THE VALIDITY OF SOME QUALITATIVE SCORING ITEMS ON THE DRAWINGS OF A HOUSE, TREE, AND PERSON

by Michael E. Cavanagh

Thesis presented to the School of Psychology and Education of the University of Ottawa as partial fulfillment of the requirements for the degree of Master of Arts

Ottawa, Canada, 1963

ACKNOWLEDGMENTS

This thesis was prepared under the supervision of Assistant Professor Gilles Chagnon, M.Ps., of the School of Psychology and Education of the University of Ottawa.

The writer is indebted to Gordon Riley, Ph.D., for his cooperation in obtaining patients of Napa State Hospital, Imola, California, as subjects in the study, and to Jack Blumenkrantz, Ph.D., for his assistance in obtaining subjects at the Veteran's Administration Hospital, Fort Miley, San Francisco, California.

CURRICULUM STUDIORUM

Michael E. Cavanagh was born April 29, 1937, in San Francisco, California. He received his Bachelor of Arts degree in Philosophy from Saint Patrick's College, Menlo Park, California, in 1960.

TABLE OF CONTENTS

LIST OF TABLES

INTRODUCTION

Although the use of free hand drawing as a clinical tool has gained increasing repute during the past twenty-five years, experimental verification of assumptions concerning specific interpretive items of the drawings has not been entirely satisfactory or complete.

Among the projective drawing techniques that are in the process of validation is Buck's[1] House-Tree-Person Test (henceforth referred to as the H-T-P). It was designed as an instrument to elicit data concerning the dynamics and degree of integration of an individual's personality. The qualitative system of the H-T-P and particularly the interpretation assigned individual scoring items is the least objective of the interpretive aspects, although it can also be the most revealing.[2] Hammer states that

> (...) the H-T-P is presently embarking on the thorny path of being transformed from an empirical and practical technique to a scientifically controlled and experimentally rooted method of personality analysis. The chief need at this time is for further validation studies.[3]

---

1 John N. Buck, "The H-T-P Technique: A Qualitative and Quantitative Manual", Journal of Clinical Psychology, Monograph Supplement No. 5, Vol. 4, No. 4, October, 1948, p. 320-396.

2 --------, Administration and Interpretation of the H-T-P Test, Beverly Hills, California, Western Psychological Services, 1950, p. 36.

3 Emanuel F. Hammer, The H-T-P Clinical Research Manual, Beverly Hills, California, Western Psychological Services, 1955, p. 1.

Since the H-T-P is being used with continued frequency as a clinical and research tool the writer feels that testing "hunches" concerning some interpretive aspects of the H-T-P is one step in the direction of rendering the theory and effective application of the test more tenable. Therefore the aim of this study is threefold: 1) to compile a list of items from the H-T-P literature that are inferred as possible indicators of schizophrenia and to objectify these items as much as possible to decrease subjective interpretation; 2) to establish if these items significantly differentiate between a group of schizophrenics and a group of non-schizophrenics as a means of validating[4] these items; and 3) to compare this detailed atomistic approach to the interpretation of drawings to a total impressionistic approach as regards their respective efficiency in differentiating the groups. This is done to ascertain the relative value of using such a checklist in a clinical setting.

The study has theoretical importance in terms of providing some accumulative evidence for or against the acceptance of some quantifiable aspects of the qualitative system as a basis for further objective personality evaluation.

The first portion of this thesis is concerned with a review of the literature concerning the qualitative aspects of drawings in general and the H-T-P in particular.

---

4 Validity in this sense refers to concurrent validity.

The formulation of the basic hypothesis is followed by a description of the experimental design, which includes a discussion of the sample, reliability, and statistical procedures applied to the data.

The results obtained are then presented and discussed. Implications for further research are indicated, and in the appendices are given the list of items and the objectification of these items.

CHAPTER I

REVIEW OF THE LITERATURE

In order that the research on the qualitative aspect of the H-T-P may be more meaningful, it is necessary to place it in its proper perspective in the continuum of previous research on the clinical use of drawings. Since it would be impractical to include the entire research on drawings, the method of selection for the presentation of studies is based on an attempt to give a panoramic view of the various theories that have been postulated and tested with emphasis on studies that are pertinent to the design and purpose of this study.

The first section of this chapter will deal with the historical foundations of the diagnostic use of drawings. In it will be mentioned studies that laid the groundwork for research in this area and which were heuristically the most influential. The early 1940's is the period which separates the section on historical literature from that on related studies. This division is not entirely arbitrary since there appears at this time a slight change in trend in the literature from "open-end" type research on drawings in general to a more systematized attempt at validating specific hypotheses in human figure drawings. The next section will deal with related studies--those that use human figure

drawings as tools and seek to validate interpretive hypotheses in the clinical use of drawings. Theoretical implications, the majority of which arose from the historical foundations and developed through the related studies, will be discussed in the third section in order to better understand the H-T-P and its literature which are presented in the fourth section of this chapter. The final section summarizes the review of the literature and states the hypothesis for this study.

1. Historical Foundations.

Although one can trace the mention of "the art of the insane" as far back as Pinel's Treatise on Insanity[1] in 1806, it appears that Max Simone[2] in 1876, was the first to emphasize the diagnostic value of drawings and attempt correlations between clinical syndromes and certain characteristics of drawings. His work exerted a definite influence upon subsequent studies in Europe. In 1906, Mohr[3] gave a

---

1 A. Anastasi and J.P. Foley, Jr., "A Survey of the Literature on Artistic Behavior in the Abnormal: I. Historical and Theoretical Background", Journal of General Psychology, Vol. 25, First Half, July, 1941, p. 114.

2 Ibid., p. 115.

3 --------, "A Survey of the Literature on Artistic Behavior in the Abnormal: IV. Experimental Investigations", Journal of General Psychology, Vol. 25, First Half, July, 1941, p. 189.

systematic approach to the study of the psychological components of expressive movements. One approach he used is exemplified in a study with eighteen patients including schizophrenics, manic depressives, catatonics, paretics, epileptics and mental defectives. Mohr drew a picture of a church-like object in view of the patient and asked the patient to copy it. His results confirm those of other investigators who obtained spontaneous drawings and suggest specific differences in the drawings for various forms of abnormality. The worst cases of catatonia and schizophrenia produced only a scribble. The manic depressive drawings contained additions and ornamentations. The drawings of paretics showed complete disintegration, and those by epileptics showed perseveration. The approach Mohr used in this study is one of several he suggested for the study of psychotic drawings; others include copying a complicated geometric figure, drawing a simple object from nature, drawing from memory (e.g. drawing hallucinations), illustrating a simple story, coloring drawings and paintings, and completing partial drawings. He classified the functions involved in simple drawing behavior into optico-physiological activity, perception, apperception, combination of visual perceptions with kinesthetic and volutional impulses, purposiveness, attention, motor memory, general memory and training, and practice in drawing.

Kürbitz,[4] in 1912, repeated Mohr's study of copying a church-like drawing. He used what seems to be eleven patients (though he does not state the number). Like Mohr, Kürbitz concluded that the drawings of psychotics are closely related to their disease and are therefore diagnostically useful. In addition, he found certain similarities between the drawings by excited schizophrenics and those of manic depressives in the manic state. Paretics and schizophrenics in the stupor state showed extreme poverty of performance, their productions often resembling those of mental defectives. Kürbitz also pointed out certain resemblances between psychotic drawings and those of children and primitive groups, as illustrated by "transparent drawings" and the combination of side and front views (double perspective).

In 1933, Earl[5] using the Goodenough Scale with 420 feebleminded adults, found thirty-four cases clearly psychotic. He reported that emotionally unstable persons who were not psychotic showed some typical psychotic features in their drawings. On the other hand, not all psychotics exhibited these characteristics. Of the thirty-four psychotic cases, nine produced drawings of the type ordinarily made by non-psychotics of the same mental level.

---

4 Anastasi and Foley, "A Survey of the Literature on Artistic Behavior in the Abnormal: IV. Experimental Investigations", Op. Cit., p. 190.

5 Ibid., p. 210.

Satô[6] obtained drawings from seventy-four patients (44 male and 30 female) and thirty-five control subjects (19 male and 16 female). He reported that "primitive" or crude drawings were not restricted to the psychotics but were also found in the control group. Satô stated further that the "primitive" manner of drawing could not be attributed to educational level since three highly educated schizophrenics produced completely primitive drawings. In spite of some apparent inconsistencies in the drawings Satô concluded that, in general, schizophrenia tends to evoke primitive drawing. Among the special characteristics, which Satô observed in the drawings of psychotics, are stereotypy, individual mannerisms, negativism, and fluctuation of performance over a period of time.

Becker,[7] in 1934, used the technique of copying geometric figures with seventy-five schizophrenic patients (50 male and 25 female) of whom sixty responded, the remainder being too negativistic or catatonic to cooperate. So called controlled data were obtained on forty-one normal subjects (29 male and 12 female) consisting chiefly of patients attending a clinic on venereal and skin diseases and a few attendants and nurses. Each of three simple

---

6 Anastasi and Foley, "A Survey of the Literature on Artistic Behavior in the Abnormal: IV. Experimental Investigations", Op. Cit., p. 197.

7 Ibid., p. 191.

geometric figures was exposed for ten seconds to a subject who was required to reproduce it immediately on removal. The schizophrenic drawings showed a greater percentage of errors than the control drawings. From his results Becker discussed three general characteristics which he considers most nearly typical of schizophrenic drawing: the introduction of idiosyncratic details not present in the example copy; reproductions that appear as if the subject were feeling his way along the outline of the example without grasping the meaning of the object; and lack of integration. Becker stated that the motor factor tends to outweigh the visual in schizophrenic drawings, the movements being impulsive and even "gesture-like". Schizophrenic drawing is compared to the "doodling" of normal persons when distracted.

In a more extensive study Pfister,[8] using psychoanalytic theory, studied drawings from 411 patients (207 male and 204 female). The majority of the patients were schizophrenic, the remainder being distributed among psychopaths, manic depressives, epileptics, paretics, senile psychotics, and mental defectives. In addition data were obtained from one hundred subjects in an out-patient department, including normals, psychopaths, neurotics, and a few

---

8 Anastasi and Foley, "A Survey of the Literature on Artistic Behavior in the Abnormal: IV. Experimental Investigations", Op. Cit., p. 192.

incipient psychotics. The only mention of a control group was a reference to a normal rural Swiss population. Each subject was given five crayons and one piece of paper divided into eight equal parts with the name of an object at the top of each section. Among the eight objects to be drawn were a house, a tree and a man. Beyond findings related to movement and color, Pfister found that schizophrenics often drew the tree or the man sidewise on the sheet, a reaction which Pfister attributes to their ego-centric attitude since it is easier to draw in that manner. In reference to content he found that about ten per cent of the schizophrenics drew something other than the object directed. Special peculiarities also occurred in the schizophrenics' drawing of the eyes, some omitting them, others drawing them in a profile view. This study encompasses most of the points found in earlier investigations and is typical of a larger group of studies in both procedure and results.

In 1938, Bender[9] investigated the operation of Gestalt principles in mental defectives and various types of psychoses, using what is now called The Bender-Gestalt Perceptual-Visual Motor Test. The population consisted of twenty-five schizophrenics and twenty-five non-psychotics,

---

9 Anastasi and Foley, "A Survey of the Literature on Artistic Behavior in the Abnormal: IV. Experimental Investigations", Op. Cit., p. 203.

the non-psychotics ranging in intelligence from low grade idiots to superior adults. Among the specific distortions found in the reproduced figures are perseverations, the influence of the preceding sample figure, the tendency to use compact and enclosed "energy saving" units, reduplication of parts, flight of ideas as when the figure suggests a number of familiar objects into which the sample figure is converted, and a tendency to substitute letters, numbers, and personal material such as the patient's initials for parts of the figure.

Most of these studies, while they are valuable as frontier work, suffer from various contaminations and privations that often befall embryonic scientific attempts. Some general criticisms have been levelled at these first studies, and each study mentioned is vulnerable to at least one of these criticisms. There are little or no pertinent data regarding the subjects which may influence the results of the drawings. For example, such variables as age at onset of illness, length of institutionalization, education, and degree of remission are not mentioned. Often the interpretations are made in the jargon of specific schools of thought; the terminology is generally vague and ill-defined and hampers a clear understanding of the results. Few attempts at standardization in the control of the experimental conditions are present; for example, the setting in which the test was taken,

administration procedures, time allotment, and presence of distractions. Perhaps the salient criticism is that most of the data are not reported in quantitative form but merely described in terms of general statements supported by illustrative examples.[10] However, these studies effected a growing interest in drawing as a diagnostic tool and precursed the way for testing hypotheses in a more experimental climate.

2. Related Studies.

Investigations with human figure drawings apropos to clinical situations generally began in the later years of World War II. Anastasi and Foley[11] collected drawings of a man from 340 mental patients and 340 normal control subjects with an equal number of males and females in each group. The study purposed to differentiate the groups by a criterion of twenty-three items. Fourteen items differentiated between the groups at better than the one per cent level of confidence. Because there was no attempt to cross validate

---

10 Anastasi and Foley, "A Survey of the Literature on Artistic Behavior in the Abnormal: IV. Experimental Investigations", Op. Cit., p. 232.

11 --------, "An Experimental Study of the Drawing Behavior of Psychotics in Comparison with that of a Normal Control Group", Journal of Experimental Psychology, Vol. 34, No. 3, June, 1944, p. 169-194.

or to determine the chance expectancy for interrelated scoring items, the real meaning of the study cannot be assessed.

Waehner,[12] in 1946, asked fifty-five female college students to: 1) draw a series of drawings with a free choice of kinds of material (paints, chalk, paper, etc.) and theme; 2) draw human faces in the following order: a) any kind of face, b) an ugly and a beautiful face, and c) a self-portrait without a mirror; 3) a group of human figures. Four hundred and twenty-two pictures were analyzed and scored. Rorschachs were also administered to the students. From the drawings and the Rorschach tentative personality sketches were made and given in sets of four, five, or six to teachers of the students, to a psychologist, and to a Rorschach expert. One hundred and sixteen matching judgments from teachers were received on forty-one students (fourteen students could not be matched for technical reasons). In one hundred and three cases the teachers recognized the students and in thirteen cases referring to thirteen students the teachers did not recognize them. The agreement arrived at in individual statements was extremely high. The matching with the Rorschach was eighty-seven per cent.

---

12 T.S. Waehner, "Interpretation of Spontaneous Drawings and Paintings", Genetic Psychology Monographs, Vol. 33, First Half, February, 1946, p. 3-70.

In 1949, Machover[13] published the Draw-a-Person Test and several hypotheses regarding this drawing technique. Machover considered three sources in which certain organs come to have specific meanings which are noted in drawings of the human figure. This, she felt, occurs regardless of age, skill, or culture. The first is that physical attributes tend to acquire social meanings in the course of social participation and expression. The second is that psychosomatic correlation of body expression may be revealed by overemphasizing or deemphasizing certain parts of the body in human figure drawings. And, thirdly, psychic datum may be noticed in the symbol value projected in the drawing. From this point the majority of human figure drawing studies were done under the spectrum of Machover's claims.

Fiedler and Siegel[14] used the elaboration of facial features of human figure drawings in predicting improvement in psychotherapy; it was felt that a subject's drawing might indicate the extent of his ability to relate to other persons and that this ability was a prerequisite to successful therapy. The thirty-four clinic patients used were given

---

13 Karen Machover, Personality Projection in the Drawing of the Human Figure, Springfield, Thomas, 1949, ix-181 p.

14 F.E. Fiedler and S.M. Siegel, "The Free Drawing Test as a Predictor of Non-Improvement in Psychotherapy", Journal of Clinical Psychology, Vol. 5, No. 4, October, 1949, p. 392-395.

psychological tests prior to therapy, had at least five therapy interviews, and were diagnosed psychoneurotic. Fifteen improved patients and nineteen unimproved patients were studied. The predictor sign was reliable at the three per cent level of confidence using the Chi Square test.

Royal[15] compared the human figure drawings of eighty anxiety neurotics and one hundred dental patients using twenty-eight scoring items as a criterion of difference. Both groups were equated as far as possible for age and intelligence. The selection of items for the checklist depended partially on clinical "hunches" derived from superficial examination of the drawings while other items were adapted from suggestive differential characteristics reported in the literature. Eight items approached significance in differentiating the two groups; however, these eight items did not hold up under cross validation. None of the twenty-eight items was successful differentia on both trials.

Stonesifer[16] felt that, although the Goodenough Scale was developed to measure the intelligence of children, it

---

15 Robert E. Royal, "Drawing Characteristics of Neurotic Patients Using a Drawing-of-a-Man-and-Woman Technique", Journal of Clinical Psychology, Vol. 5, No. 4, October, 1949, p. 392-395.

16 Fred A. Stonesifer, "A Goodenough Scale Evaluation of Human Figures Drawn by Schizophrenics and Non-Psychotic Adults", Journal of Clinical Psychology, Vol. 5, No. 4, October, 1949, p. 396-398.

might also differentiate between schizophrenics and non-psychotic adults. He based this belief on the rationale that schizophrenic thought was felt to resemble a childish mode of thought, a turning away from reality, and personality disintegration. He hoped that the childish thinking and general regressive behavior associated with the effects of schizophrenia might be reflected in a lower point score for the psychotic group. The control groups consisted of one hundred males, and the experimental group consisted of fifty-seven schizophrenic males. The mean age of the control group was .25 higher than the experimental group, and the mean educational level was .46 of a grade higher in the control group than in the experimental group. One hundred fifty drawings were scored because some subjects were classified as unsuitable for this study, and fifteen were rescored giving .96 as the reliability of scoring. A critical ratio was used to investigate the significance of difference in the Goodenough point score which was not significant and so low that it is not likely that the Goodenough Scale will discriminate between larger groups of subjects.

Albee and Hamlin,[17] also in 1949, studied the reliability and validity of judgments of adjustment made from

---

17 G.W. Albee and R.M. Hamlin, "An Investigation of the Reliability and Validity of Judgments of Adjustment Inferred from Drawings", Journal of Clinical Psychology, Vol. 5, No. 4, October, 1949, p. 389-392.

human figure drawings. They had fifteen clinical psychologists rank ten human figure drawings by the method of paired comparisons on global evaluation of adjustment. These drawings were by subjects who ranged from "normal" to hallucinated delusional psychotics. In terms of validity the rank order correlation between the composite rank of the judges' evaluation and the rank of adjustment from case records was .64 which is significant at the .05 level of confidence. The interjudge reliability was .95, significant at the .01 level. The authors state that from other data not completely analyzed clinical psychologists not experienced in projective techniques can make as reliable judgments as the experienced clinical psychologists. The question of validity in this study rests on more precarious assumptions than does the reliability. The original clinical records achieve as wide a range of adjustment as possible. However, the objection as to the validity of the judgments of adjustment, even though the selection was agreed upon by all judges, may reasonably be raised. Also the study's use of the term "adjustment" appears to be relative. Perhaps a psychotic in partial remission is "better adjusted" than an anxiety neurotic. Moreover, the small number of cases suggests caution in interpreting the results.

A year later the same investigators[18] collected human figure drawings from twenty-one schizophrenic, twenty-one neurotic, and thirty normal subjects, matched for age, education, and sex. The ten drawings mentioned in the previous study were used as reference points on a rating scale. The authors pooled the rating scale values of four judges (.89 interjudge reliability) which yielded a significant difference at the .01 level of confidence between the normal group and both groups of patients, but not between the neurotic and schizophrenic groups.

In a study which was primarily an attempt to validate human figure drawings as a diagnostic tool and secondarily to provide some basis for an objective analysis of drawings, Holzberg and Wexler[19] used a checklist of 174 items to differentiate between a control group of seventy-eight student nurses and thirty-eight schizophrenic women (18 paranoid, 12 hebephrenic, 6 catatonic, 1 simple, and 1 mixed). Twenty-seven items differentiated between the two groups at the .01 level of confidence; twenty-three items

---

18 G.W. Albee and R.M. Hamlin, "Judgment of Adjustment from Drawings: the Applicability of Rating Scale Methods", Journal of Clinical Psychology, Vol. 6, No. 4, October, 1950, p. 363-365.

19 J.D. Holzberg and M. Wexler, "The Validity of Human Form Drawings as a Measure of Personality Deviation", Journal of Projective Techniques, Vol. 14, No. 4, December, 1950, p. 343-361.

differentiated the control group from hebephrenics; and eighteen items differentiated the control group from the paranoids, at the .01 level of confidence. Age was not controlled for, and there was a gap of eight years between the mean ages of each group in favour of the schizophrenics. Reliable differences among the three schizophrenic sub-groups were not demonstrated.

Fisher and Fisher,[20] in 1950, tested two of the more explicit assumptions of Machover's work; namely, that the detection of strong paranoid attitudes in drawings is not too difficult, and that one can almost by direct inspection "feel" the expression in a figure drawing face or the attitude in a stance. The investigators used thirty-two drawings of paranoid schizophrenics ranging in age from twenty to forty-five years. The criteria were Machover's six "Paranoid" signs: eye emphasis, large grandiose figure, speared fingers, large head, rigid stance, and large ears. The drawings were evaluated by both a detailed atomistic approach and a total impressionistic analysis of paranoid trends. The majority of these drawings did not fall under the category of paranoid in terms of the criteria implicit in each method. Furthermore, the two separate methods of evaluating paranoid

---

20 S. Fisher and R. Fisher, "Test of Certain Assumptions Regarding Figure Drawing Analysis", Journal of Abnormal and Social Psychology, Vol. 45, No. 4, October, 1950, p. 727-732.

trends did not correlate significantly. The drawings were evaluated as regards figure, facial expression and stance. Little agreement was found among the raters in the description of these characteristics. No comprehensive statistical analysis was presented in the study.

In 1952, Lehner and Gunderson[21] attempted to determine the consistency of sets of ratings on twenty-one graphic traits on the Draw-a-Person Test of ninety-one subjects (66 male, 25 female, ranging in age from 18 to 26 years). The study consisted of three parts: 1) A check of the authors' consistency was made by rating a sample of ninety drawings and rerating the sample one week later. 2) The interrater reliability in the application of the authors' rating system involved the use of three different raters each independently rating a sample of thirty drawings. 3) A study of the consistency with which the figures were drawn by the same person in the two different administrations which were approximately four months apart was made on the two sets of the 181 drawings (one female figure was omitted by one subject). For number one, the authors obtained a total of 90.5 per cent in all their ratings (all traits combined). In the second portion, three ratings rendered a per cent of agreement of

---

21 George F.J. Lehner and Eric K. Gunderson, "Reliability of Graphic Indices in a Projective Test (The Draw-a-Person)", Journal of Clinical Psychology, Vol. 8, No. 2, April, 1952, p. 125-128.

83.4. For the last section, the per cent of agreement for all traits was 64.1, with the body-type trait highest at 92.5 per cent.

The authors conclude that, given an objective and explicitly formulated rating system, a relatively high per cent of agreement may be achieved in evaluating indices commonly used in dealing with the Draw-a-Person Test. They state that nevertheless it indicates that there is a tendency for many traits to remain constant over a period of months. The more constant traits include both formal and content aspects which, the authors feel, seems to contradict Machover's claim that content aspects are less constant than formal aspects; however, it is possible that certain formal aspects may be vulnerable to transient disturbances in personality.

In 1952, Fisher and Fisher[22] studied the predictive value of drawings as indicators of sexual adjustment. They divided the drawings of seventy-six female psychiatric patients into those showing low, average, and high femininity. The three groups were compared on a number of ratings of past sexual behavior based on case history records. Of fifty-four comparisons, eight differentiated the groups at the .05 level of confidence.

---

22 S. Fisher and R. Fisher, "Style of Sexual Adjustment in Disturbed Women and Its Expression in Figure Drawings", Journal of Psychology, Vol. 34, Second Half, October, 1952, p. 169-179.

Berman and Laffal[23] with Schilder's body image concept in mind attempted to establish if people, in drawing a human figure of the same sex, tend to represent their own bodies in such drawings, to draw idealized figures representing how they would like to look, or to draw figures which have no relationship at all to their body build or to idealized figures. Thirty-nine male neuropsychiatric patients ranging in age from twenty-three to thirty-five years were classified as to body build according to the Stevens and Tucker method. Significant correlations (Pearson correlation r = .35, significant between .05-.01 level of confidence) were found between body type and the figure drawn. The authors felt that this result supports the validity of figure drawing tests as a projective device, since it shows a relationship between one aspect of the individual (his body type) and of the figure drawn by him. The authors also stated that the findings support the hypothesis that the figure drawn is, in part, a projection of the body image.

Also in 1953, Kotkov and Goodman,[24] investigating the basic premise that one's body image is projected in his

---

23 Sidney Berman and Julius Laffal, "Body Type and Figure Drawing", Journal of Clinical Psychology, Vol. 9, No. 4, October, 1953, p. 368-370.

24 B. Kotkov and M. Goodman, "The Draw-a-Person Tests of Obese Women", Journal of Clinical Psychology, Vol. 9, No. 4, October, 1953, p. 362-364.

drawings, obtained human figure drawings from a sample of 101 obese and ideal weight women matched for age, education, intelligence, marital, and employment status. The drawings of forty-five subjects were scored for forty-three items, and of 129 Chi Square comparisons, thirty-two items were significant at better than .02 level of confidence. Cross validation of these thirty-two items on another group of fifty subjects yielded seven items which differentiated the groups at better than the .05 level of confidence. Reliability of the scoring was not stated. Although the measurements tested were used on both the male and female drawings of the subjects the most striking aspect of the results was that significant differences were found exclusively in comparing the female drawings. The authors concluded that, since obese women express their variation from normal weight women through their perception of the female image, projection of differences into drawings is expressed quite exclusively through the drawing of the same sex.

Witkin et al.,[25] in 1954, compared a perceptual orientation test of fifty-two male and fifty-one female college students to their human figure drawings. Two sets of drawings were obtained; the first set was done in a group environment under the direction of several different

---

25 H.A. Witkin, et al., Personality through Perception, New York, Harper, 1954, p. 238.

examiners; the second set was obtained four to nine months later under the direction of one personal interviewer for all subjects. Each set of drawings was scored on eighty-eight items for males and ninety-eight items for females. These items were assumed to be related to perceptual orientation in terms of field dependence or independence. Then a short scale of graphic items (forty for men and forty-five for women) was composed from the original numbers of eighty-eight for men and ninety-eight for women. The correlation between the two forms was .97 for men and .95 for the women. Both scales were used in all the analyses. A total score was given which, when correlated with an index representing the subject's average performance on three perceptual orientation tasks, yielded product-moment correlations of .73 for the male group and.69 for the female group. The authors felt that, in general, the results showed that persons who are field dependent in their perception produce figure drawings reflecting a low evaluation of their bodies, infantile defenses against anxiety, lack of self-assurance, passivity coupled with uncontrolled expression of hostility, and difficulty in accepting an adult role. On the other hand, people who are not influenced by the prevailing field and who are capable of dealing with it in an active, analytic fashion produce drawings expressing a high degree of narcissistic investment in the body, sophisticated defenses

against anxiety, self-assurance, identification with desirable characteristics of both sexes, strong drive, and manipulative tendencies controlling their drives. Gruen[26] criticizes the study because the techniques used deal with the addition of discontinuous signs or with the presence or absence of constellations of signs. The correlations do not permit inference about the personal determinants of such perception. The existence of a perception-personality relationship does not by itself give a basis for inferring causality for purposes of reflecting on "personality through perception".

In 1955, Lewis[27] checking on some of Machover's claims collected Draw-a-Person records of twenty male subjects from each of ten diagnostic categories, ranging from normals to catatonic schizophrenics. Although Machover observed only presence signs, the results of this study show both presence and absence signs significant at the .01 level of confidence. Of the thirty-one significant absence signs twenty-four were found in the three schizophrenic groups and two in the involutional melancholia group. With the

---

26 Arno Gruen, "A Critique and Re-Evaluation of Witkin's Perception and Perception-Personality Work", Journal of General Psychology, Vol. 56, First Half, January, 1957, p. 73-93.

27 Augustus Lewis, "A Check on the Validity of Some of Machover's Claims", in unpublished doctoral thesis, School of Psychology and Education, University of Ottawa, 1955, p. vi-79.

exception of one sign in the catatonic group the significant presence signs occurred in the non-psychotic groups and normal groups. The total significant signs (thirty-one absence and thirty-eight presence) were found on thirty-five of the 214 scoring items. Of these thirty-five items, eight were present in some of Machover's diagnostic patterns, but the results of this study did not show complete agreement with Machover's observations on any of these items.

Swenson,[28] in 1955, investigated the differentiation of male and female figures as indicating the degree to which the subject has adequately identified himself sexually. Seven judges sorted 101 drawings of males who were patients at a mental hygiene clinic as a basis for constructing a scale to measure sexual differentiation. The drawings were sorted into five piles of approximately twenty drawings each with those showing the poorest sexual differentiation at one end and those showing the best sexual differentiation at the other. The twelve drawings agreed upon by the judges as showing the poorest sexual differentiation were selected. These were ranked for sexual differentiation and the three with the lowest average rank were chosen to represent the bottom of the scale. The same procedure was followed for

---

28 Clifford H. Swenson, "Sexual Differentiation on the Draw-a-Person Test", Journal of Clinical Psychology, Vol. 11, No. 1, January, 1955, p. 37-41.

the selection of three drawings representing the best sexual differentiation. The remaining sixty drawings were sorted into three equal piles, one pile to be made up of drawings falling midway between the best and poorest. The twelve drawings on which there was the greatest agreement as to their falling midway between the extremes of sexual differentiation were given to the judges and the three on which there was the greatest agreement were chosen to represent the midpoint. The same procedure was followed to select three drawings falling exactly between the poor end of the scale and the midpoint and three falling exactly between the midpoint and the best end of the scale. This produced a scale consisting of five points with three drawings as examples of each point, giving a total of fifteen drawings on the scale. The judges were asked to write down the characteristics of each drawing. The descriptions were condensed and used to describe each point on the scale. A reliability study showed the scale to have a reliability coefficient of .84. Sexual differentiation as measured by this scale did not appear to correlate significantly with another similar visual motor test, the Bender-Gestalt (.24 correlation at the .05 level of confidence). Hospitalized mental patients (fifty-eight of them) had significantly poorer sexual differentiation (at the .05 level of confidence) on the Draw-a-Person than fifty-eight clinic patients treated in an out-patient clinic.

Using the above-mentioned scale of sexual differentiation, Swenson and Newton[29] judged the Draw-a-Person drawings of 165 grade school students (grades one to eight) and twenty-two college students. Sexual differentiation was shown to increase with age. Up to the eighth grade the girls differentiated significantly better than the boys. Beyond the eighth grade there was no significant difference between the groups in sexual differentiation. The tendency of subjects to draw their own sex first did not appear to be related to increased sexual differentiation on the drawings.

Although the H-T-P was the test administered, Waxenberg[30] used only human figure drawings in his comparison of the drawings of three groups of physically ill women with twenty subjects in each group. The asthma group had a median age of 36.5; the colitis group, 32.5; and the tumor group, 40.8. After being administered the H-T-P, subjects were asked to draw a figure of the opposite sex to that first drawn. In psychosomatic formulations asthmatics are regarded as passive, dependent, and very much identified with

---

29 Clifford H. Swenson and Kenneth R. Newton, "The Development of Sexual Differentiation on the Draw-a-Person Test", Journal of Clinical Psychology, Vol. 11, No. 4, October, 1955, p. 417-419.

30 Sheldon E. Waxenberg, "Psychosomatic Patients and Other Physically Ill Persons: A Comparative Study", Journal of Consulting Psychology, Vol. 19, No. 3, June 1955, p. 165-169.

supportive, maternal figures, while women afflicted with ulcerated colitis are regarded as assertive, dominating, and often career-orientated. The personal projections obtained in the human figure drawings were used to test these psychosexual differentiations. It was hypothesized that women asthmatics in carrying out the figure drawing task would draw the female figure before the male figure and tend to draw the female larger in vertical dimension than the male figure, while women with ulcerative colitis would reverse the process. The results showed that the female was drawn first by thirteen asthma, sixteen colitis, and fourteen tumor subjects. The female figure was drawn larger by eleven asthmatics, as against four who drew the male larger; by six colitis patients as against six who drew the male larger; and by ten tumor patients as against two who drew the male larger. The Chi Square method rendered no significant differences between any of the groups.

Silverstein and Robinson,[31] in 1956, attempted to discover if children with poliomyelitis reflected their physical deformity in human figure drawings. The drawings of twenty-two children (ages nine through twelve years and equally divided by sex) in the chronic stage of the disease were studied by three methods. The first method was

---

31 A.B. Silverstein and H.A. Robinson, "The Representation of Orthopedic Disability in Children's Figure Drawings", Journal of Consulting Psychology, Vol. 20, No. 5, October, 1956, p. 338-341.

inspection; more than three fourths appeared to represent their disability either directly or indirectly. Comparison of the drawings with those of forty-four normal children was the second method. The two groups, which were equated for age, sex, and intelligence, showed no significant differences on a series of fifty-five scoring items. Finally, judges experienced in the diagnostic use of figure drawings were unable to differentiate the drawings of the disabled and the normal subjects at a level better than chance. The youth of the subjects may have contributed to the negative results because Harris and Goodenough[32] state that, on the basis of large scale investigations with children's drawings, it is somewhat doubtful whether children project their body image in figure drawings.

In 1956, Graham[33] studied the reliability of human figure drawings. Twenty-three subjects in a graduate school (12 male and 11 female) were asked to perform the Draw-a-Person Test. They were then given a two hour lecture on the interpretation of human figure drawing and told to repeat the test. The majority of the subjects made little or no change in the second figure drawing. In instances where

---

32 Silverstein and Robinson, Op. Cit., p. 339.

33 Stanley R. Graham, "A Study of Reliability in Human Figure Drawings", Journal of Projective Techniques, Vol. 20, No. 4, December, 1956, p. 385-386.

radical changes were attempted no subject was able to "improve" his production. Evaluated as a group the means of the before and after drawings were shown to be statistically identical, and the correlation of paired drawings demonstrated a positive relationship significant at better than the .01 level of confidence.

Goldstein and Rawn[34] tested the hypothesis that experimentally induced feelings of aggression interpolated between two figure drawing presentations will elicit the following in the second drawing: 1) greater line pressure; 2) greater figure size; 3) greater number of specific details generally interpreted as representing aggression; and 4) greater overall subjective impression of aggression. Aggression was experimentally induced in thirty-nine male and female attendants at a state mental hospital by telling them that their work would be increased from forty-four to forty-eight hours per week with no pay increase. All results were negative (less than the .05 level of confidence) except for seven drawing details which as a group did relate to aggression. The rationale for these results offered by the authors is that aggression is shown in a symbolic expression rather than a graphomotor expression.

---

34 Arnold P. Goldstein and Moss L. Rawn, "The Validity of Interpretive Signs of Aggression in the Drawings of the Human Figure", Journal of Clinical Psychology, Vol. 13, No. 2, April, 1957, p. 169-171.

In 1958, Lorge et al.[35] tested the hypothesis that drawings of older people show increasing loss in intactness. The younger group was composed of seventy-five graduate students (mean age of thirty-two years), and the older group (average age of seventy-three years) was composed of 104 subjects from institutions for the aged and recreational day centers. In contrast to the human figure drawings by the younger adults the drawings of the older people were characterized by incompleteness, lack of integration, lack of proportion, bizarreness, evidence of inadequate motor coordination, and difficulty in identifying the sex of the respondent. Also the drawings of the younger group were quite homogeneous in comparison to those of the older group. It appears that age would not be the only variable responsible for these results.

In 1959, Starr and Marcuse[36] investigated the reliability of human figure drawings in three ways: the reliability of using the same and different examiners, the effects of different time intervals between drawings, and the reliability of the test-retest method. The characteristics

---

35 Irving Lorge, et al., "Human Figure Drawings by Younger and Older Adults", Journal of Clinical Psychology, Vol. 14, No. 1, January, 1958, p. 54-56.

36 S. Starr and F.L. Marcuse, "Reliability in the Draw-a-Person Test", Journal of Projective Techniques, Vol. 23, No. 1, March, 1959, p. 83-86.

analyzed for reliability were: 1) placement on the page; 2) sex drawn first; 3) presence or absence of sexual symbols; 4) perspective: front profile; 5) incompletions; 6) height of the drawing; and 7) the size of the head in relation to body height. The authors utilized the Chi Square method for characteristics one to five and product-moment correlation coefficients for characteristics six and seven. The results showed that a change in examiner was not important and the time interval between test and retest had no effect. Five factors were found to be reliable at the .01 level of confidence for both sexes: perspective, position on the page, incompletions, height of the figure, ratio of head size to the height of the figure; one factor was found unreliable for both males and females (presence or absence of sexual symbols), and one factor was found reliable for males only (sex drawn first).

Hunt and Feldman,[37] in 1960, investigated temporal and sex differences in attitudes towards parts of the body and the relationship between these attitudes and human figure drawings. A modification of the Secord and Jourard Body-Cathexis Scale was administered to sixty-five college students (39 male and 26 female) who were then instructed

---

37 Raymond G. Hunt and Marvin J. Feldman, "Body Image and Ratings of Adjustment on Human Figure Drawings", Journal of Clinical Psychology, Vol. 16, No. 1, January, 1960, p. 35-38.

to draw an unclothed human figure. In the Body-Cathexis scale the subjects were asked to rate the twenty-five body parts twice--as they felt about them at the time of the test and as they felt about them during early adolescence. Three clinicians rated the figure drawings on presence or absence of disturbance for each body part. The results showed that the present Body-Cathexis ratings of these students were consistently more favorable (t of 2.73 was significant at the one per cent level of confidence); however, the individual body parts maintained the same relative position in the two sets of ratings. Women's ratings were found to be more variable at both time periods. The correlation between Body-Cathexis scores and signs of disturbance on the human figure drawings was minimal.

In 1960, Bodwin and Bruck[38] using thirteen drawing characteristics thought to reveal self concept examined the Draw-a-Person tests of sixty subjects aged ten to seventeen years and rated them on a five point rating scale. The subjects were also independently rated by judges after a psychiatric interview. The judges' rating and the self concept scale correlated significantly at the .01 level of confidence.

---

38 Raymond F. Bodwin and Max Bruck, "The Adaptation and Validation of the Draw-a-Person Test as a Measure of Self-Concept", Journal of Clinical Psychology, Vol. 16, No. 4, October, 1960, p. 427-429.

Haworth and Normington[39] adapted the Swenson scale for children which can be scored against objective criteria. Pairs of drawings were secured from 312 children, ages seven to twelve, with approximately fifty children at each age level. Four levels of sex differentiation were recognized ranging from no apparent sex in either figure to clear and adequate differentiation in both figures. Normative data showed a gradually increasing ability in differentiation, with girls consistently superior to boys not only in the level used but in the emphasis on own sex figure as well. Chronological age was a more pertinent factor than mental age in determining the use of the higher levels. Reliability between the scores was .88. The authors felt that the scale provides a developmental index of psychosexual maturity, but stated that further investigation is needed to determine to what extent the ability to differentiate between the sexes in figure drawings can be considered a measure of sex-role identification.

Laird[40] administered the H-T-P but used only the Person to test the hypothesis that there is no significant

---

39 M.R. Haworth and C.J. Normington, "A Sexual Differentiation Scale for the D-A-P Test", Journal of Projective Techniques, Vol. 25, No. 4, December, 1961, p.441-450.

40 James T. Laird, "A Comparison of Male Normals, Psychiatric Patients, and Alcoholics for Sex Drawn First", Journal of Clinical Psychology, Vol. 18, No. 3, July, 1962, p. 302.

difference between male normals, male psychiatric patients, and male alcoholics for drawing the opposite sex first. Of the 132 normals tested, 125 (94.7 per cent) drew their own sex first; of the seventy-one psychiatric patients, 60 (84.5 per cent) did so; and of one hundred alcoholics, 81 (81 per cent) drew their own sex first. Chi Square comparisons indicated that the psychiatric patients drew the opposite sex significantly more often than did the normals ($P < .02$); that alcoholics drew the opposite sex significantly more than normals ($P < .01$); and that the psychiatric patients did not draw the opposite sex significantly more often than the alcoholics ($P < .50$). The results also indicated that while psychiatric patients and alcoholics differ from normals for drawing the opposite sex first, they do not differ from one another.

In a similar study, Laird[41] used females as subjects but used the same experimental design. Of 175 normal women, 56 per cent drew their own sex first; seventy-five per cent of 20 psychiatric patients did so; and of twenty alcoholics, 50 per cent drew their own sex first. Chi Square comparisons indicated that neither the psychiatric patients and the normals, the alcoholics and normals, nor the alcoholics

---

41 James T. Laird, "A Comparison of Female Normals, Psychiatric Patients and Alcoholics for Sex Drawn First", Journal of Clinical Psychology, Vol. 18, No. 4, October, 1962, p. 473.

and psychiatric patients differed significantly in drawing their own sex first.

From these studies have arisen hypotheses and theories related to the phenomenon of drawing as a projective technique. Before the H-T-P is discussed it is appropriate to mention the theoretical constructs which have evolved from many of the previously mentioned studies and upon which the H-T-P is based.

3. Theoretical Considerations.

While a categorically agreed upon theory of projective tests is lacking at present, some general statements can be made. The theoretical origin of projective techniques emanates from Freud's discovery of the importance of the unconscious element in personality. Freud viewed projection as a defense mechanism by which the source of anxiety was attributed to the external world rather than to the individual's own primitive impulses or to the threats of conscience. However, the concept of projection as it is applied to techniques today takes on a wider scope. The belief is that the place to look for deep motivation, for basic character structure, and for conflicts is in a person's fantasies; these are best revealed not by direct questions but by veiled, circuitous projective probes. The unconscious does not assert itself directly or patently but rather it emits

messages which reach consciousness only after being clothed in a symbolic disguise to merit acceptance by the individual. This disguise can have either private or universal meaning and often appears in the creative aspects of behavior. Part of the function of projective techniques is to interpret the conscious manifestations in the light of unconscious needs, wishes, and drives.

It is generally agreed upon that certain accepted concepts of personality underlie the use of projective techniques. One is that personality is a dynamic process which must be measured by instruments that can detect modifications over a period of time as well as the present personality. Such measurement is felt to be possible because, while the personality is not static, it is structured. Despite disagreement as to how organization takes place, the general consensus is that personality is structured and reveals itself in behavior since behavior is functional. The manifest responses of the individual in specific situations are psychologically, if not logically, consistent with the expression of the personality in other situations. Projective techniques are attempts to glean from a sample of behavior an overview of the personality.

The field of projective drawing interpretation rests on several sources. One is the use of common psychoanalytic and folklore meanings of symbols derived from clinicians'

study of dreams, art, mythology, and fantasy. Gestalt theory has also contributed to interpretation with its theory of perception for the Gestalt idea of perception is used in the belief that individual responses are organized into a total pattern the discovery of which is the goal of the experimenter. From Gestalt theory can also be obtained the content of motivations, frustration, adjustment and the fields in which they occur. Cultural anthropology is also a source of interpretation in that its study of the effect of an individual's culture and environment has rendered insight into the individual as a whole person and helped remove the mechanistic bias of earlier personality theories. Learning theory has contributed the concepts of conditioning, reward and punishment as motivating forces, and the analysis of the physiological processes of habit formation. The most basic source of interpretation is clinical experience, especially in the realms of neurosis and psychosis, while experimental studies help limit exaggerated or biased interpretations or "hunches".[42]

Regarding projective drawings in particular, Machover[43] felt that each deviation in the drawing of a

---

42 John Elderkin Bell, Projective Techniques, New York, Longmans, Green, 1948, p. 7-11.

43 Machover, Op. Cit.

person has a significant meaning in terms of conscious and unconscious forces, of which the unconscious predominate. She stated that because drawing is a creative experience it reveals the conflicts, fantasies, and inadequacies of the creator. Machover also felt that in drawing a human figure the individual is compelled to draw his body image as it has developed from past personal experience. This concept was substantiated by Schilder[44] and Fisher and Cleveland[45] who stated that emotional influences, organic pathology, or physical deformities could cause a person to make distortions in the perception of identical parts in the bodies of others and that the body image and parts are saturated with the emotional and ideational experiences associated with their development in the life space of the individual.

Until the time of Buck's[46] H-T-P test in which he considered each picture drawn as a self portrait as well as the drawing of a specific object, drawings other than that of the Person were not considered psychological and physical

---

44 Paul Schilder, Image and Appearance of the Human Body, New York, International University Press, 1950, p 5-353.

45 Seymour Fisher and Sidney E. Cleveland, Body Image and Personality, Princeton, D. Van Norstrand, 1958, v-420 p.

46 John N. Buck, "The H-T-P Technique: A Qualitative and Quantitative Manual", Journal of Clinical Psychology, Monograph Supplement No. 5, Vol. 4, No. 4, October, 1948, p. 320.

self portraits. In addition to this, Buck[47] offered eleven other theoretical postulates for the H-T-P, the validity of which he considered well demonstrated in clinical usage during ten years prior to the first presentation of the majority of these postulates in 1948. The postulates related to this study are: 1) The H-T-P is a projective device. 2) Each drawing is believed to arouse conscious, subconscious, and unconscious associations. 3) Any emotion exhibited by the subject while drawing is presumed to present his emotional reaction to the relationships, situations, needs, or presses that he feels directly or symbolically represented by the drawing or some phase of the drawing. 4) In the H-T-P details do not necessarily have universal and absolute meaning. 5) Interpretation of a given item can be made accurately only after its relationship to the entire configuration has been determined. 6) In drawings of the H-T-P personality characteristics are represented as they appear in the effector processes in contrast to the perceptual processes.

With the first three sections of this chapter serving as an introduction more meaningful examination of the H-T-P can now be considered. First, the test itself with its

---

47 John N. Buck, Administration and Interpretation of the H-T-P Test, Beverly Hills, California, Western Psychological Services, 1950, p. 1-67.

various facets and details will be discussed followed by a review of the H-T-P literature.

4. The H-T-P.

The free hand drawing of the House, Tree, and Person was published by Buck[48] as a standardized test in 1948. The H-T-P can be administered achromatically or chromatically. Achromatically the test is administered with No. 2 lead pencils on seven by eight and one-half inch sheets appropriately labelled and, according to proper procedure, is followed by a Post Drawing Interrogation consisting of sixty questions. The chromatic H-T-P requires a set of eight colored wax crayons instead of the lead pencil. Buck[49] regards the H-T-P as an instrument which elicits data concerning the dynamics and degree of integration of an individual's personality and which reflects strengths, weaknesses, and the extent to which a person can mobilize his inner resources to handle his psychodynamic conflicts. It is felt that the H-T-P is also useful as a forewarner of pathological onset and as a prognostic indicator.

---

48 John N. Buck, "The H-T-P Technique: A Qualitative and Quantitative Manual", Op. Cit., p. 320-396.

49 Ibid., p. 320.

There are three systems by which data from the H-T-P can be studied. The Quantitative approach[50] is used to establish the general intelligence of a person and renders four scores purporting to include different aspects of intelligence. The Qualitative system[51] was established to identify and evaluate those items which, in most instances, did not appear to differentiate regarding intelligence, but did differentiate between drawings of persons who did not exhibit a major personality maladjustment and those who were maladjusted, psychopathic, neurotic, or psychotic. In the standardization work on this system, 150 subjects were tested and included seven categories of maladjustment: adult maladjustment, epilepsy with personality maladjustment, psychopathic personalities, psychoneurosis, prepsychotic state, mental deficiency with psychosis, and psychosis. Buck states that while the population was not a well balanced one "it did serve to indicate very definitely that the H-T-P productions of subjects with personality disorders would differ in many respects from drawings produced by subjects who were not maladjusted".[52] The third system, the Quality

---

50 John N. Buck, "The H-T-P Technique: A Qualitative and Quantitative Manual", Op. Cit., p. 321.

51 Ibid., p. 325.

52 Ibid., p. 326.

of the Quantity,[53] is regarded as bringing into closer and more efficient relationship the previous systems by providing a relatively objective check upon the quantitative appraisal and an opportunity for the qualitative extension of quantitative scores. This is done by extending the quantitative scoring method.

Regarding the Qualitative system, with which this study is concerned, Buck[54] stated in 1948 that many qualitative differential points were to be discovered and that the points already identified were not necessarily correct or complete. Since then, authors have been attempting to improve the situation. The review of the H-T-P studies, unlike the previous studies that were presented chronologically, will be presented as much as possible according to content.

The first area to be considered is studies dealing with children. Jolles,[55] in 1952, obtained 2560 sets of drawings from children ranging from five to 18 years of age

---

53 John N. Buck, "The Quality of the Quantity of the H-T-P", Journal of Clinical Psychology, Vol. 7, No. 4, October, 1951, p. 352-356.

54 ---------, "The H-T-P Technique: A Qualitative and Quantitative Manual", Op. Cit., p. 327.

55 Isaac Jolles, "A Study of the Validity of Some Hypotheses for the Qualitative Interpretation of the H-T-P for Children of Elementary School Age: I. Sexual Identification", Journal of Clinical Psychology, Vol. 8, No. 1, January, 1952, p. 111-118.

and living in a cross-section of various socio-economic groups in Illinois. The results showed that Buck's hypothesis that subjects tend to draw their own sex applied to this population and that Buck's hypothesis that the sex of the drawn Person represented felt sex role is applicable to children. The author states that one has to be particularly careful in evaluating the significance of the drawn Person of the opposite sex among five, six and seven year old males and among eleven and twelve year old females since the results suggest that such trends are typical at these ages. Jolles' study has been criticized in the literature for making inferences about psychosexual development and sexual identification which are not clear. Sloan[56] states that it is a moot question whether the subject identifies with the drawn Person since not all projective psychologists subscribe to this notion. Jolles presents no evidence to justify his conclusion that the sex of the drawn Person tends to represent the felt sex role of the subject. He fails to make explicit the inferences and to point out the steps that need to be taken in going from a bit of observed behavior to a hypothetical construct such as psychosexual development.

---

56 William Sloan, "A Clinical Review of H-T-P Validation Studies", Journal of Clinical Psychology, Vol.10, No. 2, April, 1954, p. 143-148.

In a second study in 1952, Jolles[57] investigated Buck's hypothesis that a phallic Tree denotes phallic preoccupation of some type. Jolles' aim in this study was to discover whether this hypothesis would hold for children, particularly young children or older children with mental ages of young children since this type of Tree occurred frequently in the drawings of mental defectives. Trees were divided into three classes: 1) absolute phallic Tree; 2) phallic Tree suggested; 3) phallic Tree not implied. These results were checked with other items in the drawings which might be indicative of psycho-sexual conflict. The population consisted of 1393 males and 1308 females from five to twelve years old. A negative correlation was found between the age of the subject and the absolute phallic Tree and positive correlation between age and the non-phallic Tree. It was suggested from this that as the child grows older he is less likely to draw an absolute phallic Tree and conversely more likely to draw a non-phallic Tree. Results also showed that girls tend to draw more phallic Trees than boys. The check for sexual significance of the phallic Tree by comparing its presence or absence to other

---

57 Isaac Jolles, "A Study of the Validity of Some Hypotheses for Qualitative Interpretation of the H-T-P for Children of Elementary School Age: II. The 'Phallic Tree' as an Indicator of Psycho-sexual Conflict", Journal of Clinical Psychology, Vol. 8, No. 3, July, 1952, p. 245-255.

hypothesized symptoms of psychosexual disturbance in the drawings yielded a Chi Square of 284.5 significant at less than the one per cent level of confidence. This study is criticized[58] for supposedly investigating Buck's statement concerning the meaning of a phallic Tree as a hypothesis and then beginning the study by postulating that a phallic Tree indicates an overconcern with sexual matters. Jolles seems to accept and take for granted that such an interpretation is correct and is concerned with determining if this interpretation applies to a young child without considering that Buck's statement may not be correct. It is felt that since experimental evidence does not clearly show that psychosexual development occurs in exactly the way in which psychoanalytic theory postulates, the tying in of the drawings is synthetic as a validity criterion.

Jolles'[59] third children's study attempted to establish the psychological center of horizontal placement in drawings for children from five to twelve years old and to validate the assumption that horizontal placement is a sign of an individual's degree of intellectual control over his affect. Drawings of 1046 males and 1037 females were used,

---

58 Sloan, Op. Cit.

59 Isaac Jolles and Harry S. Beck, "A Study of the Validity of Some Hypotheses for the Qualitative Interpretation of the H-T-P for Children of Elementary School Age: III. Horizontal Placement", Journal of Clinical Psychology, Vol. 9, No. 2, April, 1953, p. 161-164.

and the results indicated that Buck's hypothesis was generally correct--that the psychological center is to the left of the geometric center. However, in this study the psychological center was not as far left as Buck has indicated for adults. Horizontal placement tended to vary with age and with the drawn whole. The relation between horizontal placement and age was significant according to an analysis of variance at less than the .05 level of confidence. One would expect this if Buck's hypothesis that the further to the right the drawing is the greater the degree of intellectual control is true. However, an analysis of variance does not indicate in which direction age is a factor in horizontal placement. Sex differences in relation to the horizontal placement of the three drawings were not significant according to an analysis of variance. This tends to contradict Buck's hypothesis that drawings to the left of center indicate femininity and to the right, masculinity. The results also showed that the males were more free with their emotions in interpersonal relations than girls. Girls tend to react emotionally to the environment in general to a greater degree than boys as would be indicated by the higher negative correlation of horizontal placement with age in the drawing of the Tree. This study

is criticized[60] for reasoning by analogy. One could almost set up alternative hypotheses and proceed to justify them with as much reasoning as has been done with the hypothesis presented.

In the final article of the series, Jolles and Beck[61] investigated the relationship between age and vertical placement on the H-T-P as a followup of the third study. The authors also wished to investigate the validity of Buck's hypothesis that vertical placement of the drawing above the psychological center suggests an attempt to gain satisfaction from fantasy and/or striving and that vertical placement below the psychological center suggests a tendency to be reality bound, a tendency to feel insecure and/or a depressive mood tone. On the basis of clinical experience with the Rorschach, Jolles and Beck felt that the ability to seek satisfaction in fantasy increases to a certain extent with age (older children expected to produce more M). Jolles and Beck applied this to Buck's interpretation of vertical placement and expected a positive correlation of height

---

60 Sloan, Op. Cit.

61 Isaac Jolles and Harry S. Beck, "A Study of the Validity of Some Hypotheses for the Qualitative Interpretation of the H-T-P for Children of Elementary School Age: IV. Vertical Placement", Journal of Clinical Psychology, Vol. 9, No. 2, April, 1953, p. 164-167.

of drawings on the page with age. The drawings of 1013 male and 1022 female children were studied, and an analysis of variance indicated that age is a significant factor in vertical placement at less than the one per cent level of confidence although obtained means failed to verify Buck's hypothesis that the psychological center for vertical placement is above the geometrical center of the page. Jolles stated that the results were in keeping with his clinical experience with the Rorschach, and the authors also felt they showed that the hypothesis that placement above the psychological center suggests seeking of satisfaction and degree of striving is sound. According to the results of the analysis of variance the sex of the subject was a factor in vertical placement. The mean vertical placement was higher for girls on all three drawings. This the authors interpret as suggesting a greater tendency for girls to seek satisfaction in fantasy or a greater degree of striving among them. This study, Sloan[62] states, does not clarify why the "M" on the Rorschach is used as a criterion of fantasy life since it is just as readily a criterion of other factors. The reasoning concerning the relationship is weak since it might logically be suggested that the finding that higher placement of the drawings on the page increased with

---

62 Sloan, Op. Cit.

age is directly proportional to the mental age of the subjects. Just why it would have to be fantasy is not clear.

Hammer,[63] in studying frustration and aggression, administered H-T-P's to four hundred Negro and white children ranging in grade level from first to eighth. Six clinicians without knowing whether they were rating drawings by Negro or white subjects rated each set of drawings on a scale of aggression consisting of three points: none, mild, and severe. The criterion used for the judgment of aggression was a list of twelve qualitative items which were hypothesized to indicate aggression, placement of the drawing on the page, and an examination of the entire drawing to take the total constellation into consideration. It was assumed by the author that the Negro group would show more aggression because the outer white world is often found to be full of disappointment, frustration, and threat. The results showed that the hostility rating for all eight grades of white children was .308. The mean hostility rating for all eight grades merited by the drawings of the Negro children was .823. A t score of 12.56 was statistically significant at far better than the one per cent level of confidence

---

63 Emanuel F. Hammer, "Frustration-Aggression Hypothesis Extended to Socio-Racial Areas: Comparison of Negro and White Children's H-T-P's", The Psychiatric Quarterly, Vol. 27, No. 4, October, 1953, p. 597-607.

and indicated that a difference exists between the degree of need for aggression in the Negro and white groups, with the incidence being higher in the Negro group. Frustration was measured according to Buck's hypothesis that drawings made conspicuously too large for the page without adequate space framing them indicate frustration. When the incidence of such drawings was tabulated quantitatively, it was found that 10.3 per cent of the white children presented such drawings whereas 28.3 per cent of the Negro children did so. A *t* score of 3.84 indicated a statistically significant difference at the one per cent level of confidence in regard to this factor. Correlations among the judges ranged from .74 to .84. The results appear to support the assumption that Negro children possess more need of aggression than white children. However, the results of this study were obtained from students of two schools of one community; hence, it is doubtful that a cross-section of intellectual and socio-economic areas was reached.

In another study with children, Beck[64] attempted to validate Buck's definition of what constitutes essential details in the drawn House as it applies to normal, mentally

---

64 Harry S. Beck, "A Study of the Applicability of the H-T-P to Children with Respect to the Drawn House", Journal of Clinical Psychology, Vol. 11, No. 1, January, 1955, p. 60-63.

handicapped, and organic children. The normal group consisted of 212 five year old boys and girls and 593 six year olds. There were thirteen boys and girls (mean age 9.8, mean I.Q. 69.48) in the mentally handicapped group, and twenty-five boys and girls (mean age 11.2, mean I.Q. 56.77) in the organic group. The results indicated that the addition of irrelevant details apparently is a chance factor associated with the individual's need to structure the situation. Since the two experimental groups did not differ significantly, the absence of essential details or unrecognizability in the drawing of a House could not be used as a means of differential diagnosis with the mentally handicapped individuals. The author stated that it would seem that an unrecognizable House or one with three or more essential details missing in a child of six years and over and of normal intelligence would at least be a suspected sign of organicity. Hence, it was concluded that Buck's definition of what constitutes essential details in a House is applicable to the children in this study. Buck's interpretation of the general omission of essential details was also applicable to this group. It was concluded that the inclusion of details in the drawing of a House is to some extent a developmental process which seemingly matures at about the age of six years insofar as essential details are concerned, and therefore great caution should be used in

making any interpretations with children below the age of six years. Since the non-organic group was particularly small, the fact that there were no demonstrable differences in the experimental groups should not be taken as conclusive evidence.

Wawrzaszek et al.[65] administered the H-T-P to forty-one physically handicapped children and a matched group of non-handicapped children followed by a shortened version of the Post-Drawing Interrogation. The ten PDI questions which the authors hypothesized would distinguish between the two groups was the criterion used. None of the PDI questions significantly differentiated the groups.

Orgel[66] investigated the relationship of the H-T-P to a sociometric evaluation of thirty-two primary school children using a checklist of thirty-seven H-T-P items devised from Buck's qualitative evaluation procedure. The selection of items was based on the assumption that certain aspects of perspective, proportion, and detailing, as well as omissions and distortions in the drawing, would reflect

---

65 Frank Wawrzaszek et al., "A Comparison of H-T-P Responses of Handicapped and Non-Handicapped Children", Journal of Clinical Psychology, Vol. 14, No. 2, April, 1958, p. 160-162.

66 Rita G. Orgel, "The Relationship of the H-T-P to a Sociometric Evaluation of the Group of Primary Grade School Children in Determining the Degree of Social Acceptance", Journal of Clinical Psychology, Vol. 15, No. 2, April, 1959, p. 222-225.

the child's perception of and interaction with the environment. The items were correlated with a sociometric questionnaire used as a measure of personal popularity. The entire checklist correlated .271 with the sociometric evaluation which was not significant at the .05 level of confidence. The checklist items of the House and Person drawings correlated with the sociometric evaluation were significant at the .05 level of confidence. The Tree showed a negative correlation of -.173.

The next content area of studies will deal with psychosexual dynamics as they are manifested in the H-T-P. The first of four studies by Hammer[67] investigated sexual symbolism. The H-T-P drawings of twenty subjects undergoing surgical sterilization were compared with those of twenty subjects undergoing various other operations. On the assumption that subjects who are about to undergo eugenic sterilization would be prone to respond with feelings of castration, the drawings of the two groups were compared for indices of genital symbolization and castration feelings extracted from psychoanalytic theory. Significant statistical differences were found between the two groups on twenty-six of the fifty-

---

67 Emanuel F. Hammer, "An Investigation of Sexual Symbolism: A Study of H-T-P's of Eugenically Sterilized Subjects", Journal of Projective Techniques, Vol. 17, No. 4, December, 1953, p. 401-413.

four drawing items assumed to be sexual symbols which were used as a criterion. The assumption that the preoperative group would exhibit castration feelings appears weak since the operation in fact is not one of castration. Also, the small population lessens the importance of the results.

In another study, Hammer[68] compared the H-T-P drawings of sixty-four incarcerated sex offenders, thirty-one rapists (ages 24 to 60) and thirty-three pedophiles (ages 23 to 63). After drawing the Tree each subject was asked to assign an age to it. The ages assigned to the drawn Trees of pedophiles were then compared to those of the rapists with the latter serving as a control group for such variables as detection and incarceration for sex offenses and the sharing of a common prison environment. The mean age projected onto the Trees by rapists was 24.4 and by pedophiles, 10.6. After drawing the Person and a figure of the opposite sex of the first drawn Person the subjects were asked to specify the apparent ages of both the male and female drawn. Both groups drew the female person as older than the male but only with the pedophile group was the difference in assigned age statistically significant. It

---

68 Emanuel F. Hammer, "A Comparison of H-T-P's of Rapists and Pedophiles", Journal of Projective Techniques, Vol. 18, No. 3, September, 1954, p. 346-354.

was concluded from these results that the age with which the drawn Tree is endowed may serve as a rough index of the subject's felt level of psychosexual maturity. Also the author concluded that one of the motivational mainsprings operative in the pedophile's sexual behavior is a psychosexual fixation on, or regression to, a childhood level. This regression to an intermediate level encourages the pedophile to seek immature sex objects of approximately the age at which he feels psychosexually adapted. To the pedophile, adult women appear as older or maternal individuals endowed with greater relative status. Since these individuals were apprehended for sex offenses the degree of distortion in the sample brought about by this selective factor cannot be accurately determined. Whether or not there is definite involvement of guilt feelings and punishment needs in the motivational dynamics of the subjects which would be conducive to detection cannot be definitely stated which raises a question as to the selectiveness of the sample.

In a similar study, Hammer[69] investigated the relationship between psychosexual pathology and the sex of the first drawn Person. H-T-P's were administered to

---

69 Emanuel F. Hammer, "Relationship between Diagnosis of Psychosexual Pathology and the Sex of the First Drawn Person", Journal of Clinical Psychology, Vol. 10, No. 2, April, 1954, p. 168-170.

eighty-four sex offenders of which thirty-one were rapists, thirty-three were heterosexual pedophiles, and twenty were homosexual pedophiles. After drawing a House and a Tree the subject was asked to draw a Person and then a Person of the opposite sex of that first drawn. Comparisons of the three subgroups of sex offenders were made in regard to the percentage drawing the opposite sex figure first. Of the rapists, 87.1 per cent drew the male first; of the heterosexual pedophiles, 87.9 per cent drew the male first; and of the homosexual pedophiles, 75 per cent drew the male first. There were no significant differences. From these results the author stated that considerable doubt is cast on the projective drawing postulate that the sex of the first drawn figure may serve as an index of the subject's sexual identification or as evidence of psychosexual conflict. On the basis of the relatively small number of subjects employed in this study, the conclusions should be considered tentative.

Using the "dead" Tree as an index of psychopathology to test Buck's hypothesis that the psychologically sickest people see their drawn Trees as dead, Hammer[70] studied the drawings of the subjects in the previously reviewed study. After drawing the Tree, each subject was asked if he depicted

---

70 Emanuel F. Hammer, "A Comparison of the H-T-P's of Rapists and Pedophiles: III. The 'Dead' Tree as an Index of Psychopathology", Journal of Clinical Psychology, Vol. 11, No. 1, January, 1955, p. 67-69.

the Tree as dead or alive. The homosexual pedophiles had a significantly greater incidence of "dead" Trees than did the rapists. The results showed that 6.4 per cent of the rapists, 15.2 per cent of the heterosexual pedophiles, and 50 per cent of the homosexual pedophiles drew "dead" Trees. Only the results of the homosexual pedophiles attained statistical significance at the .01 level of confidence. There seemed to be a progression from rapists to heterosexual pedophiles to homosexual pedophiles in regard to the number who saw their drawn Trees as "dead", which parallels the increasing distance from an appropriate sex object. The author stated that on the one hand these parallel data tend to support Buck's hypothesis, and on the other hand tend to describe the homosexual pedophiles who deviate from the norm in both age and sex of the partner chosen as the sickest subgroup of the sex offenders studied. The logic in the author's conclusions is not too clear. Before the study he assumed that the homosexual pedophiles were the psychologically sickest group in order to prove Buck's hypothesis, and then he concluded that they are the sickest group because Buck's hypothesis is confirmed by them.

The next area that will be considered is the effect of somatic or psychosomatic debilities on the H-T-P.

Landisberg[71] studied the H-T-P's of twenty relatively high grade epileptics (15 male and 5 female) ranging in age from eleven to seventy-one years. Evaluation of the subjects' concept formation on the H-T-P helped to confirm data from other tests that showed that organic deterioration had occurred in the subjects. Body feelings of insupport and imbalance may have helped to produce such a phenomenon. Other organic signs present in the drawings were the omission of essential details, presence of detail fragments, and some details showing regression to borderline, moron, and imbecile level. Generally, patients projected body feelings of lifelessness, imbalance, robot-like propulsion, and preoccupation with sensual stimuli. Their contact with the external world was shown as rigid, non-fluid, and anxiety laden. Reality tended to be evaluated in a distorted fashion because of compensatory fantasy.

Michal-Smith[72] attempted to identify pathological cerebral functioning through the H-T-P. He used a population of fifty males ranging in age from eight to sixteen years who possessed approximately the same socio-economic status

---

71 Selma Landisberg, "A Personality Study of Institutionalized Epileptics", American Journal of Mental Deficiency, Vol. 52, No. 1, July, 1947, p. 16-22.

72 Harold Michal-Smith, "The Identification of Pathological Cerebral Function Through the H-T-P Technique", Journal of Clinical Psychology, Vol. 9, No. 3, July, 1953, p. 293-295.

and were classified in the range of dull-normal to normal intelligence. This group was divided into two groups of twenty-five subjects, one with normal cerebral electrical function and the other with abnormal function. The results indicated that "line quality" bears a significant relationship to the EEG as a predictor of psychopathology. In addition, the scoring categories taken together tended toward significance as predictors of imbalance in the electrocortical field.

Meyer, Brown, and Levine[73] administered H-T-P's to fifteen subjects before and after surgery. The preoperative drawings showed quasi-self portraits of individuals threatened by an objective danger consisting either of a disease or its surgical management. Yet, the drawings rarely supplied a direct allusion to the actual nature of the threat. The authors gained the impression that the specific danger is repressed only to return at times in a disguised form. The post-operative drawings showed an absence of impending danger reflected by the disappearance of many anxiety indications which were noted preoperatively. There was an emphasis on allusions to favourable or unfavourable changes in body and in mood wrought by the surgical procedure. The following

---

73 Bernard C. Meyer, Fred Brown, and Abraham Levine, "Observations on the House-Tree-Person Drawing Test Before and After Surgery", Psychosomatic Medicine, Vol. 17, No. 6, November-December, 1955, p. 428-454.

conclusions were made: 1) Subtle or gross paranoid attitudes were elicited by several refusals to participate in the test. 2) There was a striking dissimilarity between the pre- and post-operative drawings of the same patient. 3) Pre-operative drawings were usually characterized by multiple indications of psychological regression far exceeding clinical impression. 4) Somatic illness, the site of pathology in surgery, was rarely directly alluded to, whereas the symbolic or graphic equivalents of psychological defences were present and often most clearly expressed in the House drawing. 5) Post-operative drawings often revealed the abandonment of regressive aspects. Psychic reactions to mutilation symbolically represented by psychic defenses and changes in mood were reflected. In cases where the outcome was fortunate there was revealed a change in mood and an abandonment of several defensive measures. 6) Pre- and post-operative drawings revealed certain characterological features of the patients apart from their reaction to the acute surgical emergency. Regarding the first conclusion the fact that physically ill persons about to undergo surgery refused to participate in the test does not seem necessarily to indicate subtle or gross paranoid attitudes.

Bieliauskas and Kirkham[74] restudied some findings of Michal-Smith[75] in order to reach more conclusive evidence as to the validity of the H-T-P "organic signs", and to test some additional "organic signs". H-T-P's were obtained from twenty organic and twenty non-organic institutionalized subjects matched for sex, age, period of institutionalization, and intelligence. The criteria were eighteen signs suggested by Jolles and Buck. None of the statistical results showed that the eighteen criteria claimed to be indicative of organicity discriminated either individually or compositely the organics in the population of this study.

In an investigation dealing with the scar drawn on the Tree, Levine and Galanter[76] first tested Buck's assumptions that blemishes on a Tree denote a trauma and that the ratio between the height of the traumatic indicator and the total height of the Tree will yield the approximate age of the individual when the trauma occurred. Drawings of the Tree from twenty-seven hospitalized paraplegic veterans

---

74 Vytautas Bieliauskas and Sandra L. Kirkham, "An Evaluation of the 'Organic Signs' in the H-T-P Drawings", Journal of Clinical Psychology, Vol. 14, No. 1, January, 1958, p. 50-54.

75 Michal-Smith, Op. Cit.

76 Murray Levine and Eugene H. Galanter, "A Note on the Tree and Trauma Interpretation in the H-T-P", Journal of Consulting Psychology, Vol. 17, No. 1, February, 1953, p. 74-75.

ranging in age from eighteen to forty years were studied. It was assumed that the severity and the circumstances of the injury were sufficient to constitute a trauma. The presence of the indicators was independently rated by the authors who agreed on eighty-two per cent of the cases. Of the twenty-seven drawings only seven contained unequivical traumatic indicators. The correspondence with the trauma and the date of injury was poor; none were within two years of the actual trauma, the time limit suggested by Buck. A rank order correlation showed a correlation of .73 between actual trauma and the estimated age of the occurrence on the Tree which was significant for the seven cases at the .05 level of confidence.

In another study of the Tree trauma, Lyons[77] asked fifty subjects to put a scar on their H-T-P Tree and to report the worst event of their lives; twenty-one of these subjects were also asked for the best event of their lives. The correlation between the relative height of the scar on the Tree and the relative age of the worst event in the person's life was .54 and significant at the .01 level of confidence. Correlations between the scar location and the best event in the subject's life was .10 and not significant.

---

77 Joseph Lyons, "The Scar on the H-T-P Tree", Journal of Clinical Psychology, Vol. 11, No. 3, July, 1955, p. 267-270.

The distribution of the location of induced scars was shown to differ significantly from the results of two other similar but "neutral" tasks including drawing a horizontal line across a vertical line and marking an otherwise unidentified drawing of a Tree.

In a study of the prognostic value of the H-T-P, Hammer,[78] basing his statements on clinical experience, stated that a positive prognosis on the H-T-P is suggested by either: 1) the drawn Tree conveying a healthier impression than the drawn Person; 2) the crayon drawings indicating a better adjustment than the pencil drawings; or 3) the H-T-P representing a healthier personality picture than the Rorschach. When such relationships occurred a reactive maladjustment (for example, war neurosis, reactive depressions) in which latent positive resources were being overshadowed by the effects of an emotional upheaval later proved to be the clinical diagnosis. On the other hand, a negative prognosis, Hammer stated, is suggested by either: 1) the chromatic set showing more pathology than the achromatic set; 2) the drawn Tree conveying a sicker impression than the drawn Person; or 3) the H-T-P carrying negatively toned feelings more predominantly than the Rorschach. When such relationship existed later follow-up disclosed clinical diagnosis

---

78 Emanuel F. Hammer, "The Role of the H-T-P in the Prognostic Battery", Journal of Clinical Psychology, Vol. 9, No. 4, October, 1953, p. 371-374.

of incipient, latent, preschizophrenic or severe neurotic states.

The prognostic value of the H-T-P was also studied by Cowden et al.[79] who administered eight different tests, including the H-T-P, to fifty-eight patients and at a later date reexamined them with the same tests in an attempt to predict adjustment outside the hospital. Follow-up data of the hospital's Social Service, including a statement of re-admission, type of adjustment outside hospital, and length of time out of the hospital were used as a criterion. Ninety days out of the hospital was selected as the criterion for placing a subject in the experiment (discharge) group rather than in the control (remain in hospital) group. Patients' test data were presented to four psychologists in separate pairs, and the production on the retesting was compared with the subject's earlier testing. The judges were to evaluate the pair and state which was the better effort on the hypothesis that a discharge production should be better than an admission test battery. The total of correct judgments was summed. Where the prediction of adjustment outside the hospital was required, the tests were reassembled in the original battery and then given to five judges who

---

79 Richard C. Cowden et al., "The Prognostic Value of the Bender Gestalt, H-T-P, TAT, and Sentence Completion Test", Journal of Clinical Psychology, Vol. 11, No. 3, July, 1955, p. 271-273.

were asked to determine whether the tests were by the experimental or discharge group in order to establish if the judges could determine patients ready for discharge. In the statistical analysis of the probability value of the tests in the battery in revealing improvement, the H-T-P had a Chi Square of 12.61 significant at the .01 level of confidence.

The concluding studies deal with miscellaneous uses of the H-T-P. Hammer and Piotrowski[80] investigated hostility in the clinician's personality as it affects his interpretation of the H-T-P. The degree of a clinician's hostility as manifested in his interrelationships with patients and staff members was measured by one of the writers, then the supervisor of the three clinicians used in the study. The validity of this judgment was supported by a correlation with the clinicians' Szondi protocols. The three clinicians were asked to judge the H-T-P's of four hundred elementary school children on a scale of aggression. Three other judges rated the drawings of the third, fifth, and seventh graders in the population. The correlations of the judgments of the three principle judges were computed, and all six judges were put in rank order according to the degree of

---

80 Emanuel F. Hammer and Zygmund A. Piotrowski, "Hostility as a Factor in the Clinician's Personality as It Affects His Interpretation of Projective Drawings (H-T-P)", Journal of Projective Techniques, Vol. 17, No. 2, June, 1953, p 210-216.

hostility apperceived in the H-T-P's of the students from grades three, five and seven. The rank difference correlations among the rank of the three judges were .74, .78, and .84 with standard errors of .031, .030, and .014 respectively. The supervisor's judgment when compared to the rank order of the degree of hostility seen by the judges in the drawings yielded a rank order correlation of .94 with a standard error of .48 while the apperceived aggression rating when compared to the clinician's Szondi rating gave a rank order correlation of .94 with a standard error of .48. A comparison of the Szondi rating and the supervisor rating also yielded a rank order correlation of .94 with a standard error of .48.

Singer[81] compared a group of thirty-four male hospitalized schizophrenics with an unmatched group of forty male college students on the H-T-P by scoring the drawings on 428 separate descriptive items. Out of the 428 tests of significance which were computed, 121 (28 per cent) were found significant at better than the .05 level of confidence. This study is weakened by the fact that no judge other than the author was used and no reliability for the scoring was given.

---

81 R.H. Singer, "A Study of Drawings Produced by a Group of College Students and a Group of Hospitalized Schizophrenics", unpublished Master's thesis, Pennsylvania State College, 1950.

Tolor[82] studied the Tree drawings of the H-T-P as they related to several Rorschach signs. The specific hypothesis formulated was that an individual's degree of rigidity would be manifested in his Tree drawing and his Rorschach protocol. Sixty-one patients were instructed to make two Tree drawings as different from one another as possible. The drawings were rated on fourteen dimensions which would indicate the amount of change between the drawings of each subject. Ten signs which were assumed to indicate rigidity were the criteria for rigidity on the Rorschach. The amount of rigidity on the two drawn Trees was compared to the amount of rigidity on the Rorschach. The results did not support the hypothesis that rigidity in the drawn Trees would be reflected in the Rorschach.

Cassel et al.[83] assumed that the ego defenses of subjects will be higher when the examiner is present when administering the H-T-P and therefore that constriction in the performance will result. H-T-P's were given to 130 adults with fifty-eight in the examiner present group and seventy-two in the examiner absent group. The criteria for

---

82 Alexander Tolor, "The Stability of Tree Drawings as Related to Several Rorschach Signs of Rigidity", Journal of Clinical Psychology, Vol. 13, No. 2, April, 1957, p. 162-164.

83 Robert H. Cassel et al., "Examiner, Ego Defense, and the H-T-P Test", Journal of Clinical Psychology, Vol. 14, No. 2, April, 1958, p. 157-160.

judging constriction were forty-eight features on the H-T-P that were assumed to indicate constriction. These features were judged by three raters who after three attempts at high interrater reliability finally reached. 98. The examiner present group had a significantly smaller number of interpretable features on the House and the Person drawings; also the size of the Person and the size of all three drawings summed was significantly smaller for the examiner present group. Since the presence or absence of the examiner seemed to affect the Tree drawing least, the data were in accord with the hypothesis that the Tree represents a deep level of personality integration and the Person and the House a more superficial level.

Judson and McCasland[84] tested the hypothesis that bare Trees in the H-T-P were more likely to be drawn in winter than in summer. Twenty H-T-P records from each month of the year were collected from clinical files. The drawings by men and women were equally represented and examined for the presence or absence of foliage. Results were presented separately for men and women and grouped into four categories of three months each, corresponding roughly to the four seasons. The results for women's drawings confirmed

---

84 Abe J. Judson and Barbara W. MacCasland, "A Note on the Influence of the Season on Tree Drawings", Journal of Clinical Psychology, Vol. 16, No. 2, April, 1960, p. 171-173.

the hypothesis at the .01 level of confidence. The results for men were in the expected direction but fell short of the .05 level of confidence. The results, particularly for women, suggested that the season should be taken into account before ascribing significance to a bare Tree. Less significance should be attached to a bare Tree in the winter than in the summer.

In a followup study, Moll[85] administered H-T-P's to 269 male university students during the fall and during the winter. The drawings were sorted into two groups: those with leaves and those without. All drawings were sorted independently by two judges and agreement was achieved on all except three of the drawings. Of 150 drawings made in fall, twenty had foilage absent and 130 had foilage; of 119 drawings made in winter, thirty had foilage absent and eighty-nine had foilage. The Chi Square was 5.43 significant at the .01 level. The frequencies indicate that time of the year has a definite influence on whether the Tree drawings of male subjects are made with or without leaves; hence leafless Tree drawings should be interpreted guardedly.

---

85 Richard P. Moll, "Further Evidence of Seasonal Influences on Tree Drawings", Journal of Clinical Psychology, Vol. 18, No. 1, January, 1962, p. 109.

Lair and Trapp[86] tested the hypothesis that more esthetic deterioration from the House to the Person drawings on the H-T-P would occur in a maladjusted group as compared with a non-maladjusted group. The maladjusted group consisted of twenty-five hospital patients with psychiatric or organic diagnosis. The non-maladjusted group was composed essentially of twenty-five first admissions with physical problems. The groups were matched for age, I.Q., education, and occupational level. The judges for esthetic content were two university faculty members of an art department. They showed eighty-four per cent agreement in their ratings. Each judge selected thirteen Person drawings from the maladjusted group and twelve from the non-maladjusted group as reflecting the most esthetic appeal, which results are clearly not statistically significant.

5. Summary and Basic Hypothesis.

The research on drawings has tapped a wide range of possible theories and uses, and the results indicate several promising areas for research and clinical application. The encompassing of such a wide practical and theoretical plain is good on one hand, but less good on the other, since there

---

86 Charles V. Lair and E. Philip Trapp, "Performance Decrement on the H-T-P Test as a Function of Adjustment Level", Journal of Clinical Psychology, Vol. 16, No. 4, October, 1960, p.431.

has been very little concentrated work on one specific area. The place of this study in the literature is that of carrying accumulated empirical evidence in a specific area through the rigors of experimental testing to thoroughly test certain aspects of the sensitivity of the H-T-P for a particular population by establishing how well it will differentiate schizophrenics from a group composed of various pathological syndromes as well as normals.

In terms of validity of the qualitative H-T-P, studies show that general personality dynamics can be measured by this method, but that much work remains to render the H-T-P more critical and sensitive to specific shades of maladjustment.

Studies concerned with reliability suggest that in terms of scorer agreement, if the criteria being used is defined, the scorer reliability ranges from adequate to good, depending on the degree to which the criteria are defined. This writer has not found a study where a test-retest method was used with the qualitative system specifically to establish reliability, although in two studies[87,88] where the H-T-P was administered twice to the same sample, the results were in the expected direction.

---

87 Cowden, Op. Cit.

88 Meyer, Levine, Brown, Ip. Cit.

This investigation is concerned with the general problem of validating some qualitative items of the H-T-P. Specifically it is concerned with validating on a particular and defined population those items inferred as indicative of schizophrenia. The null hypothesis can be stated as follows: Specific interpretive items on the H-T-P do not significantly differentiate between a group of schizophrenics and a group of non-schizophrenics. The following experimental design was established to test this hypothesis.

CHAPTER II

EXPERIMENTAL DESIGN

This chapter presents the procedures which were used to explore the hypothesis proposed in the preceding chapter. It commences with a description of the tool and then presents the population and techniques of analysis.

1. The Tool.-

The specific tool used in this study was a checklist of forty-three scoring items inferred in the literature[1,2,3] to be possible indicators of schizophrenia. These items with their corresponding code numbers appear in Appendix 1. The writer attempted to objectify these items as much as possible to ensure more reliable and objective scoring. The objectified items are presented in Appendix 2.

---

1 Emanuel F. Hammer, The H-T-P Clinical Research Manual, Beverly Hills, California, Western Psychological Services, 1955, p. 12-15.

2 John N. Buck, Administration and Interpretation of the H-T-P Test, Beverly Hills, California, Western Psychological Services, 1950, p. 36-65.

3 Isaac Jolles, A Catalogue for the Qualitative Interpretation of the H-T-P, Beverly Hills, California, Western Psychological Services, 1952, p. 1-97.

An adaptation of Buck's[4] H-T-P test was used in this study. Unlike Buck's procedure, the drawings of a House, Tree, and Person, were drawn on paper measuring 8 1/2" by 11". The appropriate label for each drawing was placed at the top of the three sheets. Buck's instructions were used except for the drawing of the Person. The instructions for the House and Tree were as follows: "I want you to draw me as good a picture of a (House, Tree) as you can. You may draw any kind of a (House, Tree) you wish, it's entirely up to you. You may erase as much as you like; it will not be counted against you and you may take as long as you wish. Just draw as good a (House, Tree) as you can." The direction for the Person was the following: "Now drew me as good a picture of a Person as you can. You may draw any kind of a Person you wish. But remember do not draw a "stick-man", but do drew a whole Person." A No. 2 lead pencil with an eraser was provided each subject. The directions followed a brief preamble regarding the nature of the testing.

2. The Population.

The adaptation of the H-T-P was administered to ninety-seven male schizophrenics ranging in age from sixteen

---

4 John N. Buck, "The H-T-P Technique: A Qualitative and Quantitative Manual", Journal of Clinical Psychology, Monograph Supplement No. 5, Vol. 4, October, 1948, p. 320-396.

to sixty-four years with a mean age of 34.8 and to 134 non-schizophrenics (114 male and 20 female) ranging in age from seventeen to sixty-six years with a mean age of 38.2.

The original number of ninety-seven schizophrenics was reduced to eighty-six because five subjects tore their drawings, each when confronted with the task of drawing a Person, and six schizophrenic drawings were unrecognizable to the extent that they could not be scored. The remaining schizophrenics consisted of thirty-five paranoid types, two hebephrenics, thirteen catatonics, thirty chronic undifferentiated types, and six simple schizophrenics.

The non-schizophrenic group consisted of eleven "organics" (chronic or acute brain syndromes), seventeen manic-depressive psychoses, three mental defectives, six constitutional psychopaths, fourteen inadequate personalities, twenty-one mixed neurotics, twelve neurotic depressives, seven anxiety reactions, three phobics, five hysterical reactions, and thirty-five normals.

The diagnosis of the clinical syndromes was arrived at by a minimum of three staff psychiatrists and a psychologist, based on interviews, psychological tests, and the case history.

In an attempt to obtain two random groups, the following procedures were employed. The schizophrenic subjects were taken from Napa State Hospital, Imola, California.

Though this is a State hospital, it is not a charitable institution, and the rates are set according to each individual's financial resources. Since there are no private mental hospitals within reasonable distance, it can be assumed that this population represents a relatively good cross-section of socio-economic and educational levels. The writer tested at this hospital once every two weeks over a period of thirteen weeks, the length of time that the writer would be available. The tests were administered to groups of five subjects; each subject was seated in such a way that there would be minimal distraction and no chance for copying. The day of the week that the testing was done varied each time, and a sample of those patients which were admitted during the preceding two weeks was chosen by random numbers and tested. The patients were first admissions, were on an admitting ward for the purpose of diagnosis, and were not receiving drugs or electroconvulsive therapy.

Regarding the non-schizophrenic group, the "organics" and normals were chosen from in-patients and out-patients at the Veteran's Administration Hospital, San Francisco, California. These patients were also chosen by random numbers. This hospital is totally supported by the federal government and is open to all veterans and nurses who served in the Armed Forces during the years when war or "police action" was declared. Since the medical and mental health

facilities are on a par with other hospitals, there is reason to believe that a relatively good cross-section of people is tapped. The majority of neurotics, all of the manic depressive psychotics, and the three mental defectives were obtained from Napa State Hospital by the identical technique used for choosing the schizophrenic group. Although it is possible that neither hospital gets a representative sample of the highest socio-economic and educational levels, it is felt that an adequate cross-section was reached.

A second test for reliability was not given because during the interim most of the subjects in the clinical syndromes would have been receiving radical therapy (electroshock, drugs, etc.).

3. The Technique of Analysis.

This section presents a description of the various statistical techniques applied in the study. Subsection A discusses the procedures followed to test the reliability of the writer's judgments in scoring the H-T-P's and those used to establish the reliability of both the significant items and the global method of interpreting the H-T-P drawings. Subsection B describes the main statistic used in the study.

A. Reliability.- In order to test the reliability of the writer's judgments, preliminary comparisons were

made with other scorers.[5] This was done by having three graduate students score sixty sets of drawings individually. This task did not require any degree of knowledge of projective drawings since the items are "presence of - absence of" items. To avoid bias, drawings from both groups were unlabled and shuffled so that the scorers would be unaware of what drawings they were scoring. The scorers worked with the objectified items.

Each scorer scored the H-T-P's of twenty-six schizophrenics and thirty-four non-schizophrenics to establish the presence or absence of each of the forty-three items. In the statistical method for arriving at the scorer's reliability, each scorer's results were compared to each of the others (scorer A was compared with scorer B, A with C, and B with C) for each item, and the reliability was computed in terms of proportion of agreement. This method was used for the partial sample of the schizophrenic group and the non-schizophrenic group and finally for the combined drawings of each partial sample. The results of this method are presented in the following chapter.

An attempt was made to ensure the reliability for the scoring of all the drawings in the entire sample by

---

5 The term "scorer" is used in contrast to the term "judge" which will be used to designate those clinicians who used a global approach to interpret the drawings.

having two scorers go through each set of drawings together. When a particularly difficult judgment had to be made, the scorers discussed it, and it was scored only after the agreement of both.

After the main statistical technique established the significant items, the per cent of correct diagnosis, in terms of number of "hits", was determined for each significant item. The per cent of correct diagnosis achieved by three clinician judges who judged the drawings on the global method of interpretation was also determined. These three judges had an average of eight years experience in the clinical use and interpretation of H-T-P drawings and rated them on a total impressionistic basis derived from their clinical experience rather than any specific criteria.

B. Main Statistic.- The main statistic used in this study was the Chi Square Method. This was used for those items which appeared in sufficient frequency to merit statistical consideration. Since a four cell, 2 x 2 table, was employed the formula used in this study is as follows:

$$\chi^2 = \frac{N(ad - bc)^2}{(a+b)(a+c)(b+d)(c+d)}$$

When the cell frequency was less than ten, Yate's Correction for Continuity was employed. The formula is as follows:

$$\chi^2 = \frac{N(|ad - bc| - \frac{N}{2})^2}{(a+b)(a+c)(b+d)(c+d)}$$

The results of these various procedures and their discussion are presented in the following chapter.

CHAPTER III

PRESENTATION AND DISCUSSION OF RESULTS

The presentation of this chapter will be divided into four sections. The first section gives the results of interscorer reliability procedures. Section Two presents the statistically significant items and is followed by a third section which gives the proportion of correct diagnosis obtained by the statistically significant items and by the clinicians who evaluated the H-T-P's by the global method. Finally, the fourth section discusses the results of the study.

1. Interscorer Reliability.

Interscorer reliability was computed in terms of proportion of agreement for the forty-three items on sixty H-T-P's (of twenty-six schizophrenic and thirty-four non-schizophrenic subjects) scored independently by three scorers. The range (in terms of proportion of agreement) for the schizophrenic group was .77 to 1.00; for the non-schizophrenic group it was .82 to 1.00; and for the combined groups it was .85 to 1.00. The mean proportion of agreement was computed for each set of two scorers (AB, AC, BC) on the total forty-three items for the schizophrenic and non-schizophrenic partial samples considered separately

and combined. These nine means ranged from .96 to .98. The presentation of all the reliability scores is in Table I for the schizophrenic and non-schizophrenic partial samples and in Table II for the combined sample of sixty H-T-P's.

2. Statistically Significant Items.

In testing to establish if a list of forty-three items will differentiate between a group of schizophrenics and a group of non-schizophrenics, five items emerged significant by the Chi Square Method at better than the .01 level of confidence. Table III presents the Chi Square and level of confidence for all items which occurred in sufficient frequency to merit statistical consideration. The five items significant at better than the .01 level of confidence are presented here with their objectifying description if they have one.

G 4 Transparency: if objects or parts of objects can be seen through what would in reality be a solid substance in the House, Tree, or Person. Single overlapping lines are not scored as transparencies.

H 3 Roof drawn with heavy lines that are not general.

H13 Windows, absence of from ground floor: if no part of a window is present in the measured lowest one-third of any one wall of a House with more than one storey.

P 1 Arms wing-like.

Table I.-

Interscorer Reliability for Items from a Partial Sample of Schizophrenics and Non-Schizophrenics in Terms of Proportion of Agreement for Scorers A, B, C.

Item	Schizophrenic Sample			Non-Schizophrenic Sample		
	AB	AC	BC	AB	AC	BC
G 1	1.00	1.00	1.00	1.00	1.00	1.00
G 2	.96	1.00	.96	.97	1.00	.97
G 3	.96	1.00	.96	.91	.97	.94
G 4	1.00	1.00	1.00	.82	.85	.97
H 1	.96	.96	1.00	.97	.97	1.00
H 2	1.00	1.00	1.00	1.00	1.00	1.00
H 3	.96	.92	.96	.97	.97	1.00
H 4	1.00	1.00	1.00	1.00	1.00	1.00
H 5	.88	.92	.96	.94	.97	.97
H 6	1.00	1.00	1.00	1.00	1.00	1.00
H 7	1.00	1.00	1.00	1.00	1.00	1.00
H 8	1.00	1.00	1.00	1.00	1.00	1.00
H 9	.96	.96	1.00	1.00	1.00	1.00
H10	.96	.96	1.00	1.00	1.00	1.00
H11	1.00	1.00	1.00	.97	.97	.94
H12	1.00	1.00	1.00	1.00	1.00	1.00
H13	1.00	1.00	1.00	.91	.88	.97
H14	.96	.96	1.00	.97	1.00	.97
H15	.92	.92	1.00	1.00	.97	.97
H16	.92	.96	.96	.97	1.00	.97
H17	1.00	1.00	1.00	1.00	1.00	1.00
H18	1.00	1.00	1.00	1.00	1.00	1.00
T 1	.96	.96	.92	.88	.97	.91
T 2	.88	.83	.96	.97	.97	.94
T 3	.96	.92	.96	.88	.88	1.00
T 4	1.00	1.00	1.00	1.00	1.00	1.00
T 5	1.00	1.00	1.00	1.00	1.00	1.00
T 6	1.00	1.00	1.00	1.00	1.00	1.00
T 7	.96	1.00	.96	.97	1.00	.97
T 8	1.00	1.00	1.00	.97	.97	.94
T 9	.83	.88	.96	.97	.97	.94
P 1	.77	.77	.92	.91	.91	1.00
P 2	.92	.92	1.00	.91	.94	.97
P 3	.96	.96	1.00	.88	.91	.97
P 4	1.00	1.00	1.00	1.00	1.00	1.00
P 5	.88	.92	.88	.85	.91	.94
P 6	.96	1.00	.96	.88	.91	.88
P 7	1.00	1.00	1.00	1.00	1.00	1.00
P 8	.92	.92	1.00	1.00	1.00	1.00
P 9	1.00	1.00	1.00	1.00	1.00	1.00
P10	1.00	1.00	1.00	1.00	1.00	1.00
P11	.96	.96	1.00	.97	.97	1.00
P12	1.00	1.00	1.00	1.00	1.00	1.00
Mean	.96	.97	.98	.96	.97	.98

Table II.-

Interscorer Reliability for Items from the Combined Schizophrenic and Non-Schizophrenic Samples In Terms of Proportion of Agreement for Scorers A, B, C.

Item	AB	AC	BC
G 1	1.00	1.00	1.00
G 2	.97	1.00	.97
G 3	.93	.98	.95
G 4	.90	.92	.98
H 1	.97	.97	1.00
H 2	1.00	1.00	1.00
H 3	.97	.95	.98
H 4	1.00	1.00	1.00
H 5	.92	.95	.97
H 6	1.00	1.00	1.00
H 7	1.00	1.00	1.00
H 8	1.00	1.00	1.00
H 9	.98	.98	1.00
H10	.98	.98	1.00
H11	.98	.98	.97
H12	1.00	1.00	1.00
H13	.95	.93	.98
H14	.97	.98	.98
H15	.97	.95	.98
H16	.95	.98	.97
H17	1.00	1.00	1.00
H18	1.00	1.00	1.00
T 1	.92	.97	.92
T 2	.93	.92	.95
T 3	.92	.90	.98
T 4	1.00	1.00	1.00
T 5	1.00	1.00	1.00
T 6	1.00	1.00	1.00
T 7	.97	1.00	.97
T 8	.98	.98	.96
T 9	.92	.93	.95
P 1	.85	.85	.97
P 2	.92	.93	.98
P 3	.92	.93	.98
P 4	1.00	1.00	1.00
P 5	.87	.92	.92
P 6	.92	.93	.92
P 7	1.00	1.00	1.00
P 8	.97	.97	1.00
P 9	1.00	1.00	1.00
P10	1.00	1.00	1.00
P11	.97	.97	1.00
P12	1.00	1.00	1.00
Mean	.97	.97	.98

Table III.-

Chi Square and Level of Confidence for Items Which Occurred in Sufficient Frequency to Merit Statistical Consideration.

Item	X^2	P
G 2	.748	.50
G 3	.037	.90
G 4	17.340	.001
H 1	.033	.90
H 3	6.694	.01
H 9	5.674	.02
H11	4.012	.05
H13	7.694	.01
H14	.278	.70
H15	.292	.70
H16	.130	.80
T 1	2.649	.20
T 2	1.427	.30
T 4	.107	.80
T 6	.120	.80
T 7	.154	.70
P 1	12.476	.001
P 2	18.161	.001
P 3	4.302	.05
P 4	.017	.90
P 5	2.104	.20
P 6	3.894	.05
P 8	1.052	.50
P 9	.001	.98
P10	1.222	.30
P11	3.226	.10

P 2 Ears, omission of; face emphasis upon: if there is great attention to detail in the drawing of the face which is lacking in other parts of the Person, and there is an absence of ears in the drawing.

By inspection all of these five items were shown to be significantly present in the schizophrenic sample.

On the basis of the obtained results the null hypothesis that specific interpretive items on the H-T-P do not significantly differentiate between a group of schizophrenics and a group of non-schizophrenics must be rejected for five of the forty-three items.

3. Per Cent of Correct Diagnosis.

After the statistically significant items were determined, the per cent of correct diagnosis, in terms of "hits" was established for each significant item. Item G 4 was sixty-five per cent correct in its diagnosis for the entire sample of the 220 H-T-P's (58 per cent for the schizophrenic sample and 70 per cent for the non-schizophrenic sample); H 3 was sixty-four per cent correct for the total sample (24 per cent for the schizophrenic sample, 89 per cent for the non-schizophrenic sample); H 13 was sixty-three per cent correct for the total sample (41 per cent for the schizophrenic sample, 77 per cent for the non-schizophrenic sample); P 1 was sixty-six per cent correct for the total sample (24 per cent for the schizophrenic sample, 93 per cent for the

non-schizophrenic sample); P 2 was sixty-eight per cent correct for the total sample (31 per cent for the schizophrenic sample, 91 per cent for the non-schizophrenic sample).

The per cent of correct diagnosis was established in the same manner for the three clinician Judges X, Y, and Z, in order to observe the relative efficiency of the detailed atomistic approach to the interpretation of the H-T-P and of the judges' global, impressionistic approach. The per cent of correct diagnosis for the entire sample for Judge X was sixty-five per cent (42 per cent for the schizophrenic sample, 80 per cent for the non-schizophrenic sample); Judge Y was sixty-three per cent correct for the total sample (58 per cent for the schizophrenic sample, 66 per cent for the non-schizophrenic sample); and Judge Z was sixty per cent correct for the total sample (37 per cent for the schizophrenic sample, 74 per cent for the non-schizophrenic sample).

4. Discussion of Results.

In terms of interscorer reliability, the results of this study appear to reinforce the idea suggested in previous studies that the more explicitly an item can be formulated, the less subjective is the scorer's judgment. Generally, those items which could not be explicitly defined showed the lower reliability coefficients. From the

relatively high coefficients obtained in this study, it seems justified to assume that this particular list of items could be used with similar preciseness by anyone adhering to the objective criteria. The results of this study must be viewed in reference not only to the list of items itself, but the accompanying objective criteria as well.

Regarding the validity of the items, the writer cannot offer any qualitative rationale for the fact that these particular five items were statistically significant and the other thirty-eight were not because the entire list of items was inferential of personality characteristics associated with withdrawal, fantasy or denial of reality. The significant items did not fall into any particular category; hence it is difficult to see an underlying cause for the significance of these particular items based on personality characteristics. It is possible that the fact that five of the forty-three items emerged significant is in itself a chance factor. The possibility could be supported by the fact that even though five items were significant, their per cent of correct diagnosis was relatively low.

The fact that there were not more than five significant items could be due to several factors. Since the vast majority of the schizophrenic patients were not severely deteriorated, one might expect that the distortions in the

drawings would not be too great. Another factor could be that the non-schizophrenic group quite possibly consisted of some schizophrenics although they were not diagnosed as such. Many neurotics, for example, rendered severely distorted drawings.

The results seem to indicate that such a list of items while it quite possibly may be highly effective in differentiating schizophrenics from normals becomes much less sensitive when other clinical syndromes are introduced into the population. Several times generally distorted drawings did not in fact possess any of the forty-three items; by the same token some drawings that possessed a number of these items remained relatively healthy in appearance. Also, while strict objectification aided reliability in the study, it led to several instances where strict measurement required the exclusion or scorings of an item which in less stringent definition might have allowed the tapping of more pathology. Because of the different types of schizophrenics in this sample, it should not be expected that so-called "schizophrenic" items would be given by every type; for example, an item connoting hallucinations would not necessarily be expected from a catatonic.

Any generalization of these results must be done most carefully. It was evident throughout the scoring of the drawings that it was possible to have "schizophrenic"

items in drawings of subjects who were definitely not schizophrenic, and also it was possible to have a very distorted H-T-P which did not contain any of the forty-three items. It appears safe to assume that results similar to those of this study would be obtained from any population in similar circumstances.

In terms of per cent of correct diagnosis, the absence of the statistically significant items was much more effective in "diagnosing" non-schizophrenic subjects than the presence of the items was in "diagnosing" schizophrenic subjects. Therefore it is more likely that the absence of these particular items indicates a non-schizophrenic subject than that their presence indicates a schizophrenic subject. The fact that the entire five significant items showed this trend renders the implication more cogent. The comparison of these percentages to those achieved by the clinician judges show similar trends although the significant items showed a slightly higher total per cent of correct diagnosis. This finding does not support entirely the constant warnings in the H-T-P literature against interpreting atomistically since the global approach in this case did not prove itself to be more efficient.

In general, the total number of items were not effective in differentiating schizophrenics from non-schizophrenics and the few items that appeared as statistically significant

were more effective in correctly diagnosing non-schizophrenics than schizophrenics. The global approach did not appear more efficient than the atomistic approach. It is the writer's opinion that some doubt can be attributed to the idea of such an entity as a "schizophrenic item" in H-T-P drawings, that is, an item that will with a high degree of reliability be found present in schizophrenic drawings and absent in non-schizophrenic drawings. It appears from the results of this study that because of much overlapping of personality characteristics and degrees of characteristics in our present nosological categories that such "schizophrenic" items should be interpreted with the utmost caution.

SUMMARY AND CONCLUSIONS

This paper reported on the validity of some interpretive hypotheses of the H-T-P on a particular population. A review of the literature was followed by a perspectus of the relatedness of this study to the existing literature. The problem was defined in terms of the hypothesis that a number of interpretive items on the H-T-P would not differentiate between a group of schizophrenics and a group of non-schizophrenics. The method used to test this hypothesis as well as a description of the population and statistical evaluation were presented in the context of the experimental design.

The results of the study were then reported including interscorer reliability, significant items, and the per cent of correct diagnosis of each significant item and of the clinician judges who rated the H-T-P's by the global method. The null hypothesis being rejected for five of the forty-three items, a discussion of significant findings was presented. The results appear to indicate that caution should be used in the interpretation of so-called "schizophrenic" items since even those items which attained statistical significance did not prove highly reliable in correctly diagnosing the presence of schizophrenia. Also the global approach to H-T-P drawings was shown not to be more efficient in its

per cent of correct diagnosis for the total population of this study than the atomistic approach of the statistically significant items.

The possibilities for future research concerning the problem discussed are varied. It would be beneficial to do a cross-validation study with these items on a different and larger population. To carry the testing of the diagnostic sensitivity of the H-T-P items one step further than this study using larger and relatively equal numbers of patients from each schizophrenic and non-schizophrenic category, intragroup differences might be established on the basis of these items as a means of further validation. Regarding the items which did not occur in sufficient frequency to merit statistical consideration, it would be beneficial to establish if any or some of these items could be "critical signs", i.e., signs that seldom appear but when they do it is always in a particular diagnostic group.

BIBLIOGRAPHY

Anastasi, A., and J.P. Foley, Jr., "A Survey of Literature on Artistic Behavior in the Abnormal: I. Historical and Theoretical Background", Journal of General Psychology, Vol. 25, First Half, July, 1941, p. 111-142.

A comprehensive review of the literature of drawings used in the study of mental health in the 19th and early 20th century. It gives the historical foundations necessary for research in the clinical use of drawings.

--------, "A Survey of the Literature on Artistic Behavior in the Abnormal: IV. Experimental Investigations", Journal of General Psychology, Vol. 25, First Half, July, 1941, p. 167-237.

A review of investigations carried on under modified experimental conditions which also revealed the need for better experimental control.

Bell, John Elderkin, Projective Techniques, New York, Longmans, Green, 1948, p. v-533.

A source which discusses various theoretical assumptions involved in projective techniques.

Buck, John N., "The H-T-P Technique: A Qualitative and Quantitative Manual", Journal of Clinical Psychology, Monograph Supplement No. 5, Vol. 4, October, 1948, p. 320-396.

The basic source of H-T-P information: the two phase approach to personality with validation studies and hypotheses, Buck's method of administration, scoring and interpretation. The sine qua non for H-T-P research.

--------, Administration and Interpretation of the H-T-P Test, Beverly Hills, California, Western Psychological Services, 1950, p. 1-67.

Further elaboration of the H-T-P with more hypotheses concerning placement of drawings and specific items. Another essential source for H-T-P research.

Hammer, Emanuel F., "An Investigation of Sexual Symbolism: A Study of H-T-P's of Eugenically Sterilized Subjects", Journal of Projective Techniques, Vol. 17, No. 4, December, 1953, p. 401-413.

A study which attempted to validate some psychosexual signs in the H-T-P using a checklist of items based on psychoanalytic thought. The design of this study is most useful for concurrent validation of items, although the study was based on a weak assumption.

Hammer, Emanuel F., The H-T-P Clinical Research Manual, Beverly Hills, California, Western Psychological Services, 1955, p. 1-58.

The setting forth of more hypotheses in terms of needs, traits, strengths, and defenses along with three studies showing how these hypotheses can be tested. A source of research implications.

Jolles, Isaac, A Catalogue for the Qualitative Interpretation of the H-T-P, Beverly Hills, California, Western Psychological Services, 1952, p. 1-97.

This is a catalogue containing interpretive hypotheses for specific H-T-P items. It is a "cookbook" based on empirical evidence. In itself it is more dangerous than valuable.

Machover, Karen, Personality Projection in the Drawing of the Human Figure, Springfield, Thomas, 1949, ix-181 p.

This text contains a description of Machover's rationale for drawings: her method of administration, scoring, and interpretation.

Sloan, William, "A Clinical Review of H-T-P Validation Studies", Journal of Clinical Psychology, Vol. 10, No. 2, April, 1954, p. 143-146.

An article in which some validation studies with the H-T-P are criticized for their lack of design.

APPENDIX 1

SCORING ITEMS

APPENDIX 1

SCORING ITEMS

The following "Presence of, Absence of" items are to distinguish between schizophrenics and non-schizophrenics. The Item Code is used throughout the research to designate the item.

Code letter "G" designates general items which may be found in any one or all of the H-T-P drawings.

Code letter "H" designates items limited to the drawing of the House.

Code letter "T" designates items limited to the drawing of the Tree.

Code letter "P" designates items limited to the drawing of the Person.

G 1 Spontaneously written comments.
G 2 Paper-chopping, top of page.
G 3 Paper-topping.
G 4 Transparency.

H 1 Loss of perspective sign.
H 2 Daisies or other flowers drawn about the House.
H 3 Roof drawn with heavy lines that are not general.
H 4 Roof and fence only drawn for the House.
H 5 Roof walling.
H 6 Smoke blowing simultaneously to left and right.
H 7 Steps leading to blank wall.
H 8 Wall, absence of baseline to.
H 9 Walls, double perspective.
H10 Wall, single perspective (absolute profile).
H11 Wall, vertical dimension overemphasized.
H12 Windows, absence of.
H13 Windows, absence of from ground floor.
H14 Windows, absent from upper storey but present on ground floor.
H15 Windows, curtained, shaded, or shuttered.
H16 Windows, lack of conformity.
H17 Windows, with locks.
H18 House in motion.

T 1	Bark, depicted by vine-like vertical lines rather well separated.
T 2	Roots, talon-like, not penetrating ground surface.
T 3	Roots, thin-lined, making tenuous contact with ground.
T 4	Trunk dead.
T 5	Movement in Tree.
T 6	Open base.
T 7	Excessively wide base.
T 8	Split Tree.
T 9	White shading.
P 1	Arms, wing-like.
P 2	Ears, omission of; face emphasis upon.
P 3	Ears, over-emphasis upon.
P 4	Eyes, omission of.
P 5	Feet, on tiptoe.
P 6	Head, emphasis on peripheral line of.
P 7	Head with back to viewer.
P 8	Neck, long and thin.
P 9	Absolute profile.
P10	Trunk, long and narrow.
P11	Feet, large or heavy.
P12	Hat brim low over the eyes.

APPENDIX 2

SCORING ITEMS OBJECTIFIED

APPENDIX 2

SCORING ITEMS OBJECTIFIED

The following items are those which 1) were felt to need objectifying and 2) which could be objectified in order to clarify their meanings as much as possible.

G 2 Paper-chopping, top of page: if the drawing appears incomplete due to lack of space at the top of the page.

G 3 Paper-topping: if the drawing uses the top edge of the page as an integral part of the drawing or if the top of the drawing reaches within one-fourth of an inch of the top of the page.

G 4 Transparency: if objects or parts of objects can be seen through what would in reality be a solid substance in the House, Tree, or Person. Single overlapping lines are not scored as transparencies.

H 1 Loss of perspective sign: if an end wall and roof in good perspective are drawn at one end of the House but inadequate representation of depth at the opposite end with the result that a vertical endline for both roof and wall is drawn perpendicular to the baseline of the House. The result is in a sense incongruous: at one end the House shows depth, proper angulation, etc., at the other end it looks as if it had been chopped off abruptly.

H 5 Roof walling: if there is a lack of distinction between the sidewall and roof because there is no drawn line separating them.

H 8 Wall, absence of baseline to: if there is no groundline or baseline drawn to which the walls of the House may be connected. This item is scored unless the House is paper-based.

H 9 Walls, double perspective: if both endwalls and a sidewall are presented simultaneously.

H10 Walls, single perspective: if only one sidewall is shown.

H11 Wall, vertical dimension overemphasized: if the height of the House is more than 2½ times the measured width of the front end of the House.

H12 Windows, absence of: if no windows are present in the drawing of the House.

H13 Windows, absence of from ground floor: if no part of a window is present in the measured lowest one third of any one wall of a House with more than one storey.

H14 Windows, absence from upper storey but present on ground floor: if no part of a window is present in the measured highest one third of any one wall of a House with more than one storey.

H15 Windows, curtained, shaded, or shuttered: if curtains, shades or shutters are closed over any window in the House or cover one-half or more of the viewing space of any window.

H16 Windows, lack of conformity from wall to wall and floor to floor: if a window differs one eighth of its height in placement in relation to any other window of the same type on the same storey regardless of wall.

H18 House in motion: if any integral part of the House appears to be moving, such as the roof flying off or the walls toppling over.

T 1 Bark, depicted by vine-like vertical lines rather well separated: if the lines are not heavily drawn nor straight, but light and curved.

T 2 Roots, talon-like, not penetrating ground surface: if claw or hook-like roots are drawn (curved but not wiggly) and do not penetrate the drawn ground-line.

T 4 Trunk dead: if the trunk only is drawn and/or signs of decay are present on the drawing.

T 5 Movement in Tree: if the Tree is almost bent double down to the ground or wind is blowing from ground level to the treetop.

T 6 Open base: if no groundline or baseline is present. This item is scored unless the Tree is paper-based.

T 7 Excessively wide base: if the base of the trunk is more than three times the width of the trunk as measured from the approximate midpoint of the Tree's height.

T 8 Split Tree: if the trunk presented is actually two one-dimensional trunks because the sidelines of the Tree never meet at the top or the bottom. This item is scored unless the Tree is paper chopped.

T 9 White shading: if the white space is given an implied solidarity by having two-dimensional branches showing through the white space, in effect, at intervals.

P 2 Ears, omission of; face emphasis upon: if there is great attention to detail in the drawing of the face which is lacking in other parts of the Person, and there is an absence of ears in the drawing.

P 3 Ears, overemphasis upon: if the ears are made prominent by emphasis on their size, shading or details. Size is to be considered emphasized if the length of the ear equals more than one-fourth of the head's measured length or if the width of the ear is more than one-eighth of the measured length of the head.

P 5 Feet, on tiptoe: if the posture indicates that the Person is on tiptoe; if it is evident the base of the foot does not reach the groundline; if the front part of the foot only shows it is pointed so that the Person appears to be on his toes.

P 8 Neck, long and thin: if the length of the neck is greater than one-twelfth of the length of the complete figure and is greater than the width of the neck.

P 9 Absolute profile: if the side view of the Person is drawn and only one arm and leg are visible.

P10 Trunk, long and narrow: if the length of the trunk is a measured one-half or more of the complete figure and more than twice the width of the trunk as measured at the midpoint between the shoulders and waist.

APPENDIX 3

ABSTRACT OF

The Validity of Some Qualitative Scoring Items On The Drawings of a House, Tree, and Person

APPENDIX 3

ABSTRACT OF

The Validity of Some Qualitative Scoring Items On The Drawings of a House, Tree, and Person[1]

This study investigated the validity of specific qualitative items inferred in the H-T-P literature as possible indicators of schizophrenia. The first aim of the study was to compile a list of forty-three items and objectify these items as much as possible to limit subjective scoring. Secondly, the study sought to establish if these items would significantly differentiate a group of schizophrenics from a group of non-schizophrenics. Finally, the study investigated the respective efficiency of the significant items and a total impressionistic approach to the H-T-P's in correctly "diagnosing" the presence or absence of schizophrenia.

The H-T-P's of eighty-six schizophrenics and 134 non-schizophrenics were rated on the presence or absence of forty-three items coded and objectified by the author. A partial sample of sixty H-T-P's was scored by three independent scorers to establish interscorer reliability which ranged from .77 to 1.00. The scoring of the total sample

---

1 Michael E. Cavanagh, Master's thesis presented to the School of Psychology and Education at the University of Ottawa, Ontario, May 1963, viii-103 p.

of 220 H-T-P's was followed by a Chi Square analysis of those items which occurred in sufficient frequency to merit statistical consideration. Five items were found to be significantly present in the schizophrenic group at better than the .01 level of confidence. The per cent of correct diagnosis was established for the items, and the range for the total sample was from sixty-three per cent to sixty-eight per cent. Three judges with an average of eight years experience in the clinical use and interpretation of the H-T-P also rated the drawings on a total impressionistic basis derived from their clinical experience rather than any specific criteria. The per cent of correct diagnosis for the three judges on the total sample of 220 drawings ranged from sixty to sixty-five per cent.

The relatively high reliability results appear to substantiate the suggestion of the literature that the greater the objectification of the scoring item, the higher the degree of reliability. In terms of validity, although there is no definite rationale that can be offered for the fact that the particular five items were statistically significant and the remaining thirty-eight items were not, several factors could account for the results. These include the degree of non-deterioration in the schizophrenic sample, the possibility of incorrect diagnosis in the non-schizophrenic sample, and the possible overlapping

of some personality characteristics among the pathological syndromes. The global method of interpretation by the three clinician judges in terms of per cent of correct diagnosis did not prove more effective than the significant items in differentiating the two groups.

The results appear to indicate that caution should be used in the interpretation of so-called "schizophrenic" items and that further work is necessary in order to more safely discuss these items. It would be beneficial to do a cross-validation study as well as to further test the diagnostic sensitivity of H-T-P items by establishing whether intragroup differences exist within different pathological syndromes.