, the tapping examinations require five rhythmic uniform taps per four seconds. It is rather difficult to maintain uniformity and also to keep in mind the sequence which must be precisely followed, especially at the upper levels.

Another criticism of the Ontario is that almost half of the items are timed. It is difficult to convey to a deaf child the need to work with speed. At times, he may misinterpret the tester, think he is wrong, and stop working. More time is lost and results are lowered. Thus, a deaf child is
frequently at a disadvantage on timed items.

The response pattern is generally set without much difficulty at the beginning of the Leiter and Raven examinations. However, on the Ontario the testee is expected to respond in a different manner from one portion to another. Some children, especially those with little if any school experience tend to perseverate and have difficulty adjusting to the change. Frequently a lowered score results.

Besides being expensive, the Leiter has the additional disadvantage of being heavy. Three large boxes can be placed in a canvas bag for transporting from one office to another. If no automobile is available, the bag becomes a burden to carry. Both the Ontario and Raven are compact and light.

These are the advantages and disadvantages of the Ontario, Leiter, and Raven as viewed by the writer. The final judgment of which is better or more suitable should be made by each examiner who is aware of the needs of his particular testing situation.

Selection of Subjects

The subjects used in this study were seventeen boys and seventeen girls who attend day schools for the deaf. These schools are members of the hearing program of the Department of Special Services. The chronological ages of the children ranged between seven and ten years. Their mean age was eight years, six months, the standard deviation was 10.9 months. The particular age group was selected because the children were presumed to have

11 See Table II, page 27.
### TABLE II

**ANALYSES OF ONTARIO, LEITER, AND RAVEN**

<table>
<thead>
<tr>
<th>Categories</th>
<th>Ontario</th>
<th>Leiter</th>
<th>Raven</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reliability</td>
<td>No r stated</td>
<td>.92</td>
<td>under 7 yrs: .65; age 9: .80</td>
</tr>
<tr>
<td>Validity</td>
<td>No r stated</td>
<td>.77</td>
<td>No r stated</td>
</tr>
<tr>
<td>Scoring</td>
<td>standardized</td>
<td></td>
<td>standardized</td>
</tr>
<tr>
<td>Norms</td>
<td>questionable</td>
<td></td>
<td>questionable</td>
</tr>
<tr>
<td>Cost of Materials</td>
<td>reasonable ($17)</td>
<td>expensive ($180)</td>
<td>reasonable ($2.60)</td>
</tr>
<tr>
<td>Type of Scoring</td>
<td>hand scored</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ease of Scoring</td>
<td>somewhat difficult</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time for Scoring</td>
<td>satisfactory</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ease of Administration</td>
<td>somewhat difficult</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time for Administration</td>
<td>20 min. to 1 hr.</td>
<td>20 min. to 1 hr.</td>
<td>15 min. to 30 min.</td>
</tr>
<tr>
<td>Comparable forms Available</td>
<td>none</td>
<td>none</td>
<td>none</td>
</tr>
</tbody>
</table>
sufficient social maturity and school training to respond satisfactorily in an individual testing situation. Also, it was believed that each child could reach the limits of his mental ability without exceeding the limits of the three tests utilized. Both of these suppositions proved correct. Each subject demonstrated good cooperation and effort in the testing situation. In no case were the test ceilings insufficient.

Contacting the Subjects

A form letter was sent to the parents of all children between the ages of seven and ten. The letter advised of the nature and purpose of the proposed study and requested their cooperation. Subsequently, individual contacts were made by telephone or letter to arrange for appointments. Of the thirty-eight letters sent, successful contacts were made with the parents of thirty-four children. The parents were cooperative and eager to be of assistance. Occasionally, appointments were cancelled for reasons such as illness on part of child or parent, school party, over-sleeping, etc. A second scheduled appointment was kept in all cases.

Administration of Tests

The writer individually administered the three tests to the thirty-four deaf children. Most were examined at the central office of the Catholic Charities Guidance Center. Several were seen in suitable testing rooms at two of the schools. The requirements for satisfactory testing conditions were met. The examinations were administered and scored in accordance with the specifications set down by the manuals. All scoring was checked a second time.
All three tests were administered in one session. These testings ranged in time from one hour to two hours and fifteen minutes. The majority of the examinations were completed within two hours. Short rest periods were given appropriately. The administration of the three tests was staggered: Ontario, Leiter, Raven; Raven, Ontario, Leiter; Leiter, Raven, Ontario; Ontario, Leiter, Raven; Raven, Ontario, Leiter; etc.

The Sample

As far as can be judged from a survey of accepted applicants, the subjects of this study are representative of children accepted into the educational program for deaf children conducted by the Department of Special Services of Catholic Charities. Children are referred by clinics, doctors, nurses, social workers, or relatives of other deaf children. The following information is obtained on each applicant: otological diagnosis, hearing measurements, educational history and mental and social evaluations. On the basis of these data the decision is made as to whether or not an applicant would profit from formal, special education. If so, he is enrolled in one of the four day school centers located in different sections of the city.

All thirty-four subjects were white. Thirty-one were considered to be in good health by physicians. Several children had visual defects corrected by prescribed lenses. One girl had slight chorea, another girl had slight athetoid movements, a boy had bilateral foot deformities and another boy had a collapsed lung. None appeared handicapped on the performance tests by these physical defects.
All children were born deaf or became deaf in early childhood. In fourteen cases, the onset and cause of deafness was undetermined. Twelve subjects had congenital deafness. Childhood illnesses, whooping cough, meningitis, yellow jaundice, and chicken pox, which occurred between the ages of two to fifteen months, were presumed responsible for the deafness of seven subjects. In one case, the cause was considered to be hereditary. Two children had two siblings who were also deaf. Four had one deaf sibling. In three instances, a sibling and/or a relative was hard of hearing and/or deaf. In twenty-six cases there was no indication of deafness in other members of the family.

Of the thirty-four children, the records indicated that fourteen had had nursery school or kindergarten experience prior to their entrance in one of the Catholic schools for the deaf. Five had had instruction at the primary level in a public school or a regular Catholic school. The remaining nineteen children had received all their formal education from one of the day centers for the deaf. Four of these had transferred from one day center to another.

**Statistical Analysis**

A statistical analysis of the test results was made. For the Ontario and Leiter, the intelligence quotient scores were utilized. For the Raven, raw scores were used.

Since \( N \) was below fifty, small group statistics were utilized. Mean scores and the standard deviations of the performances on the three scales were calculated. The Pearson product-moment correlation was used to calculate coefficients of correlation between the Ontario, Leiter, and the Raven. Finally, the reliabilities of the coefficients of correlation were determined.
CHAPTER IV

RESULTS

The thirty-four subjects, when tested with the Ontario School Ability Examination, obtained a mean I.Q. of 104.2; the median I.Q. was 105.5; the standard deviation was 12.5. The same group, examined with the Leiter International Performance Scale, received a mean I.Q. of 98.4, a median I.Q. of 99.5, and a standard deviation of 9.62. The raw score results obtained on the Raven Progressive Matrices were: mean, 19.1; median 18.5; and standard deviation, 5.74.

The Ontario mean I.Q. of 104.2 obtained by the present group of thirty-four children is similar to the findings reported by Kirk and Perry (17), with forty-nine deaf children whose ages ranged from five to eleven years. Their mean I.Q. was 102.9. Ross (34), in her study of thirty-six deaf children who ranged from two years to ten years, also reported an Ontario mean I.Q. of 104.6. This is a slight deviation from present results.

Since Birch and Birch (5 and 6) did not report a mean, median, or standard deviation in either study, their Leiter findings cannot be directly compared with those of the present study.

Martin and Wiechers (27) administered the Raven to a group of one hundred hearing, nine-year-old children and obtained a mean score of 25.3. This is higher than the mean score, 19.1, presently obtained. Although not specifically mentioned, it is presumed that the mean is derived from raw
scores. The thirty-four subjects in this study ranged from seven years to nine years, ten months of age. Since raw scores are generally intended to vary in proportion with the age of subjects it is expected that a group of nine year olds will tend to score somewhat higher. Thus, the deviation of scores is in the proper direction. In addition, Martin and Wiechers subjects were normal, whereas present subjects were deaf.

Levine and Iscoe (23) administered the Raven to seventy-three deaf adolescents. The boys obtained a mean raw score of 34.1 while the girls had a mean raw score of 31.5. These results are considerably higher than mean results presently obtained. It is probable that the major part of the discrepancy is due to the ages of subjects involved. Adolescents, provided that they are not mentally retarded, generally tend to obtain higher raw scores than average seven to nine-year-olds.

The Ontario intelligence quotient scores ranged from 81 to 134, or 53 points, while those of the Leiter ranged from 72 to 123, or 51 points. The Ontario's range is only two points more than the Leiter's. However, the former scores tend to be higher than the latter scores.

The Raven's raw scores ranged from raw score of 11 to raw score of 36, or 25 points.

The Pearson product-moment method of correlation was applied to the results of the three scales. There was a correlation of .78 between the Ontario and Leiter results, significant at the one per cent level of confidence. This correlation would indicate that the two scales measure to a great extent the same factors. The coefficients of correlation between the Ontario and Raven was .28 and between the Leiter and Raven was .13, neither of which was
significant. It would seem that the Raven measure few of the qualities measured by the Ontario and Leiter scales.

In analyzing individual test results, some variations between scores obtained by the Ontario and by the Leiter were noted. Table IV presents the variations between scores obtained on these two tests.

From Table IV it can be seen that approximately two-thirds of the scores did not vary by more than ten points. In one case the same score was maintained on both tests. Six children received a higher Leiter than Ontario score. The Leiter intelligence quotient was lower than the Ontario intelligence quotient in twenty-seven cases. The amount of increase ranged from one to fifteen points with a mean increase of 6.3; the amount of decrease ranged from one to twenty-one points with a mean decrease of 9.3 points.

**TABLE IV**

**CHANGES BETWEEN LEITER AND ONTARIO I.Q. POINTS**

<table>
<thead>
<tr>
<th>Points</th>
<th>Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - 5</td>
<td>12</td>
</tr>
<tr>
<td>6 - 10</td>
<td>10</td>
</tr>
<tr>
<td>11 - 15</td>
<td>10</td>
</tr>
<tr>
<td>16 - 20</td>
<td>1</td>
</tr>
<tr>
<td>21 - 25</td>
<td>1</td>
</tr>
</tbody>
</table>
CHAPTER V

SUMMARY AND CONCLUSIONS

Among the instruments employed by psychologists in evaluating the mental ability of deaf children are the Ontario School Ability Examination, the Leiter International Performance Scale and the Raven Coloured Progressive Matrices. Few studies have been conducted on the applicability of these scales to deaf children. The purpose of this study was to explore the relationships between them when administered to a group of deaf children. In addition, the tests were analyzed and judged on the basis of technical and practical criteria.

The following hypotheses were tested in the study.

(1) It is hypothesized that there will be a significant relationship between Ontario and Raven results when the two tests are administered to a group of deaf children from seven to nine years of age.

(2) It is hypothesized that there will be a significant relationship between Leiter and Raven results when the two tests are administered to a group of deaf children from seven to nine years of age.

(3) It is hypothesized that there will be a significant relationship between Ontario and Leiter results when the two tests are administered to a group of deaf children from seven to nine years of age.
A survey of literature indicated that the following studies employed the Ontario, Leiter, or Raven in the testing of deaf children.

Kirk and Perry (17) administered the Ontario and Nebraska to forty-nine deaf children and to forty-nine hearing children. The latter group were also given the Binet. The children’s ages ranged from five to eleven years. A statistically significant difference was found between the Ontario mean I.Q. of 102.9 and a Nebraska mean I.Q. of 95.8, with the group of deaf children. In the other group the mean scores were: Binet, 107.3, Ontario, 102.9, and Nebraska, 95.7. The differences between the Binet and Ontario were not statistically significant while those between the Ontario and Nebraska were significant.

Ross (35) studied the relative effectiveness of the Ontario, Nebraska, and Vineland. The thirty-six subjects obtained the following mean scores: Ontario, 104.6, Nebraska, 104.8, and Vineland, 94.7. The correlation between the Ontario and Vineland was .91; between the Nebraska and Ontario was .81; and between the Nebraska and Vineland was .65. The children may have ranged from three to ten years of age.

Birch and Birch conducted two studies with the Leiter. In the 1951 study (5) they compared Leiter results with those obtained by four other tests of intelligence. The subjects were composed of fifty-three deaf children who were experiencing learning difficulties. They concluded if the Leiter scores were considerably below the four other test scores, one could predict that the child would find it difficult to learn.

In the 1956 study (6) Birch and Birch compared Leiter scores with teachers’ intelligence and achievement ratings. The thirty-five subjects
ranged from three to eight years of age. The correlations between Leiter I.Q.'s and teachers' ratings of general intelligence was .86; between the Leiter and teachers' achievement ratings was .71 and between teachers' intelligence and achievement ratings was .82.

Levine and Iscoe (23) administered the Raven (138) to seventy-three deaf adolescents. They compared results with those previously obtained on the W.B., I and C.N.V.E. The correlation between the C.N.V.E. and Raven was .41, which was not significant. The correlation between the W.B., I and Raven was .55, significant at the one per cent level of confidence. They considered results sufficiently high for the Raven to be included in a test battery for screening purposes.

The Raven was also used by Martin and Becher in their testing of one hundred hearing children, nine years of age. (27) Raven results were compared with WISC I.Q.'s and following correlations were obtained: Verbal Scale .81, Performance Scale .83, Full Scale .91. The mean Full Scale I.Q. was 107, the Raven mean was 25.3. The investigators inferred that the two tests measure the same factors to a considerable extent.

The final study reviewed was that of Ceron. (28) His eighty-five subjects ranged from nine to thirteen years of age. The medians of each deaf group were compared with the medians of the corresponding age in the standardization group. The deaf tended to do poorer than the standardization group.

The thirty-four children who served as subjects for this study attended day schools for the deaf. These schools are a part of the hearing program of the Department of Special Services, Catholic Charities, Archdiocese of
Chicago. The ages of the children ranged from seven years to nine years and ten months. Each child was administered the Ontario, Leiter, and Raven on the same day. Appropriate statistical procedures were applied to obtain the range, mean, median, standard deviation, and Pearson product-moment coefficient of correlation. These statistics are summarized in the following table.

**TABLE V**

**SUMMARY OF STATISTICAL RESULTS**

<table>
<thead>
<tr>
<th></th>
<th>Ontario</th>
<th>Leiter</th>
<th>Raven</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Range</strong></td>
<td>81 to 134</td>
<td>72 to 123</td>
<td>11 to 36</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>104.2</td>
<td>98.4</td>
<td>19.1</td>
</tr>
<tr>
<td><strong>Median</strong></td>
<td>105.5</td>
<td>99.5</td>
<td>18.5</td>
</tr>
<tr>
<td><strong>Standard Deviation</strong></td>
<td>12.5</td>
<td>9.62</td>
<td>5.74</td>
</tr>
<tr>
<td><strong>Correlations</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ontario</td>
<td></td>
<td>.78</td>
<td>.28</td>
</tr>
<tr>
<td>Leiter</td>
<td></td>
<td></td>
<td>.13</td>
</tr>
</tbody>
</table>

On the basis of these calculations the following results were obtained:

1. The group of thirty-four children performed on an average level with other groups which used the Ontario and Leiter. The Raven score tended to be somewhat lower.
(2) A Pearson product-moment correlation of .78 was obtained between the Ontario and the Leiter. This was significant at the one percent level of confidence. Thus hypothesis number three can be accepted.

(3) The correlation between the Ontario and Raven was .26 and between the Leiter and Raven was .13. Neither correlation was significant. Hypotheses one and two can be rejected.

(4) A variation of intelligence quotients on the Ontario and Leiter was noticed. Approximately two-thirds of the Ontario results were within ten points of the Leiter results. Six children increased their Leiter score whereas twenty-seven children decreased it. The mean increase was 6.3 points. The mean decrease was 9.3 points.

It appears that the following conclusions can be drawn from these findings. Since the coefficient of correlation between the Ontario and Leiter was highly significant it seems that either test would be acceptable for testing deaf children. However, notice should be made of the tendency for Ontario results to be somewhat higher than Leiter results. Thus, it is imperative that the instrument utilised to obtain an estimate of mental ability be also indicated with the intelligence quotient.

The Raven tended to have a low correlation with the Leiter and Ontario tests of intelligence when it was administered to a deaf group from seven to ten years of age. However, the mean percentile rank for each age group compared favorably with the standardization group except for those who were in the nine and one-half years age group.
Much more research with the Ontario, Leiter, and Raven appears necessary. Fields open for study would include data regarding the reliability and validity of the instruments and also normative data based both on a normal and deaf sample of the population.

The tendency of researches appears to be the utilization of an available test as a validity criterion for a test under study. If the validity is sufficiently high one could be considered an approximation of the other one. Anastasi emphasizes, "that unless the new test represents a simpler or shorter substitute for the earlier test, the use of the latter as a criterion is indefensible . . . If the new test correlates too highly with an already available test, without such added advantages as brevity or ease of administration, then the new test represents needless duplication." (2, 1:3-4)

Of the studies examined, all but one compared results of one test with those of one or more other tests. From such information one can estimate to what extent one test approximates another. Another aspect of the test may also be of interest to examiners and teachers of the deaf. That is, how valid is this test result as an estimate of the deaf child's ability to learn? To obtain some data towards a satisfactory reply to this question a criterion different from that of another test appears needed. A satisfactory one seems to be that utilized by Birch and Birch in their 1956 study. (6) They correlated Leiter results with teachers' intelligence ratings and teachers' achievement ratings and obtained a high positive correlation. It was regrettable that a technique similar to that of Birch and Birch was not employed in the present study. As is often said, "Hindsight is better than foresight." It may also be helpful to employ counseling, in the form of
play therapy, with those deaf who do not respond adequately to test situations.


APPENDIX I

TABLE III

FREQUENCY DISTRIBUTION OF RAVEN'S SCORES

<table>
<thead>
<tr>
<th>Raw Scores</th>
<th>f</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 - 12</td>
<td>4</td>
</tr>
<tr>
<td>13 - 15</td>
<td>6</td>
</tr>
<tr>
<td>16 - 18</td>
<td>7</td>
</tr>
<tr>
<td>19 - 21</td>
<td>5</td>
</tr>
<tr>
<td>22 - 24</td>
<td>7</td>
</tr>
<tr>
<td>25 - 27</td>
<td>3</td>
</tr>
<tr>
<td>28 - 30</td>
<td>0</td>
</tr>
<tr>
<td>31 - 33</td>
<td>1</td>
</tr>
<tr>
<td>34 - 36</td>
<td>1</td>
</tr>
</tbody>
</table>

N=34

TABLE IV

CHANGES BETWEEN LEITTER AND ONTARIO I.Q. POINTS

<table>
<thead>
<tr>
<th>Points</th>
<th>Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - 5</td>
<td>12</td>
</tr>
<tr>
<td>6 - 10</td>
<td>10</td>
</tr>
<tr>
<td>11 - 15</td>
<td>10</td>
</tr>
<tr>
<td>16 - 20</td>
<td>1</td>
</tr>
<tr>
<td>21 - 25</td>
<td>1</td>
</tr>
<tr>
<td>-----</td>
<td>------</td>
</tr>
<tr>
<td>1.</td>
<td>7-8</td>
</tr>
<tr>
<td>2.</td>
<td>9-10</td>
</tr>
<tr>
<td>3.</td>
<td>7-0</td>
</tr>
<tr>
<td>4.</td>
<td>8-8</td>
</tr>
<tr>
<td>5.</td>
<td>7-8</td>
</tr>
<tr>
<td>6.</td>
<td>7-10</td>
</tr>
<tr>
<td>7.</td>
<td>7-5</td>
</tr>
<tr>
<td>8.</td>
<td>9-9</td>
</tr>
<tr>
<td>9.</td>
<td>9-8</td>
</tr>
<tr>
<td>10.</td>
<td>9-9</td>
</tr>
<tr>
<td>11.</td>
<td>9-8</td>
</tr>
<tr>
<td></td>
<td>Ontario</td>
</tr>
<tr>
<td>---</td>
<td>---------</td>
</tr>
<tr>
<td>12.</td>
<td>C.A. M.A. 6-4</td>
</tr>
<tr>
<td></td>
<td>7-3</td>
</tr>
<tr>
<td>13.</td>
<td>C.A. M.A. 9-2</td>
</tr>
<tr>
<td></td>
<td>9-8</td>
</tr>
<tr>
<td></td>
<td>9-4</td>
</tr>
<tr>
<td>15.</td>
<td>C.A. M.A. 10-6</td>
</tr>
<tr>
<td></td>
<td>7-10</td>
</tr>
<tr>
<td>16.</td>
<td>C.A. M.A. 8-10</td>
</tr>
<tr>
<td></td>
<td>8-4</td>
</tr>
<tr>
<td>17.</td>
<td>C.A. M.A. 7-6</td>
</tr>
<tr>
<td></td>
<td>7-2</td>
</tr>
<tr>
<td>18.</td>
<td>C.A. M.A. 9-8</td>
</tr>
<tr>
<td></td>
<td>8-2</td>
</tr>
<tr>
<td>19.</td>
<td>C.A. M.A. 7-8</td>
</tr>
<tr>
<td></td>
<td>8-7</td>
</tr>
<tr>
<td>20.</td>
<td>C.A. M.A. 6-6</td>
</tr>
<tr>
<td></td>
<td>7-11</td>
</tr>
<tr>
<td>21.</td>
<td>C.A. M.A. 8-10</td>
</tr>
<tr>
<td></td>
<td>7-3</td>
</tr>
<tr>
<td>22.</td>
<td>C.A. M.A. 10-5</td>
</tr>
<tr>
<td></td>
<td>9-10</td>
</tr>
<tr>
<td>23.</td>
<td>C.A. M.A. 9-0</td>
</tr>
<tr>
<td></td>
<td>8-7</td>
</tr>
<tr>
<td>#</td>
<td>Ontario</td>
</tr>
<tr>
<td>----</td>
<td>---------</td>
</tr>
<tr>
<td>22</td>
<td>C.A. 8-8</td>
</tr>
<tr>
<td>26</td>
<td>C.A. 8-10</td>
</tr>
<tr>
<td>27</td>
<td>C.A. 9-7</td>
</tr>
<tr>
<td>28</td>
<td>C.A. 8-4</td>
</tr>
<tr>
<td>29</td>
<td>C.A. 9-9</td>
</tr>
<tr>
<td>30</td>
<td>C.A. 9-10</td>
</tr>
<tr>
<td>31</td>
<td>C.A. 8-0</td>
</tr>
<tr>
<td>32</td>
<td>C.A. 9-4</td>
</tr>
<tr>
<td>33</td>
<td>C.A. 8-2</td>
</tr>
<tr>
<td>34</td>
<td>C.A. 8-8</td>
</tr>
</tbody>
</table>

Psychometric examinations:
- Ontario School Ability Examination
- Leiter International Performance Scale
- Raven's Coloured Progressive Matrices
APPROVAL SHEET

The thesis submitted by Angela Lucille Chemazar 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.

March 15, 1919
Date

Signature of Adviser