1931

Tests for the prediction of reading success and reading failure on the primary level..

Storer, Elizabeth Wood
Boston University

https://archive.org/details/testsforpredicti00stor

Boston University
Thesis

A STUDY OF TESTS FOR THE PREDICTION OF READING SUCCESS AND READING FAILURE ON THE PRIMARY LEVEL

by

Elizabeth Wood Storer

(B.A., Wellesley, 1929)

submitted in partial fulfilment of the requirements for the degree of Master of Arts

1931
OUTLINE

I. Introduction: Summary of previous studies made on the prediction of reading ability. page 1
   A. Prediction on the basis of mental age. 2
   B. Prediction on the basis of reading readiness. 3

II. Part I: Study of the Stanford-Binet sub-tests 11
   A. The Problem 11
      1. Allied Research 11
   B. Collection of data. 14
   C. Study of data. 17
   D. Interpretation of results. 21
   E. Recommendations. 22

III. Part II: Study of a reading-readiness test battery. 24
   A. Selection of tests 24
      1. Test of attention 24
      2. Test of word-matching ability 24
      3. Test of vocabulary background 30
   B. Administration and scoring of tests 38
      1. Test of attention 38
      2. Test of word-matching ability 38
      3. Test of vocabulary background 40
   C. Study of test results 43
      1. Collection of data 43
      2. Treatment of data 43
      3. Interpretation of data 49
   D. Criticism of tests, and recommendations 54
      1. Test of word-matching ability 54
      2. Test of vocabulary background 56
      3. Test of attention 57

IV. Summary 58

V. Bibliography 60
I.

INTRODUCTION
SUMMARY OF PREVIOUS STUDIES MADE ON
THE PREDICTION OF READING ABILITY.

Within recent years educators have shown an increasing interest in, and have become more seriously concerned with, finding a satisfactory means of determining children's ability to achieve success in first-grade work. Evidence of this fact may readily be found in Reed's inquiry into the first-grade entrance and promotion requirements of various school systems throughout the country. Her study of the responses made by 220 educational administrators to a detailed questionnaire on this subject reveals the fact that admission standards vary considerably. The oldest and most prevalent is chronological age, but this is now almost universally recognized as a poor criterion of ability to do school work. Next to chronological age the standard reported most frequently was mental ability, the most common figure being six years. Other factors mentioned by these administrators but indicated much less frequently than either chronological age or mental ability were general ability to do first grade work, general health, physical maturity and emotional stability. The methods of determining these various factors varied all the way from teachers' estimates to the more objective psychological and physical tests. Regarding the standards for promotion to the second grade, Reed found progress in reading to be the most important

1. Reed, Mary M. An Investigation of Practices in First Grade Admission and Promotion.
criterion. If this be the case, the determination of children's ability to achieve success in the first grade as it is now organized is largely dependent on the prediction of reading success. Consequently the greater part of the research in the field of predicting first-grade success has concerned itself with reading. These studies resolve themselves chiefly into two types of problems, namely the determination of the relationship between mental age and the ability to do school work, and the formulation of tests of reading readiness.

A. Prediction on the basis of mental age.

Arthur's study of 171 first-grade children grouped by half-years of mental age indicated that mental age is a very important factor in reading achievement and that the optimum mental age for beginning to teach reading for comprehension is \(\frac{3}{4}\) to 7 years. Zornow and Pechstein, Thiesen, and Brooks find chronological age a very poor indicator of ability to succeed in reading or first-grade work in general, and produce evidence of a positive correlation between mental age and reading ability. Zornow and Pechstein state that in their study those children having a normal intelligence quotient and a mental age of at least six years "have uniformly been equal to the work of the first grade". They conclude that the Stanford-Binet Test is a "reasonably reliable method of classifying young children".

3. Zornow and Pechstein. An Experiment in the Classification of First Grade Children.
4. Thiesen, W. W. Factors Affecting Results in Primary Reading.
5. Brooks, Fowler. Chronological Age as a Factory in Reading Achievement.
Dickson made a study of first-grade children in the Oakland, California, schools and found that a "very high percentage of those who test below six years, mentally, fail to do the work which has been set up as standard for first-grade pupils". He also conducted an experiment on the prediction of school progress on the basis of mental age. After 4½ years the predictions were found to be true in 90% of the cases. He, therefore, concludes that the mental test furnishes a valuable index of a child's chances for success in school work" and he offers a set of standards in terms of intelligence quotients. Thus, there appears to be a general agreement as to the value of the mental age as a criterion of first-grade and later school success and the importance of making the minimum entrance requirement a mental age of six years.

B. Prediction on the basis of reading readiness.

A survey conducted by the International Kindergarten Union's Committee on Reading Readiness revealed some significant data. The responses to a questionnaire sent to teachers of first grades throughout the country led the committee to reach the following conclusions: that approximately one-fifth of the children entering the first grade are not ready to read in September, and that the outstanding causes of reading failure are lack of mental efficiency, lack of interest, and lack

7. Dickson, Virgil E. What First Grade Children Can Do in School.
9. Dickson, Virgil E. The Use of Mental Tests in School Administration.
With the first-grade curriculum founded so largely on reading and providing for the early introduction of the child to reading, this large percentage is most significant both from the point of view of the child and his attitude toward school, and that of the school authorities and their expenditure of time, money, and effort. Not only is the immature child apt to develop faulty reading habits which may require much unnecessary effort on the part of both teacher and child to overcome, but he is also in danger of developing a most unwholesome attitude toward school work in general. The satisfactory correction of the latter condition quite frequently requires the help of a psychiatrist. In other words, premature introduction to reading may give the child a most unsatisfactory beginning and result in needless maladjustments the correction of which will necessitate the expenditure of much time and effort which might have been used in exploring new fields.

There appear to be two ways of meeting this problem. One is the reorganization of the curriculum to provide for a more thorough preparation before the teaching of reading is attempted. The other point of attack is the formulation of satisfactory tests of reading readiness which can be used to exclude from first-grade reading classes all those who are not psychologically mature enough to make a satisfactory approach to reading. Probably the real solution lies in a combination of the two, that is, the reorganization of the curriculum to provide for individual differences and the use of tests to determine those differences, especially in the field of reading readiness. Curriculum reorganization is too large a field to be included in this study.
Tests of reading readiness are as yet in the experimental stage, but much of the material which has been developed should prove valuable in the construction of a standardized test. Rockwell reports the use of the Cleveland Kindergarten Achievement Test which serves as a "check on minimum attainments in motor coordination, manual skill, music, and English". A Picture Cross-out Test known as a General Information Test examines the child's knowledge of color, number, form, and general information. Miss Rockwell questions its universal applicability, since kindergarten curricula and customs differ quite widely in various sections of the United States; however it is quite possible that certain portions of it at least could be made sufficiently general and adapted for use as a measure of the child's vocabulary and experience. Such a test would be more valuable if designed as a test not of kindergarten achievement but of reading readiness, since the kindergarten has not yet become a part of many school systems.

In describing the Baltimore Reading Readiness Test, Berry states that because "experienced teachers of the entering IB children have found that the chronological or mental age of six does not determine this ability" (readiness to begin reading), a test was devised for this purpose. This test is composed of two parts: (1) a word discrimination test, based on the Thordike Word List and designed to measure the ability to find like-

nesses and differences in printed words and (2) a picture vocabulary test composed of the nouns, adjectives, and verbs occurring most frequently in primers. While the predictive value of the test has not yet been determined, its correlation with the Detroit First Grade Intelligence Test has been found to be .796.

Gates made a study of the relationship between reading achievement and several kinds of mental ability. His results are based on the following group of tests:

A. Visual perception.

Test 1: Detection of small differences between pairs of (a) geometrical figures, (b) digits, and (c) words.

Test 2: Matching a sample with one of a group of five or more similar geometrical figures or words.

Test 3: Speed of naming common objects represented by drawings.

B. Capacity for auditory-visual associative learning.

Test: Learning to associate certain geometric figures with spoken words.

C. Capacity for visual-visual associative learning.

Test: Learning to associate certain geometric figures with the nonsense word printed beneath.

D. Capacity for general linguistic and abstract learning.
   Test: Stanford-Binet Test.

E. Reading Abilities.
   Test 1. Speed of recognition and pronunciation of words of varying degrees of difficulty.
   Test 2. Level of recognition and pronunciation of words of varying degrees of difficulty.
   Test 3. Rate of comprehension of easy material.
   (Courtis Silent Reading Test, No. 2.)

F. Spelling ability.
   Test: Spelling words of graded difficulty.

This material, of course, was not used for purposes of prediction, but portions of it quite possibly might be included in a test of reading readiness or of capacity to achieve success in reading. On the basis of results obtained from tests of 310 children in grades one through seven, Gates concluded that "of the several abilities studied, that termed 'word perception' is most closely associated with achievement in reading and spelling"; that "intelligence yields the next highest correlation"; and that the other tests show a much smaller degree of correlation.

20 Smith, Nila B. Matching Ability as a Factor in First Grade Reading.
that the difficulties of matching capitals with lower case letters "do carry over in recognizing the capitalized forms of word wholes". A comparison of the ability of 200 children in matching lower case letters during the first week of school with their progress in reading measured by the Detroit Word Recognition Test during the twelfth week of school revealed a correlation of .87. Miss Smith also found a significant relationship between intelligence and the ability to match lower case letters. These results suggest the possibility of using a letter and word matching test as a means of predicting reading success in the first grade.

Probably the most recent investigation of reading readiness is that of Deputy based on the tests of 103 first-grade children. During the first four weeks of school they were given the Pintner-Cunningham Primary Mental Test and the following four individual tests:

1. A test of word selection consisting of letter and word matching.

2. A test of visual-auditory association consisting of associating a spoken word with a simple geometric figure.

3. A test of visual-visual association consisting of associating the printed name of a familiar object with its picture.

22. Idem. p. 62
23. Idem. p. 60
24. Idem. p. 62
25. Deputy, E. C. Predicting First-Grade Reading Achievement.
4. A test of content comprehension and recall in which the child listens to a story and repeats in his own words as much of it as he can remember. At the end of the thirteenth week the children were given an especially constructed group test consisting of (1) a test of word recognition, (2) a test of word recall, and (3) two types of sentence comprehension. During the eighteenth week they were given the Detroit Word Recognition Test and another especially constructed group test of the true-false type designed "to measure the pupils' ability to comprehend complete sentences". Borderline cases were also tested individually. Deputy concludes from a study of the correlations between the tests of reading readiness and those of reading ability that "the Pintner-Cunningham Mental Test gives the best single means of predicting first-grade reading achievement", that the other tests of reading readiness "raised the predictive power of the mental test appreciably", and that "the combined scores on the individual tests of reading readiness correlate with reading achievement practically as well as the scores on the standardized mental test." He finds the best predictive measure to be a weighted score obtained by combining the scores on all five tests of reading readiness according to a specified formula.

Other less extensive studies of reading readiness have been made, among them those by Winch and by Minor. Winch,

27. Idem. p. 44.
29. Minor, Ruby. Experience as a Basis for Beginning Reading.
after using a test of associating nonsense syllables with simple geometric figures, concluded that it "furnished an adequate basis for dividing pupils into groups that were later found to be equal in reading achievement". Minor, in an investigation not scientifically controlled, produced convincing evidence that experience is a factor in reading readiness. Thus, some very definite attempts have been made to formulate tests for predicting success in reading, but none have resulted in a satisfactorily standardized measure.

II.

PART I.

STUDY OF THE STANFORD-BINET SUB-TESTS.
THE PROBLEM

The studies made on the relationship of mental age to reading ability were concerned with the intelligence test only insofar as it served as a means of providing a fairly reliable measure of mental age. They have used the results of the test as a whole rather than of its several sub-tests. Such data are probably sufficient for investigations of that nature; however, in predicting future success or failure on the basis of mental age, it is frequently more valuable to know the details of how the mental age score was obtained. For example, on the Binet test a child might pass several of the more advanced tests of memory for digits yet fail noticeably on tests requiring a greater degree of apperception, such as the comprehension, picture description, or ball and field tests. The fact that his auditory memory span is high but that his apperceptive powers as measured by the above mentioned tests are low may possibly be sufficient to predict for him a reading difficulty in the field of reading for comprehension. It is the purpose of the present study to investigate this problem of the value of the Stanford-Binet sub-tests for the prediction of reading success or failure.

Allied Research.

There appears to be but one investigation of this nature. This is the work of Durrell on the effect of reading disability on performance on the Stanford Binet sub-tests. Al-

though the abilities investigated by Durrell are the same as those of this study, the points of attack are quite different. He studied the Stanford-Binet test results of 134 children having special disability in reading paired with a control group of an equal number of children having normal reading ability. In order to obviate the possibility of including marginal cases, he confined his reading disability group to those "who presented a median reading score of at least one grade below their obtained mental age as well as one or more grades below their actual grade placement". The following tests were found to be "significantly harder for the reading disability group than for the control group:

- Year X - Reading and report.
- Year XII - Abstract words.
- Year XII - Dissected sentences.
- Year XIV - Vocabulary.

The vocabulary tests of the ten and twelve year levels seem to be harder for the reading disability group, the index of significance falling barely short of the criterion." On the other hand, the reading disability group showed successes significantly higher on the following tests:

- Year XIV - Problems of fact.
- Year XIV - Arithmetical reasoning.
- Year XVI - Problem of enclosed boxes.
- Year XVI - Digits backwards.

From this data Durrell evolved an Abbreviated scale for Reading Disability Cases. When this scale was used on the normal reading group and the I.Q.'s of the abbreviated and of the complete form compared, the abbreviated scale was found to have a reliability coefficient of .968 - .002; when used on the reading disability group, there was a "significant raise in median I.Q. of 4.57 points".

Durrell, op. cit., pp. 34-5.
Idem., p. 35.
Idem., p. 36.
COLLECTION OF DATA.

The group selected for study comprised all the second grades of the public school system of Belmont, Massachusetts, a total of approximately 250 children. Since the purpose of the investigation was to compare the results of Stanford-Binet tests given in kindergarten with those of the Gates Reading Tests given in the second grade, all children who had not been so tested were excluded. Some had not been tested by the Binet until the end of kindergarten or the beginning of the first grade; some were absent during part or all of the reading test program. In order that a low reading score might not be interpreted as resulting from low mentality rather than from actual reading disability, only those individuals whose obtained intelligence quotients were 100 or over were included. This necessitated the exclusion of a fairly large group of otherwise acceptable test results. The group was decreased still further by reason of the fact that many of those tested in kindergarten have since moved out of town. Thus, several factors have very definitely influenced the size of the study group.

The tests used for this study were the starred form of the Stanford Revision of the Binet-Simon Test and the Gates Primary Reading Test. The former was given during the first six weeks of kindergarten; the latter, during October of the second grade. Both were administered by the school psychologist. The Gates Primary Reading Test is divided into three parts: (1) word recognition, (2) word, phrase, and sentence reading, and (3) reading of paragraphs of directions. In eight of the cases included in the study this test was found to be inadequate as a
measure of the child's reading ability; consequently it was necessary to use the more advanced test by the same author, the Gates Silent Reading Test (Grades 3 to 8). It is divided into four sub-tests, namely (1) reading to appreciate the general significance of a paragraph, (2) reading to predict the outcome of given events, (3) reading to understand precise directions, and (4) reading to note details. Gates has standardized these two series in such a way that each is a true extension of the other and the age and grade scores of one are comparable with those of the other.

The total number of cases having both intelligence and reading test results as outlined above was 112. Of these the group selected for actual study comprised those whose reading scores, when compared with their computed mental ages at the time of the reading test, showed a difference of six months or more. This was taken as the criterion for either disability or superior ability in reading, it being the standard proposed by Durrell for the interpretation of second-grade reading results. The mental ages used for this selection were obtained by two different methods. Therefore the data for the resulting two groups have been treated separately. One set of mental ages was procured by multiplying the chronological ages by the intelligence quotients obtained on the Binet tests given in kindergarten, thereby assuming a constant rate of mental development in proportion to chronological age. The other set of mental ages was also computed by multiplying chronological age by intelligence quotient, but in this case the intelligence quotients were

38. Durrell, op. cit.
based on the results of the Binet given in kindergarten, the results of the Kuhlmann Anderson Group Test of Intelligence given in March of the first grade, and the teacher's and psychologist's estimates of the more correct figure. For the sake of avoiding confusion, these groups will be referred to as Group I and Group II, respectively. The number of cases resulting from this method of grouping is as follows:

<table>
<thead>
<tr>
<th></th>
<th>Normal</th>
<th>Disability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group I</td>
<td>36</td>
<td>32</td>
</tr>
<tr>
<td>(mental age based on Stanford-Binet)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group II</td>
<td>29</td>
<td>44</td>
</tr>
<tr>
<td>(mental age based on Stanford-Binet, Kuhlmann Anderson, and teacher's and psychologist's judgments)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
STUDY OF DATA

The successes and failures on the various Stanford Binet sub-tests were tabulated for each child. All tests below the basal level (the year in which all tests were passed) were assumed to have been passed successfully. Likewise, all tests beyond the upper level (the year in which no tests were passed) were considered as failures. At this point, another unfortunate limitation was encountered, namely the fact that the first alternative tests in the fifth and seventh year levels were substituted for test six in these two levels, but occasional exceptions to this rule make the data on these particular tests incomplete. Consequently, the results of these two tests were discarded. The percentage of successes on the remaining sub-tests of the starred form of the Stanford-Binet are recorded in the table on the following page.
<table>
<thead>
<tr>
<th>Test No.</th>
<th>Title</th>
<th>Percentage of Successes</th>
<th>Group I</th>
<th>Group II</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Normal</td>
<td>Disability</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Normal</td>
<td>Disability</td>
</tr>
<tr>
<td>III-1</td>
<td>Pointing to parts of body</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>III-2</td>
<td>Naming familiar objects</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>III-3</td>
<td>Enumeration of objects</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>III-6</td>
<td>Repeating 6 to 7 syllables</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>IV-1</td>
<td>Comparison of lines</td>
<td>97.22</td>
<td>100.00</td>
<td>96.55</td>
</tr>
<tr>
<td>IV-3</td>
<td>Counting 4 pennies</td>
<td>94.44</td>
<td>100.00</td>
<td>96.55</td>
</tr>
<tr>
<td>IV-4</td>
<td>Copying square</td>
<td>88.88</td>
<td>93.75</td>
<td>93.10</td>
</tr>
<tr>
<td>IV-5</td>
<td>Comprehension, first degree</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>V-1</td>
<td>Comparison of weights</td>
<td>80.55</td>
<td>96.88</td>
<td>86.21</td>
</tr>
<tr>
<td>V-2</td>
<td>Naming colors</td>
<td>80.55</td>
<td>81.25</td>
<td>82.76</td>
</tr>
<tr>
<td>V-3</td>
<td>Aesthetic comparison</td>
<td>69.44</td>
<td>93.75</td>
<td>62.07</td>
</tr>
<tr>
<td>VI-1</td>
<td>Distinguishing right and left</td>
<td>72.22</td>
<td>90.63</td>
<td>75.86</td>
</tr>
<tr>
<td>VI-2</td>
<td>Finding omissions in pictures</td>
<td>55.55</td>
<td>68.75</td>
<td>55.17</td>
</tr>
<tr>
<td>VI-3</td>
<td>Counting 13 pennies</td>
<td>47.22</td>
<td>62.50</td>
<td>48.28</td>
</tr>
<tr>
<td>VI-4</td>
<td>Comprehension, second degree</td>
<td>69.44</td>
<td>81.25</td>
<td>75.86</td>
</tr>
<tr>
<td>VII-1</td>
<td>Giving the number of fingers</td>
<td>16.66</td>
<td>9.69</td>
<td>24.14</td>
</tr>
<tr>
<td>VII-2</td>
<td>Description of pictures</td>
<td>36.11</td>
<td>78.13</td>
<td>41.38</td>
</tr>
<tr>
<td>VII-5</td>
<td>Giving differences from memory</td>
<td>33.33</td>
<td>68.75</td>
<td>37.93</td>
</tr>
<tr>
<td>VIII-2</td>
<td>Counting backwards from 20 to 1</td>
<td>00.00</td>
<td>00.00</td>
<td>00.00</td>
</tr>
<tr>
<td>VIII-3</td>
<td>Comprehension, 3rd degree</td>
<td>13.88</td>
<td>56.25</td>
<td>24.14</td>
</tr>
<tr>
<td>VIII-4</td>
<td>Similarities, two things</td>
<td>00.00</td>
<td>15.73</td>
<td>10.34</td>
</tr>
<tr>
<td>VIII-6</td>
<td>Vocabulary</td>
<td>00.00</td>
<td>00.00</td>
<td>00.00</td>
</tr>
<tr>
<td>IX-1</td>
<td>Giving the date</td>
<td>00.00</td>
<td>00.00</td>
<td>00.00</td>
</tr>
<tr>
<td>IX-2</td>
<td>Arranging 5 weights</td>
<td>00.00</td>
<td>00.00</td>
<td>00.00</td>
</tr>
<tr>
<td>IX-4</td>
<td>Repeating 4 digits reversed</td>
<td>00.00</td>
<td>00.00</td>
<td>00.00</td>
</tr>
<tr>
<td>IX-5</td>
<td>Using 3 words in a sentence</td>
<td>00.00</td>
<td>00.00</td>
<td>00.00</td>
</tr>
</tbody>
</table>
This same data has been represented graphically (figures I and II) in order that the relationships between the tests may be more readily perceived.
The fact that in almost every instance the tests were passed by a larger percentage of the disability group than of the normal group led to the belief that mental ability rather than reading ability was the controlling selective principle in determining these groups. Inspection of the data confirmed this, as may be seen in the following table of mental ages:

<table>
<thead>
<tr>
<th></th>
<th>Group I</th>
<th>Group II</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>7-11</td>
<td>7-11</td>
</tr>
<tr>
<td>Disability</td>
<td>8-6.5</td>
<td>8-7</td>
</tr>
<tr>
<td>Median mental age</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>7-11</td>
<td>7-11</td>
</tr>
<tr>
<td>Median intelligence quotient</td>
<td>109.5 124</td>
<td>109 122.5</td>
</tr>
</tbody>
</table>

This indicates that the groups are not composed of children having similar mental ability yet differing in reading ability. Instead, the normal group represents those whose mental ability is relatively low but whose reading achievement is high, and the "disability group" represents those whose mental ability is relatively high but whose reading achievement is low, undoubtedly due in large degree to the fact that the more brilliant children in the class have not been forced to achieve the maximum of which they are capable.
INTERPRETATION OF RESULTS

The data of this study being so small in amount and being further complicated by the introduction of a second variable, only the most tentative conclusions could be drawn. It was therefore deemed unwise to attempt to rule out the variations in intelligence as would be necessary to enable one to draw such conclusions.
RECOMMENDATIONS

1. In view of the limitations of this study the writer would suggest that another investigation might well be made of the same problem with a much larger collection of data. The groups should be made comparable by pairing the cases on the basis of chronological age, mental age, and intelligence quotient; otherwise one will be confronted with the same difficulty of dealing with two variables, mental ability and reading ability, rather than with reading ability alone. The value of such a study would be greatly increased by the use of the complete form rather than the starred form of the Stanford-Binet test.

2. It is to be hoped that attempts will soon be made to establish more conclusively the validity and reliability of such tests of reading readiness as those reported by Rockwell, Berry, Gates, Smith, and Deputy. This would mean the saving of much waste time, money, and effort now expended by both the child and the school authorities in cases where children are not psychologically ready for reading instruction. If, as Reed found, there are approximately 20% of the children entering first grade who are not ready, it can readily be seen that the saving would more than compensate for the cost of standardizing the tests.

3. It seems probable that a fairly satisfactory measure of reading readiness could be obtained by the following test battery:

   a. A test of vocabulary background.

   As suggested by Durrell, this should be of the

non-verbal motor response type, such as the three commissions test of the Stanford-Binet. A picture cross-out test such as that described by Rockwell might very well answer this purpose and would have the advantage of being adaptable for group use.

b. A test of letter- and word-matching ability.

Gates' finding of a low correlation between reading achievement and the ability to match geometrical figures or digits, and of a high correlation between reading ability and word-perception is suggestive of the value of such a test. Smith found it to have a high predictive value and gave the following explanation of this relationship: "One of the basic elements in gaining a percept of a new word is that of obtaining a clear-cut visual impression of this particular word as being distinctive from other words...General ability in word recognition cannot progress very far until pupils have developed the particular ability of making visual discriminations." A test of this ability might very easily be adapted to group use.

c. A test of attention.

Durrell found the Kohs Block Test to be a good measure of ability to attend to interesting material. The chief objections to it are the long time required to administer the test and the fact that it does not lend itself to group use. It is to be hoped that a satisfactory group test or a less time-consuming individual test will be made available for this purpose.

41. Smith, op. cit. pp. 58-9
42. Durrell, op. cit.
III.

PART II.

STUDY OF A READING-READINESS TEST BATTERY.
STUDY OF A READING-READINESS TEST BATTERY

Since the study of the Stanford-Binet sub-tests proved to be so inconclusive, it was deemed advisable to investigate the predictive value of certain other types of tests, namely the three suggested in recommendation 3 of Part I. It is with these tests that this portion of the study is concerned.

A. Selection of tests

1. Test of attention.

Other than Durrell's statement that the Kohs Block Test serves as a good measure of attention, there appear to be no reports on materials which could be used for such a purpose. Since the Kohs Block Test, as previously stated, has the disadvantages of requiring individual administration and of being too time-consuming, another test was sought as a substitute. At the suggestion of Durrell, the Pintner-Cunningham Primary Mental Test was selected for this purpose, since the materials are highly motivated and the test is not highly indicative of mental ability.

2. Test of word-matching ability.

The only available material for this type of test is that presented by Smith in her article on the matching-ability of first grade children. She shows the relative difficulty of matching the various letters and of matching a limited number of sense and nonsense words. There being insufficient material for a word-matching test, it was necessary to construct one, using Smith's findings as a basis and supplementing them

44. Smith, Nila B. Matching Ability as a Factor in First Grade Reading.
with such other material as might be considered valuable.

Smith reports that the letters presenting the greatest difficulty in matching are b, p, q, and d and that those with the least difficulty are g, s, m, o, w, c, and a. Combining these letters in various ways with a constant vowel, i, she found the most difficult words to match were those beginning and ending with a hard letter, the next in difficulty those beginning with an easy letter and ending with a hard one, and the least difficult those beginning with a hard letter and ending with an easy one.

Through the use of the Iota Word Test and Gray's Oral Reading Paragraphs as means of diagnosing difficulties among reading disability cases in the primary grades, the writer observed certain types of errors which occurred more or less frequently. For instance, some children tended to interchange the positions of letters, reading form instead of from, was instead of saw, of instead of for, etc.; some confused letters of similar form, such as b - d - p - q, h - n - r, n - u, and m - n; others tended to add or omit one or more letters in a word; still others appeared to notice only the beginning or ending of a word.

These observations and Smith's findings constituted the governing principles in the construction of the majority of the items in this test. However, in a few instances, notably items 7, 12, 13, 17, 18, 20, 21, 29, and 39, the words were included because they appeared to the writer to present situations which

45. See test on pages 28 and 29.
45. Monroe, Marion. Methods for Diagnosis and Treatment of Cases of Reading Disability
might prove difficult to those whose powers of discrimination were poor. These difficulties included t's, l's, and f's, especially when doubled, h and b, m and w, l1, l1, l1, l1, 1f, a, e, and o, s and a, oo and oc, ee and ec, and cl and d.

With these various difficulties as guides, the writer constructed 378 items of four words each. Each group presented one or more of the matching difficulties. These items were then examined critically and the poorest ones discarded. The remaining 292 were then divided into 4 corresponding groups of 41 items each, leaving 128 which were discarded because they were inferior or because they represented the same type of difficulty as that of some item already included in the test. The result was a test of 41 items with three supposedly comparable forms.

In an attempt to arrange the items in the order of difficulty, the following principles were used.

1. The two-letter words were placed at the beginning, then followed words of three letters, four letters, etc.

2. In each of these groups, those items having only one difficulty (initial letter, middle letter, final letter, etc.) were placed first, then followed those having 2, 3, 4, etc. types of difficulties.

In the determination of the test word in each item an attempt was made to select that word which appeared to be most easily confused with the other three words and therefore to present the greatest matching difficulty. As is prescribed for all multiple-choice tests, the correct word appears approxi-
mately the same number of times in each of the four possible positions, and the sequence of these positions follows no set pattern.

Since the usual reading situation, especially that of young children, is the printed rather than the type-written page, this test was printed. The material was put on four small pages rather than on two larger ones in order to minimize the danger of discouraging the child before undertaking the test. The original plan was to use large boldface type since that is like the usual reading situation in the first grade, but through a misunderstanding with the printer, somewhat smaller type was used. Inasmuch as there appeared to be no particular advantage in using nonsense syllables, only real English words were included. However, errors in the printing of the test resulted in the unintentional inclusion of two nonsense syllables, namely bleud and gabe instead of blend and gable in rows 26 and 27 respectively.

The following is a sample of this test:
<table>
<thead>
<tr>
<th>as</th>
<th>ax</th>
<th>as</th>
<th>an</th>
<th>am</th>
</tr>
</thead>
<tbody>
<tr>
<td>of</td>
<td>of</td>
<td>off</td>
<td>if</td>
<td>or</td>
</tr>
<tr>
<td>put</td>
<td>put</td>
<td>pot</td>
<td>pit</td>
<td>pet</td>
</tr>
<tr>
<td>odd</td>
<td>ebb</td>
<td>old</td>
<td>odd</td>
<td>add</td>
</tr>
<tr>
<td>ben</td>
<td>hen</td>
<td>ben</td>
<td>pen</td>
<td>den</td>
</tr>
<tr>
<td>rib</td>
<td>rid</td>
<td>rip</td>
<td>nib</td>
<td>rib</td>
</tr>
<tr>
<td>fad</td>
<td>fad</td>
<td>fed</td>
<td>led</td>
<td>lad</td>
</tr>
<tr>
<td>tan</td>
<td>ten</td>
<td>ant</td>
<td>tan</td>
<td>net</td>
</tr>
<tr>
<td>oar</td>
<td>oat</td>
<td>out</td>
<td>oar</td>
<td>our</td>
</tr>
<tr>
<td>tame</td>
<td>team</td>
<td>tame</td>
<td>mate</td>
<td>meat</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ball</th>
<th>bull</th>
<th>bill</th>
<th>ball</th>
<th>bell</th>
</tr>
</thead>
<tbody>
<tr>
<td>skin</td>
<td>skim</td>
<td>slim</td>
<td>akin</td>
<td>skin</td>
</tr>
<tr>
<td>tall</td>
<td>talk</td>
<td>tail</td>
<td>tall</td>
<td>tell</td>
</tr>
<tr>
<td>zeal</td>
<td>weal</td>
<td>veal</td>
<td>seal</td>
<td>zeal</td>
</tr>
<tr>
<td>oven</td>
<td>even</td>
<td>over</td>
<td>ever</td>
<td>oven</td>
</tr>
<tr>
<td>part</td>
<td>part</td>
<td>pert</td>
<td>past</td>
<td>pest</td>
</tr>
<tr>
<td>bail</td>
<td>ball</td>
<td>bail</td>
<td>hail</td>
<td>hall</td>
</tr>
<tr>
<td>lock</td>
<td>look</td>
<td>lock</td>
<td>tock</td>
<td>took</td>
</tr>
<tr>
<td>mine</td>
<td>wine</td>
<td>nine</td>
<td>mine</td>
<td>mire</td>
</tr>
<tr>
<td>hail</td>
<td>hail</td>
<td>hall</td>
<td>halt</td>
<td>half</td>
</tr>
</tbody>
</table>
loss  lass  less  toss  loss
some  same  some  come  came
grid  grid  grip  gird  girl
every  every  very  ever  even
loss  lose  toss  loose  loss
bland  bleud  blind  blond  bland
sable  cable  able  sable  gabe
queen  queer  queen  green  preen
clove  dove  love  clove  cove
couch  cough  couch  coach  conch

skies  shies  skier  shier  skies
rogue  vogue  rogue  rouge  vague
squir  squint  spurt  squirt  spirit
farmer  former  framer  fanner  farmer
lighter  lighter  tighter  tighten  lighten
purpose  porpoise  suppose  purpose  surprise
docker  doctor  backer  decker  docker
babble  dabble  babble  pebble  paddle
loller  loller  lofter  latter  letter
butter  bitter  butter  bitten  button
clearly  cleanly  clearing  clearly  dearly
3. Test of vocabulary background.

While a greater amount of work has been done in this field than in testing either attention or word-matching ability, there are as yet no satisfactory tests available for first grade use. The Merrill Palmer Scale of Mental Tests for pre-school children includes a vocabulary test, but it has the disadvantage of requiring a verbal response from the child in answer to questions such as "What cries?" "What boils?" "What stings?" Baldwin and Stecher devised a test of fifty nouns represented on cards by magazine pictures. Because of the pictures it is probably more highly motivated than the Merrill-Palmer test, but it still has the disadvantage of requiring a verbal response from the child, this time in answer to the question, "Of what is this a picture?" Madorah Smith built a vocabulary test of 203 words representing a regular sampling of the Thorndike Word List. However, it is not satisfactory for our purpose, because it requires considerable subjective scoring and because it is largely verbal, involving in some instances a comprehension not only of the test word but also of the question asked by the examiner. Probably the most satisfactory test yet devised is the Van Alstyne Picture Vocabulary Test for Pre-School Children which was formu-

lated as part of a test battery for the study of three-year-old children. Miss Van Alstyne notes the following advantages over the other vocabulary tests: (1) it requires no oral response; (2) the scoring is absolutely objective; (3) the questions asked the child are easily understood; (4) it is short; and (5) it is highly motivated.

All the tests described thus far have the disadvantage of being designed primarily for pre-school children. If used in the first grade, they would probably produce scores too closely clustered for purposes of diagnosis. Of the vocabulary tests intended for first grade use the writer finds reference to only two, the Stormzand Vocabulary Test for Pupils Entering First Grade and the Baltimore Reading Readiness Test. Regarding the former no information is available, nor of the latter except that it is composed of a word discrimination and a picture vocabulary test. Apparently it was found to be inadequate, for it is no longer used in the Baltimore schools and as yet no copies have been obtainable for examination.

As a result of this survey the Van Alstyne test was selected as being the most useful for our purpose. When given to a few of the least capable children in a Belmont first grade, almost perfect scores were obtained. This indicated that the test was too easy and would therefore fail to differentiate among the children sufficiently to serve as a diagnostic measure. There-

50. Van Alstyne. op. cit. p. 9.
52. Idem.
fore, an attempt was made to extend the test to include more difficult words.

Various difficulties were encountered in the selection of words to incorporate in this extension. Miss Van Alstyne had used a list of 7000 words "compiled by Madeline Darrough Horn in her study of the oral vocabulary of over 3,000 kindergarten children" and from this selected her test words at regular intervals throughout the list. Since only the first 1,000 words of Mrs. Horn's list are available, it was necessary to refer to other word lists. Consequently, the Van Alstyne test words were checked with the Thorndike Word List, the Gates Word List, Wheeler and Howell's list of 453 words found in 10 primers and 10 first readers, Packer's list of the vocabularies of 10 first readers, Kircher's list of the vocabularies of 37 primers and first readers, and the list compiled by the International Kindergarten Union. No single list contained all of the test words, although each word was found in at least one of the lists. Therefore, it was impossible to ex-

54. Van Alstyne, op. cit., p. 10
55. Horn, Madeline Darrough. The Thousand Three Words Most Frequently Used by Kindergarten Children, pp. 118-123.
58. Wheeler and Howell. A First Grade Vocabulary Study.
59. Packer, J. L. The Vocabularies of Ten First Readers.
tend the Van Alstyne test by using any one list exclusively. This difficulty and the fact that the vast majority of the test items were too simple for measuring first grade vocabulary led us to construct a test modelled after the Van Alstyne test but composed of entirely new test items.

The first problem was the selection of a vocabulary upon which to base the test. At first it was thought advisable to use one of the lists based on studies of primers and first readers, but there were two objections to this plan. In the first place, the child's reading vocabulary is noticeably smaller than his comprehension vocabulary except in the case of children coming from homes where a foreign language is used instead of English. The second objection lay in the fact that only the studies of old series of readers presented full vocabulary lists whereas reports on readers now in use in the schools present only the words occurring most frequently or those common to all readers studied. The Thorndike list was considered to be impracticable for use as a basis for the test because the method of selecting the words suggested the inclusion of much material not ordinarily found in the reading or comprehension vocabularies of primary grade children. On the other hand, the Gates list was thought to be confined too closely to a strictly reading vocabulary and therefore to be lacking in words found in the comprehension vocabulary of lower grade children. The list compiled by the International Kindergarten Union was therefore chosen as being best adapted to the purpose in mind. It is composed of 2596 words used by kindergarten children while attending kindergarten, when stimulated by pictures, and when at home.

Attempts were made to select the words at regular intervals throughout the list, but this led to the inclusion of extremely common words known by all children and also such words as the, than, and of, which would be practically impossible to represent pictorially. This being the case, a more subjective method was adopted, namely the selection of words which were listed as having a relatively low degree of frequency, which might conceivably be unknown to a first grade child and yet be more or less valuable for the comprehension of conversations in the first grade, and which would reasonably lend themselves to representation in pictures. The resulting list comprised at least 400 words. From this was constructed a group of 60 test items of the multiple-choice type. In every instance an effort was made to follow the same principles as those set up by Van Alstyne for the building of her test. They are as follows:

"1: At least one word to be of an equal level of frequency with the test word (and therefore probably of difficulty also).

2. At least one word to be associated with the test word in life situations, e.g., thimble-needle.

3. At least one word to have a sound similarity, of the first letter, with that of the test word."

The resulting 60 items were then grouped according to the part of speech or the type of situation represented by the test word, and the number of items reduced to 55 by discarding such groups as tended too nearly to duplicate other items.

64. Van Alstyne. op. cit., p. 11.
This final group of items, with the exception of 6, was then represented in pictures drawn in India ink on white paper by four art students more or less familiar with educational tests. The drawings were mounted on manila cardboard of a somewhat larger size and incorporated in a loose-leaf notebook. They were arranged presumably in the order of difficulty from easy to hard, the criterion used being the frequency of use in kindergarten as indicated in the vocabulary list upon which the test was based. In the following table are listed the 49 test words (underlined) and the multiple-choice words appearing on the same sheet.
<table>
<thead>
<tr>
<th>No.</th>
<th>Upper Left</th>
<th>Upper Right</th>
<th>Lower Left</th>
<th>Lower Right</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>over</td>
<td>between</td>
<td>in</td>
<td>under</td>
</tr>
<tr>
<td>2.</td>
<td>straight</td>
<td>crooked</td>
<td>curved</td>
<td>wavy</td>
</tr>
<tr>
<td>3.</td>
<td>letter</td>
<td>flower</td>
<td>farmer</td>
<td>painter</td>
</tr>
<tr>
<td>4.</td>
<td>skates</td>
<td>cloud</td>
<td>stairs</td>
<td>kite</td>
</tr>
<tr>
<td>5.</td>
<td>against</td>
<td>outside</td>
<td>inside</td>
<td>on top of</td>
</tr>
<tr>
<td>6.</td>
<td>horseshoe</td>
<td>snake</td>
<td>monkey</td>
<td>mitten</td>
</tr>
<tr>
<td>7.</td>
<td>skates</td>
<td>circle</td>
<td>square</td>
<td>spade</td>
</tr>
<tr>
<td>8.</td>
<td>canoe</td>
<td>cradle</td>
<td>bed</td>
<td>music</td>
</tr>
<tr>
<td>9.</td>
<td>top</td>
<td>(carrying)</td>
<td>middle</td>
<td>bottom</td>
</tr>
<tr>
<td>10.</td>
<td>grapes</td>
<td>cherries</td>
<td>gloves</td>
<td>glasses</td>
</tr>
<tr>
<td>11.</td>
<td>purse</td>
<td>carpenter</td>
<td>cupboard</td>
<td>dishes</td>
</tr>
<tr>
<td>12.</td>
<td>see-saw</td>
<td>carriage</td>
<td>pancakes</td>
<td>porridge</td>
</tr>
<tr>
<td>13.</td>
<td>mop</td>
<td>spider</td>
<td>frog</td>
<td>flag</td>
</tr>
<tr>
<td>14.</td>
<td>apron</td>
<td>bib</td>
<td>acorn</td>
<td>lion</td>
</tr>
<tr>
<td>15.</td>
<td>ahead of</td>
<td>after</td>
<td>around</td>
<td>beside</td>
</tr>
<tr>
<td>16.</td>
<td>woodpecker</td>
<td>windmill</td>
<td>wagon</td>
<td>wheelbarrow</td>
</tr>
<tr>
<td>17.</td>
<td>hatchet</td>
<td>cage</td>
<td>cabbage</td>
<td>corn</td>
</tr>
<tr>
<td>18.</td>
<td>parade</td>
<td>picnic</td>
<td>pitcher</td>
<td>pineapple</td>
</tr>
<tr>
<td>19.</td>
<td>calf</td>
<td>gate</td>
<td>goat</td>
<td>bird</td>
</tr>
<tr>
<td>20.</td>
<td>tower</td>
<td>turtle</td>
<td>snail</td>
<td>rocking horse</td>
</tr>
<tr>
<td>21.</td>
<td>sailing</td>
<td>paddling</td>
<td>sliding</td>
<td>riding</td>
</tr>
<tr>
<td>22.</td>
<td>scissors</td>
<td>country</td>
<td>city</td>
<td>kitten</td>
</tr>
<tr>
<td>23.</td>
<td>fireplace</td>
<td>stove</td>
<td>flowerpot</td>
<td>football</td>
</tr>
<tr>
<td>24.</td>
<td>lantern</td>
<td>electric</td>
<td>light</td>
<td>engine</td>
</tr>
<tr>
<td>25.</td>
<td>basket-lunch</td>
<td>castle</td>
<td>knight</td>
<td>cottage</td>
</tr>
<tr>
<td>No.</td>
<td>Upper Left</td>
<td>Upper Right</td>
<td>Lower Left</td>
<td>Lower Right</td>
</tr>
<tr>
<td>-----</td>
<td>-------------</td>
<td>-------------</td>
<td>------------</td>
<td>-------------</td>
</tr>
<tr>
<td>26.</td>
<td>racing</td>
<td>reaching</td>
<td>sleeping</td>
<td>jumping</td>
</tr>
<tr>
<td>27.</td>
<td>skiing</td>
<td>sweeping</td>
<td>paddling</td>
<td>swimming</td>
</tr>
<tr>
<td>28.</td>
<td>boat</td>
<td>dough</td>
<td>bread</td>
<td>drum</td>
</tr>
<tr>
<td>29.</td>
<td>fairy</td>
<td>angel</td>
<td>fern</td>
<td>shell</td>
</tr>
<tr>
<td>30.</td>
<td>paddling</td>
<td>eating</td>
<td>peeling or paring</td>
<td>pouring</td>
</tr>
<tr>
<td>31.</td>
<td>postman</td>
<td>bugle</td>
<td>medicine</td>
<td>butcher</td>
</tr>
<tr>
<td>32.</td>
<td>basket</td>
<td>flowerpot</td>
<td>bucket</td>
<td>trough</td>
</tr>
<tr>
<td>33.</td>
<td>measuring</td>
<td>marching</td>
<td>dusting</td>
<td>decorating</td>
</tr>
<tr>
<td>34.</td>
<td>bakery</td>
<td>library</td>
<td>bridge</td>
<td>lightning</td>
</tr>
<tr>
<td>35.</td>
<td>bridge</td>
<td>stool</td>
<td>flag</td>
<td>bench</td>
</tr>
<tr>
<td>36.</td>
<td>woods</td>
<td>ocean</td>
<td>mountain</td>
<td>meadow</td>
</tr>
<tr>
<td>37.</td>
<td>fruit</td>
<td>tools</td>
<td>vegetables</td>
<td>animals</td>
</tr>
<tr>
<td>38.</td>
<td>ready for bed</td>
<td>bareheaded</td>
<td>well-wrapped</td>
<td>barefooted</td>
</tr>
<tr>
<td>39.</td>
<td>bank</td>
<td>carrot</td>
<td>beet</td>
<td>arrow</td>
</tr>
<tr>
<td>40.</td>
<td>fairy</td>
<td>giant</td>
<td>dwarf</td>
<td>jam</td>
</tr>
<tr>
<td>41.</td>
<td>king</td>
<td>trunk</td>
<td>cane</td>
<td>queen</td>
</tr>
<tr>
<td>42.</td>
<td>pump</td>
<td>faucet</td>
<td>bubbler or fountain</td>
<td>jack-in-the-pulpit</td>
</tr>
<tr>
<td>43.</td>
<td>hammering</td>
<td>sewing</td>
<td>kneading</td>
<td>knitting</td>
</tr>
<tr>
<td>44.</td>
<td>sled</td>
<td>apple</td>
<td>sailboat</td>
<td>ache</td>
</tr>
<tr>
<td>45.</td>
<td>chopping</td>
<td>reading</td>
<td>clapping</td>
<td>stirring</td>
</tr>
<tr>
<td>46.</td>
<td>chain</td>
<td>shell</td>
<td>shepherd</td>
<td>soldier</td>
</tr>
<tr>
<td>47.</td>
<td>saw</td>
<td>thread</td>
<td>yarn</td>
<td>yawn</td>
</tr>
<tr>
<td>48.</td>
<td>pushing</td>
<td>waving</td>
<td>hopping</td>
<td>hauling</td>
</tr>
<tr>
<td>49.</td>
<td>stepping</td>
<td>stooping</td>
<td>bowing</td>
<td>waving</td>
</tr>
</tbody>
</table>
B. Administration and scoring of tests

1. Test of Attention.

This test was administered to each of the three first grades as a group test with the teacher and, in one instance, a practice teacher assisting. The regular procedure for giving and scoring the test as prescribed in the revised manual was followed.

2. Test of Word-Matching Ability.

This test was given to the children in groups of six or less. Each child was provided with a pencil and a sample test like the following:

cat  boy  dog  cat  pig
will  here  hope  were  will
The following directions were then read slowly:

Here are some words. Look at the first row of words, and put your finger on the first word (indicate it). Now look all along this row until you find another word that looks just like it. Draw a line under it. (Wait until all are finished.)
Now, put your finger on the first word in the next row. Look at all the words in this row until you find a word that looks just like the first one. Draw a line under it. (Wait until all are finished.)
Now, there are four more pages that have rows of words like these. I want you to draw a line under the right word in each row, but don't wait for me to tell you about each row. Start with the first row. After you have drawn a line under the right word, go right on to the next row and the next row and the next row until you have done them all. Do them just as fast as you can. When you have done all the words on the page, raise your hand and I'll give you another page. Remember! Draw a line under the right word in each row; work as fast as you can; and raise your hand when you have finished.

Page 1 of the test was then presented to each child and the signal to begin given. Pages 2, 3, and 4 were presented to each child just as rapidly as he finished the preceding page. At the end of the fourth page the total time required for the test was recorded as indicated by a stop watch.

In the scoring of the test, omitted items and those in which the wrong word or more than one word was underlined were counted as errors. In nine instances, the directions were very poorly followed, that is, all the words in an item were underlined, all the words in one column and only that column were underlined, or many items were omitted. After an interval of at least a week, these nine individuals were retested and especial care was taken to insure the complete understanding of the directions. Two children still failed to follow directions
correctly; so their scores were excluded, because it was evident that the test was measuring not matching ability alone but also ability to follow directions or ability to attend. The scores on the remaining seven tests were substituted for the original scores. It is probable that these scores were not noticeably affected by practice. At least, in the cases of two other children who were unnecessarily retested the results were almost identical as may be seen in the following table.

<table>
<thead>
<tr>
<th>Errors</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Case A</td>
<td>6</td>
</tr>
<tr>
<td>Case B</td>
<td>3</td>
</tr>
</tbody>
</table>

3. Test of Vocabulary background.

This test was given individually one page at a time and the choices made by each child noted on a chart designed for the purpose. The recording was done as inconspicuously as possible, for it tended to influence the responses of some children. The following set of directions was used:

Card No.                      Direction.
1. Show me the birds that are under the trees.
2. Show me the straight line.
3. Show me the farmer.
4. Show me the kite.
5. Show me the ladder which is inside the house.
6. Show me the monkey.
7. Show me the square.
8. Show me the cradle.
<table>
<thead>
<tr>
<th>Card No.</th>
<th>Direction</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.</td>
<td>Show me the boy who is at the <strong>bottom</strong> of the ladder.</td>
</tr>
<tr>
<td>10.</td>
<td>Show me the <strong>grapes</strong>.</td>
</tr>
<tr>
<td>11.</td>
<td>Show me the <strong>cupboard</strong>.</td>
</tr>
<tr>
<td>12.</td>
<td>Show me the <strong>porridge</strong>.</td>
</tr>
<tr>
<td>13.</td>
<td>Show me the <strong>frog</strong>.</td>
</tr>
<tr>
<td>14.</td>
<td>Show me the <strong>apron</strong>.</td>
</tr>
<tr>
<td>15.</td>
<td>Show me the boy who is running <strong>ahead</strong> of the girls.</td>
</tr>
<tr>
<td>16.</td>
<td>Show me the <strong>wheelbarrow</strong>.</td>
</tr>
<tr>
<td>17.</td>
<td>Show me the <strong>cabbage</strong>.</td>
</tr>
<tr>
<td>18.</td>
<td>Show me the <strong>picnic</strong>.</td>
</tr>
<tr>
<td>19.</td>
<td>Show me the <strong>goat</strong>.</td>
</tr>
<tr>
<td>20.</td>
<td>Show me the <strong>turtle</strong>.</td>
</tr>
<tr>
<td>21.</td>
<td>Show me who is <strong>sailing</strong>.</td>
</tr>
<tr>
<td>22.</td>
<td>Show me the <strong>city</strong>.</td>
</tr>
<tr>
<td>23.</td>
<td>Show me the <strong>fireplace</strong>.</td>
</tr>
<tr>
<td>24.</td>
<td>Show me the <strong>lantern</strong>.</td>
</tr>
<tr>
<td>25.</td>
<td>Show me the <strong>castle</strong>.</td>
</tr>
<tr>
<td>26.</td>
<td>Show me who is <strong>reaching</strong>.</td>
</tr>
<tr>
<td>27.</td>
<td>Show me who is <strong>swimming</strong>.</td>
</tr>
<tr>
<td>28.</td>
<td>Show me the <strong>dough</strong>.</td>
</tr>
<tr>
<td>29.</td>
<td>Show me the <strong>fairy</strong>.</td>
</tr>
<tr>
<td>30.</td>
<td>Show me who is <strong>pouring</strong>.</td>
</tr>
<tr>
<td>31.</td>
<td>Show me the <strong>butcher</strong>.</td>
</tr>
<tr>
<td>32.</td>
<td>Show me the <strong>bucket</strong>.</td>
</tr>
<tr>
<td>33.</td>
<td>Show me who is <strong>measuring</strong>.</td>
</tr>
<tr>
<td>34.</td>
<td>Show me the <strong>library</strong>.</td>
</tr>
<tr>
<td>35.</td>
<td>Show me the <strong>bench</strong>.</td>
</tr>
</tbody>
</table>
36. Show me the meadow.
37. Show me the tools.
38. Show me the barefooted boy.
39. Show me the beet.
40. Show me the giant.
41. Show me the queen.
42. Show me the pump.
43. Show me who is knitting.
44. Show me who has an ache.
45. Show me who is chopping.
46. Show me the shepherd.
47. Show me the yarn.
48. Show me who is hauling.
49. Show me who is stooping.
C. Study of test results.

1. Collection of data.

The three tests already described and the Kuhlmann-Anderson Test of Intelligence were given by the writer to the three first grades of one of the public schools in Belmont, Massachusetts, during April and the first week in May, 1931. Each class was tested by the school psychologist at the beginning of February with the Word Recognition Test and at the end of April with the Word, Phrase, and Sentence Reading Test of the Gates Primary Reading Test. The reading ages obtained on the two tests were averaged in order to secure a measure of reading ability with which to compare the other test results. In several instances, the child failed to achieve a plus score and was therefore listed as below norms. In such cases the reading age was regarded as 6.00, since at that age the average child is not expected to be reading.

Results on some or all of the tests were obtained for the 79 children in the three first grades. Twelve of this number have been excluded because they were absent when one of the reading tests was given. Of the remaining 67, three had no vocabulary score, one no Kuhlmann-Anderson score, and four no word-matching score.

2. Treatment of data.

Although it is frequently considered valuable to determine the degree to which test scores correlate with some particular criterion such as reading ability, it was thought that the relationship of the data collected in this study could be more clearly indicated by means of the charts included in the following pages. Accordingly, two types of charts have been con-
structured. One shows in a distribution graph the range of scores obtained on a particular test by the three first grades taken as a whole and by the A, B, and C grades taken separately. The other shows the relationship between the averaged reading age scores and the scores on the test in question.

a. Test of Intelligence.

This type of graph should be read as follows:

Of those whose reading scores were among the highest fifth of the group, 3 individuals obtained mental age scores which were among the highest fifth of the group, 7 obtained mental age scores which were among the second highest fifth of the group, etc.
b. Test of Attention

**Mental Age Scores**

<table>
<thead>
<tr>
<th>Cases</th>
<th>Number of Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

**A, B, and C Groups**

**Distribution Graphs**

- **A Group**
- **B Group**
- **C Group**

**Scores**

<table>
<thead>
<tr>
<th>Mental Age Scores</th>
<th>9:4-</th>
<th>9:3-</th>
<th>9:2-</th>
<th>9:1-</th>
<th>8:9-</th>
</tr>
</thead>
<tbody>
<tr>
<td>Highest fifth</td>
<td>III</td>
<td>****</td>
<td>**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High fifth</td>
<td>III</td>
<td>**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle fifth</td>
<td>III</td>
<td>**</td>
<td>**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low fifth</td>
<td>II</td>
<td>****</td>
<td>**</td>
<td>**</td>
<td></td>
</tr>
<tr>
<td>Lowest fifth</td>
<td>I</td>
<td>****</td>
<td>***</td>
<td>**</td>
<td></td>
</tr>
</tbody>
</table>

**Median Score**

- 8:10
- 8:10
- 7:7
- 7:5
- 6:6

\(\text{V} = \text{median}\)
a. Test of Vocabulary Background.

Number of Errors

A, B, and C Groups

A Group

B Group

C Group

Distribution Graphs

v = median

Number of Errors

Median Number of Errors

<table>
<thead>
<tr>
<th>3-5</th>
<th>5-7</th>
<th>9-10</th>
<th>10-14</th>
<th>14-28</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Highest fifth</td>
<td></td>
<td></td>
<td></td>
<td>***</td>
</tr>
<tr>
<td>High fifth</td>
<td></td>
<td></td>
<td></td>
<td>***</td>
</tr>
<tr>
<td>Middle fifth</td>
<td></td>
<td></td>
<td></td>
<td>***</td>
</tr>
<tr>
<td>Low fifth</td>
<td></td>
<td></td>
<td></td>
<td>***</td>
</tr>
<tr>
<td>Lowest fifth</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
d. Test of Word Matching

Number of Errors

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20 +</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Number of Cases

A, B, and C Groups

A Group

B Group

C Group

Distribution Graphs

A = median

Reading Scores

Number of Errors

0-3 3-6 6-10 10-13 13-

Highest fifth

High fifth

Middle fifth

Low fifth

Lowest fifth

Median Number of Errors

5.5

9.5

7

9.5

10
Distribution Graphs.

<table>
<thead>
<tr>
<th>Time</th>
<th>4' 5' 6' 7' 8' 9' 10' 11' 12' 13' 14' 15' 16' 17' 18' 19' 20'</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Cases</td>
<td></td>
</tr>
<tr>
<td>A Group</td>
<td></td>
</tr>
<tr>
<td>B Group</td>
<td></td>
</tr>
<tr>
<td>C Group</td>
<td></td>
</tr>
</tbody>
</table>

**Scores**

- Highest fifth
- High fifth
- Middle fifth
- Low fifth
- Lowest fifth

**Time**

- Median Time
- 4' 35"
- 4' 34.5"
- 9' 53"
- 11' 1.6"
- 10' 36.5"
3. Interpretation of data.

a. Test of intelligence.

The distribution graph indicates a grouping of scores which more nearly approaches the normal curve of distribution than do the scores on any of the other tests. The median mental age score in the C group is lower than that in the B which in turn is lower than that in the A group. This is to be expected, since the classes were grouped on the basis of mental ability.

There appears to be a general tendency for those obtaining a high mental score to do well on the reading tests and for those obtaining a lower mental score to do correspondingly poorer on the reading test. It is more noticeable among those whose obtained mental ages are low than among those whose ages are high. In other words, the majority of those whose reading scores were in the lowest fifth of the class obtained mental age scores of 7 years 5 months or less: on the other hand, several children who scored higher on the reading tests also had mental age scores of 7 years 5 months or less. The relationship between these two test scores is not strongly marked. Therefore, a low mental age as measured by the Kuhlmann-Anderson Intelligence Test may be a factor in determining low reading ability, but it does not preclude reading success.
b. Test of attention.

While the distribution graph of the entire group is somewhat similar to the normal distribution curve, the graphs of the A, B, and C groups are decidedly irregular. If this test is a valid measure of attention, these graphs indicate that there is a normal variation in the attention ability of the group as a whole but that the A group is not necessarily the most, and the C group the least, attentive. However, the medians for the groups show a degree of correspondence between attention ability and the general ability of the group.

The relationship between reading ability and ability to attend is quite clearly indicated. Assuming that the results are representative of first-grade children in general and that the test is valid, one may draw the following conclusions. (1) Those whose attention is poor are more apt than not to have trouble with reading. (2) A high score on this test, however, does not insure reading success. (3) Poor attention and high reading ability, and good attention and low reading ability are rare occurrences.

c. Test of vocabulary background.

The distribution graph is decidedly irregular, it being more nearly like two normal
distribution graphs placed side by side. Possibly a detailed analysis of the test results would reveal the cause of such a situation, but on the surface there appears to be no adequate explanation of it. The graphs for each class group are also noticeably irregular. They indicate a fairly wide scatter of scores and a marked difference between the ability of the C group and that of either the B or A group, as indicated by the medians.

There is a degree of correlation between the scores on this test and those on the reading test. It is not as pronounced as for the attention test results, but is more definite than in the mental test results. Apparently, those whose scores in reading are excellent do well on this test, but those whose reading ability is less pronounced do not necessarily do well or poorly. Even some of the poorest readers were quite successful on this test. The median scores for each reading group indicate a real relationship between reading ability and ability to succeed on this test. In other words, the test is probably useful in predicting the reading ability of groups but is not sufficiently diagnostic for individual prognoses.
d. Test of word-matching ability.

When scored on the basis of errors, the test does not appear to be valuable. The distribution graphs are extremely irregular; the class medians do not produce any evidence of value; and no decided correspondence is shown between reading ability and matching ability. This latter fact is most noticeable in the table of the median scores obtained by each reading group, the medians for the second highest and the second lowest groups being identical.

When the time required to complete the test was taken as the criterion of matching ability, more valuable data resulted. The distribution graph is skewed to the left, indicating a greater differentiation among the slower children than among those who finished quickly. The medians for each class group show that a noticeably shorter time was required by the A, than by either the B or C, group. They also show a less pronounced difference in the time required by the B group and that required by the C group.

When the time scores are compared with reading ability, the test is found to be rather more diagnostic of good than of poor readers. In other words, those who read well required only a short time to complete the test. However, the converse of this statement is not true, namely that those who complete the test in a short time read well.
There appears to be no marked relationship between the time required and the number of errors made on this test, as is indicated in the following graph.
D. Criticism of tests, and recommendations.

In cases of all untried tests, means, or at least the need, of improving them become evident after a comparatively small number of the tests have been administered. The word-matching and vocabulary tests are no exceptions.

1. Test of word-matching ability.

The word-matching test is especially in need of remodelling. In the first place larger type should be used, since it would then parallel the sort of material used by the child in actual reading situations. The smaller sized type demands of the child a certain degree of adjustment to a rather novel situation.

Another improvement would be the separation of the test items by horizontal lines. This would help to overcome the tendency of some children to look for a word in more than one row at a time.

Another device which would assist the majority of the children is the underlining of the test words in the left hand column. Many scores were lowered because the children took time to underline the test word before trying to find the correct word. This may have been due to inability to work without this guide or to the fact that some of their school tasks may have required a double underlining process of this nature. With the test word underlined, the directions could then be made to read somewhat as follows:

Look at all the words in the row until you find one that looks like the word with the line under it; then draw a line under that word so that you will have two words that look just alike.
Also, in the directions some mention should be made of crossing out instead of erasing errors, and greater emphasis should be laid on the necessity of underlining only one word in each row.

The administration of the test would be greatly simplified if the samples and test items could be compiled in book form. It would eliminate the confusion caused by passing out the separate sheets as the child is ready for them.

The frequency of the automatic underlining of all the words in one column would doubtless be somewhat reduced if, in the revision of the test, care was taken to avoid the occurrence of the correct word in the same position in consecutive rows. In several instances, this condition was apparently the explanation of a series of correct responses followed by a series of apparently automatic responses.

Probably the most glaring weakness of the test as it now stands is its lack of motivation. Some children worked with apparent interest until the presentation of the fourth page when a sigh or some other act indicated discouragement or a feeling of being over-worked. Other children, from the very beginning of the test gave the impression of being forced to work. Still others displayed either mild enthusiasm or no particular emotion. It is to be hoped that some means may be devised for motivating this test. The most promising suggestion made thus far is the use of a simple card sorting or filing system.

The reactions pointed very definitely to the advisability of omitting the last page of the test, because in too many instances it proved to be an arduous task and became more
nearly a test of persistence than of matching ability. It is extremely doubtful whether words as long as those on that page occur in first-grade reading with sufficient frequency to warrant their inclusion in the test.

Possibly the test becomes a truer measure of matching ability when used individually. In group tests, especially when so poorly motivated, lapses of attention tend to affect the final score both as to time and errors.

2. Test of vocabulary background.

This test has fewer obvious fallacies, doubtless because it was patterned after another test in which various defects had been noted and corrected. The most outstanding criticism is that some pictures in the test items are too much alike to be fair tests. This is particularly true of tests 21, 43, and 49. In test 21, the drawing for sailing should have been of a man in a sailboat instead of a child sailing a toy boat; and paddling should not have been included in the same item, because it led to frequent confusion among children who knew the meaning of sailing. Likewise, in test 43, there is confusion because the drawing of sewing depicts a spool of thread which has fallen to the floor, a situation which is more common because in knitting. Also confusion arises/in the picture the hand is in the attitude of taking a stitch, whereas the more usual picture of sewing is of the arm outstretched in the act of completing a stitch. In test 49, the drawings of stooping and bowing are too much alike and lead to confusions among children who know the meaning of the test word.

In test 44, there is an element of verbal confusion in that ache sounds very much like egg. Many children
chose the picture of the child eating an apple, because they thought the apple was an egg. It would be well, therefore, to substitute a picture of an object with a name less easily confused with ache, for instance a sandwich.

In test 28, although the drawing of the fairy is very good, it is apparently not the usual picture which children have of fairies. Consequently, the angel was more frequently chosen. The uncertainty of the children's responses and their remarks on this item indicated that very few of the many incorrect choices were due to a lack of understanding of the word, fairy. The item, therefore, is not a good vocabulary test.

Some of the pictures, although not noticeably confusing, could probably be made more satisfactory. For example the drawing of the shepherd would doubtless be better if, instead of being a small silhouette, it were made more nearly the size of the soldier on the same page.

3. Test of attention.

Whether this is or is not a true test of attention cannot be determined from the results of this study. To ascertain its validity it would be well to test a group of children with both the Pintner-Cunningham and the Kohs Block tests and to compare the scores obtained by each individual.
IV.

SUMMARY.
SUMMARY

There is a definite need for a satisfactory means of predicting school success, especially in the field of reading. School administrators and teachers have indicated this lack and their attempts to find a solution to the problem. Research investigators have also recognized the deficiency and have set about meeting it in two main ways. One group has tried to establish a minimum mental age for admission to first grade and to make predictions on the basis of I.Q.'s. The other group has attempted to construct tests of reading-readiness. In so doing they have presented much valuable material but have not yet arrived at a well standardized test.

The first part of the present study was intended as an investigation of the value of the Stanford-Binet sub-tests in predicting reading ability. The only other research of this nature is that in which a study was made of the effect of reading disability on performance on these tests and which resulted in a Stanford-Binet Scale for Reading Disability Cases. In the present study a comparison was made of the sub-test results of the Stanford-Binet test given in kindergarten with the reading abilities of the children as tested in the second grade by the Gates Reading Tests. The study groups were found to be extremely small and to vary not only in reading ability but also in mental ability. Since conclusions based on such questionable data would be very unsatisfactory, the study was carried no further. Although no definite conclusions were reached, the study served to indicate several problems for further research, namely a more complete investigation of the problem
undertaken in this study, the standardization of the various
tests designed to measure reading readiness, and the formulation
of a reading readiness test battery.

The second part of the present study was concerned with
the investigation of a reading-readiness test battery such as
was indicated in the recommendations of part I. After all the
available tests of this nature had been examined, the Pintner-
Cunningham Primary Mental Test was selected as a test of atten-
tion, and tests of word-matching ability and vocabulary back-
ground were formulated by the writer. The tests were then
given to the A, B, and C groups of a first grade and the results
compared by means of graphs with the scores obtained by the
same grade on the Kuhlmann-Anderson Intelligence Test and the
first two types of the Gates Primary Reading Test. Mental age
scores appeared to correlate rather moderately with reading
ability. Attention varied not so much with the class group as
with reading ability. The vocabulary scores appeared to vary
with reading ability but not so decidedly as did the scores on
the attention test. The time required to complete the word-
matching test also correlated rather well with reading ability,
but the number of errors made on the test was found to have no
predictive value. In general, these tests are probably more
valuable as prognoses of reading success than of reading failure.
As yet they are too inaccurate for use in diagnosing individual
cases with any degree of certainty. Possibly their value will
increase as the defects noted in this study are remedied.
V.

BIBLIOGRAPHY


Dickson, Virgil E. The Use of Mental Tests in School Administration. Monograph No. 4, Board of Education, Berkeley, California, June, 1922.


Gray, William S. Summary of Investigations Relating to Reading. Supplementary Educational Monograph, No. 28, June, 1925.


