THE VALIDITY OF SOME QUALITATIVE SCORING ITEMS,
IN TERMS OF HOSTILITY AND AGGRESSION,
ON THE DRAWINGS OF A HOUSE, TREE
AND PERSON

by Paul J. Williams

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

Ottawa, Canada, 1964
INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI Microform EC55378
Copyright 2011 by ProQuest LLC
All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.
ACKNOWLEDGEMENTS

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

The writer is indebted to M.M. Le Corre, Director of St. Vincent de Paul Penitentiary, for his co-operation in permitting the use of inmates of the penitentiary and neighbouring Farm Camp as subjects in the study.
CURRICULUM STUDIORUM

Paul J. Williams was born October 4, 1937, in Montreal, Quebec. He received his Bachelor of Arts degree from Loyola College, Montreal, in 1959.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>INTRODUCTION</strong></td>
<td>vi</td>
</tr>
<tr>
<td>I.- REVIEW OF THE LITERATURE</td>
<td>1</td>
</tr>
<tr>
<td>1. Projective Techniques: Theory</td>
<td>1</td>
</tr>
<tr>
<td>2. Specific Technique: The H-T-P</td>
<td>4</td>
</tr>
<tr>
<td>3. The Qualitative Approach</td>
<td>9</td>
</tr>
<tr>
<td>4. Statement of Problem: The Hypothesis</td>
<td>13</td>
</tr>
<tr>
<td>II.- EXPERIMENTAL DESIGN</td>
<td>15</td>
</tr>
<tr>
<td>1. The Tool</td>
<td>15</td>
</tr>
<tr>
<td>2. The Population</td>
<td>19</td>
</tr>
<tr>
<td>3. Method of Data Analysis</td>
<td>25</td>
</tr>
<tr>
<td>III.- PRESENTATION AND DISCUSSION OF RESULTS</td>
<td>28</td>
</tr>
<tr>
<td>1. Interjudge Reliability</td>
<td>28</td>
</tr>
<tr>
<td>2. Significance of Statistically Frequent Items</td>
<td>31</td>
</tr>
<tr>
<td>3. Discussion of Results</td>
<td>34</td>
</tr>
<tr>
<td>SUMMARY AND CONCLUSIONS</td>
<td>38</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>40</td>
</tr>
</tbody>
</table>

**Appendix**

1. SCORING ITEMS | 41 |
2. JUDGMENT SHEET | 43 |
3. ABSTRACT OF The Validity of Some Qualitative Scoring Items, in Terms of Hostility and Aggression, on the Drawings of a House, Tree, and Person | 45 |
**LIST OF TABLES**

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. - Biographical Data on Experimental Group and Control Group Obtained from Inmate Records.</td>
<td>23</td>
</tr>
<tr>
<td>II. - Interjudge Reliability for Items within the Experimental Group and the Control Group in Terms of Proportion of Agreement.</td>
<td>29</td>
</tr>
<tr>
<td>III. - Interjudge Reliability for Items within the Combined Groups in Terms of Proportion of Agreement.</td>
<td>30</td>
</tr>
<tr>
<td>IV. - Raw Frequencies of the Presence and Absence of the Items in the Experimental Group and the Control Group.</td>
<td>32</td>
</tr>
<tr>
<td>V. - Chi-Square Values and Level of Confidence for Items which Occurred in Sufficient Frequency to Merit Statistical Consideration.</td>
<td>33</td>
</tr>
</tbody>
</table>
INTRODUCTION

Projective techniques in general and free-hand drawings in particular have been in increasingly wide use for the past three decades. The goal of free-hand drawings, to obtain information concerning personality dynamics and their resulting behaviour, is aptly confirmed by Anastasi and Foley, "the artistic behaviour of emotionally or intellectually abnormal subjects can be investigated by the psychologist as a means of gaining a more complete understanding of the behaviour itself".¹

Buck's House-Tree-Person test (referred to in the present study as H-T-P) was designed for that purpose; to help the clinician obtain information concerning the integration of personality and its interaction with the environment.² Though the H-T-P serves as a principal projective tool in the clinical setting, empirical validation of the various interpretive hypotheses has been, for the most part, neglected. Hammer points directly to the problem, as it affects all projective techniques, and stresses the needed

---
remedy, "The clinical use of projective techniques has rapidly widened the gap between the use of projectives and their validation(...) . The chief need at this time is for further validation studies". 3

In the present study, an attempt is made at validating specific qualitative scoring items on the H-T-P. The literature 4, 5 suggests certain interpretive items as indicative of hostility and aggression. Two groups of penitentiary inmates, representing the extremes in overt aggressive behaviour within an inmate population, were selected to test the validity of these items.

Chapter one presents a review of the literature. From the theory and rationale behind projective techniques, through the specific technique of the H-T-P, it arrives at the qualitative approach proposed by certain authors. The chapter concludes with a statement of the problem at hand and the formulated hypothesis.


The experimental design used to test the hypothesis is explained in chapter two. The procedures used in applying the tool, a description of the sample and the method employed in analyzing the obtained data are included.

In the third chapter, the results of the above procedures are presented and discussed. The ensuing summary and conclusions include implications for further research.
CHAPTER I

REVIEW OF THE LITERATURE

In this chapter, the original concept of projection is related to the present day usage of projective techniques. The various techniques are classified under three headings and the underlying rationale is explained. From the broad theory of projective techniques, the discussion then turns to the specific technique used in the present study. Both the rationale and administration of the H-T-P are considered, and the qualitative approach to research with the H-T-P is discussed. Previous research concerned with the qualitative approach to hostility and aggression is critically reviewed. The chapter concludes with a statement of the problem at hand and the hypothesis to be tested.


The psychological concept of projection was introduced by Freud\(^1\) and defined as a primary defense mechanism. It was understood as a means by which the ego, in Freud's personality superstructure, alleviated anxiety and made it

an objective fear by attributing its source to the external world. The mechanism was believed to be unconscious and, consequently, not recognized as a projection by the person himself.

In projective techniques, this concept of projection is broadened somewhat. The techniques or tests, as they are also termed, elicit not only the defense mechanism of projection but other mental mechanisms and symbols of human relationships. The unconscious element in personality structure, the importance of which was stressed by Freud, is believed to be the seat of deep motivation. The aim of the techniques is to elicit the unconscious processes in personality and better understand the resultant behaviour. A common principle underlies the various projective techniques: the use of vague or ambiguous stimuli, inciting the subject to reveal unconscious levels of personality.

It has been suggested that projective techniques be classified under the three general headings of perceptive, apperceptive and productive. The perceptive technique requires the subject to tell what certain, ambiguous visual or auditory stimuli mean. The apperceptive technique, though similar to the perceptive, demands a more elaborate

interpretation of the meaning of the stimuli. The productive
technique relies on the performance of the subject, as in
drawing, painting or acting.

Whatever the technique employed, perceptive, appercepti­ve or productive, it is considered a means of obtaining
information from the unconscious which may not be elicited
by more direct questioning. This information from the uncon­scious, symbolic in content, must be interpreted. It may be
said that projective techniques interpret, in part, the
conscious manifestations of behaviour in the light of uncon­scious needs, wishes and drives.

Underlying all projective techniques is a basic
premise which lends credence to a rationale. The "projective
hypothesis" as proffered by Rapaport is, "behaviour manifes­
tations of the human being, from the least to the most signi­ficant ones, are revealing of his personality". It is upon
this premise that all interpretation of projective material
is based and the framework within which the results of H-T-P
drawings may be understood.

3 Gordon W. Allport, Pattern and Growth in Persona­

4 David Rapaport, Diagnostic Psychological Testing,
2. Specific Technique: The H-T-P.

After ten years of controlled study, Buck's method, the H-T-P, was published as a standardized test in 1948. The method, designed in a clinical setting, was primarily for clinical use. It was to serve as an aid to the clinician in obtaining information concerning the integration of personality and its interaction with the environment. The basic postulate of the technique is that each of the drawn wholes is a self-portrait.  

The administration of the test involves a twofold approach. In the first phase, the non-verbal phase, the subject is provided with three sheets of white paper, size seven by eight and one half inches and appropriately labelled House, Tree and Person. According to instructions, the subject is then asked to draw, in turn, a House, a Tree, and a Person. It is to be stipulated for the drawing of the Person that it be of the whole body. Throughout the non-verbal phase, conversation is to be limited. Questions may be parried with exhortations as, "do as well as you can, it is not necessary to be an artist". Since the test may be


6 Ibid., p. 320.
administered chromatically or achromatically, the drawing implements may be a set of eight coloured wax crayons or a No. 2 lead pencil.

The second phase of the administration is the Post Drawing Interrogation. This consists of a series of sixty standard questions to be asked and recorded by the examiner. Since the test may also be applied as a group test, the one innovation is that the subjects record the answers themselves. Buck, however, designed the test primarily as an individual test, on the assumption that the drawings would elicit an emotional reaction which may result in the verbal expression of previously suppressed material. The graphic recording of the answers by the subjects themselves may lessen the value of the responses.

Prior to the advent of the H-T-P, psychological investigation of the artistic behaviour of the abnormal had been extensive. Anastasi and Foley have conducted a seemingly exhaustive survey of the literature up to 1941 concerning the development of interest in the field of various European countries. For present purposes, only those studies considered pertinent to this research will be reviewed.


In 1949, Machover \(^9\) gave impetus to the study of drawings of the human figure with the publication of her Draw-a-Person Test. The basic rationale of the test is that, regardless of such variables as age, skill and culture, certain organs have specific meanings which may be elicited by drawings of the human figure. An elaboration of this rationale may be expressed in three statements: 1) physical attributes acquire social meanings in the course of social participation and expression; 2) psychosomatic correlation of body expression is revealed by overemphasizing or deemphasizing certain parts of the body; 3) psychic data are evident in the symbol value projected. Numerous studies, including those with the H-T-P, have been conducted under these basic assumptions.

In a study by Fisher, \(^{10}\) certain assumptions concerning the analysis of human figure drawings were critically evaluated. The figure drawings of thirty-two, previously diagnosed paranoid schizophrenics, were studied to determine three factors which supposedly underlie these assumptions: 1) if one can objectively detect paranoid trends in figure

---


drawings; 2) how much agreement exists among trained personnel in evaluating facial expression and stance; 3) if a group of trained raters agree more than a group of untrained raters in evaluating facial expression and stance.

To ascertain whether one can objectively detect paranoid trends in figure drawings, two trained psychologists, with full knowledge that the drawings were those of paranoid schizophrenics, were to determine the presence or absence of six paranoid signs. Though there was total agreement as to the presence of the signs, only thirteen of the thirty-two drawings contained two or more signs.

The second factor was studied by three trained psychologists. Unaware beforehand that the drawings were the product of paranoid schizophrenics, the psychologists were asked to independently choose the drawings with strong paranoid trends. The author reports that the percentage of agreement among the three judges was eighteen per cent for choice of paranoid schizophrenic drawings and fifteen per cent for non-schizophrenic drawings.

The result of the comparison between trained raters and untrained raters is recorded as follows: the poor agreement among the latter is no greater than the poor agreement among the former. The author concludes from these results that the assumptions regarding figure drawing analysis are in need of empirical validation.
A study concerning psychosexual pathology and the sex of the first drawn person was undertaken by Hammer, as part of a larger study. The subjects were administered the H-T-P and asked to draw a second Person of the sex opposite to that of the first drawn Person. The assumption under study was that if the sex of the first drawn Person is opposite to that of the subject performing the task, it is indicative of psychosexual pathology.

The population consisted of eighty-four sex offenders at Sing Sing Prison and divided into three groups according to the type of sex crime. The first group included thirty-one rapists, four of whom drew a Person of the opposite sex as the first drawn Person. Thirty-three heterosexual pedophiles comprised the second group, of whom four drew a Person of the opposite sex as the first drawn Person. The third group included twenty homosexual pedophiles, five of whom drew a Person of the opposite sex as the first drawn Person.

A t test indicated no significant difference between any of the three groups. The author concludes that, in the light of his study, the sex of the first Person is no index of psychosexual conflicts or sexual inversion.

This latter study had arisen from the theory of certain authors concerning psychossexual development and sexual identification. From Buck's hypothesis that subjects tend to draw their own sex and that subjects identify with the drawn person, unwarranted inferences had been drawn and subsequently criticized by Sloan.

Three further studies concerning elementary school children are critically evaluated by Sloan. The principal bone of contention with each study is the blind acceptance of certain hypotheses of Buck, which have not been previously validated, and their application to a particular age group. The hypotheses are supposedly to be tested but are postulated at the outset.


The brief review of the aforementioned studies does not pretend to be complete. The purpose has been to attempt to point out the main fallacy in much of the previous research with the H-T-P: the blind acceptance of interpretive hypotheses and their application to various fields of study.

In the following section, the writer will discuss a particular approach to research with the H-T-P. The two previous studies pertinent to the present research project will be critically evaluated.

3. The Qualitative Approach.

The H-T-P was designed as a projective tool to aid the clinician in eliciting data concerning the sensitivity, maturity, flexibility, efficiency, and the degree of integration of a subject's personality. The drawing of the dynamic and familiar concepts of House, Tree and Person was thought to help project conscious and unconscious structural elements of personality, particularly the subject's deepest fantasies, conflicts and fears. It was to serve as a canvas upon which the subject would paint his personality strengths and weaknesses.18

Since its inception, the H-T-P has remained primarily an intuitive tool rather than a scientifically validated one. The need for experimentally controlled validation studies has been stressed by several authors and is aptly stated by Holzberg and Wechsler:

(...)

Various authors have suggested certain drawing elements and their interpretations as a particular approach to the understanding of figure drawings. This is known as the Qualitative Approach. Hammer feels it is expedient to break down the many elements and their interpretations into categories of needs, traits, strengths, conflicts and psychiatric symptoms and, defences. He has


20 John K. Buck, Administration and Interpretation of the H-T-P Test, V.A. Hospital, Richmond, Virginia, 1950, 1-67 p.


compiled the various interpretive items and listed them under particular categories according to their hypothesized interpretation.

In the present study, the items inferred as indicative of hostility and aggression are examined. The aim of the study is to test the validity of each individual item. Previous research with these specific items has apparently been limited to a study of the items as a group.

Hammer administered H-T-P's to four hundred Negro and White school children, ranging in grade level from first to eighth, to study the frustration-aggression hypothesis as applied to the socio-racial area. The drawings were rated by six clinicians on a three point scale of aggression: none, mild and severe. The drawings were placed in random order of Negroes and Whites by grade level so that none of the judges knew whether they were rating drawings by Negro or White subjects. The criteria for judgment were the qualitative items hypothesized to indicate hostility and aggression.

The author's assumption, based on the frustration-aggression hypothesis, was that the Negro group would show

---


more aggression because of the frustration incurred by social discrimination. The mean results of the hostility and aggression rating for all the grades of white children was .308 and .823 for the Negro children. A $ score of 12.56 was statistically significant at far better than the one per cent level of confidence. Indications are, concludes the author, the drawings of the Negro children suggest greater feelings of frustration produced by a restraining environment, with concomitant feelings of hostility and a desire to react aggressively.

It appears that, in the study cited above, the author has selected two groups on the assumption that they differ according to the frustration-aggression hypothesis and then concludes that the hypothesis may be applied to the socio-racial area because of his findings with interpretive items which have been hypothesized rather than validated.

Goldstein and Rawn tested the validity of certain interpretive signs of aggression by means of artificially induced feelings of aggression. Thirty-nine male and female attendants at a state mental hospital were asked to draw the human figure both before and after being told that their work week would be increased by four hours with no increase in pay.

---

The drawings completed before and after the observed aggressive reaction were studied for: 1) greater degree of line pressure; 2) greater figure size; 3) greater number of specific details interpreted as representing aggression; and 4) greater overall subjective impression of aggression by two clinical psychologists.

The results were that the first, second and fourth signs of aggression were not significantly different in the two sets of drawings. The specific interpretive details, however, were significantly different in the two sets of drawings at the .01 level of confidence. The statistical method employed was the chi-square technique. The authors conclude that aggression is expressed symbolically rather than in a graphomotor manner.


The foregoing, brief review of the literature was presented as an orientation toward the present study which attempts to validate specific scoring items on the H-T-P. Previous research has been effected, in part, on the assumption that the interpretive items are valid. Other studies have shown that certain items when considered as a group may be validly interpreted as indicative of a particular personality trait.
In the present study, an attempt is made to validate individual interpretive items suggested as indicators of hostility and aggression. The population is composed of two groups of inmates who had been sentenced to a maximum security institution. The subjects in the experimental group are characterized by their overt aggressive behaviour during incarceration which resulted in their placement in a segregated area of the penitentiary. The subjects in the control group, on the other hand, are characterized by their good behaviour which resulted in their placement within a minimum security camp.

The following chapter will present a more complete description of the tool employed in the study, the population and the statistical techniques used in analyzing the obtained data.

The hypothesis under study, presented in the null form, is as follows: Individual qualitative scoring items on the H-T-P are not significantly different between a group of segregated, maximum security penitentiary inmates and a group of minimum security inmates.
CHAPTER II

EXPERIMENTAL DESIGN

The present chapter concerns itself with the procedures followed in carrying out the experiment itself. It contains a description of the specific tool employed, an investigation of the population and a presentation of the techniques used in the statistical analysis of the obtained data.

1. The Tool.

Certain qualitative scoring items are suggested in the H-T-P literature¹,² as indicative of hostility and aggression. A list of these items was compiled by the writer and appears in Appendix 1. The list is comprised of seventeen items each prefixed by a code number to indicate in which drawing or drawings the particular item may be found. The letter G refers to the one general item which may be found in any or all of the drawings. The letter H refers to those

---


items which may be found in the drawings of the House; the letter T, to those items which may be found in the drawings of the Tree; the letter P, to those items which may be found in the drawings of the Person.

Hammer offers a rationale for a number of the qualitative items. Item G1, drawings made conspicuously too large for the page:

(...) tends to indicate a feeling of great frustration produced by a restraining environment with concomitant feelings of hostility and a desire to react aggressively, either against the environment, the self or both.

---


4 Ibid., p. 39.

5 Ibid., p. 9.
Item P4, well-outlined but unshaded hair, "suggests hostile phantasy concerning sexual matters".\(^6\) Item P7, sharply squared shoulders, "connotes over-defensive, hostile attitudes".\(^7\)

The rationale, according to Hammer,\(^6\) for the remaining items is that hostility and aggression is "clearly indicated" or "strongly implied".

It appeared to the writer as though some of the qualitative items lacked clarity and may have proved susceptible to varied interpretation. An attempt was made to objectify those items which seemed ambiguous. A judgment sheet was composed including all the items and an explanatory note beside those which were not clear. A uniform set of instructions was drawn up and placed at the top of the sheet. The judgment sheet is illustrated in Appendix 2.

Buck\(^8\) designed the H-T-P primarily as an individual test. After the drawings had been completed by the subject, an interrogation form was to be completed by the examiner asking the subject questions. It is Buck's contention that the drawings often arouse a strong emotional reaction which

---


7 Ibid., p. 39.

8 Ibid., p. 39.

may result in the verbal expression of previously suppressed material.

For the present study, however, Buck's method was modified. Since part of the testing was in a segregated area of a maximum security penitentiary, security measures had to be taken. It was not feasible to remove the inmates from their cells, so the testing was conducted along the cell range. A range contained twelve cells with one inmate per cell. As this study dealt with specific qualitative items in the drawings, the post-drawing interrogation period was dispensed with.

Each subject was provided with the standard No. 2 pencil with eraser and an unmarked sheet of white paper, size 8½" by 11". The instructions, given in both French and English, were as follows: "Simply by free hand, without copying, tracing or using a ruler, I want you to draw a house (tree, person)"; "Simplement de longue main, sans copier, calquer ou l'emploi de la règle, je veux que vous dessinez une maison." For the drawing of the person, it was stipulated that the drawing was to be of the whole body. Questions were parried with the phrase: "Just do as well as you can"; "Faites votre possible."
2. The Population.

The entire population of the study consisted of inmates who had been sentenced to the maximum security penitentiary, St. Vincent de Paul. Criterion for sentencing to a penitentiary is conviction of an indictable offence punishable by a minimum penalty of two years incarceration.

An attempt was made to select two groups representative of the extremes in overt aggressivity within an inmate population. Since the classification and placement of inmates is decided upon their behaviour during incarceration, there was easy access to an "aggressive group" and a group relatively free from aggression.

The term "aggressive group" is applied to the experimental group and operationally defined as follows: those inmates who, because of their overt aggressive behaviour during incarceration, were considered a threat to the security and well-being of the institution and, consequently, placed in a segregated area of the penitentiary. This segregated area is not to be construed as a punishment sector for those inmates who have been reported for a misdemeanour. On the other hand, it is a punitive detention for those extreme cases who have actively participated in the physical destruction of the institution, those who are known to instigate others in undermining the authorities and, finally,
EXPERIMENTAL DESIGN

those who have exerted physical violence upon an officer or other inmate.

An individual may not be placed in Segregation arbitrarily. A Segregation Committee, comprised of three to five members, weighs the evidence and gravity of a particular offence and decides the placement of the inmate concerned. Unanimous consent of the committee members is required to place the inmate in Segregation.

The experimental group encompasses the entire population of Segregation, a total of sixty-three inmates. Since the writer had been in direct contact with these inmates for some six months prior to the testing, their co-operation was elicited verbally. The consent of each inmate was obtained.

The control group, those relatively free from aggression, is operationally defined as follows: that group comprised of inmates who, because of their good behaviour during incarceration had been considered suitable transferees to a medium security institution and subsequently to a minimum security institution. A total of seventy-one inmates, the entire population of the minimum security Farm Camp, was tested. Groups of twelve inmates were brought to the refectory by the Director of the Camp and introduced to the writer. The main point stressed to both inmate groups was the fact that none of the obtained information would be for institutional use and that their co-operation or lack of it would
have no bearing on their present status. It is felt that this facilitated matters for obtaining and maintaining the co-operation of all inmates.

The physical environment of Segregation and the Farm Camp would seem to have some bearing on the aggressivity of the inmates. It has been suggested that aggression is always a consequence of frustration and the intensity of aggressivity should vary directly with the intensity of frustration. It would seem as though all forms of incarceration would frustrate at least some of the physical needs and desires of the individual. Since all the inmates in the study were originally sentenced to a maximum security penitentiary, yet one group were eventually placed in Segregation and the other group in a minimum security setting, it may be well to explain the principal differences between the two places.

Segregation and the Farm Camp exemplify the extremes in security and restrictive measures within a penal environment. The Farm Camp, on the one hand, is an open setting in which inmates work an eight-hour day and are entitled to many privileges including television, hobbies and private visits of one hour duration, bi-monthly. No armed guards are present and eating and sleeping accommodations are communal.

---

On the other hand, the segregated inmates are not accorded the above-mentioned privileges. Visits are restricted to one a month of a half-hour duration and in the presence of two custodial officers. Living quarters are a cell, within a walled-in area, inside the maximum security institution. The inmates are confined to the cell permanently, except for a half-hour walk each day in a specially guarded enclosure.

Though both settings are restrictive and apparently constitute a frustration of certain desires, it would seem as though the extreme measures adopted in Segregation are indicative of a more intense frustration. According to Dollard this should result in more intense aggressivity.

To further describe the two groups, the inmate files were perused and the information is recorded in Table I. The mean number of penitentiary terms served, including the present sentence, is 2.5 terms in the experimental group and 1.9 in the control group. The mean length of present sentence is 7.9 years in the experimental group and 2.8 years in the control group. The discrepancy here may be explained by the convictions, previous and present, for assault or violence which tend to increase the length of sentence. In the experimental group the mean for previous convictions including assault or violence is .67 and for present convictions

Table I.-
Biographical Data on Experimental Group and Control Group Obtained from Inmate Records.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Experimental Group (N:63)</th>
<th>Control Group (N:71)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of penitentiary terms</td>
<td>2.5</td>
<td>1.9</td>
</tr>
<tr>
<td>Length of present sentence</td>
<td>7.9 years</td>
<td>2.8 years</td>
</tr>
<tr>
<td>Age</td>
<td>26.1 yrs.(R:18-44)</td>
<td>34.2(R:20-56)</td>
</tr>
<tr>
<td>Admitted school grade</td>
<td>7.1 grade</td>
<td>6.0 grade</td>
</tr>
<tr>
<td>Educational level</td>
<td>3.4 grade</td>
<td>2.7 grade</td>
</tr>
<tr>
<td>Previous convictions(^a)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>including assault or violence</td>
<td>.57</td>
<td>.28</td>
</tr>
<tr>
<td>Present convictions(^a)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>including assault or violence</td>
<td>.49</td>
<td>.10</td>
</tr>
<tr>
<td>Time served on present sentence</td>
<td>2.3 years</td>
<td>1.3 years</td>
</tr>
</tbody>
</table>

\(^a\) Information obtained from R.C.M.P. report on file.
including assault or violence, .49. In the control group the means are .28 and .10 respectively.

The mean admitted school grade is 7.1 grade, experimental group; 6.0 grade, control group. The educational level, as established by the educational supervisor when the inmate enters the penitentiary, is 3.4 grade, experimental group; 2.7 control group. The time served on present sentence up to testing period is 2.3 years, experimental group; 1.3 years, control group.

There is a wide discrepancy between the mean age of the two groups: 26.1 years, experimental group; 34.2 years, control group. The age range is 18 - 44 years and 20 - 56 years respectively. While there is a discrepancy in the mean age between the two groups, there is nothing in the literature to suggest that this would be an influential variable.

The above-mentioned variables were compiled from inmate files and are meant as an added description of the groups. The two variables concerning convictions including assault or violence indicate behaviour before incarceration and point in the same direction as the indicated differences between the groups with respect to criteria for inmate placement: behaviour during incarceration.
3. Method of Data Analysis.

The drawings of a House, a Tree and a Person from each inmate within the experimental and control groups consisted of one hundred and thirty-four sets of drawings. One subject in the experimental group refused to draw the Person and two subjects in the control group refused to draw the Tree and the Person. These drawings were to be closely scrutinized and judged as to whether each item was present or absent.

To ensure the reliability of the judgments, three judges, including the writer, were to study all the sets of drawings independently. A judgment sheet, illustrated in Appendix 2, was provided for each judge with every set of drawings. The sets of drawings were then distributed in random order. No judge knew from which group the particular set of drawings came.

Since the judgments were to be based only on the presence or absence of each sign, familiarity with the H-T-P or projective techniques in general was not necessary. The other two judges, both college graduates, were chosen by the writer. Twenty sets of drawings, apart from those in the present study were judged as a training period. The judgment sheets for the twenty sets of drawings were completed independently, after which the three judges discussed
the reasons for their particular judgments. This trial period, together with the objectification of the judgment sheet, attempted to establish a uniform approach to the items.

The drawings in the present study were then judged independently and without discussion. The total frequencies of presence and absence for every item on the three sets of judgments were tabulated and inter-judge reliability expressed in terms of proportion of agreement. The results are recorded in the following chapter.

To establish one final set of judgments for each item on all the drawings, the presence or absence of the items was determined on a two out of three basis. The chi-square technique was used to establish the significance of difference between the frequencies of each item in the two groups whose theoretical frequency was five or over. A four cell table was employed and the formula was as follows: 12

\[ X^2 = \frac{N (ad - bc)^2}{(a+b) (c+d) (a+c) (b+d)} \]

For those items in which the theoretical cell frequency was less than twenty-five, Yate's Correction for Continuity formula \(^{13}\) was used:


13 Ibid., p. 396.
The following chapter will present the results of the procedures outlined in the present chapter. A discussion of those results together with the implications for further research will be included.

\[ \chi^2 = \frac{N \left( (ad - bc) - \frac{N}{2} \right)}{(a+b) (c+d) (a+c) (b+d)} \]
CHAPTER III

PRESENTATION AND DISCUSSION OF RESULTS

This chapter is divided into three sections. The results of interjudge reliability are recorded in the first section. The second section presents the items which occurred frequently enough to be considered statistically. The final section discusses the results of the procedures used in the study.

1. Interjudge Reliability.

Interjudge reliability was computed in terms of proportion of agreement for the seventeen items on all sets of drawings. One subject in the experimental group refused to draw the Person and two subjects in the control group refused to draw the Tree and the Person. The range, in terms of proportion of agreement, for the experimental group was .92 to 1.00; for the control group, .90 to 1.00; for the entire population of the study, .91 to 1.00.

The mean proportion of agreement was computed for each combination of two judges (A-B, B-C, A-C) on the total seventeen items for the experimental group, the control group and the entire population. Table II illustrates the reliability proportions for the experimental and control group; Table III for the combined groups.
Table II. -

Interjudge Reliability for Items within the Experimental Group and the Control Group in Terms of Proportion of Agreement.

<table>
<thead>
<tr>
<th>Experimental Group</th>
<th>Item</th>
<th>Control Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-B</td>
<td>B-C</td>
<td>A-C</td>
</tr>
<tr>
<td>.94</td>
<td>.97</td>
<td>.97</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>
Table III.-

Interjudge Reliability for Items within the Combined Groups in Terms of Proportion of Agreement.

<table>
<thead>
<tr>
<th>Item</th>
<th>A-B</th>
<th>B-C</th>
<th>A-C</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1</td>
<td>.94</td>
<td>.96</td>
<td>.97</td>
</tr>
<tr>
<td>H1</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>H2</td>
<td>.98</td>
<td>.98</td>
<td>1.00</td>
</tr>
<tr>
<td>H3</td>
<td>.98</td>
<td>.98</td>
<td>1.00</td>
</tr>
<tr>
<td>T1</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>T2</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>T3</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>T4</td>
<td>.98</td>
<td>.95</td>
<td>.96</td>
</tr>
<tr>
<td>T5</td>
<td>.95</td>
<td>.95</td>
<td>.97</td>
</tr>
<tr>
<td>P1</td>
<td>.98</td>
<td>.97</td>
<td>.97</td>
</tr>
<tr>
<td>P2</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>P3</td>
<td>.99</td>
<td>.99</td>
<td>1.00</td>
</tr>
<tr>
<td>P4</td>
<td>1.00</td>
<td>.99</td>
<td>.99</td>
</tr>
<tr>
<td>P5</td>
<td>.99</td>
<td>.98</td>
<td>.99</td>
</tr>
<tr>
<td>P6</td>
<td>.91</td>
<td>.93</td>
<td>.93</td>
</tr>
<tr>
<td>P7</td>
<td>.98</td>
<td>.95</td>
<td>.97</td>
</tr>
<tr>
<td>P8</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>
2. **Significance of Statistically Frequent Items.**

Of the seventeen interpretive items under study, eight occurred in sufficient frequency to be considered statistically:

- **Item G1**, Drawings made conspicuously too large for the page, without adequate space framing them;
- **Item H2**, Windows without panes, curtains or shutters;
- **Item T4**, Two dimensional branches resembling clubs;
- **Item T5**, Sharply pointed branches or leaves;
- **Item P1**, A mutilated Person;
- **Item P5**, Prominent teeth;
- **Item P6**, Sharply pointed fingers or toes;
- **Item P7**, Sharply squared shoulders.

Table IV lists the raw frequencies for all the items.

Table V presents the chi-square values and levels of confidence for the eight items described above.
Table IV.-

Raw Frequencies of the Presence and Absence of the Items in the Experimental Group and the Control Group.

<table>
<thead>
<tr>
<th>Item</th>
<th>Experimental Group</th>
<th></th>
<th>Control Group</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Presence</td>
<td>Absence</td>
<td>Presence</td>
<td>Absence</td>
</tr>
<tr>
<td>G1</td>
<td>9</td>
<td>54</td>
<td>10</td>
<td>61</td>
</tr>
<tr>
<td>H1</td>
<td>0</td>
<td>63</td>
<td>0</td>
<td>71</td>
</tr>
<tr>
<td>H2</td>
<td>3</td>
<td>60</td>
<td>7</td>
<td>64</td>
</tr>
<tr>
<td>H3</td>
<td>2</td>
<td>61</td>
<td>3</td>
<td>68</td>
</tr>
<tr>
<td>T1</td>
<td>1</td>
<td>62</td>
<td>0</td>
<td>69</td>
</tr>
<tr>
<td>T2</td>
<td>0</td>
<td>63</td>
<td>0</td>
<td>69</td>
</tr>
<tr>
<td>T3</td>
<td>1</td>
<td>62</td>
<td>0</td>
<td>69</td>
</tr>
<tr>
<td>T4</td>
<td>21</td>
<td>42</td>
<td>26</td>
<td>43</td>
</tr>
<tr>
<td>T5</td>
<td>36</td>
<td>27</td>
<td>54</td>
<td>15</td>
</tr>
<tr>
<td>P1</td>
<td>13</td>
<td>49</td>
<td>26</td>
<td>43</td>
</tr>
<tr>
<td>P2</td>
<td>4</td>
<td>58</td>
<td>0</td>
<td>69</td>
</tr>
<tr>
<td>P3</td>
<td>2</td>
<td>60</td>
<td>0</td>
<td>69</td>
</tr>
<tr>
<td>P4</td>
<td>2</td>
<td>60</td>
<td>2</td>
<td>67</td>
</tr>
<tr>
<td>P5</td>
<td>7</td>
<td>55</td>
<td>3</td>
<td>66</td>
</tr>
<tr>
<td>P6</td>
<td>20</td>
<td>42</td>
<td>28</td>
<td>43</td>
</tr>
<tr>
<td>P7</td>
<td>33</td>
<td>29</td>
<td>28</td>
<td>41</td>
</tr>
<tr>
<td>P8</td>
<td>0</td>
<td>62</td>
<td>0</td>
<td>69</td>
</tr>
</tbody>
</table>
### Table V.

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

<table>
<thead>
<tr>
<th>Item</th>
<th>$X^2$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1</td>
<td>.046</td>
<td>.80</td>
</tr>
<tr>
<td>H2</td>
<td>2.1</td>
<td>.20</td>
</tr>
<tr>
<td>T4</td>
<td>.300</td>
<td>.50</td>
</tr>
<tr>
<td>T5</td>
<td>7.78</td>
<td>.01</td>
</tr>
<tr>
<td>P1</td>
<td>5.2</td>
<td>.03</td>
</tr>
<tr>
<td>P5</td>
<td>3.32</td>
<td>.10</td>
</tr>
<tr>
<td>P6</td>
<td>.245</td>
<td>.50</td>
</tr>
<tr>
<td>P7</td>
<td>2.1</td>
<td>.20</td>
</tr>
</tbody>
</table>
3. Discussion of Results.

The high proportion of agreement among the three judges, in determining the presence or absence of the qualitative items, is not to be attributed solely to the objectification of the items. It was obvious from the discussion during the training period that the judges were not in agreement as to the exact meaning of the items, even with the accompanying explanatory note. It was only after discussion of each item that they were able to understand the item in approximately the same way.

Since the present study attempted to validate the items, it was mandatory to first have reliable judgments. The training period, described in the previous chapter, was designed to meet this end. It is the writer's opinion that the items, even when objectified, are subject to varied interpretation and would likely result in unreliable judgments.

Of the seventeen interpretive items under study, only eight had frequencies high enough to be considered by the chi-square technique. Of the eight items with sufficient frequency, Item T5, sharply pointed branches or leaves, is significant at the .01 level of confidence, and Item P1, a mutilated Person, is significant at the .03 level of confidence. None of the other six items proved significant.
The chi-square method was employed to determine the significance of difference. However, to understand the differences which occurred it is necessary to see the direction of the raw frequencies.

As illustrated in Table IV, three items did not appear in any drawing of either group. Item H1, a degraded house, Item T2, Tree with looping line as branch structure and two vertical lines as trunk, and Item P8, arms folded across the chest, all were judged as being absent from every drawing.

The groups in the present study were differentiated by their behaviour, either overtly aggressive or non-aggressive, during incarceration. However, both groups comprise a criminal population. All the individuals within this population have been convicted of breaking an established law, serious enough to warrant at least two years incarceration. By definition, this type of behaviour is aggressive behaviour, and items, suggested as indicative of hostility and aggression, would be expected to appear in the drawings of such a population.

In the light of this study, the three above-mentioned items are not indicative of hostility and

---

aggression. The validity of the other six items, which were not of sufficient frequency, is likewise doubtful. The extremely low frequencies of items present would counterindicate even a combination of these signs as indicative of hostility and aggression.

Of the eight items which occurred in sufficient frequency to merit statistical consideration, five of them point in the direction opposite to the expected. Along with items T5 and P1, which are significantly different between the two groups, the frequency of items H2, T4, and P6 is greater in the control group than in the experimental group. The frequencies of the remaining three items, G1, P5, and P7, are greater in the experimental group than in the control group.

Though the results of the eight items are conflicting, they may still be interpreted as indicative of hostility and aggression. A plausible explanation may be that the items do not distinguish between overt and covert aggression and may be more sensitive to covert aggression. As mentioned before, the criminal behaviour of both groups is indicative of some previous aggressive behaviour. In the control group, this aggressivity has apparently become repressed or suppressed and may be expressed symbolically in the projective drawings. The experimental group, on the other hand, is composed of individuals who have continually expressed
much overt hostility and aggression. Perhaps the physical expression of aggression replaces, for the greater part, any symbolic expression of the same.

However, any interpretation of these eight items as indicative of hostility and aggression, either as a group or individually, must be tempered by the fact that only two items were judged to be present more often than absent. Item T5, in both the experimental and the control group, was present more often than absent. Item P7 was judged present more often than absent in the experimental group only. Hammer stresses that no one sign may be taken alone as a definite sign of a strength or weakness in a subject. However, if signs, considered as a constellation, are to have any weight their individual validity should first be established.

---

SUMMARY AND CONCLUSIONS

The present paper reported on the validity of specific qualitative items interpreted on the H-T-P as indicative of hostility and aggression. A review of the literature included the rationale underlying projective techniques in general and the H-T-P test in particular. Previous research with the qualitative approach to the area of hostility and aggression was critically evaluated. The hypothesis that individual qualitative scoring items on the H-T-P are not significantly different between a group of segregated maximum security penitentiary inmates and a group of minimum security inmates was tested and the procedures recorded, along with a description of the population, in the presentation of the experimental design.

The results of the study included interjudge reliability, the raw frequencies of all the items and the significance of the statistically frequent items. The extremely low frequencies on a number of items indicated their invalidity in light of the present study. Though the null hypothesis was rejected by two of the seventeen items, the raw frequencies indicated the direction of the items and the true interpretation of their results. The two items which were significantly different between the two groups may not be interpreted as indicative of aggression as stipulated by
the operational definitions. On the other hand, an explanation for their interpretation as indicative of covert aggression is preferred.

The possibilities for further research are based upon a more explicit set of items to ensure reliable judgments. Since a training period for the judges was necessary to maintain a similar approach to the items, a complete reliability study is suggested.

A cross-validation study with an inmate group and a non-inmate group may further clarify the discriminatory value of the items. The idea of overt as opposed to covert aggression and their susceptibility to the H-T-P technique may be better understood with more divergent groups.
BIBLIOGRAPHY


Prime source of the H-T-P rationale, two-phased approach to personality evaluation, administration, scoring and interpretation.


Extensive treatment of hypotheses for interpretation: horizontal and vertical placement, specific items. Essential for research.


An attempt at validation of interpretive signs of aggression by means of artificially induced aggression.


Listing of qualitative scoring items and tentative rationale.


Breakdown of drawing elements and their suggested interpretation into needs, traits, strengths and defences. Includes pertinent studies with qualitative item approach.


Compendium of interpretive items and their interpretation. Dangerous because it appears as a definitive answer to interpretation.


Rationale for projective techniques. Explains underlying assumptions.
APPENDIX 1

SCORING ITEMS
APPENDIX 1

SCORING ITEMS

To facilitate the calculation of the final judgments, the scoring items are prefixed by a code number. G refers to the one item which may be present in any of the three drawings in each set. H refers to the items which may be present in the drawings of the House. T refers to the items which may be present in the drawings of the Tree. P refers to the items which may be present in the drawings of the Person.

G1 Drawings made conspicuously too large for the page, without adequate space framing them.

H1 A degraded House;
H2 Windows without panes, curtains or shutters;
H3 Attic windows which are open.

T1 A degraded Tree;
T2 Tree with looping line as branch structure (unclosed at junction with trunk) and two vertical lines as trunk;
T3 Side of the page used as the edge of the Tree;
T4 Two dimensional branches resembling clubs;
T5 Sharply pointed branches or leaves.
P1  A mutilated Person;
P2  Person carrying a weapon;
P3  Person in a threatening attitude;
P4  Well outlined but unshaded hair;
P5  Prominent teeth;
P6  Sharply pointed fingers or toes;
P7  Sharply squared shoulders;
P8  Arms folded across the chest.
APPENDIX 2

JUDGMENT SHEET
APPENDIX 2

JUDGMENT SHEET

CHECK EACH ITEM LISTED BELOW, WITH THE ACCOMPANYING SET OF DRAWINGS AND MARK ON THE DOTTED, MARGINAL LINE WHETHER THE PARTICULAR ITEM IS PRESENT OR ABSENT. SHOULD THE ITEM BE PRESENT IN A DRAWING, MARK A PLUS SIGN (+); SHOULD THE ITEM BE ABSENT, MARK A ZERO (0). MAKE A MARK (+ or 0) FOR EACH ITEM.

......G1. Any drawing conspicuously too large for the page (without at least 1 inch space framing it).
......H1. Degraded House (details drawn on or around the House which would be degrading to the style or form of House drawn).
......H2. Windows without panes, curtains and shutters (simply an enclosed white space).
......H3. Attic windows which are open (those windows drawn within the roof structure).
......T1. Degraded Tree (details drawn on or around the Tree which would be degrading to the style or form of Tree drawn).
......T2. Tree with looping line as branch structure (unclosed at junction with trunk) and two vertical lines as trunk.
......T3. Side of the page used as the edge of the Tree (side of page replacing drawn line).
......T4. Two dimensional branches resembling clubs.
......T5. Sharply pointed branches or leaves.
......P1. A mutilated Person (any dismembering or disfiguration; artistic quality of figure drawn is not to be considered).
......P2. Person carrying a weapon.
.....P3. Person in a threatening attitude (fist upraised, etc., etc...).

.....P4. Well outlined but unshaded hair.

.....P5. Prominent teeth (not necessarily protruding teeth).

.....P6. Sharply pointed fingers or toes.

.....P7. Sharply squared shoulders. (including shoulders cut of proportion with body drawn).

.....P8. Arms folded across the chest.
APPENDIX 3

ABSTRACT OF

The Validity of Some Qualitative Scoring Items, in Terms of Hostility and Aggression, on the Drawings of a House, Tree, and Person
APPENDIX 3

ABSTRACT OF

The Validity of Some Qualitative Scoring Items, in Terms of Hostility and Aggression, on the Drawings of a House, Tree, and Person

This study investigated the validity of certain qualitative items on the H-T-P suggested in the literature as indicative of hostility and aggression. Seventeen items were compiled and objectified in an attempt to decrease the subjectivity of the scoring. The aim of the study was to establish whether these individual items were significantly different between a group of segregated, maximum security penitentiary inmates and a group of minimum security inmates.

The H-T-P drawings of the population of 134 inmates were independently judged by three judges as to the presence or absence of the seventeen items. Interjudge reliability was determined in terms of proportion of agreement and ranged from .90 to 1.00. The chi-square technique of analysis was employed to determine the significance of difference of those items which occurred in sufficient frequency to merit statistical consideration.

1 Paul J. Williams, master's thesis presented to the School of Psychology and Education of the University of Ottawa, Ontario, May 1964, viii-46 p.
Eight of the seventeen items occurred frequently enough to be considered statistically. One of these items was significant at the .01 level of confidence; a second at the .03 level. The direction of five of the eight items including the two significant ones was opposite to that expected, as the raw frequencies were higher in the control group than in the experimental group.

The interpretation of all results was in the light of the raw frequencies. The extremely low frequencies cast doubt on the value of certain items whether considered alone or within a group of items. The direction of the frequencies brought to question the discriminatory value of the qualitative approach as to covert and overt aggression. Suggestions for further research centered around a need for an extensive reliability study and a further cross-validation study.