THE DRAW-A-PERSON TECHNIQUE AS A MEASURE
OF SEXUAL CONFLICT

APPROVED:

Merl E. Bonney
Major Professor

__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\__\_
THE DRAW-A-PERSON TECHNIQUE AS A MEASURE OF SEXUAL CONFLICT

THESIS

Presented to the Graduate Council of the North Texas State University in Partial Fulfillment of the Requirements

For the Degree of

MASTER OF SCIENCE

By

Arthur Courtney Caspary, B.S.
Denton, Texas
January, 1968
# TABLE OF CONTENTS

**LIST OF TABLES** ........................................ iv

**Chapter**

I. **PRESENTATION OF THE PROBLEM** ....................... 1
   - Introduction
   - Statement of the Problem
   - Background Knowledge and Theory
   - Hypothesis and Assumptions
   - Definitions and Procedure

II. **RELATED RESEARCH** ................................. 8

III. **METHODOLOGY** .................................... 22
   - Subjects
   - Description of Instruments and Procedure

IV. **RESULTS AND DISCUSSION** ......................... 26

V. **SUMMARY** ........................................ 32

**BIBLIOGRAPHY** ......................................... 35
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Number and Percentage of Agreement Between Mean Psychologist Ratings and Psychiatrists' Ratings</td>
<td>27</td>
</tr>
<tr>
<td>II. Number and Percentage of Agreement Between Scorer Level Means</td>
<td>30</td>
</tr>
</tbody>
</table>
CHAPTER I

PRESENTATION OF THE PROBLEM

Introduction

The drawn human figure has long been used to determine personality characteristics. In 1928 Goodenough (1) published instructions for the interpretation of a drawn human figure as a measure of intelligence in children. Machover (3), in 1949, suggested the use of the method as a means of personality assessment and since that time the Draw-a-Person technique has gradually gained in popularity until today, some authors consider it to be the second most widely used projective clinical instrument in the country (5).

The use of this technique is predicated on the assumption that an individual's perceptions are deeply rooted in his physical and psychological experiences and, when the unstructured stimulus of "draw a person" is presented, these experiences are integrated and interpreted in the resultant production. This drawing represents some major aspects of the individual's self-image, including his feelings and attitudes toward the opposite sex (1, p. 10).

The nonverbal nature of the task theoretically reduces the possibility of evasion since, culturally, Western peoples are conditioned to the use of verbal symbols (7, p. 1). Its
main benefits seem to be in data yielded concerning self and opposite sex concepts.

The use of the Draw-a-Person technique as a method of making clinical inferences about personality variables has come in for a good deal of criticism from those who point to the lack of empirical evidence backing up its validity (6, p. 609); however, considering its wide acceptance by clinicians and that many studies do support its validity, further research would seem to be warranted.

Statement of the Problem

Since much of the existing research which points to the low validity of the technique has used the differentiation of diagnostic categories as a criterion, this study will be restricted to a consideration of the Draw-a-Person technique as a means for discriminating between a group of individuals who manifest conflicts in the area of sexual adjustment and a control group which does not show these conflicts in any appreciable amount.

The rationale for the choice of sexual conflicts as opposed to other areas of maladjustment is that sexuality is closely related to total physical functioning and hence is probably more than most other aspects of personality closely tied to body concept. It should therefore be most ably demonstrated by the use of the D-A-P technique. Further, many workers in the field state that the technique should
not be used except within the context of an entire battery of clinical tests. Perhaps the focus of research in this area should be narrowed so as to provide an instrument with increased validity in specific areas of personality and used for ascertaining the dynamics of behavior rather than for gross differentiations among diagnostic categories.

The problem, then, is to investigate the D-A-P's ability to yield data in a rather specific area of conflict, the area of sexual adjustment, and to determine if clinicians of varying levels of experience with the technique can use these data to differentiate individuals having sexual conflicts from those who do not.

**Background Knowledge and Theory**

Goodenough, in her early studies of the use of drawings of the human figure as a measure of intellectual development, was aware of the possibilities of discerning personality variables in the productions of the children she observed (2, p. 343). As the Goodenough technique came into more widespread use, other investigators began to notice what seemed to be systematic deviations in the drawings of maladjusted children, and began to suspect that predictions could be made on the basis of these deviations.

Machover, noting that although the intellectual development of two children may be the same, their drawings may be quite different in any number of ways, stated that
Children literally grasped this opportunity to project any of their anxieties, feelings and aspirations upon the figure drawings. It was made obvious to me that the body image was a natural figure for such projections (2, p. 344).

Her publication is credited with starting the current trend in using the drawn human figure as a projective technique.

Concerning the idea of "projection" of the personality, Ruch states

The essence of projective methods is presenting the subject with stimulus situations which are "ambiguous" or neutral—that is, they have no particular meaning except the special, private meaning which the individual, by virtue of his own needs, projects into it (4).

Since sexuality is so closely tied to body function, it could be assumed that the psychological component of sex would be most readily expressed when the stimulus is draw a person and then to draw a person of the opposite sex.

Machover adds

We have in the course of growth, come to associate various sensations, perceptions, and emotions with certain body organs. This investment in body organ, or the perception of the body image as it has developed out of personal experience, must somehow guide the individual who is drawing in the specific structure and content of a "person." Consequently, the drawing of a person, in involving a projection of the body image, provides a natural vehicle for the expression of one's body needs and conflicts (3, p. 5).

Hypothesis and Assumptions

It was hypothesized that the individual, when presented with the stimulus of "draw a person" and then "draw a person of the opposite sex" will, in his productions reveal data
which can be validly interpreted so as to determine degree of conflicts in the sexual sphere.

It was assumed that the individual's personality structure is affected by his experiences and that each response to new stimuli is affected to some degree by previous responses to old stimuli. Thus, it is assumed that sexual attitudes and conflicts, both conscious and unconscious, will be expressed in response to the unstructured stimulus of "draw a person." It is acknowledged that the technique requires at least a minimum of conscious cooperation on the part of the subject although it is assumed that the information divulged in his productions may include that which he is not aware of or does not wish to reveal.

Definitions and Procedure

Sexual conflicts were operationally defined as

(a) Any report of sexual conflicts made by the individual when a social history was taken prior to a psychological testing.

(b) Estimations of sexual conflict along a five point rating scale made by the psychiatrist or psychologist who treated the subject in psychotherapy.

Sexual conflicts were considered to be any adjustments in the area of psychosexuality which, either consciously or unconsciously, have resulted in maladaptive responses on the part of the individual. The individual may or may not be aware of these problem areas.
The drawings were interpreted by scorers at three levels of sophistication with the Draw-a-Person technique. The first level of sophistication was represented by two Ph.D. level professional psychologists with at least five years experience in using the technique in a clinical setting. The second level was represented by two master's level psychologists who have had at least two years experience with the technique in a clinical setting, and the third level consisted of two individuals who were in a psychology internship and have used the technique in a clinical setting more than once a week for a period of three months. Each scorer was presented with a typewritten copy of the above definition of sexual conflicts and was asked to rate each of the drawings for the existence of "sexual conflicts" and to quantitatively rate the degree of conflict indicated on a five point rating scale.

Other problems considered in this study were the relationship of scorer experience to the validity of the technique and interjudge reliability.
CHAPTER BIBLIOGRAPHY


CHAPTER II

RELATED RESEARCH

The area of drawing analysis has been considered important enough for study and research since the nineteenth century (23). The types of techniques used for the projective investigation of personality are quite varied, including such diverse techniques as the Bender Visual-Motor Gestalt test which is used for rough screening of organicity as well as functional disorders, and the use of finger painting both as a diagnostic measure and therapeutic technique.

The Draw-a-Person technique is perhaps the most widely used of all the graphic projective techniques (24), yet a survey of the literature shows that there is considerable disagreement about its validity.

In 1952, Lehner and Gunderson (15) administered Draw-a-Person tests to ninety-one college students and attempted to determine the consistency of ratings on twenty-one graphic traits. The traits used were shading, reinforcement, erasures, detail, body type, hands, eyes, mouth, nose, hair, feet, breast line, body details, transparency, extraneous drawing, position on page, perspective, symmetry, position of hands, posture, and lines. Each trait had ten descriptive categories such as 0, none of the following apply;
1, no shading; 2, hair shaded; 3, parts of face shaded; and so forth. The findings indicate that a relatively high percentage of agreement in evaluating the indices may be achieved when an objective scoring system is used.

In an important study concerned with global versus insightful impressions and their respective validity and reliability, Albee and Hamlin (1) found that global judgments were reliable but that validity was questionable. In this study experienced clinicians used any means they wished to discriminate different levels of psychosexual maturity, good interpersonal relationships, and cultural-syntonic perception of reality. Interjudge reliability was .977 and validity of judgments was .64. The coefficients were significant at the .01 and .05 levels of confidence, respectively.

The validity was questionable due to the small N and the assumption of a unidimensional continuum of adjustment resulting from obtaining the mean scale values for each drawing. Consequently the same investigators (2) further studied the use of a criterion scale when evaluating adjustment and found that such a scale was effective in discriminating a group of normals from groups of neurotics and schizophrenics. They concluded that inferences about adjustment drawn from D-A-P production, "... possess a reasonable degree of validity" (2, p. 365).

In a study that differed in results from Albee and Hamlin, Whitmyre (28) tried to evaluate overall adjustment
from figure drawings. In this study, clinicians used drawings to rate overall adjustment for twenty-five psychiatric patients and twenty-five normal males. The correlation, while in the appropriate direction, was not significant. Also, the study found that artistic ability tended to correlate with judgments of adjustment.

Fisher (8) tested eleven of Machover's indices thought to be able to differentiate the schizophrenic from the normal but found no significant difference. Later, Fisher and Fisher (9) used both the sign and global approach but still found the technique unable to differentiate between diagnostic groups.

While some investigators such as Holzberg and Wexler (13) found that the D-A-P could validly differentiate between diagnostic groups, the majority of the literature indicated that the technique was not effective as a diagnostic instrument (25).

With evidence mounting that the D-A-P was ineffectual in discriminating between gross diagnostic categories, research turned more and more to limited uses of the test. Berman, Klein and Lippman (4) used the sign approach to gain a general impression of the personality structure and were moderately successful.

Royal (21), however, found that none of twenty-eight signs discriminated anxiety neurotic Ss from controls. He did note that eight signs, taken as a whole, did discriminate.
These characteristics were more than two inches between figures drawn on the same page, absence of heavy lines, tendency to rectangular or circular head and trunk, obvious erasures, a tendency for body and clothing lines to be continuous, one figure placement higher than the other, and a tendency for one figure to be longer than the other.

In checking a number of homosexual discrimination indices which had been referred to in psychiatric literature, Barker, Mathis and Powers (3) found that many of the indices would not discriminate a group of fifty male homosexuals in the Army from a control group of thirty-five non-homosexuals. In analyzing form and content it was found that only two areas allowed significant differentiation between the groups: (1) delay in identification of the self-sex figure which occurred in thirty-nine of the homosexuals' records, while it occurred in only six of the control group records; (2) distortion of the female figure was seen in thirty-eight of the homosexual records and two of the control group records. Twenty-three of the thirty-eight homosexuals who distorted the female figure gave spontaneous hostile responses.

Fisher and Fisher (10) found a slightly significant relationship between the degree of femininity of figure drawings and style of past behavior in a group of hospitalized female mental patients. The more maladjusted of the subjects tended to draw figures with either very marked femininity or very little femininity. A close relationship
between reproductive system problems in the subjects and low femininity in the drawings was noted.

Reed (20), in comparing fifty psychotic women with fifty normals, found that a greater number of masculine body-image attributes were drawn by the psychotics.

Witkin (29), on the other hand, found a high relationship between personality characteristics and attitudes towards the body in male subjects. He concluded that men are more body-conscious while women are more concerned with relationships with the outside world.

Berman and Laffal (5) found that a group of male psychiatric patients showed a significant correlation between the patient's actual body type and the body type of the drawn figure. This, of course, tends to support the thesis that projection of one's own body does take place in figure drawing.

A number of studies, however, show findings that weaken the Berman and Laffal findings. Woods and Cook (30) found that the manner of representing the hands was significantly a function of proficiency in drawing in a group of 138 eighth-graders, thus limiting the degree to which a figure drawing can be considered a projection of the body image.

Using a group of twenty-two children who had poliomyelitis and who manifested paralytic residuals of the lower extremities, Silverstein and Robinson (22) obtained same-sex, opposite-sex, and self drawings. It was found that judges
were unable to distinguish this group from forty-four normal children using either the sign approach or global judgments.

Blum (4) compared a battery of tests, psychiatrists' ratings, and ward supervisors' ratings with the D-A-P technique and found no significant relationship.

In spite of much negative evidence about the technique, some studies indicate that it is valuable in discerning specific personality traits. Griffith and Peyman (12) drew 18 subjects who had obvious eye-ear emphasis on their drawn figures from a hospital population of 745 males. An 8 percent control sample was then drawn and hospital files, exclusive of psychometric data, were reviewed by two independent psychologists who made judgments as to the presence or absence of ideas of reference. Results indicate a significant difference between the two groups ($X^2$, 7.754 for 1 df; $P < .01$). It was noted that if diagnostic categories had been used, as was the case with Swensen's study, the results would have been statistically insignificant.

Hoyt and Baron (14), in investigating the sign approach to scoring, gave 112 females the Taylor Manifest Anxiety Scale and then the D-A-P technique. The upper and lower 27 percent of the anxiety scores was used to determine the high and low anxious groups. Then the drawings of these 60 subjects were scored for the presence or absence of the following anxiety indices: placement, type of line, reinforcement, shading, erasing, size, size of head and
relative size of head to figure, omissions and body areas out of proportion. Each index received a score value of from one to three as rated by two psychologists. $X^2$ values indicated that only two indices differentiated between groups: placement in the upper left hand corner ($P < .05$) and small size ($P < .05$).

Hiller and Nesvig (12) obtained D-A-P's from thirty adolescents from a state hospital and from thirty "successful" adolescents. Six psychologists were asked to discriminate between the normals and the pathological group. They then were asked what criteria they had used. It was found that the criteria that discriminated beyond the 1 per cent level of confidence were definitely bizarre, incomplete, and for normals, nothing pathological, and happy, pleasant expression. Many inaccurate criteria were used and it was determined that if only empirically derived criteria were used the prediction rate improved from 65 per cent accurate to 79 per cent accurate.

Another approach was taken by Lewinsohn (16), who examined the psychological correlates of overall quality of the drawing. Drawings of seventy-five female and sixty-two male psychiatric patients were rated for OVQ on a nine point scale. Relationships were established between OVQ ratings and psychological test data, interview and behavioral ratings and psychological test data, interview and behavioral ratings, and information regarding social background and adjustment.
The results indicated that OVQ was not related to specific aspects of psychopathology, improvement in clinical condition, nor to a wide range of personality trait ratings. Low, but statistically significant relationships were found between OVQ and three independently obtained indices of adjustment in specific situations. It was also related to verbal IQ and father's occupational level. There were several significant and interesting correlations found however. OVQ was related to ratings on the Hospital Adjustment Scale made by nurses who had observed behavior in the wards. It was also related to follow-up information and ratings of test cooperativeness. These three indices all involve assessments of global behavior in reality testing. The authors suggest that, while OVQ is not related to specific aspects of psychopathology (i.e., symptoms), it may reflect, to some degree, reality testing ability.

Money and Wang (18) compared normal and clinical groups with regard to sex of first figure drawn in the D-A-P. The relationship of male to female first drawings of male transvestites versus men with Klinefelter's Syndrome (10:11 versus 5:4) did not differ from each other, contrary to expectations. A comparison of these two groups with normals (1:1 versus 9:3) revealed a difference that was significant \( (X^2 = 38.92, P < .001) \). A group of sixteen precocious puberty boys compared with the transvestite group (ratio 13:3 versus 10:11) yielded a difference at the .001 level of confidence; with
the Klinefelter group (ratio 13:3 versus 5:4), a difference between the .001 and .01 levels. Sixteen effeminate juveniles differed from normal ten year old males (ratio 3:5 versus 4:1) so that $X^2 = 16.70$ and $P < .001$.

Blank and Roth (6) studied aspects of voyeurism and exhibitionism and their influence on behavior. Nudists, college females, engineering students and suburban couples were given the Minnesota Multiphasic Personality Inventory and the D-A-P. The only conclusive findings were that nudists were shown to have greater body-image distortion as well as sexual preoccupation, conflicts, and inhibitions.

Oxhorn (19) found that sixty eleventh-grade boys who consistently drew the female larger than the male perceived greater conflict with their mothers than did the group that drew the male consistently larger.

In a study by Litt and Margoshes (17) the D-A-P was administered to 341 school children. On a second testing 65 opposite sex drawers produced 39 same-sex and 26 opposite sex drawings. The results indicate that there is considerable variability on retesting and it is recommended that caution be exerted in interpreting sex drawn first from single drawings.

Whitaker (27) used an extended D-A-P technique to identify homosexual and effeminate men. Drawing from a population of 236 court clinic referrals, those rated as homosexual and/or effeminate produced significantly more
female figures as a free choice on a third drawing. However, the improvement over base rates was minimal.

In investigating the level of scorer sophistication to D-A-P diagnostic accuracy, Watson (26) asked twenty psychologists to categorize forty-eight D-A-P protocols as being the productions of organics, paranoid schizophrenics, non-paranoid schizophrenics, or normals, and to rate the drawings for distortion. Of the psychologists, ten were D-A-P users, ten were non-users and four were projective test experts. The accuracy of the judges' diagnostic acuity did not vary with the judges' D-A-P experience or projective test sophistication. Although drawing distortion ratings were very highly related to every judge's diagnostic impressions, they were not correlated with hospital-record diagnosis in any of the twenty-four cases. The latter finding suggests that many psychologists have seriously overestimated the extent to which drawing distortion is useful as a diagnostic indicator. The author concluded that the technique has usefulness as a trait index and for various other limited goals but is ineffective as a diagnostic instrument.
CHAPTER BIBLIOGRAPHY


5. Berman, Sidney and Julius Laffal, "Body Type and Figure Drawing," Journal of Clinical Psychology, IX (October, 1953), 368-70.


12. Hiller, E. Wesley and David Nesvig, "An Evaluation of Criteria Used by Clinicians to Infer Pathology from Figure Drawings," *Journal of Consulting Psychology*, XXIX, No. 6 (December, 1965), 520-29.


16. Lewinsohn, Peter M., "Psychological Correlates of Overall Quality of Figure Drawings," *Journal of Consulting Psychology*, XXIX, No. 6 (December, 1965), 504-12.


18. Money, John and Christine Wang, "Human Figure Drawing: Sex of First Choice in Gender-Identity Anomalies, Klinefelter's Syndrome and Precocious Puberty," *Journal of Nervous and Mental Disease*, CXLIII (August, 1966), 157-162.


CHAPTER III

METHODOLOGY

Subjects

The sixty subjects in this study were taken from a population of patients and ex-patients from a neuropsychiatric hospital. Both in-patients and out-patients were in the population and all diagnostic categories were theoretically represented as the clinic and hospital has no selectivity as to type of disorder taken for treatment. Only subjects who had been given psychological tests were included in the sample, and all had been seen at least two times for evaluation by one of the three psychiatrists who made up the clinic medical staff. In addition, most patients had been followed in out-patient psychotherapy for a period of more than one month.

The subjects ranged in age from fourteen years to fifty-four years with a mean age of thirty-four years. There were forty-one females and nineteen males in the sample.

The sample was divided into two groups on the basis of the presence or absence of mention of sexual conflict in the psychological report. Group I (G₁) consisted of thirty Ss in whose psychological report it was mentioned that sexual conflicts constituted a portion of the clinical picture.
$G_1$ consisted of twenty-one females and nine males with a mean age of 32.6 years and ranging in age from fourteen years to fifty-two years.

The thirty Ss who comprised Group II ($G_2$) were selected on the basis of no mention of sexual conflict in the psychological report. This group consisted of twenty females and ten males, with an age range of fourteen to fifty-four and a mean age of 35.3 years.

Description of Instruments and Procedure

All of the Ss in the study were tested as a result of referral for evaluation and/or treatment at a neuropsychiatric clinic and thus none were considered asymptomatic. Each S was presented with several sheets of paper, eight and one-half by eleven inches, a pencil and an eraser. He was then instructed to draw a person. Questions asked by the subject as to the desired nature of the drawing were answered in such a way as to give no guides or biases to his production. After the first drawing was finished he was then requested to draw a person of the opposite sex.

The drawings were then used as part of a battery of psychological tests from which a report concerning the dynamics and diagnosis of each individual was derived.

For purposes of this study any interpretations or notes written on the drawings at the time of their original use were covered by stapling other sheets of paper over them so
as to provide no cues for the scorers. Thus, each subject was represented by a same-sex and opposite-sex drawing. The only information given the scorers was the age of the subject, the sex of the subject and the sequence of his drawings.

The drawings from $G_1$ and $G_2$ were shuffled into a random order and presented to the scorers with instructions to rate each drawing along a five-item rating scale for the degree of sexual conflict shown. Each scorer was given all sixty sets of drawings and a typewritten copy of the definition for "sexual conflicts" as used for this study. The definition was worded thus:

Sexual conflicts are considered to be any adjustments in the area of psychosexuality which, either consciously or unconsciously, have resulted in maladaptive responses on the part of the individual. The individual may or may not be aware of these conflicts.

For each set of drawings rated the scorer was asked to check one of the following degrees of sexual conflict:

None  Some  Little  Much  Very Much

No further instructions were given nor were any specific scoring criteria indicated as it was wished to simulate as closely as possible a clinical situation. Thus, theoretically each scorer globally rated each set of drawings according to whatever criteria he normally used in clinical work.

A total of six scorers was used. They were divided into three levels of two scorers each, based on educational level and experience with the D-A-P. The first level
consisted of two private practitioners with Ph.D.'s in clinical psychology who had used the technique in a clinical setting for over five years each. The second level consisted of two Master's degree level professional psychologists who had used the D-A-P in a clinical setting for over two years each. The third level was comprised of two intern psychologists who were in the process of fulfilling the requirements for their master's degree and who had used the technique in a clinical setting for about two months.

Three criterion measures were used. The first was a rating made by the psychiatrist who treated each patient, along a five item scale like the one used by the drawing scorers. Each psychiatrist was presented with only the patient's name and asked to subjectively rate that person from memory as to degree of sexual conflict. The same definition of sexual conflicts was given to the psychiatrist as was given to the scorers.

Not only were the scorer's ratings of the drawings related to the psychiatrists' ratings but also they were related to the presence or absence of mention of sexual conflicts on a social history and to whether the subject was a member of $G_1$ or $G_2$. 
CHAPTER IV

RESULTS AND DISCUSSION

In the first chapter it was hypothesized that the D-A-P technique would yield data that could be validly interpreted so as to discriminate individuals having sexual conflict from individuals who do not manifest sexual conflict in any significant degree.

It was also noted in Chapter I that this technique was best used within the context of an entire battery of psychological tests and therefore need not have extremely high individual validity. With this latter statement in mind, it is suggested that despite the fact that some of the results of this study are contradictory, the research hypothesis must be accepted.

The evidence for this assumption is presented in Table I. It will be noted that the percentage of agreement between the mean ratings given by the intern psychologists (Level I) and the ratings given by the psychiatrists are within ±1 point of each other 50 per cent of the time. The ratings of both the Master's level psychologists (Level II) and the Doctoral level psychologists (Level III) fell within ±1 point of the psychiatrists' ratings 61.7 per cent of the time. Thus the overall mean percentage of the time that the
### TABLE I

**NUMBER AND PERCENTAGE OF AGREEMENT BETWEEN MEAN PSYCHOLOGIST RATINGS AND PSYCHIATRISTS' RATINGS**

<table>
<thead>
<tr>
<th>Deviations from Psychiatrists Ratings</th>
<th>±0</th>
<th>±.5</th>
<th>±1.0</th>
<th>Over ±1.0</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean for Level I (Interns)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number</td>
<td>6</td>
<td>8</td>
<td>16</td>
<td>30</td>
<td>60</td>
</tr>
<tr>
<td>Percentage</td>
<td>10.0</td>
<td>13.3</td>
<td>26.7</td>
<td>50.0</td>
<td>100</td>
</tr>
<tr>
<td><strong>Mean for Level II (Master's Level)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number</td>
<td>11</td>
<td>13</td>
<td>13</td>
<td>23</td>
<td>60</td>
</tr>
<tr>
<td>Percentage</td>
<td>18.3</td>
<td>21.7</td>
<td>21.7</td>
<td>38.3</td>
<td>100</td>
</tr>
<tr>
<td><strong>Mean for Level III (Ph.D. Level)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number</td>
<td>4</td>
<td>14</td>
<td>19</td>
<td>23</td>
<td>60</td>
</tr>
<tr>
<td>Percentage</td>
<td>6.7</td>
<td>23.3</td>
<td>31.7</td>
<td>38.3</td>
<td>100</td>
</tr>
</tbody>
</table>

The psychologists' ratings of the D-A-P fell within ±1 point of the psychiatrists' ratings was 57.8 per cent.

When the rather gross nature of the rating scale is considered, it can be seen that ±1 point, or one category, could easily be the deviation produced by different psychological sets concerning the rating scale on the part of the various raters, slight differences in defining "sexual conflicts," or simple human error. Thus, for test results to indicate "some" sexual conflict and for a psychiatrist to make an evaluation of "much" sexual conflict based on his
observation of a patient is really to say that there is probably not too much difference between the two evaluations in a clinical sense.

It is interesting to note that other statistical methods did not often yield results indicating validity. Correlations between the mean psychologists' ratings and the psychiatrists' ratings was only .1326. This small correlation was probably the result of the small range necessitated by the use of the rating scale. When the relationship between the mean psychologists' ratings and psychiatrists' ratings was tested using $X^2$, the results, while in the predicted direction, were not significant ($X^2 = 6.64644, P > .05$). This could be due to the fact that $X^2$ did not allow for close deviations between ratings.

Since previous research on the D-A-P shows a history of low relationships between D-A-P users' results and the criterion, it might be well in future research to establish a broader criterion since the technique is generally used as part of a battery of tests and therefore needs only moderate validity.

Other validity relationships between the mean psychologists' ratings and criterion measures were not significant. The relationship between mean ratings and mention or non-mention of sexual conflicts on a social history was negligible ($X^2 = 1.10941, P > .05$), as was the relationship between the mean psychologists' ratings and mention of neurotic problems.
conflict on the psychological report. Both of these results are, however, subject to the same statistical criticism as is noted above and cannot therefore be considered highly valid.

Results concerning inter-rater reliability indicate that the reliability of judgments concerning sexual conflict is quite high. Correlations between mean ratings for Rater Levels I, II, and III respectively with the mean for all raters were for Level I, .7732; for Level II, .7873; and for Level III, .7024.

The inter-rater reliability coefficient for Levels I and II was .3451; for Levels I and III, .3337; and for Levels II and III, .4056.

Since all correlations were lowered due to the constricted range, a percentage of agreement between judges was computed. The results are shown in Table II. It is noted that means for Levels I and II fall within ±.5 points of each other 61.7 per cent of the time and between ±1.0 points 86.7 per cent of the time. Mean ratings for Levels I and III fall within ±.5 of each other 61.6 per cent of the time and between ±1.0 91.6 per cent of the time. For Levels II and III, mean ratings fall within ±.5 of each other 66.7 per cent of the time, and between ±1.0 of each other 98.4 per cent of the time. Thus the mean percentage of agreement between all levels is 92.3 within ±1.0.
TABLE II

NUMBER AND PERCENTAGE OF AGREEMENT BETWEEN SCORER LEVEL MEANS

<table>
<thead>
<tr>
<th>Deviations</th>
<th>±0</th>
<th>±.5</th>
<th>±1.0</th>
<th>Over ±1.0</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean for Levels I and II</td>
<td>Number</td>
<td>12</td>
<td>25</td>
<td>15</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Percentage</td>
<td>20.0</td>
<td>41.7</td>
<td>25.0</td>
<td>13.3</td>
</tr>
<tr>
<td>Mean for Levels I and III</td>
<td>Number</td>
<td>14</td>
<td>23</td>
<td>18</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Percentage</td>
<td>23.3</td>
<td>38.3</td>
<td>30.0</td>
<td>8.0</td>
</tr>
<tr>
<td>Mean for Levels II and III</td>
<td>Number</td>
<td>12</td>
<td>28</td>
<td>19</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Percentage</td>
<td>20.0</td>
<td>46.7</td>
<td>31.7</td>
<td>1.7</td>
</tr>
</tbody>
</table>

The results of this study would indicate that, while the validity of the D-A-P technique in determining the presence of sexual conflicts is not extremely high, it does have sufficient validity to be of value when used within the context of an entire battery. It should be strongly noted that a study of this nature can not give an accurate picture of the validity of the technique simply because the "blind" nature of the ratings excluded a considerable portion of the available information that would ordinarily be used in a clinical situation. Since the rater had no way of noting the subject's behavior during the drawing he could not utilize the complex set of cues that is often called "clinical intuition." Thus, for example, it is assumed that if
the rater could have noted at which points in the drawing process the subject made anxiety-laden remarks, engaged in nervous movements, etc. he would then have had a great deal more information on which to base his rating.

Another problem was noted in a study of this nature, that of interdiscipline communications. It could very well be that "sexual conflicts" meant one thing to the psychologists and something slightly different to the psychiatrists even though a written definition was given to both. For example, an individual who has mild ambivalence toward women due to an overprotective mother may show definite signs of "sexual conflict" on his drawings, such as differential treatment of figures, etc., but not actually manifest any difficulties in the psychosexual sphere. It is possible that this additional hypothetical group could have obscured the correlations by adding a dimension to one or the other of the ratings that was not being considered by the other raters.

From inspecting the results of Table I it may be concluded that the technique can be used with valid results by the relatively inexperienced as well as those more sophisticated with the technique. It was further noted that the Level II raters were within ±.5 of the criterion 40.0 per cent of the time as opposed to only 30.0 per cent of the time for the Level III raters.
CHAPTER V

SUMMARY

This study investigated the use of the Draw-a-Person technique as a means of measuring sexual conflict. A limited area of the personality was chosen because a survey of the literature indicated that the technique had proven relatively useless in differentiating between diagnostic categories, but might have some value in determining conflict areas, traits, and dynamics of behavior.

The area of sexual conflict was chosen because psychosexuality is so closely linked to physical function and hence to body concept. Since one projects via body representation with the D-A-P technique, it was thought that this technique might best express conflicts in the area of psychosexuality.

A sample of sixty Ss was randomly drawn from a clinical population who had been given a battery of psychological tests, including a social history. The sample was divided into two groups according to whether mention was made of sexual conflicts in the psychological report.

The self-sex and other-sex drawings of the Ss were pulled from the files, randomized and rated by six scorers representing three levels of education and experience with the D-A-P. Level I consisted of two psychology students in
the process of completing a clinical internship who had
used the technique for two months. Level II consisted of
two Master's degree level professional psychologists who had
used the D-A-P in a clinical setting for over two years.
Level III consisted of two Ph.D. level professional psychol-
ogists who had used the technique for over five years.

The only information given the scorers was a definition
of "sexual conflicts," the sex of the subject, the age of
the subject, and the sequence of the drawings. They were
asked to rate each person for the degree of sexual conflict
along a five category rating scale ranging from "none" to
"very much."

The primary criterion measure was a similar rating
scale completed by the psychiatrist who treated the subject
in psychotherapy.

The results showed that the mean psychologists' ratings
agreed within ±1 degree of severity with the ratings made by
the psychiatrists 57.8 per cent of the time.

The correlational computation of the same relationships
yielded only low values, as did X^2 computations. An examina-
tion of the nature of the data revealed that the low
correlation coefficients were the product of the rather
constricted range and that the X^2 results could be explained
on the basis of the inappropriateness of the statistical
technique due to the leptokurtic distribution of the data.
It is suggested that this may be considered as a possible explanation for some of the past findings indicating low validity.

Results of $X^2$ and correlational comparisons between scorers' means and presence or absence of mention of conflict on the social history and in the psychological report are insignificant; however, the same reasoning as applied to the former computations is appropriate here.

Inter-rater reliabilities were quite high, with mean inter-rater correlation coefficients being .3615. The mean percentage of the time that the mean for one level fell between $\pm 1$ degree of severity was 92.2.

It is concluded that the Draw-a-Person technique has good inter-judge reliability and is valid enough to be useful in a battery of psychological tests in discerning sexual conflicts.

It is suggested that further research attempt to measure validity of the technique under conditions more closely resembling a clinical setting and it is also noted that results are quite variable, depending on the statistical technique used.
BIBLIOGRAPHY

Books


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

Ruch, Floyd L., Psychology and Life, Dallas, Texas, Scott, Foreman and Co., 1953.


Articles


Berman, Abraham B., Alexander A. Klein, and Abbott Lippman, "Human Figure Drawing as a Projective Technique," The Journal of General Psychology, XLV (July, 1951), 57-70.
Berman, Sidney and Julius Laffal, "Body Type and Figure Drawing," *Journal of Clinical Psychology*, IX (October, 1953), 368-70.


Fisher, "Style of Sexual Adjustment in Disturbed Women and its Expression in Figure Drawings," *Journal of Psychology*, XXXIV (1952), 169-79.


Hiller, E. Wesley and David Nesvig, "An Evaluation of Criteria Used by Clinicians to Infer Pathology from Figure Drawings," *Journal of Consulting Psychology*, XXIX, No. 6 (December, 1965), 520-29.


Lewinsohn, Peter M., "Psychological Correlates of Overall Quality of Figure Drawings," *Journal of Consulting Psychology*, XXIX, No. 6 (December, 1965), 504-12.
Litt, S. and A. Margoshes, "Sex-change in Successive Draw-
a-Person Tests," Journal of Clinical Psychology, XXII
(October, 1966), 471.

Money, John and Christine Wang, "Human Figure Drawing: Sex
of First Choice in Gender-Identity Anomalies, Kline-
felder's Syndrome and Precocious Puberty," Journal of
Nervous and Mental Disease, CXLIII, 2 (August, 1966),
157-162.

Oxhorn, Joseph L., "The Relation of Figure Drawing to
Masculine-Feminine Orientation and Academic Achievement,"

Royal, Robert E., "Drawing Characteristics of Neurotic
Patients Using a Drawing-of-a-Man-and-Woman Technique,"

Schaeffer, Robert W., "Clinical Psychologists' Ability to
Use the Draw-a-Person Test as an Indicator of Personality
Adjustment," Journal of Consulting Psychology, XXVIII,
No. 4 (August, 1964), 383.

Silverstein, A. B. and H. A. Robinson, "The Representation
of Orthopedic Disability in Children's Figure Drawings,"
Journal of Consulting Psychology, XX (October, 1956),
333-41.

Sundberg, N. D., "The Practice of Psychological Testing in
Clinical Services in the U. S.," American Psychologist,
XVI (February, 1961), 79-83.

Swensen, Clifford H., Jr., "Empirical Evaluations of Human
Figure Drawing," Psychological Bulletin, LIV, No. 6

Watson, Charles G., "The Relationship of Distortion to
Draw-a-Person Diagnostic Accuracy Among Psychologists
at Three Levels of Sophistication," Journal of Consulting

Whitaker, L., Jr., "The Use of an Extended Draw-a-Person
Test in Identifying Homosexual and Effeminate Men,"
Journal of Consulting Psychology, XXV, 6 (December,
1961), 482-85.

Whitmyer, J. W., "The Significance of Artistic Excellence in
the Judgement of Adjustment Inferred from Human Figure
Drawings," Journal of Consulting Psychology, XVII
(December, 1953), 421-24.

Publications of Learned Organizations


Unpublished Materials