A Study of the Feasibility of the W.A.I.S. as a Self-Administered Pencil-Paper Test

Mason
A STUDY OF THE FEASIBILITY
OF THE W.A.I.S. AS A SELF-
ADMINISTERED PENCIL-PAPER TEST

by

Marlene Mason

A Thesis
Submitted to the
Faculty of The Graduate College
in partial fulfillment
of the
Degree of Master of Arts

Western Michigan University
Kalamazoo, Michigan
December 1974
ACKNOWLEDGMENT

My sincere appreciation goes to Drs. George Sidney, Malcolm Robertson and Frank Fatzinger for their suggestions, information, guidance and criticism in the completion of this thesis. My thanks also go to Dr. Bradley Huitema for his assistance in the statistical aspects of this study. Lastly, I must also thank the subjects, for without their cooperation, the results would not have been possible.

Marlene Mason
INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

Xerox University Microfilms
300 North Zeeb Road
Ann Arbor, Michigan 48106

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
MASTERS THESIS

MASON, Marlene Sue,
A STUDY OF THE FEASIBILITY OF THE W.A.I.S. AS
A SELF-ADMINISTERED PENCIL-PAPER TEST.

Western Michigan University, M.A., 1974
Psychology, clinical

Xerox University Microfilms, Ann Arbor, Michigan 48106

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I  INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>II METHOD</td>
<td>8</td>
</tr>
<tr>
<td>III RESULTS</td>
<td>14</td>
</tr>
<tr>
<td>IV DISCUSSION</td>
<td>25</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>30</td>
</tr>
</tbody>
</table>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
CHAPTER I
INTRODUCTION

For centuries it was known that people differed in their intellectual ability; however, it was not until the early 20th century that a systematic measure of intelligence was developed. Alfred Binet in France published the Measuring Scale of Intelligence for use with school children. This instrument, with slight modifications, was also used for testing adult intelligence.

It was not until the United States entered World War I that an instrument for testing adults was required. The Army Alpha and Beta were group tests developed for this purpose. Since World War I a number of group instruments have been developed, many of them based upon the Army Alpha.

For individual testing, however, a form of the Binet continued to be used for testing many individuals until David Wechsler in 1939 developed his own tests for adults. His reasons for developing a test for adults were based upon problems encountered in using tests for adults that were originally developed for children, i.e., adult norms had not been established, there was an emphasis on speed as compared to accuracy and the material was not suitable for testing adults (Wechsler, 1944).

A number of tests in use today measure aptitudes and achievement rather than I.Q. Many of these were designed for high school and college students and are not suitable
for use with the general population.

Wechsler's measures of adult intelligence, the Wechsler-Bellevue series (W-B I and W-B II) and the Wechsler Adult Intelligence Scale (W.A.I.S.) are generally considered to be the best measures of adult intelligence available today. One of the major criticisms generally cited about the test, however, is that not enough research has been done on correlating W.A.I.S. scores and personality traits. This doesn't seem to be important because it is basically a test of intelligence, not personality. Another criticism offered is that it discriminates against non-white and foreign born individuals, in that it contains questions that are culturally more meaningful to whites who are native born. Other criticisms levied against the test, although not against the test per se, but against the scoring manual, seem to affect everyone who uses the W.A.I.S. There is nothing contained in the manual relating to Wechsler's theory of intelligence. Some of the items require frequent updating. Also, the foreign born may have a cultural disadvantage on some questions.

The format of the manual is the most troublesome to those who are learning how to administer the W.A.I.S. and probably to those who give it only occasionally. Scoring criteria are usually separate from administration instructions, particularly for the most important verbal subtests, i.e., Vocabulary, Comprehension and Similarities. If the examples followed the directions or items, it would be easier to
determine the point level of the response simultaneously, aiding inquiry if the Examiner is in doubt about point value at any time. The recording blank might be modified to include a statement on each subtest about how many failures are required to terminate a subtest and to make it more clinically useful, perhaps by allowing more recording space for behavior observations (Buros, 1972).

In the standardization of the Wechsler-Bellevue, Form I, Wechsler developed norms for different age groups, different levels of education, both sexes, and for different occupational groups. Two disadvantages of the standardization group of W-B Form I were the exclusion of non-whites and inclusion of a basically urban population centered about New York (Wechsler, 1944). The standardization for W-B Form II was the same except Wechsler used only adult males (Wechsler, 1946).

The standardization of the W.A.I.S. in 1956 was similar to that of the W-B, but it did overcome many of the W-B disadvantages.

The W.A.I.S. was standardized on 1,700 subjects of both sexes, ranging in age from 16 to 64, plus an additional group of subjects ranged 60 to 75 and over. Subjects were assigned to groups according to geographical region, rural-urban areas, white and non-white ratios, occupation and educational levels based upon the proportions given in the 1950 census (Wechsler, 1955).
Today the United States population is approximately 210 million. It is estimated that approximately 10 to 30 percent^1 will require or desire testing each year. The W.A.I.S. is one of the best and most widely preferred instruments available for measuring adult intelligence. Its administration, however, requires from one to two hours per subject, and another 30 minutes are required to score and interpret, with report writing adding to that time. It can only be administered individually, not in small groups, and its administration requires the undivided attention of the Examiner and Subject.

If the W.A.I.S. could be given in a written or group form, using the same questions and scoring criteria as the traditionally administered form, one might save much of the time needed for individual administration alone. It could be administered to groups, or to an individual, when an Examiner does not have time to individually administer the items in face to face situations. Although one might lose some of the behavioral observations that could prove useful later, and the subjects might lose the benefit of some time bonus points given to increase basic scores, this has been determined to be of minimal importance. On the verbal section, the W.A.I.S. verbal scale allows only a maximum of four extra points, while on the performance sections,

the subject may lose as many as 27 points.

Three experimenters have investigated group administration and self or automated administration of the W.A.I.S. The results of these should be interpreted with caution, in that the subjects for the most part, were college students or individuals having an above average intelligence level and the results may not generalize to individuals that are below average in ability.

Eme and Walker (1971) administered the W.A.I.S. to 60 undergraduate students at three Chicago colleges. The subtests Information, Similarities, Picture Completion and Digit Span were given in a group administration. Questions for Information and Similarities were read to the subjects and Picture Completion and Digit Span were presented by means of an overhead projector. Subjects' responses were written on a sheet of paper. The subtests Vocabulary, Arithmetic, Block Design and Picture Arrangement were presented to the subjects individually in the standardized manner. The total I.Q.'s were pro-rated for the group and individual administrations and were compared. The correlations between group and individual administration were significant at the .01 level ($r = .53$). There were no significant differences between the mean I.Q. scores for the group and individual administration tests.

Mishra (1971) investigated examiner versus machine administration of the verbal scale of the W.A.I.S. on 40
graduate students at the University of Oregon. Two forms were created, one using odd numbered items and the other using even numbered items. These forms were administered by a tape recorder or an examiner. The order of administration was alternated between tape recorder and examiner, and form one and form two. The tape recorder was modified to permit the subject to turn it on and off and regulate the volume but he could not reverse or play back the tape. A second recorder was turned on at the beginning of the session and continuously recorded the items and the subject's responses. The results indicated that the two forms of presentation were comparable and the differences in performance was not significant.

Elwood and Griffin (1972) have devised an automated method of administering the W.A.I.S. A room was specially constructed for the testing program. The subtests Information, Comprehension, Similarities and Vocabulary were presented one item at a time when the subject pressed a "ready" button. Responses were tape recorded for transcription later. The Arithmetic subtest item 1 was presented on a projection screen. The remainder of the Arithmetic items were presented by recorder and the responses were tape recorded. Digit Span was presented by recorder and the subject was required to push the appropriate buttons on the response panel. The Digit Symbol subtest key was projected onto a screen, the numbers were projected onto a digit
readout device on the response panel. Buttons contained the symbols that were used for responding. To respond, the subject looked at the digit readout, then looked at the screen to determine which symbol was associated with the number shown and then pressed the button with the appropriate symbol. Picture Completion was administered using the projection screen and verbal responses were recorded.

Block Design, Picture Arrangement and Object Assembly were administered individually from small drawers. The subject was required to arrange the pieces in the drawer and later a technician scored the responses and rearranged the pieces according to instructions in the W.A.I.S. manual. The recorded responses were transcribed verbatim and scored by a technician according to instructions in the W.A.I.S. manual. Subjects were retested a few days later using the automated system again. The correlations on test-retest using the automated method were significant at the .01 level with $r = .95$ for Performance I.Q., $r = .97$ for Verbal I.Q. and $r = .98$ for Full Scale I.Q.

To determine the feasibility of administering the W.A.I.S. as a paper-pencil type test, the W.A.I.S. and Wechsler-Bellevue Form II were administered to 32 subjects. Each subject received an oral and written form of the test with specially constructed instructions for the latter.

A paper-pencil form was chosen, because a search of the literature failed to provide any reference that this type of administration had been attempted before.
CHAPTER II

METHOD

The Wechsler-Bellevue Form II (W-B II) and Wechsler Adult Intelligence Scale (W.A.I.S.) were chosen as the instruments to test the question of whether or not a written form of the W.A.I.S. yields approximately the same verbal I.Q. scores as does the traditionally administered form.

The Wechsler-Bellevue Form I is generally considered to be the equivalent form of the W.A.I.S., but it was decided not to use this because many of the items of the W-B I also appear on the W.A.I.S. in the same terms. Originally, Wechsler had intended the W-B II to be the equivalent form of the W-B I but because the means and standard deviations were found to be different, it was not used to a great extent. The W-B II was standardized on an adult population, and its questions are similar to but not identical with those appearing on the W.A.I.S.

Gibby (1949) reported correlations of Form I with Form II ranging from .20 to .93 for individual subtests, with Comprehension being lowest and Vocabulary being highest. Verbal I.Q. correlated .76, Performance I.Q. .82 and Full Scale I.Q. .87. Quereshi and Miller (1970) found correlations of the W-B II and W.A.I.S. ranging from .24 to .71 for individual subtests and .78 for Full Scale I.Q.'s.
The W-B II and W.A.I.S. were administered to 32 male and female subjects. Sixteen of these were patients of the William Upjohn DeLano Memorial Clinic, either as an outpatient (three males and two females) or as an inpatient (four males and seven females) of the Borgess Hospital psychiatric Unit. The other sixteen subjects were non-patients (seven males and nine females) whose occupations included housewives, secretaries, high school and college students, salesmen and lawyers. Patients ranged in age from 16 to 58. Non-patients ranged in age from 16 to 30. The subjects in each group were not matched for age, sex or education. It should be pointed out that the educational level for subjects was no less than ninth grade and the educational range was ninth grade through graduate school. None reported having a reading difficulty of any appreciable amount. One non-white subject was included in each group. Because of the restricted sample of subjects, results of this study should not be generalized to groups or individuals that are considered to be below average in ability.

There were four possible administration combinations for each group, with four subjects per administration combination:

1) Oral W.A.I.S. first, written W-B II second;
2) Oral W-B II first, written W.A.I.S. second;
3) Written W.A.I.S. first, oral W-B II second;
4) Written W-B II first, oral W.A.I.S. second.
The second test was administered within 1 to 24 days after the test, under identical physical and emotional conditions wherever possible.

The subtests given were: Information, Comprehension, Similarities, Arithmetic, Vocabulary, Picture Arrangement and Picture Completion. These subtests were chosen because they individually correlate highest with the W.A.I.S. full scale I.Q. (Maxwell, 1957)

None of the subtests in either administration were timed. The Verbal I.Q. on the W.A.I.S. loses a maximum of only four points by this method, and these are earned only on the Arithmetic subtest. The greatest problem is with the performance subtests because more points can be lost if time is disregarded; however, only Picture Arrangement and Picture Completion were used to compute Performance I.Q. This resulted in a maximum loss of four points on the Performance I.Q., and these are lost on the Picture Arrangement only. Picture Arrangement and Picture Completion were chosen as the subtests to measure Performance I.Q. because the written administration form differed the least from the traditionally administered form.

The Verbal I.Q. was pro-rated using the subtests Information, Comprehension, Similarities, Arithmetic and Vocabulary. The total number of weighted points were used for determining W-B II Verbal I.Q. For determining the W.A.I.S. Verbal I.Q. the total weighted score was multiplied by 6/5.
The Performance I.Q. was pro-rated using Picture Arrangement and Picture Completion. Weighted scores for both W-B II and W.A.I.S. were multiplied by 5/2 to obtain an estimate of the I.Q. possible if all five subtests were given.

The Full Scale I.Q. for both W-B II and W.A.I.S. was determined using the pro-rated values of the verbal and performance subtests, per traditional instructions.

Digit Span, Block Design and Object Assembly were also given, but were not included in determining I.Q.'s. These subtests were given to determine if they were possibly equivalent in a group administration form to traditional individual oral administration form, because the instructions for these subtests differed the most from traditional administration.

Each subtest in the written form consisted of all the items for that subtest, and were presented in the same form and order as the item in the manual with the following exceptions:

All items were presented for each subtest including some initial items that are usually omitted if the subject answers more difficult items;

On the W.A.I.S. Arithmetic subtest, item 1 was omitted. In Picture Completion, item 11 was difficult to duplicate and the equivalent item 12 from the Wechsler Intelligence Scale for Children was substituted. Block Design item 10
was rotated for better fit on the page.

On the W-B II Information subtest, the lettered items were omitted;

Picture Arrangement items for both tests used the example from the W-B II, for instruction purposes;

Digit Span on both the W-B II and W.A.I.S. consisted of all possible number series;

Object Assembly was made of buff colored poster board pieces that were waxed on the back and adhered to tape strips on sheets of paper, which allowed the subject to remove and rearrange the pieces.

A copy of each test is not included with this thesis because of the recognized need for continued confidentiality of the test items.

Instructions for the written tests were as follows:

Information: "The following questions pertain to everyday knowledge. Answer them as fully as you can."

Comprehension: "These questions deal with everyday judgment situations. Write out what you would do in each situation or how you would interpret, or deal with each situation."

Arithmetic: "Read the following arithmetic problems and do the figuring in your head. Then write your answer in the space provided. Do not do any figuring on the paper."

Similarities: "The following items in each pair are the same or most alike in certain ways. Write the way or ways they are alike, or the same. As an example, for the pair WAGON-BICYCLE, any or all of the following would be correct: both have wheels; both can be made of metal; both can be toys; more importantly, both can be means of transportation."
Vocabulary: "Write the meanings of the following words. Be as specific as possible. Don't worry about spelling or grammar because this will not be considered in the scoring."

Digit Span Forward: "Read each line of numbers once, turn the paper over and write the numbers in the space provided on the back. Do the same for each line. If you make any changes, draw a single line through the response and write your new response underneath."

Digit Span Backward: "Read each line of numbers once, turn the paper over and write the numbers backwards in the space provided. For example, if the numbers on the front are 8, 2, 7 on the back you write 7 2 8. Do the same for each line. If you made any changes, draw a single line through the response and write your new response underneath."

Picture Arrangement: "The pictures on each of the following pages tell a story. As the pictures are arranged now, they are not as they should be. Each picture has a number below it. You are to rearrange the numbers of the pictures in the order that you think makes the most sensible story. As an example, the pictures below tell a story about a man who is in a fight. The man took a beating. If the pictures were rearranged to make the most sensible story, the order would be 3, 2, 1 and you would write 3-2-1 in the space provided."

Picture Completion: "In each of the following pictures there is an important part missing. Look at each picture and in the blank below it, write the name of the missing part. If you do not know the name of the missing part, circle the specific area where it is missing. Make as small a circle as possible."

Block Design: "Below are 8 designs and 3 large squares (made up of 4 or 9 smaller squares) for each design. Study each design and then darken in the areas in the first square that would duplicate each design. The other squares are to be used in case you make a mistake. The designs have been created by using a completely dark square, and/or completely white square, and/or half dark and half white square."

Object Assembly: "When these pieces are put together they will make something. Put them together as quickly as you can."
CHAPTER III
RESULTS

The subtests in the patient group yielded generally higher correlations (see Table I) with $r$ ranging from .3678 to .7571 than the correlations for the non-patient group. This was very encouraging, since all patients were included, without exception, and one subject had manifested over a 40 point difference between written and oral I.Q. scores. The non-patient group yielded few significant correlations (see Table II) with $r$ ranging .1572 to .7287. One reason for this may be that there is less variability in the weighted scores in the subtests for the non-patient group than for the patient group.

As shown in Table II, the statistical correlations for the individual subtest scores of the written form with the oral form were found to be fairly high.

Correlations of subtests were calculated for all subjects using an average of Fisher $Z'$ transformation scores for the patient and non-patient group. The only subtests which did not attain a .01 significance level ($r$ ranged .322 to .740) were Comprehension, ($r = .322$); Arithmetic, ($r = .333$); and Object Assembly, ($r = .368$) (see Table II).

It was originally decided to use Picture Arrangement and Picture Completion as a measure of performance I.Q.
Table I

Correlations Between Written and Oral Forms of The W.A.I.S. and W-B II with Significance Levels for All Subtests for Each Group

<table>
<thead>
<tr>
<th>Test</th>
<th>Patient</th>
<th>Signif. Level</th>
<th>Non-Patient</th>
<th>Signif. Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information</td>
<td>.7571</td>
<td>.01</td>
<td>.6593</td>
<td>.01</td>
</tr>
<tr>
<td>Comprehension</td>
<td>.3678</td>
<td>---</td>
<td>.2741</td>
<td>---</td>
</tr>
<tr>
<td>Similarities</td>
<td>.7538</td>
<td>.01</td>
<td>.3107</td>
<td>---</td>
</tr>
<tr>
<td>Arithmetic</td>
<td>.4877</td>
<td>.05</td>
<td>.1572</td>
<td>---</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>.7518</td>
<td>.01</td>
<td>.7287</td>
<td>.01</td>
</tr>
<tr>
<td>Picture Completion</td>
<td>.5824</td>
<td>.05</td>
<td>.4885</td>
<td>.05</td>
</tr>
<tr>
<td>Picture Arrangement</td>
<td>.7642</td>
<td>.01</td>
<td>.4217</td>
<td>---</td>
</tr>
<tr>
<td>Digit Span</td>
<td>.5417</td>
<td>.05</td>
<td>.4346</td>
<td>---</td>
</tr>
<tr>
<td>Block Design</td>
<td>.6422</td>
<td>.01</td>
<td>.6110</td>
<td>.01</td>
</tr>
<tr>
<td>Object Assembly</td>
<td>.5017</td>
<td>.05</td>
<td>.2174</td>
<td>---</td>
</tr>
</tbody>
</table>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Table II

Correlations Based Upon Fisher's Z' Transformation Between Written and Oral Forms of the W.A.I.S. and W-B II With Significance Levels for All Subtests and for All Subjects

<table>
<thead>
<tr>
<th>Subtest</th>
<th>r</th>
<th>Significance Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information</td>
<td>.711</td>
<td>.01</td>
</tr>
<tr>
<td>Comprehension</td>
<td>.322</td>
<td>---</td>
</tr>
<tr>
<td>Similarities</td>
<td>.573</td>
<td>.01</td>
</tr>
<tr>
<td>Arithmetic</td>
<td>.333</td>
<td>---</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>.740</td>
<td>.01</td>
</tr>
<tr>
<td>Picture Arrangement</td>
<td>.621</td>
<td>.01</td>
</tr>
<tr>
<td>Picture Completion</td>
<td>.537</td>
<td>.01</td>
</tr>
<tr>
<td>Digit Span</td>
<td>.513</td>
<td>.01</td>
</tr>
<tr>
<td>Block Design</td>
<td>.626</td>
<td>.01</td>
</tr>
<tr>
<td>Object Assembly</td>
<td>.368</td>
<td>.05</td>
</tr>
</tbody>
</table>
On the basis of the correlations of the written and oral form for Block Design, perhaps it would have been a better choice for measuring performance I.Q. than Picture Arrangement.

The mean weighted score and standard deviation for each subtest and I.Q. scale for the patient group is shown in Table III. An examination of this table shows that there is an average difference of less than 1.2 points in mean weighted scores between the written and oral forms. The significance levels of the t values indicate that there is no difference in the written and oral forms; however, because of the wide difference in written and oral I.Q. means and large standard deviations, it is recommended that the written form not be used with psychiatric patients.

Table IV gives the mean weighted score and standard deviation for the subtests and I.Q. scales for the non-patient group. There is less than 1 point difference in mean weighted scores between the written and oral forms. The t values for the non-patient group indicate that there is little difference in the written and oral forms.

Correlations of written I.Q.'s with oral I.Q.'s are given in Table V for patients and non-patients. The significance levels between the written and oral scores were found to be at the .01 level (r ranged .6438 to .8085) for the patients and non-patients.

Full Scale I.Q. scores for patients ranged from 68 on W-B II written, to 124 on W-B II oral; for non-patients the range was 92 on W-B II written to 142 on W.A.I.S. oral.
Table III
Mean Weighted Scores and Standard Deviations of Individual Subtests and I.Q. Scales for Patient Group for Written Form and Oral Form

<table>
<thead>
<tr>
<th></th>
<th>Written</th>
<th></th>
<th>Oral</th>
<th></th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Standard Deviation</td>
<td>Mean</td>
<td>Standard Deviation</td>
<td></td>
</tr>
<tr>
<td>Information</td>
<td>9.06</td>
<td>2.89</td>
<td>9.06</td>
<td>3.02</td>
<td>-0.1</td>
</tr>
<tr>
<td>Comprehension</td>
<td>8.94</td>
<td>3.59</td>
<td>9.81</td>
<td>1.97</td>
<td>-0.87</td>
</tr>
<tr>
<td>Similarities</td>
<td>9.26</td>
<td>2.06</td>
<td>9.75</td>
<td>3.17</td>
<td>-0.52</td>
</tr>
<tr>
<td>Arithmetic</td>
<td>9.12</td>
<td>2.55</td>
<td>8.62</td>
<td>1.89</td>
<td>0.63</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>8.93</td>
<td>3.19</td>
<td>10.31</td>
<td>2.30</td>
<td>-1.41</td>
</tr>
<tr>
<td>Picture Arrangement</td>
<td>8.81</td>
<td>1.81</td>
<td>9.31</td>
<td>2.10</td>
<td>-0.72</td>
</tr>
<tr>
<td>Picture Completion</td>
<td>9.63</td>
<td>3.27</td>
<td>11.87</td>
<td>2.92</td>
<td>2.04</td>
</tr>
<tr>
<td>Digit Span</td>
<td>12.56</td>
<td>5.00</td>
<td>10.19</td>
<td>3.97</td>
<td>1.43</td>
</tr>
<tr>
<td>Block Design</td>
<td>8.38</td>
<td>3.65</td>
<td>9.81</td>
<td>1.84</td>
<td>-1.40</td>
</tr>
<tr>
<td>Object Assembly</td>
<td>8.31</td>
<td>2.08</td>
<td>9.12</td>
<td>1.22</td>
<td>-1.33</td>
</tr>
<tr>
<td>Full Scale I.Q.</td>
<td>96.75</td>
<td>16.33</td>
<td>105.44</td>
<td>12.35</td>
<td>-1.70</td>
</tr>
<tr>
<td>Verbal I.Q.</td>
<td>96.75</td>
<td>16.05</td>
<td>103.63</td>
<td>13.26</td>
<td>-1.33</td>
</tr>
<tr>
<td>Performance I.Q.</td>
<td>97.19</td>
<td>16.16</td>
<td>106.68</td>
<td>16.28</td>
<td>-1.65</td>
</tr>
</tbody>
</table>

n = 16
<table>
<thead>
<tr>
<th></th>
<th>Written</th>
<th></th>
<th>Oral</th>
<th></th>
<th>( t )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Standard Deviation</td>
<td>Mean</td>
<td>Standard Deviation</td>
<td></td>
</tr>
<tr>
<td>Information</td>
<td>13.19</td>
<td>1.55</td>
<td>12.81</td>
<td>2.62</td>
<td>.50</td>
</tr>
<tr>
<td>Comprehension</td>
<td>12.87</td>
<td>2.69</td>
<td>11.75</td>
<td>2.82</td>
<td>1.14</td>
</tr>
<tr>
<td>Similarities</td>
<td>13.50</td>
<td>2.52</td>
<td>13.50</td>
<td>2.12</td>
<td>-0-</td>
</tr>
<tr>
<td>Arithmetic</td>
<td>11.94</td>
<td>1.27</td>
<td>9.69</td>
<td>2.36</td>
<td>3.36*</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>13.44</td>
<td>1.94</td>
<td>13.00</td>
<td>1.80</td>
<td>.66</td>
</tr>
<tr>
<td>Picture Arrangement</td>
<td>11.63</td>
<td>1.31</td>
<td>12.18</td>
<td>1.59</td>
<td>-1.08</td>
</tr>
<tr>
<td>Picture Completion</td>
<td>12.25</td>
<td>2.33</td>
<td>12.18</td>
<td>2.24</td>
<td>.09</td>
</tr>
<tr>
<td>Digit Span</td>
<td>13.81</td>
<td>3.89</td>
<td>12.06</td>
<td>2.07</td>
<td>1.59</td>
</tr>
<tr>
<td>Block Design</td>
<td>10.75</td>
<td>1.14</td>
<td>10.75</td>
<td>1.25</td>
<td>-0-</td>
</tr>
<tr>
<td>Object Assembly</td>
<td>9.38</td>
<td>.60</td>
<td>9.38</td>
<td>.50</td>
<td>-0-</td>
</tr>
<tr>
<td>Full Scale I.Q.</td>
<td>118.06</td>
<td>10.30</td>
<td>118.63</td>
<td>9.70</td>
<td>- .16</td>
</tr>
<tr>
<td>Verbal I.Q.</td>
<td>120.81</td>
<td>10.20</td>
<td>117.75</td>
<td>9.34</td>
<td>.88</td>
</tr>
<tr>
<td>Performance I.Q.</td>
<td>115.13</td>
<td>10.10</td>
<td>115.13</td>
<td>9.10</td>
<td>-0-</td>
</tr>
</tbody>
</table>

\( n = 16 \)

*level of significance .01
Table V

Correlation of Written I.Q.'s With Oral I.Q.'s for Patients and Non-Patients

<table>
<thead>
<tr>
<th></th>
<th>Patient $r$</th>
<th>Non-Patient $r$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full Scale</td>
<td>.8085</td>
<td>.6920</td>
</tr>
<tr>
<td>Verbal</td>
<td>.6438</td>
<td>.7094</td>
</tr>
<tr>
<td>Performance</td>
<td>.7958</td>
<td>.6561</td>
</tr>
</tbody>
</table>
Full Scale I.Q. for patients had a written mean value of 96.75 and an oral mean value of 105.44. The mean written Full Scale I.Q. for non-patients was 118.06 and the oral mean value was 118.63 (see Tables III and IV).

Verbal I.Q. scores ranged from 67 on W-B II written, to 128 on W-B II oral for patients; and from 100 on W-B II written to 141 on W.A.I.S. oral and W-B II written for non-patients. For patients the mean written Verbal I.Q. was 96.75 and mean oral Verbal I.Q. was 103.63. For non-patients the mean written Verbal I.Q. was 120.81 and the mean oral Verbal I.Q. was 117.75 (see Tables III and IV).

Performance I.Q. scores ranged from 65 on W.A.I.S. written, to 129 on W.A.I.S. oral for patients; and from 101 on W.A.I.S. oral, to 140 on W.A.I.S. written for non-patients. The mean written Performance I.Q. for patients was 97.19, the mean oral Performance I.Q. was 106.68. For non-patients the mean written Performance I.Q. was 115.13 and the mean oral Performance I.Q. was 115.13 (see Tables III and IV).

Although the correlations for the patient group appear to be higher than those for the non-patient group, there is a much wider discrepancy in mean I.Q.'s found for the patient group than for the non-patient group. This may suggest that the written form may not be an acceptable substitute for the traditional W.A.I.S., especially with a psychiatric population.
In comparing means on the various I.Q.'s to determine whether there was a difference in scores if a written form was administered first or if an oral form was administered first, Table VI gives the means for the patient group. The majority of the mean I.Q.'s were higher for the oral form, whether it was administered first or second. The t values do not indicate that there is a difference in whether an oral or written form is administered first.

Table VII (page 21) shows the mean I.Q.'s for the non-patient group. No differences in mean scores were obtained with the written or oral form.

The correlations obtained in this study, although significantly greater than zero are not always as high as some of those discussed in literature on short forms of the W.A.I.S.

The results of the non-patient group demonstrated that the written and oral forms are approximately equivalent, permitting the written form to be administered when a large number of individuals require testing in a short period of time, where not enough examiners might be available, or when an examiner cannot afford to provide the one to two hours necessary to administer the W.A.I.S. individually.
Table VI

Mean I.Q. Values for Patient Group for Written Form Administered First and Oral Form Administered Second and Oral Form Administered First and Written Form Administered Second

<table>
<thead>
<tr>
<th></th>
<th>Written First</th>
<th>Oral Second</th>
<th>t</th>
<th>Oral First</th>
<th>Written Second</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full Scale</td>
<td>99.25</td>
<td>107.00</td>
<td>-1.37</td>
<td>103.87</td>
<td>94.25</td>
<td>1.10</td>
</tr>
<tr>
<td>Verbal</td>
<td>98.00</td>
<td>104.37</td>
<td>-1.35</td>
<td>102.87</td>
<td>95.50</td>
<td>.84</td>
</tr>
<tr>
<td>Performance</td>
<td>101.12</td>
<td>109.62</td>
<td>-1.86</td>
<td>103.75</td>
<td>93.25</td>
<td>1.18</td>
</tr>
<tr>
<td></td>
<td>n = 8</td>
<td>n = 8</td>
<td></td>
<td>n = 8</td>
<td>n = 8</td>
<td></td>
</tr>
</tbody>
</table>
Table VII

Mean I.Q. Values for Non-Patient Group for Written Form Administered First and Oral Form Administered Second and Oral Form Administered First and Written Form Administered Second

<table>
<thead>
<tr>
<th></th>
<th>Written First</th>
<th>Oral Second</th>
<th>t</th>
<th>Oral First</th>
<th>Written Second</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full Scale</td>
<td>117.25</td>
<td>121.12</td>
<td>-.66</td>
<td>116.00</td>
<td>118.87</td>
<td>1.09</td>
</tr>
<tr>
<td>Verbal</td>
<td>118.87</td>
<td>118.25</td>
<td>.13</td>
<td>117.25</td>
<td>122.75</td>
<td>1.17</td>
</tr>
<tr>
<td>Performance</td>
<td>120.12</td>
<td>118.87</td>
<td>2.45*</td>
<td>111.37</td>
<td>110.12</td>
<td>.43</td>
</tr>
<tr>
<td></td>
<td>n = 8</td>
<td>n = 8</td>
<td></td>
<td>n = 8</td>
<td>n = 8</td>
<td></td>
</tr>
</tbody>
</table>

*level of significance .05
Although the *t* values were not significant for the patient group for written and oral I.Q. means, it is recommended that the written form should not be administered to patients unless used with extreme caution, and only if a measure of verbal ability is required under extenuating circumstances because of the wide discrepancy in I.Q. scores.

The correlations for the performance subtests were not satisfactory for either group, and much more research is needed before the written form of the performance subtests can be administered with any reliability to all individuals, patient and non-patient alike.

For patients Verbal I.Q.'s correlated .6438, Performance I.Q.'s correlated .7953 and Full Scale I.Q.'s correlated .8085. For non-patients Verbal I.Q.'s correlated .7094, Performance I.Q.'s correlated .6561 and Full Scale I.Q.'s correlated .6920. The correlated individual subtests were significant for the patient group (*r* ranged .3678 to .7642). A few subtests were significant for the non-patient group with *r* ranging .1572 to .7287.

Although the correlations are not as high as those reported in the literature on short forms of the W.A.I.S. they are significantly greater than zero. It must be remembered that the instruments used were not identical.
A number of studies use a test-retest version or split half version of the W.A.I.S. If W-B I had been used with the W.A.I.S. or a test-retest of the W.A.I.S. had been given, the correlations might have been higher. Additional research on a written form of the W.A.I.S. is needed to help determine if a test-retest of the W.A.I.S. were given would the correlations be higher than those of this study.

In scoring the written forms, none of the subjects appeared to have any difficulty in following directions on the verbal subtests that were used to pro-rate I.Q. There were various problems with the performance subtests, some of which may have been the result of subjects not reading the instructions.

Identifying less essential parts on the Picture Completion subtest occurred frequently. Some subjects did not draw a circle around the missing part, but rather drew the part in.

On the Picture Arrangement subtest a few subjects did not follow directions for recording the numbers of the pictures in the order to make a sensible story.

Carelessness on the Block Design subtests resulted in lower scores for some subjects. Sketching in, but not darkening in, the designs occurred frequently. Not dividing the squares into diagonal halves (or anywhere near half) was much more common.
A number of subjects on the Digit Span subtest failed to write the numbers backwards for the second page. One subject wrote all numbers backwards for both pages.

These mistakes all point to the necessity for making the directions and examples much more explicit in future administration or experimentation.

Another major drawback to the written form is the inability to question responses or caution subjects to give the most important response. This can result in lower scores on a number of subtests.

Block Design and Object Assembly, although not used in determining I.Q., appeared to be equivalent on the written and oral forms. A pro-rating system could be used to give extra points on these subtests to individuals that score higher on the other performance items. Another possibility is to time the performance subtests and allow a maximum period of time. Subjects who complete the subtests in less time may be given bonus points.

Although Digit Span correlated highly with the written and oral form, it is not recommended that this subtest be given in a written form. There is not sufficient control over how many times a subject may read the numbers, in the order that they read the numbers, or in how many times they may turn the paper over for one series. This may result in higher scores on the written form.
The results of this study must be interpreted with caution and should not be generalized to those that are below average in ability. Also, the use of a written form with psychiatric patients is not recommended. A subject that may be experiencing adverse reactions to medications or one that appears to be more out of touch with reality should not be tested with a written form. Table VI gives the mean I.Q.'s for patients for the written and oral forms. The I.Q.'s were higher for oral administrations. This may be due to the individual attention given the subject during the testing.

Additional directions for research on the question of equivalency of written W.A.I.S. scores to oral scores is definitely suggested. Changing instructions for some of the subtests may help to eliminate subjects failure to follow directions. The failure to follow directions was not confined to either subject group. For the researchers that are interested in recording behavior observations, the use of a portable video-tape unit could be studied. Administering only the verbal items in written form and individually administering the performance subtests then comparing the scores with all subtests being individually administered is another possibility.

The largest area for testing the hypothesis of equivalency of written and oral forms would be in the schools normally using the Wechsler Intelligence Scale for
Children. It is not feasible to use a written form in primary grades, because of the reading comprehension problem, but in middle and high school grades this might be possible.
BIBLIOGRAPHY


Wechsler, D. *The measurement of adult intelligence*. (3rd ed.) Baltimore: Williams & Wilkins, 1944.
