A PARTIAL TEST OF VALIDITY OF THE MACHOVER DRAWING-OF-A-HUMAN-Figure TECHNIQUE

by

Paul W. Wanner
A. B., Sacramento State College, 1953

THESIS

Submitted in partial satisfaction of the requirements for the degree of

MASTER OF ARTS

AT THE

SACRAMENTO STATE COLLEGE

Approved:

Robert A. Good, Chair
Lawrence M. Brammer
Raymond W. Witt

Advisory Committee

Date July 22, 1953
ACKNOWLEDGMENTS

Acknowledgment is made to Dr. Robert A. Good who suggested the problem and gave valuable assistance and a great deal of his time and energy to the thesis. Many thanks are expressed to Dr. Lawrence M. Brammer, Dr. Edwin L. Klingelhofer, and Mr. Raymond W. Witt for their fine cooperation and guidance. It was gratefully appreciated.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. THE PROBLEM AND DEFINITIONS OF TERMS USED</td>
<td>1</td>
</tr>
<tr>
<td>The problem</td>
<td>2</td>
</tr>
<tr>
<td>Statement of the problem</td>
<td>2</td>
</tr>
<tr>
<td>Importance of the study</td>
<td>2</td>
</tr>
<tr>
<td>Definitions of terms used</td>
<td>3</td>
</tr>
<tr>
<td>Projection</td>
<td>3</td>
</tr>
<tr>
<td>Projective technique</td>
<td>3</td>
</tr>
<tr>
<td>Impediment or speech defect</td>
<td>4</td>
</tr>
<tr>
<td>Organization of remainder of the thesis</td>
<td>4</td>
</tr>
<tr>
<td>II. REVIEW OF THE LITERATURE</td>
<td>5</td>
</tr>
<tr>
<td>Projective techniques</td>
<td>5</td>
</tr>
<tr>
<td>Limitations of previous studies</td>
<td>13</td>
</tr>
<tr>
<td>Trends of psychodiagnostic</td>
<td>14</td>
</tr>
<tr>
<td>Origin and history of the Machover</td>
<td>16</td>
</tr>
<tr>
<td>Theoretical framework</td>
<td>17</td>
</tr>
<tr>
<td>Administration</td>
<td>17</td>
</tr>
<tr>
<td>Principles of interpretation</td>
<td>18</td>
</tr>
<tr>
<td>III. DESIGN OF THE EXPERIMENT</td>
<td>20</td>
</tr>
<tr>
<td>IV. JUDGES, MATERIALS USED, AND PROCEDURE</td>
<td>22</td>
</tr>
<tr>
<td>Judges</td>
<td>22</td>
</tr>
<tr>
<td>Test materials</td>
<td>22</td>
</tr>
<tr>
<td>Procedure</td>
<td>23</td>
</tr>
<tr>
<td>V. RESULTS OF THE GROUP EXPERIMENTS</td>
<td>24</td>
</tr>
<tr>
<td>Implications for further study</td>
<td>26</td>
</tr>
<tr>
<td>CHARTER</td>
<td>PAGE</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>VI. SUMMARY AND CONCLUSIONS</td>
<td>31</td>
</tr>
<tr>
<td>Summary</td>
<td>31</td>
</tr>
<tr>
<td>Conclusions</td>
<td>32</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>34</td>
</tr>
</tbody>
</table>
LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Frequency Count of Classification by Each Rater</td>
<td>29</td>
</tr>
<tr>
<td>II. Chi-square Values Obtained for Various Classifications of Disabilities, Individual Raters, and Ratings Rendered</td>
<td>30</td>
</tr>
</tbody>
</table>
CHAPTER I

THE PROBLEM AND DEFINITIONS OF TERMS USED

Projection and its underlying concepts have been known since the time of Freud. In 1939 K. K. Frank coined the term:

Projective method—a projective method for the study of personality involves the presentation of a stimulus situation designed or chosen because it will mean to the subject not what the experimenter has arbitrarily decided it ought to mean (as in most psychological experiments using standardized stimuli in order to be "objective," but rather whatever it must mean to the personality who gives it, or imposes upon it, his private idiosyncratic meaning and organization.\(^1\)

Since this time enthusiasm has been high in this area of projective tools. Especially following World War II, multifarious tests were born—a great number, unfortunately, of highly questionable theoretical parentage. Journals have included hundreds of these tests which could offer little or no validation data—empirical or experimental. However, since there has been more attention recently given to the importance of experimental validation—in viewing both new and old tests with a highly critical eye—fewer novelties have appeared on the market. The area of projective techniques seems to be headed in the direction of consolidation.

To date, clinicians are using projective techniques in analysis, prediction, and treatment. One can readily

understand that due to advantageous qualities of ease, brevity, and simplicity, these tests can hold an enviable position in practice. The results of these tests, however, must necessarily depend upon the wisdom, intelligent experience, and sensitivity of interpreters. If these techniques are to become valid and useful instruments, both their implicit and explicit hypotheses will have to be tested in controlled experiments.

I. THE PROBLEM

Statement of the problem. Using the Machover Drawing-of-the-Human-Figure Technique, a group of speech-defective children will not show differences in drawings of the oral area to a significant degree as compared with a control group of non-speech-defective children.

Importance of the study. Machover, in her discussion of drawing the human figure as a projective technique, offers little by way of experimental validity. She states that her principles of interpretation have received empirical justification and verification. She refers to some case files with drawings and clinical notes, and mentions some "blind matching" studies which came out "much better than chance."\(^2\) The test is inexpensive, simple to administer, takes little time, and offers hypothetically much information. For expediency

\(^2\)Karen Machover, Personality Projection in the Drawing of the Human Figure (Springfield, Illinois: Charles C. Thomas, Publisher, 1949), p. 25.
this technique holds much attraction. One who is inexperi-
enced in projective theory and in the knowledge of projective
test limitations might very well do a great deal of harm, in
attempting to work with the suggested inferences, both to his
patients and to the field of psychology in general through
consequent adverse public opinion. To combat misuse, the
present study is one of many needed to test Machover's numerous
hypotheses.

II. DEFINITIONS OF TERMS USED

**Projection.** Schachtel describes projection as that
psychic mechanism by which one attributes qualities, feelings,
attitudes, and strivings of his own to objects (people or
things) of his environment.  

**Projective technique.** The projective technique consists
of a number of ambiguous stimuli which are presented to a
subject who is then invited to respond to these stimuli. By
such means it is assumed that the subject projects his own
needs and press and that these will appear as responses to
the ambiguous stimuli.

---

3E. G. Schachtel, "Projection and its Relation to
Character Attitudes and Creativity in the Kinesthetic Respon-
ses: Contributions to an Understanding of the Rorschach

4Lawrence Edwin Abt and Leopold Bellak, Projective
Impediment or speech defect. Van Riper says that "speech is defective when it deviates so far from the speech of other people that it calls attention to itself, interferes with communication, or causes its possessor to be maladjusted."5

III. ORGANIZATION OF REMAINDER OF THE THESIS

The remainder of the thesis will be divided into five major parts: (1) review of the literature; (2) design of the experiment; (3) judges, materials used, and groups studied; (4) results of the group experiments; and (5) summary and conclusions. The review of literature will attempt to bring together information available at the present time in this area of projective techniques in general and figure drawing specifically. It will serve as a take-off point for this study and for future studies. Design of the experiment will tell how the study was set up. Judges, materials used, and groups studied will tell how the judges were selected, delineate the technique used in the present study, and define the groups. Results of the group experiments will show whether or not the null hypothesis taken at the onset of the experiment is rejected or retained for the groups studied. Consistency among the judges will be noted; implications for further study will be mentioned. Summary and conclusions will envelop the significance of the study and will point in the direction of needs for future investigation.

CHAPTER II

REVIEW OF THE LITERATURE

**Projective techniques.** Bellak puts the developing projective techniques into five categories of study: (1) content—(what the subject says): TAT, MAPS, Rorschach Inquiry, finger painting; (2) expressive and structural data—(how the subject says or does something): Mira, Mosaic, Rorschach, and Graphology—subsemantic levels of myoneural functioning; (3) gestalt formation: Bender-Gestalt, Mosaic, Rorschach, TAT—(rarely to any extent); (4) body image: figure drawing, Rorschach, TAT; (5) a study of choices or preferences: Szondi, color choice in finger painting, selection of figures in doll play—MAPS.  

Bell applies the most common meaning to "projection" stemming from the Latin roots to the use of the word:

In this sense it means "to cast forward," which is the action involved in the techniques. The subject manifests his personality in them by "thrusting it out" where it may be inspected. In the "throwing" the personality is not grossly modified; it is only externalized in behavior that is typical of the individual.

Hunt states that:

Freedom for idiosyncratic and idiomatic expression allowed the subject by projective techniques, in

---


contrast with objective inventory-style personality tests, appears to be a major advantage . . . personality in action . . . integrated and combined to provide the clinician with an understanding of the manner in which a patient's needs, habits, values, and the like articulate to produce his total personality. There seems to be general agreement that projective test results are rich in the sense of providing many fruitful hypotheses for diagnostic formulation, therapy and personality research. Agreement also seems general that projective devices provide more in these areas than do objective tests, though this point is carried largely by implication. 6

In reference to the problem of psychodiagnosis, Magaret states that:

Diagnostic investigation presumably begins with a question regarding the patient and ends with a more or less precise answer which advances one's understanding of the patient as a person. Direction which diagnostic investigation takes depends largely upon the comprehensiveness of the clinician's goal . . . the preferred theory may require the measurement of individual traits, the specification of needs, or the description of conflict among so-called "levels" of personality organization. There is currently wide diversity among psychologists in both the goals they accept for diagnosis and the theories they prefer. A diagnostician does not share with others and may not even specify to himself the view of personality organization which directs his study . . . this points up the importance of the subject's training, set, or pervasive attitude in determining selectively his perceptions. 9

In relation to goals of diagnosis, Miller shows a trend away from Kraepelinian and Bleulerian categories in the

---


direction of dynamic descriptions in terms of flux or process which are gradually replacing psychiatric categories.\textsuperscript{10} Grayson and Tolman show plainly the extent of disagreement and consequent misunderstanding among psychologists and psychiatrists in the use of professional terminology in report writing.\textsuperscript{11} Harris presents the consequences of the gradual shift in diagnostic emphasis. He contrasts limited "criterion group" method with a broader approach. He shows that an ego strength interpretation is possible in showing the results of research findings.\textsuperscript{12} Magaret states that:

\begin{quote}
It is just as likely that some critical thinking about the goals of diagnosis would lead to important developments in the prediction or individual behavior, in the appraisal of patients' treatability, and in the understanding of normal and deviant personality organization . . . in a good many situations, diagnosis as we traditionally conceive of it is unnecessary.\textsuperscript{13}
\end{quote}

Kelly comments that "Clinicians and personality psychologists, while more creative with respect to noval hunches for test construction seem to have been remarkably

\begin{flushleft}


\textsuperscript{13} Magaret, op. cit., p. 313.
\end{flushleft}
negligent in the utilization of rigorous procedures to separate good hunches from bad.\textsuperscript{14} This brings up a present issue concerning the degree to which the clinician is an integral part of the test. Kelly states further that:

One cannot properly speak of the validity of a technique, but only of a "technique-user combination." It would follow that for any given assessment situation, there should be a prior determination of the person-technique-criterion validity of each staff member . . . to the degree the human enters into the assessment process itself, he is merely an added source of error variance.\textsuperscript{15}

Haufmann pleads for objectivity of projective techniques. She believes that the essence of projective psychology lies in "what the psychologist does with the data obtained and in the nature and goal of treatment he accords them."\textsuperscript{16}

Harrower lists inadequacies on the part of the clinician as follows: (1) over-experience in specific areas--dealing with a particular population, one may have a distorted picture when dealing with a general population; (2) exaggerated zeal for one particular test--without sufficient validation, any particular test might be discarded leaving the clinician out on a limb; (3) insufficient attention paid


\textsuperscript{15}Ibid., p. 290.

to discrepancies between the findings on certain of the projective methods—possible false assumptions never specifically tested; (4) insufficient preparatory study by clinicians for tasks undertaken—either insufficient knowledge in the area or for a particular task; (5) lack of empirical knowledge—test records may not describe actual functioning of an individual; (6) lack of individual environmental knowledge—test record as related to actual psychological environment; (7) lack of longitudinal insight—value of longitudinal versus cross-sectional sampling.\(^\text{17}\)

With these comments in mind we will advance to studies in the area of projective techniques which are particularly relevant to the present study.

Stonsifer, using Goodenough's standards,\(^\text{18}\) scored drawings-of-a-man of 39 schizophrenics and 39 non-psychotic adults roughly matched for age and education. Reliability of scoring was high; however, there were no significant differences between the means.\(^\text{19}\)

Royal received better results when he attempted to get


objective characteristics of drawings-of-a-man and of-a-woman which would differentiate neurotics and non-neurotics. He used diagnosed anxiety state or anxiety neurosis versus non-hospitalized dental clinic volunteers. After cross validation, the original 28 scores were reduced to eight. Using a cutting score, he was able to identify 47 per cent of neurotics while misclassifying only about six per cent of the normals.20

Albee and Hamlin selected ten cases from a V. A. mental hygiene clinic files ranging from normal adjustment to hallucinated delusional psychotic. Fifteen clinical psychologists were asked to make paired comparison judgments of a patient's adjustment, basing their opinions on the patient's drawings-of-a-man and drawings-of-a-woman. Reliability between two groups of judges was .96, and the validity was .64.21

Fiedler and Siegel studied drawings-of-a-man of 15 improved and 19 unimproved psychoneurotics who had received therapy at a V. A. mental hygiene clinic. According to Goodenough's criteria of scoring the drawing of the head, they were able to separate the groups at the .01 level of confidence. Twelve out of 19 unimproved patients could be


detected by the use of one cutting score which included only
one improved patient. A lower cutting score predicted the
failure to improve on the part of ten patients and included
none of the patients who improved.22

Fisher and Fisher tested two of Machover's assumptions
concerning paranoid attitudes and "feeling" of expression and
stance. Judges failed to find any sizeable number of para-
noid indicators in the patients' drawings and could not agree
in their evaluations of expression and stance.23

Holzberg and Wexler found that 27 out of 124 scoring
items distinguished, at the .05 confidence level, the female
figure drawings of schizophrenic women from those of student
nurses.24

Weider and Noller found differences in the sex charac-
teristics drawn by children representing different socio-
economic levels.25 They also found that 61 per cent of
younger as compared with 48 per cent of older children use

22 F. E. Fiedler and S. M. Siegel, "The Free Drawing
Test as a Predictor of Non-improvement in Psychotherapy,"

23 S. Fisher and R. Fisher, "Test of Certain Assumptions
Regarding Figure Drawing Analysis," Journal of Abnormal and

24 J. D. Holzberg and M. Wexler, "The Validity of Human
Form Drawings as a Measure of Personality Deviation," Journal

25 A. Weider and P. A. Noller, "Objective Studies of
Children's Drawings of Human Figures: I. Sex Awareness and
Socio-economic Level," Journal of Clinical Psychology,
the upper left quadrant of the paper.  

Ochs offers evidence that Goodenough drawings change with alterations in the social adjustment of children suffering from primary behavior disorders. An unimproved group of child patients showed a decline in Goodenough I.Q., while an improving group increased in I.Q.  

Whitmyre found that adjustment as rated from drawings of the human figure did not discriminate normals from psychiatric patients. Adjustment ratings were highly correlated with artistic excellence.  

Blum produced no consistent agreement between ratings from the drawings and those made by the psychiatrists and others. "Machover DAP technique has highly questionable validity, but proves to be no worse than any of the other common clinical personality assessment procedures."  

Hammer and Piotrowski found a very high relationship


between the rated hostility of the clinician and the amount of hostility he inferred from the drawing-test protocols.\footnote{E. F. Hamer and Z. A. Piotrowski, "Hostility as a Factor in the Clinician's Personality as it Affects his Interpretation of Projective Drawings (HTP)," \textit{Journal of Projective Techniques}, 17:210-216, 1953.}

\textbf{Limitations of previous studies.} Challman in 1951 proposed three weaknesses in validation studies: (1) criteria for validation—psychiatric diagnoses are unreliable and often do not furnish a sufficiently homogeneous group in any one category for validational purposes. There is no clear lead from personality theory as to which characteristics or traits to select, if these are to be used; (2) lack of suitable statistical methods—present ones are inadequate with the results of projective techniques and for the analyses of patterns; (3) inadequate use of available knowledge of experimental design and statistical methods.\footnote{Robert C. Challman, "Clinical Methods: Psychodiagnosics," \textit{Annual Review of Psychology}, Volume II, Calvin Stone, Editor (Stanford, California: Annual Reviews, Inc., 1951), p. 255.}

Magaret in 1952 stated that:

Too many investigations are haphazard, isolated, or fragmentary. Too many stop where they should begin; too many seem to appear in our journals like rejected orphans, with no meaningful past and no promising future. Clinicians have three unfulfilled needs: (1) critical thinking about philosophy and goals of diagnosis—"diagnosis for what?" more meaningful, more dynamically oriented systems of classifying patients can and should be developed; (2) diagnostic instruments should be constructed and employed in terms of explicit goals and philosophies—one instrument which will turn out to be of greatest importance is the clinician himself; he
cannot be excluded from any diagnostic situation—even "blind" analyses; and (3) more comprehensive, planned research programs in diagnosis, each carefully oriented toward an important target—program orientation versus novel problem orientation.32

Windle adds that "Many of the studies failed to demonstrate empirical justification for the conclusions drawn. Further, comparisons among studies revealed little agreement among findings." He emphasizes a need for better research studies. He states that many studies do not contribute significantly for the following reasons: (1) failure to specify the conditions of the experiment—e.g., description of patient population, conditions of therapy, and criteria of outcome; (2) failure to deal with homogeneous populations and conditions; (3) failure to report their results in terms amenable to statistical evaluation; and (4) failure to subject reported findings to cross-validation.33

**Trends in psychodiagnosis.** Rotter in 1953 indicated six trends in psychodiagnosis: (1) in general, when "global" tests of personality are put to a specific predictive test, the results are poorer than those of tests devised for specific purposes, even if these tests have little but face validity; (2) when psychiatric diagnostic entities are used as criterion or predictor, experimental results seem to lack prediction

32Magaret, op. cit., p. 313.

except for purposes of gross screening; (3) several studies are concerned with using the test data to determine the subject's approach to the situation or generalized set, in contrast to attempts to use specific responses as indicators of traits. This trend is apparent in both projective and objective tests; (4) in general, prediction is poor for all kinds of instruments when put to careful experimental tests without other information regarding the subject. Findings indicate that a single test is rarely relied on for individual clinical prediction; (5) negative studies have resulted in fewer new unvalidated tests. Principles of cross-validation, elimination of bias, use of controls are relatively well accepted. The importance of the examiner and situational variables is shown in a number of studies; (6) research in personality theory has resulted in tests which measure a specific theoretical construct, such as rigidity, hostility, tolerance for ambiguity, level of aspiration and humor.34

Butler and Fiske in 1955 note that (1) studies pertinent to assessment are becoming restricted for the most part to particular criteria, specific instruments, and even to separate derived scores; (2) there is a continued and stronger trend in the direction of testing clinical hypotheses; (3) rather frequent omission of critical data and table often

makes evaluation of studies difficult; (4) insufficient systematic work on criterion behavior continues to be seen in much of the research. Psychometric properties of many tests, especially free-response tests, are far from being fully understood; (5) few investigations have been concerned with the really basic problem of the subjects' perception and definition of the testing situation.35

The foregoing should serve as ample justification for the fact that now the clinical hypotheses of Machover, among others, are being tested experimentally in controlled situations. At this time, there is little experimental validation to support the many "empirical truths" set forth.

**Origin and history of the Machover.** Human figure representation as an outlet for projection is relatively new to publication. As is often the case, theory followed practice: empirical validation for 15 years in clinics and hospitals preceded theoretical framework for the Machover technique. Machover credits the Goodenough technique for bringing her attention to differences in drawings, different comments with drawings. In routinely using the Goodenough, Machover noticed that children with the same I.Q. were producing quite varied drawings. She inquired about the drawings

and found that their comments were rich in body-image projection. This was the beginning of her draw-a-person technique.

**Theoretical framework.** It is assumed by Machover that all creative activity will reflect conflicts and needs of its creator. Machover's experience with human figure drawings shows a body-image involvement with the artist. She asserts that one draws from a known referent. One associates certain sensations, perceptions, and emotions with specific body organs. Through this personal experience, it is believed, comes the requested person in the drawings. Machover states that this graphic communication occurs regardless of age, skill, or culture.

Through selection by the subject of past experience and imagery--movement and perception--in organizing a "person," comes individualization which is amenable to personality study.

**Administration.** The administration of the test consists of simply asking the individual to "draw a person." The subject is given a medium-soft pencil with an eraser on it and a blank sheet of paper approximately 8½ by 11 inches. During the drawing the clinician observes as inconspicuously

---

36 Machover, op. cit., p. 5.

37 Ibid., p. 7.
as possible. When the figure is drawn, the subject is given another blank sheet of paper and asked to draw a person of the opposite sex with "Now draw a man," or "Now draw a woman." The subject is told, "This has nothing to do with your ability to draw. I am interested in how you try to make a person."

**Principles of interpretation.** Much of the interpretation is based on established projective methods of personality analysis and on psychoanalytical theory.\(^{38}\) To reiterate, the basic assumption is that an individual will graphically function in relation to his impulses, anxieties, conflicts, and compensations. When the individual erases his arms and changes the position of them several times, it may be literally interpreted that the subject does not know what to do with his arms in his behavior.\(^{39}\)

An adolescent, when asked why he drew pin-points for feet, promptly associated it with his own deep insecurity and inability to stand on his feet. Another subject omitted the arm from the mother-image figure, because, "She used to throw things," and the absence of the mouth being explained, "She used to curse." In still another instance, the omission of the arm is explained, "Because you might do harm with it," and a large head is drawn with the comment, "You can see what a brainy fellow he is."\(^{40}\)

The size of the figure, the placement, rapidity of graphic movement, pressure, solidity and variability of line

\(^{38}\)Ibid., p. 34.

\(^{39}\)Ibid., p. 35.

\(^{40}\)Ibid., p. 10.
used, succession of parts drawn, stance, use of background or grounding effects, relationship of arms to body and their direction, spontaneity or rigidity, profile or front view are all meaningful as clues to the subjects' personality. Proportions, incompletions, detailing, reinforcements, erasures, symmetry, treatment of midline, and mood are all to be considered.

Content--body details and clothing treatment--are interpreted in relation to the functional significance attached to it. These assumptions are presented as empirical truths.

41 Ibid., p. 36.
CHAPTER III

DESIGN OF THE EXPERIMENT

Cognizant of the salient need for tests of clinical hypotheses, the author proceeded to collect data for a partial test of validity of the Machover Draw-a-Human-Figure technique. It was decided that since the author works with speech-defective children, a rather objective experiment might be run in this area. If Machover's body-image hypotheses are valid, it would seem that as a group, speech-defective children would show more under- or over-elaboration of the oral area than a control group of non-speech-defective children, matched for age and approximate intelligence. One hundred speech defectives and one hundred non-speech defectives were chosen and "Machovers" were administered. The drawings were split and coded and put into a statistically random order for judging. Judges were chosen who use figure drawing in their professional assignments.

The instructions to the judges were as follows:

Here is my problem for you to read. Please throw the drawings into three categories—under-elaboration, normal, and over-elaboration in relation to the oral areas. Use your own criteria for judgment as to which of the three categories each drawing should fall.

Each judge was aware that this was a test of Machover theory and rated the drawings accordingly. The drawings were rated by each judge, separately, requiring from two to four hours
for the rating of the 400 figure drawings. For each judge a count of the ratings was taken according to their classifications—speech defective and non-speech defective by under-elaborated, normal, and over-elaborated oral areas—in a chi-square table. A chi-square test of significance was used to test the experimental hypothesis.
CHAPTER IV

JUDGES, MATERIALS USED, AND PROCEDURE

Judges. The judges were clinical psychologists in the Sacramento area who use human figure drawings in their analysis and therapy.42

Test materials. The Machover Draw-a-Human-Figure technique was used, the materials of which consisted of two blank sheets of paper approximately 8½ by 11 inches and a medium-soft pencil with an eraser. The subject was given one sheet of paper and a pencil and asked to "Draw a person." When this figure was completed, he was handed another sheet and asked to draw the opposite sex with "Now draw a girl (or woman)" or "Now draw a boy (or man)."

In the case of speech-defective children, the test was given previous to any speech work. They were told that "It's just an experiment. This has nothing to do with your ability to draw. I am interested in how you try to make a person."

42 The author welcomes this opportunity to express his gratitude to the judges: (a) Dr. Burton M. Castner, Diplomate in Clinical Psychology, Chief of the Division of Diagnosis and Classification, California Youth Authority; (b) Mrs. Lucille Colby, M.A. Candidate, Sacramento State College; (c) Mr. Irving B. Fecher, M.Sc., private practice; (d) Mr. Hans M. Kakies, M.A., Senior Clinical Psychologist, Northern Reception Center and Clinic, Perkins Youth Authority; (e) Mr. John Ohlson, M.A., Senior Clinical Psychologist, California State Mental Hygiene Clinic, Sacramento--for their excellent cooperation and patience in rating the 400 human figure drawings.
Procedure. Two groups were used. One group was composed of 100 diagnosed speech-defective subjects and the other was a control group matched by teachers to the speech-defective children according to age and intelligence factors. The speech-defective children were administered the Machover Draw-a-Human-Figure technique before speech therapy, while the control group were called out of their regular classes one at a time for the experiment. The population of both groups was made up of students from two elementary school districts and one high school district. Children ranged in age from 6 to 21 and in grades from 1 through 12.

The judges rated 200 drawings made by children classified as speech defective and 200 drawings made by the children equated for age and intelligence who served as a control group. The drawings were mingled in a statistically random fashion and the judges had no information as to the group source of any particular drawing. The judgings were done by each judge independently. He evaluated each of the 400 drawings and then classified it as under-elaborated, normal, or over-elaborated in relation to the oral area.

Florin Elementary, Elk Grove Elementary, and Elk Grove High School districts.
CHAPTER V

RESULTS OF THE GROUP EXPERIMENT

With five judges serving, there were a total of 2,000 judgments which were distributed as follows: 343 under-elaborated drawings, 916 normal, and 741 over-elaborated (Table I). This difference is significant at the .1 percent level of confidence as given by the chi-square analysis. The null hypothesis is, therefore, rejected. (See Table II--Ratings by Disability.)

It is noteworthy that, inversely, on 418 occasions the drawings of speech-defective children were judged to fall in the normal category, as compared to 498 times for the drawings of the control group. The implications of this for further investigation are discussed in a later portion of this writing.

Since the existence of an overall trend for the drawings of speech defectives to be judged more frequently as over- or under-elaborated was demonstrated, it seems fruitful to ascertain whether this trend was consistent for all raters individually. In view of the known variability of raters in general and the previously-mentioned need for considering the place of the administrator in the validity of the instrument, a consideration of the patterns of the individual judges seems relevant. Consequently, the frequency and type of rating made by each rater for each of
the groups separately were analyzed. With regard to the speech-disabled group, the individual judges did not differ significantly from each other in the frequencies of drawings in the separate categories. The chi-square value was calculated at $14.7$ which, with eight degrees of freedom, does not achieve significance at the $5\%$ level.

In appraising non-speech-defective drawings it was found that the raters tend to not agree on the frequencies of drawings assigned to the different categories. The chi-square value obtained for eight degrees of freedom is $16.8$, which is significant at between the $2$ and $5\%$ levels of confidence. It can be assumed, therefore, that the raters disagree significantly in appraisals made of non-speech-defective children. (See Table II--Raters by Ratings.)

In measuring the consistency of individual raters with regard to their ability to separate speech defective drawings from non-speech-defective ones, the chi-square values for raters $1$, $2$, $3$, and $5$ were $13.4$, $19.4$, $19.8$, and $10.7$, respectively. For two degrees of freedom, these values are found to be significant beyond the $.1\%$ per cent level of confidence. Four of the judges, then, may be said to have differentiated the groups to a significant degree. Rater $4$, however, had a chi-square value of $.4$. For two degrees of freedom, this value does not approach significance. (See Table II, Raters by Disabilities.)
It may be concluded, then, that a group of speech-defective children will distort the oral areas of their human figure drawings significantly more than a control group of normal children, as judged by experienced raters. Some support for the validity of a portion of the Machover assumptions is given. The utility of this evidence for individual diagnosis is still problematical since there is a far from clear-cut differentiation. Part of this lack of clarity may be explained on the grounds that: (1) oral problems exist among "normals" to an unspecified degree, and (2) with some speech-defective youngsters, the speech problems may be of secondary importance.

A further finding is that, in this experiment, judges did not disagree significantly in the frequencies with which they categorized the drawings of speech defectives as under-elaborated, normal, or over-elaborated. In general, it appears that experienced judges are consistently able to identify speech-disabled drawings to a significant degree.

Implications for further research. The design of this study, as originally devised, was intended to test some of the assumptions which Machover has presented and was, therefore, constructed according to the framework of her theory. The control group was equated for age, but no separate analysis for different age levels was undertaken. It is possible that the considerable number of adolescents
included may have contributed to the variability in the drawings and confused the picture for the judges. However, since the numbers in the two groups were equal, it should not destroy the presumed predominance of the speech defective projections.

Since Whitmyre reports that adjustment ratings are highly correlated with artistic excellence, it is possible that some efforts should have been made to control this variable of drawing skill. Machover specifically denies that training in art is a determining factor in adjustment rating, however, so no such delineation was made.

Weider and Noller found differences in sex characteristics drawn by children of different socio-economic levels. Socio-economic factors were not controlled in the present study. It seems possible that differences on such a basis may have existed and confused the ratings.

Since the placement of individual drawings was not recorded in the present study, the precise reliability of ratings needs further investigation.

The failure of one judge to differentiate the categories of speech defectives and non-speech defectives in rating the oral areas is evidence of the need to consider the user with the instrument. Further studies might be aimed at identifying the characteristics of the "valid" judge as opposed to the "invalid" one.
A further possibly fruitful area for investigation might be the study of the characteristics of those children with speech problems who seem to illustrate it in their drawings of oral areas, as opposed to the children with speech problems who do not do so. It seems possible that the latter group might be found to have the speech problems as a more secondary characteristic.
### TABLE I

**FREQUENCY COUNT OF CLASSIFICATION BY EACH RATER**

<table>
<thead>
<tr>
<th>Speech defect</th>
<th>Under-elaborated</th>
<th>Normal</th>
<th>Over-elaborated</th>
<th>Rater</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-speech defect</td>
<td>11</td>
<td>90</td>
<td>99</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>23</td>
<td>122</td>
<td>55</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>79</td>
<td>81</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>100</td>
<td>60</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>27</td>
<td>107</td>
<td>66</td>
<td>5</td>
</tr>
</tbody>
</table>

| Speech defect | 26               | 82     | 92              | 1     |
| Non-speech defect | 38               | 92     | 70              | 2     |
|                | 58               | 51     | 91              | 3     |
|                | 37               | 104    | 59              | 4     |
|                | 43               | 89     | 68              | 5     |
### TABLE II

**CHI-SQUARE VALUES OBTAINED FOR VARIOUS CLASSIFICATIONS OF DISABILITIES, INDIVIDUAL RATERS, AND RATINGS RENDERED**

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Degrees of Freedom</th>
<th>Chi-square</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ratings by disability</td>
<td>2</td>
<td>18.3</td>
<td>.0001</td>
</tr>
<tr>
<td>Raters by ratings</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speech defectives</td>
<td>8</td>
<td>14.7</td>
<td>.10, .05</td>
</tr>
<tr>
<td>Non-speech defectives</td>
<td>8</td>
<td>16.8</td>
<td>.05, .02</td>
</tr>
<tr>
<td>Raters by disabilities</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rater 1</td>
<td>2</td>
<td>13.4</td>
<td>.01, .001</td>
</tr>
<tr>
<td>Rater 2</td>
<td>2</td>
<td>19.1</td>
<td>.0001</td>
</tr>
<tr>
<td>Rater 3</td>
<td>2</td>
<td>19.3</td>
<td>.0001</td>
</tr>
<tr>
<td>Rater 4</td>
<td>2</td>
<td>.4</td>
<td>.8</td>
</tr>
<tr>
<td>Rater 5</td>
<td>2</td>
<td>10.7</td>
<td>.01, .001</td>
</tr>
</tbody>
</table>
CHAPTER VI

SUMMARY AND CONCLUSIONS

Summary. The question of this investigation was as follows: by using the Machover Drawing-of-the-Human Figure technique, will a group of speech-defective children show differences in drawings of the oral area to a significant degree as compared with a control group of non-speech-defective children?

The present study was designed in strictly Machover terms. No exacting account was taken of age, skill, or culture, since Machover states that graphic communication takes place regardless of these factors. The control group was equated for age; however, no separate analysis for different age levels was made. The possibility arises that the considerable number of adolescents included in the study may have confused the picture for the judges. Nevertheless, since the numbers of adolescents in the two groups were equal, this was assumed not to be a disturbing factor for the experimental test.

One hundred children in each group were chosen and administered the Machover test. The drawings were coded and put into statistically random order for judging. Judges were given their instructions. Drawings were then rated, the ratings were recorded in chi-square tables, and chi-square tests of significance were computed.
Results of the group experiment showed 2,000 judgments to be distributed as follows: 343 under-elaborated drawings, 916 normal, and 741 over-elaborated. Speech defectives' drawings showed over- or under-elaboration of the oral area 582 times. Normal children showed over- or under-elaboration of the oral areas 502 times. The difference is significant at the .1 per cent level of confidence as given by chi-square analysis. The null hypothesis was, therefore, rejected.

A check was made to see if this trend for the drawings of speech defectives to be judged more frequently as over- or under-elaborated was consistent for all raters individually. The frequency and type of rating made by each rater for each of the groups separately was analyzed. Individual judges did not differ significantly from each other in the frequencies of drawings in the separate categories with regard to the speech-defective group. With regard to the non-speech-defective drawings, it was found that there was a significant disagreement in appraisals made.

In measuring the consistency of individual raters with regard to their ability to separate speech-defective drawings from non-speech-defective ones, it was found that four of the judges were able to differentiate the groups to a significant degree. One, however, did not differentiate them.

Conclusions. A group of speech-defective children will distort the oral areas of their human figure drawings significantly more than a control group of normal children, as judged
by experienced raters. Some support for the validity of a portion of the Machover assumptions is given. The utility of this evidence for individual diagnosis is still highly questionable and indicates a need for further investigation, since there was no clear-cut differentiation in this study. Part of the lack of clarity may be due to (1) oral problems which exist among "normals" to an unspecified degree, and (2) secondary importance of speech problems with some speech-defective children.

It was found that judges did not disagree significantly in the frequencies with which they categorized the drawings of speech defectives as under-elaborated, normal, or over-elaborated.

The failure of one judge to differentiate the drawings points up the need for considering the user with the technique. Further studies might be aimed at identifying the characteristics of the "valid" versus the "invalid" judge.

In future research it seems necessary to consider evidence of age, drawing skill, and socio-economic status as pertinent variables. Also, on the basis of the present study, further research on speech-defective children to determine personality differences between those who show concern in oral portions of their human figure drawings and those who do not seems relevant.
BIBLIOGRAPHY
A. BOOKS


Bell, J. E., Projective Techniques; A Dynamic Approach to the Study of Personality. New York: Longmans and Green, 1948.


Machover, Karen, Personality Projection in the Drawing of the Human Figure. Springfield, Ill.: Charles C. Thomas, 1949.


B. PERIODICAL ARTICLES


Hammer, E. F., and Z. A. Piotrowski, "Hostility as a Factor in the Clinician's Personality as it Affects his Interpretation of Projective Drawings (HTP)," Journal of Projective Techniques, 17:210-216, 1953.


