6-1961

Validity of the Goodenough Draw-A-Man Test with Children Aged Four, Five, and Six

Jaren Van Den Heuvel
Western Michigan University

Follow this and additional works at: https://scholarworks.wmich.edu/masters_theses

Part of the Experimental Analysis of Behavior Commons

Recommended Citation
https://scholarworks.wmich.edu/masters_theses/3689

This Masters Thesis—Open Access is brought to you for free and open access by the Graduate College at ScholarWorks at WMU. It has been accepted for inclusion in Master's Theses by an authorized administrator of ScholarWorks at WMU. For more information, please contact maira.bundza@wmich.edu.
VALIDITY OF THE GOODENOUGH DRAW-A-MAN TEST
WITH CHILDREN AGED FOUR, FIVE, AND SIX

by

Jaren Van Den Heuvel

A Thesis Presented to the Graduate
Faculty of Western Michigan University in
Partial Fulfillment of the Requirements for the
Degree of Master of Arts

Western Michigan University
Kalamazoo, Michigan
June, 1961
ACKNOWLEDGMENTS

The writer expresses appreciation to Dr. John A. Popplestone, who served as chairman of the committee, to Dr. Stanley Kuffel and Dr. Eaton J. Asher for their helpful suggestions throughout the course of this investigation.

Acknowledgment is made also to Mr. Robert Wonderlin, School Diagnostician for Kalamazoo County, for his help in obtaining subjects, to the many helpful teachers in the different schools and nurseries, and to the pupils who so generously participated.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>ii</td>
</tr>
<tr>
<td>INDEX OF TABLES</td>
<td>iv</td>
</tr>
<tr>
<td>THE GOODENOUGH DRAW-A-MAN TEST</td>
<td>1</td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Socio-economic level</td>
<td>2</td>
</tr>
<tr>
<td>Sex</td>
<td>3</td>
</tr>
<tr>
<td>Clinical Status</td>
<td>3</td>
</tr>
<tr>
<td>Administration</td>
<td>3</td>
</tr>
<tr>
<td>THE PROPOSED STUDY</td>
<td>5</td>
</tr>
<tr>
<td>Problem</td>
<td>5</td>
</tr>
<tr>
<td>Method</td>
<td>5</td>
</tr>
<tr>
<td>Subjects</td>
<td>5</td>
</tr>
<tr>
<td>Apparatus and Materials</td>
<td>6</td>
</tr>
<tr>
<td>Procedure</td>
<td>6</td>
</tr>
<tr>
<td>THE RESULTS OF THE STUDY</td>
<td>8</td>
</tr>
<tr>
<td>Findings</td>
<td>8</td>
</tr>
<tr>
<td>Discussion</td>
<td>12</td>
</tr>
<tr>
<td>SUMMARY</td>
<td>14</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>15</td>
</tr>
</tbody>
</table>
INDEX OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Total Score Data for Each Age Level of Present and Goodenough &quot;Normal&quot; Samples.</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>Standard Error of Difference Between the Goodenough and Current Norms.</td>
<td>9</td>
</tr>
<tr>
<td>3</td>
<td>Significance of Difference Between Means at Successive Age Levels on the Current Study.</td>
<td>10</td>
</tr>
</tbody>
</table>
THE GOODENOUGH DRAW-A-MAN TEST

Introduction

In 1961 the Goodenough Draw-A-Man test is thirty five years old. The ease of administration and ease of obtaining these results from the test have made it increasingly popular through the years, for example; Ansbacher (1952) mentions that in 1946, according to a survey of psychological clinics, it held third place in frequency of use among psychometric instruments. Only the Stanford-Binet and the Wechsler-Bellevue exceeded it. This popularity exists despite frequent and basic criticism of the test. The most recent study which questions the validity of a Goodenough derived I.Q. is by Popplestone (1959) who has questioned its continued use on ages seven through ten. The present study has been conducted as a continuation of Popplestone's study, extending his data to the ages of four, five, and six, thus it is conducted as a normative study.

The Draw-A-Man test is a nonverbal scale which purports to measure a child's intelligence by means of his drawing of a man. It is intended for use with the ages three and a half to thirteen and a half. The scoring is based upon the presence of details of the figure which presumably indicate the subject's intellectual level from his perceptual level of differentiation of the figure. The esthetic quality of the drawing is not considered. (Freeman, 1959)

Goodenough (1926a) developed the rational behind her test from review of a number of studies which appeared to show that the nature and content of children's drawings are dependent primarily upon
intellectual development. That is, that drawings become more complete, more differentiated, and more organized as an individual becomes chronologically older. On the other hand this task is one which at a later age may actually decline since it reaches its maximum efficiency early in life.

The present study is, in a sense, a normative one. As such, it was necessary to control those variables which have been found to influence the drawing performance. A survey of the literature indicated that certain essential factors must be controlled or explicitly defined. These are found below.

**Socio-economic level**

Goodenough (1926b) obtained normative data for her test on 3,593 children who were generally of lower socio-economic status. She tried to compensate for this by adding a small group of upper class subjects, but since she did not account for all the data, her sampling has been criticized. Popplestone (1959) compared Goodenough's original standardization sample with a higher socio-economic sample of hers—the "Rutherford" group. He found that there was a significant difference of means between the two groups at three of the four chronological age levels which he used in his experiment. This would tend to indicate that Goodenough's standardization sample represents the performance of a lower socio-economic group and that the influence of class membership on performance is a real one.

A study by Britton (1954) reports that the average scores obtained by a lower status group were slightly lower than those obtained by a
higher status group.

**Sex**

The influence of the sex of the drawer on the drawings was considered even by Goodenough who suggests that traits such as a willingness to attend to details may be better developed in girls and so will tend to raise their score. Even though she implies that there may be a difference, she does not tell us the number of boys vs. girls in the standardization group of her test.

Others, such as Britton (1954), have also reported experiments which have shown that the drawing performance of girls was better than that of boys.

**Clinical Status**

Three studies have contrasted the Goodenough performance of pathological and non-pathological children—Springer (1941), Aron (1955), and Popplestone (1958). From these studies it is to be concluded that pathological and non-pathological groups differ significantly on the Goodenough test.

**Administration**

The effect of the instructions which request the drawing has been the subject of only one paper, that of Bliss and Berger (1954). They used both the Goodenough instructions:

> On these papers I want you to make a picture of a man. Make the very best picture that you can. Take your time and work very carefully. Try very hard and see what a good picture you can make.
and the Machover instructions:

Draw a person.

The authors concluded that the Goodenough instructions elicit a more complete performance than do the Machover instructions.

It was also pointed out that the role of the examiner in influencing drawings by children seems to have been neglected, but a study by Holtzman (1952) indicated that the sex and appearance of the examiner did not influence the drawings obtained when using Machover's instructions.

Only one worker, Feather (1950) has examined the effect of individual vs. group administration of the test, and he concludes that the productions of significant material in the drawings of college students are not eliminated in group administration.
THE PROPOSED STUDY

Problem

The study of Popplestone suggests that the continued use of Goodenough's test is to be questioned not only at ages seven through ten, but at all ages. If we assume the extrapolation of these results downward for the ages of four, five, and six, we would have to pronounce Goodenough's test as being an insufficiently valid measure of intelligence at any age. This hypothesis must be tested since we cannot assume this despite our suspicions. It could be possible that Goodenough's scoring criterion might apply at a lower age level. It is certainly plausible that this could happen as her criterion would naturally seem more justifiable at a lower conceptual and verbal level. The purpose of the study is to determine if there is a significant increase of scores on the test with an increase in chronological age at ages four through six.

Method

Subjects. As was indicated earlier, few controlled normative experiments have been conducted using Goodenough's scoring criterion. Of these only one, namely Popplestone, has attempted to control all the variables assumed to effect the outcome of the drawing score. This study has attempted to emulate the strict criterion set up by him.

Drawings were obtained from 293 white subjects from the public and private schools and nurseries from the Kalamazoo, Michigan area. An attempt was made to obtain an approximately equal number of subjects in
each of the three age levels. It was also attempted to obtain an approximately equal number of each sex in each classification.

Since the influence of the socio-economic factor seemed important, the subjects were drawn from varying social-economic status by identification of them as serving upper-lower class through lower-upper class status by the county diagnostician as serving such purposes.

An effort was made to assure the normalcy of the sample by setting up a criterion of normalcy in which any child with any physical, intellectual, or emotional deficit was excluded. The criterion of normalcy was (a) lack of referral to a diagnostic or treatment agency, (b) a school record and/or teacher or experimenter report free of a notation suggesting deviations. Grade repeaters were also excluded. Such stringency resulted in the discarding of approximately one third of the drawings or 116 drawings. The remaining 177 drawings of the sample were comprised of 28 boys and 20 girls aged 3 years 6 months through 4 years 5 months, 31 boys and 34 girls aged 4 years 6 months through 5 years 5 months, and 38 boys and 26 girls aged 5 years 6 months through 6 years 5 months.

Apparatus and Materials. Each subject was given a sheet of 8½" x 11" paper and a pencil.

Procedure. A request for a drawing of a man, in pencil, upon the distributed paper was made of each subject. The instructions were patterned after those of Goodenough. The sequence of the group administration was a distribution of the pencils and the papers with a request: "I would like you to draw me a picture of a man on these
papers, using these pencils. Draw him in any way you want but be
careful and draw him as well as you can." When the children had
completed the task they were asked to sign the picture so that the
drawer could be identified. Those subjects below school age received
more individualized attention as to administration and identification
of the drawing.

To increase the reliability of the scoring all the drawings were
scored by an experienced Goodenough scorer (Popplestone 1958). His
reliability has previously been established in an estimation of
inter examiner reliability.

The hypothesis of this study was that there would be a significant
increase of scores on the Goodenough Draw-A-Man Test with an increase
in chronological age. The total score of each age group was analyzed
to see if there was a significant rise in scores on the different age
age groups with chronological age.

For the comparability of six-year old vs. seven-year old, one
sub-sample of seven-year old Kalamazoo children and data provided by
them were incorporated into the present study.
THE RESULTS OF THE STUDY

Findings

The first statistical technique used was obtaining the means and standard deviations for each age group and then comparing them with the Goodenough sample. The means, and standard deviations from both the present and Goodenough samples are shown in Table 1.

It was revealed that Goodenough's subjects and the present data are not significantly different at the five, six, and seven-year old level, but it was significantly different at the four-year old level. These results obtained by taking the standard error of difference between the Goodenough and current sample means are shown in Table 2. Although the Goodenough and current data on the three age levels was not significantly different, the current data showed slightly lower means at each age level, and a significantly lower means at the four-year old level.

The second statistical procedure used was to obtain the significance of difference between means at each age level (t-test), to determine if these total scores increased significantly with increase in chronological age. The results of the t-tests between means at each consecutive age level are presented in Table 3. This design to test the significance of differences between the means reveals that at each age level there is a significant increase in score.

In summary: the present data reveals lower means at all age levels, but only significantly so at the four-year old age level. Further the application of the t-test of differences between the means on the current data reveals significant increases in scores at each age level.
Table 1.

Total Score Data for Each Age Level of Present and Goodenough

"Normal" Samples

<table>
<thead>
<tr>
<th></th>
<th>CA 4</th>
<th>CA 5</th>
<th>CA 6</th>
<th>CA 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>M</td>
<td>O</td>
<td>N</td>
<td>M</td>
</tr>
<tr>
<td>Goodenough</td>
<td>119</td>
<td>6.9</td>
<td>3.6</td>
<td>623</td>
</tr>
<tr>
<td>Current</td>
<td>48</td>
<td>5.29</td>
<td>3.76</td>
<td>65</td>
</tr>
</tbody>
</table>
Table 2.

Standard Error of Difference Between the Goodenough and Current Norms

<table>
<thead>
<tr>
<th></th>
<th>4 vs. 4</th>
<th>5 vs. 5</th>
<th>6 vs. 6</th>
<th>7 vs. 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t$</td>
<td>$t = 2.51$</td>
<td>$t = 0.806$</td>
<td>$t = 0.616$</td>
<td>$t = 1.49$</td>
</tr>
<tr>
<td>sig. at 2% level</td>
<td>not sig.</td>
<td>not sig.</td>
<td>not sig.</td>
<td>not sig.</td>
</tr>
</tbody>
</table>
Table 3.

Significance of Difference Between Means at Successive Age Levels on the Current Study

<table>
<thead>
<tr>
<th>4 vs. 5</th>
<th>5 vs. 6</th>
<th>6 vs. 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t = 6.98$</td>
<td>$t = 4.89$</td>
<td>$t = 4.69$</td>
</tr>
<tr>
<td>sign.</td>
<td>sign.</td>
<td>sign.</td>
</tr>
</tbody>
</table>
Discussion

This is in a sense, a normative study and consequently it was attempted to control those variables which have been found to influence the drawing performance. Goodenough's procedures were followed throughout the experiment and the means and standard deviations were obtained and compared to Goodenough's. This revealed that the present data does not differ significantly from the Goodenough data at the three older-age levels, but it does at the four-year-old level. It is to be noted however that although the current data at the three older-age levels do not differ significantly from the Goodenough data, they do portray lower means at each age level. These findings indicate then that the current study shows lower norms at each age level than does the Goodenough study.

The hypothesis of this study was substantiated in that there were significant increases in scores at each age level. This was shown by the significant difference between means at each age level.

The results of this study are surprising, because although there was a significant increase in score at each age level, the norms on the current study were lower at all age levels and significantly lower on the youngest age level. Although the test does discriminate between age levels, the validity of the Goodenough norms are to be questioned at each of these age levels for current use.

Since this was, in a sense, a normative study and the findings of this study question the validity of the Goodenough norms for the ages four through six, it appears legitimate that further inquiry
should be made. This should be done not only for the purpose of setting up more accurate norms for current usage of the test, but also for the purpose of further inquiry into the question of what it is supposedly measuring.

One reason for the difference of results in the two studies may be, of course, that in the current study a stricter criterion was used to insure a more normal sample. A second reason that could be postulated is that because of the cultural change over the past 35 years, we are now measuring something slightly different. But this would only be a condemnation of the test as not being culture free as it purports to be.

In any case, the lower norms at each age level (significant or not) on the current study cause one to question the validity of the continued use of this instrument as an index of intellectual ability with the Goodenough norms for the ages four through six. Although its use with the Goodenough norms is questioned, the question concerning it’s ability to distinguish intellectual levels is further extended since there was a discrimination shown in the current study.
SUMMARY

This study was conducted to explore the hypothesis that there would be significant increase of scores on the Goodenough Draw-A-Man Test with an increase in chronological age. Drawings were obtained from 293 subjects, CA 3 years 6 months through 6 years 5 months. After rejection of those subjects suggesting a physical, intellectual or emotional deficit, the remaining 177 drawings were scored according to Goodenough's criteria. One sub-sample of seven-year-old Kalamazoo children from a previous study was included. The results of this study revealed there was a significant increase in score at each age level, however the norms were lower at all age levels and significantly lower at the four-year-old level. The lower norms at each age level on the current study raise the question of the validity of this instrument as an index of intellectual level with the Goodenough norms.
REFERENCES


Goodenough, Florence L. A new approach to the measurement of the intelligence of young children. J. genet. Psychol., 1926, 33, 183-211. (b)


