

CSP-2  
000900

A COMPARISON BETWEEN PSYCHIATRIC PATIENTS AND  
NON-PATIENTS ON THE ACHROMATIC-CHROMATIC  
H-T-P DISCREPANCY SCORES

by Ronald L. Trites

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 in  
Psychology



Ottawa, Canada, 1963

UMI Number: EC55361

### 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<sup>®</sup>**

---

UMI Microform EC55361  
Copyright 2011 by ProQuest LLC  
All rights reserved. This microform edition is protected against  
unauthorized copying under Title 17, United States Code.

---

ProQuest LLC  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106-1346

## ACKNOWLEDGMENTS

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

The writer is indebted to W.A. Blair, M.D., for his cooperation in obtaining patients for this study from the Psychiatric Department, Ottawa Civic Hospital, Ottawa, Ontario.

## CURRICULUM STUDIORUM

Ronald L. Trites was born June 23, 1936, in Quesnel, British Columbia. He received his Bachelor of Arts degree in Education from Gonzaga University, Spokane, Washington, in 1958.

## TABLE OF CONTENTS

Chapter	page
INTRODUCTION . . . . .	vi
I.- REVIEW OF THE LITERATURE . . . . .	1
1. The Achromatic-Chromatic Postulates	3
2. Reliability of the Quantitative Scoring	4
3. Summary and Basic Hypothesis	5
II.- EXPERIMENTAL DESIGN. . . . .	7
1. The Psychometric Battery	7
2. Administration of the Test Battery	10
3. The Sample	10
4. Description of Statistical Procedures	12
III.- PRESENTATION OF RESULTS . . . . .	15
1. Reliability of the Quantitative <u>H-T-P</u> Scores	15
2. Scorer Reliability of the Quantitative <u>H-T-P</u> Scores	16
3. The <u>t</u> Tests	16
4. The <u>Chi</u> Square Tests	20
5. The Variance Ratios	20
6. The Correlations	22
IV.- DISCUSSION OF RESULTS. . . . .	26
1. Reliability of the Quantitative <u>H-T-P</u> Scores	26
2. The <u>t</u> Tests	27
3. The <u>Chi</u> Square Tests	28
4. The Variance Ratios	29
5. The Correlations	29
SUMMARY AND CONCLUSIONS. . . . .	31
BIBLIOGRAPHY . . . . .	33
 Appendix	
1. <u>ABSTRACT OF A Comparison Between Psychiatric         Patients and Non-Patients on the Achromatic-         Chromatic H-T-P Discrepancy Scores</u> . . . . .	34

## LIST OF TABLES

Table	page
<p>I.- <u>t</u> Tests between the Achromatic and Chromatic <u>H-T-P</u> Quantitative Mean Raw Scores for the Patient and Non-Patient Groups . . . . .</p>	17
<p>II.- <u>t</u> Tests Between the Patient and Non-Patient, Male, Female, and Combined Groups for the Signifi- cance of Differences on the <u>H-T-P</u> Quantitative Discrepancy Scores . . . . .</p>	19
<p>III.- Chi Square Tests for the Relationship between the <u>H-T-P</u> Quantitative Discrepancy Score Distribu- tions and Patient-NonPatient Category . . . . .</p>	21
<p>IV.- Variance Ratios Between the Male, Female, and Total Group Quantitative <u>H-T-P</u> Discrepancy Score Distributions for Patients and Non-Patients.</p>	23
<p>V.- Correlation Coefficients between the Quantitative <u>H-T-P</u> Achromatic Raw Scores and the Corresponding Discrepancy Scores for the Non-Patient, Patient, and Total Groups . . . . .</p>	24

## INTRODUCTION

The clinical use of drawing techniques has increased rapidly both in scope and popularity in recent years. However, as Hammer states, "the analysis of drawings still remains more the intuitive tool of the artisan than the controlled instrument of the scientist".<sup>1</sup>

Attempts have been made to quantify various dimensions of the drawn figures in an attempt to make the clinical assessment more objective. The H-T-P which has a carefully designed quantitative scoring system is a case in point. The task remains to validate the interpretative hypotheses on the basis of the scoring system.

The quantitative scoring system, once based on experimental evidence, can better contribute to an understanding of personality dynamics. It seems logical to begin the experimental work with the grosser differentiations and particularly with the assumptions regarding group differences.

Hypotheses have been advanced concerning the relationship between degree of emotional adjustment and the ability to shift from pencil drawings to drawing with crayons. It is the aim of this study to investigate via the

---

<sup>1</sup> Emanuel F. Hammer, The H-T-P Clinical Research Manual, Beverley Hills, Western Psychological Services, 1955, p. 1.

quantitative scoring system, the relative difficulty in shifting from one mode of graphic expression to another in two groups of individuals differing in level of emotional adjustment.

This study has theoretical importance in terms of providing evidence for or against the acceptance of the quantitative discrepancy score as an indication of level of emotional adjustment.

The first part of this study is concerned with a review of the literature concerning an evaluation of the quantitative scoring system and of the achromatic-chromatic drawings.

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

The results obtained are then presented and discussed. The discussion evaluates the quantitative aspects of the test as they relate to reliability and the differential diagnostic value of the discrepancy scores. Implications for further research are indicated.

## CHAPTER I

### REVIEW OF THE LITERATURE

The free hand drawing of house, tree and person (hereafter referred to as H-T-P) was introduced in 1948 after ten years of preparatory work. The H-T-P, according to Buck,

(...) is a technique designed to aid the clinician in obtaining information concerning the sensitivity, maturity, flexibility, efficiency, and the degree of integration of a subject's personality; and the interaction of that personality with its environment - both specific and general.<sup>1</sup>

Although primarily devised as a projective technique, the aim of the author was also

(...) to devise a quantitative scoring system which would have real qualitative value; a system which would produce a score that could not be expressed immediately and without careful analysis by a single figure, but would instead provide several measures, the relationships of which would have diagnostic value and would indicate to some extent whether the scores obtained by the subject represented his customary level of function or represented some diminution of function which might or might not be irreversible.<sup>2</sup>

The quantitative scoring system, based on analysis of the drawings of 140 subjects of six levels of intelligence (ranging from moron to superior), includes items of detail,

---

<sup>1</sup> John N. Buck, "The H-T-P Technique, a Qualitative and Quantitative Manual", in Journal of Clinical Psychology, Monograph Supplement, No. 5, 1948, p. 3.

<sup>2</sup> Ibid., p. 9.

proportion and perspective. These items "appeared best to differentiate between subjects of various levels".<sup>3</sup>

An important supplement to the achromatic H-T-P test is the Chromatic drawing phase. The importance of obtaining chromatic drawings is emphasized by Buck when he states:

Theory suggests (and the validity of the belief has empirical support) that (1) the subject reveals in his achromatic H-T-P those characteristics (behavioural, attitudinal, emotional, etc.) that he presently employs in meeting and dealing with the problems of life; (2) that the characteristics he reveals in his chromatic drawings of House, Tree, and Person are likely to be more basic to his personality since (a) they are produced after he has drawn these same objects in pencil and has been questioned directly and indirectly about them at considerable length (an experience calculated to bring into active play all the non-intellective factors in the personality that might enhance or impede efficiency of function); and (b) the adult subject, who has inevitably in our culture become at least somewhat conscious of the esthetic value of form, is compelled to make use of crayons (inferior drawing tools with childhood associations) and is not permitted to erase.<sup>4</sup>

There have been very few validation studies of the quantitative scoring system reported in the literature. It was with the studies concerning the quantitative scoring system and the achromatic-chromatic performance that this study was concerned. An attempt was made to review the reports as they relate to:

1. The Achromatic-Chromatic Postulates;
2. Reliability of the Quantitative Scoring.

---

<sup>3</sup> Buck, Op. Cit., p. 7.

<sup>4</sup> -----, "The Quality of the Quantity of the H-T-P", in Journal of Clinical Psychology, Vol. 7, 1951, p. 356.

This chapter ends with a summary and the formulation of the hypothesis.

1. The Achromatic-Chromatic Postulates.

Payne, in an important theoretical formulation, states that the chromatic drawings are an important part of the H-T-P in that:

The task itself tends to promote regressive, crude, structurally simple, artistically naive usage of color, and it appears that a high degree of personality integrity and emotional adequacy are necessary to shift with equanimity from the achromatic pencil to the chromatic crayon. This factor, then, is of value in differentiation of superior from mediocre adjustment in clinically "normal" people and in the differentiation of clinically "normal" from clinically "abnormal" personality types.<sup>5</sup>

A recent study by Bieliauskas and Heffron<sup>6</sup> utilized discrepancy scores on the achromatic and chromatic drawings to measure differences in performance between groups. Two groups of high school students, twenty-five boys and twenty-five girls in each group, were matched for intelligence on the basis of the Otis Quick-Scoring Mental Ability Tests.

---

<sup>5</sup> John T. Payne, "Comments on the Analysis of Chromatic Drawings", in Appendix of "The H-T-P Technique, A Qualitative and Quantitative Scoring Manual", in Journal of Clinical Psychology, Monograph Supplement No. 5, 1948, p. 119.

<sup>6</sup> V.J. Bieliauskas and A. Heffron, "Differences in Performance on the Chromatic vs. Achromatic H-T-P Drawings", in Journal of Clinical Psychology, Vol. 16, 1960, p. 334-335.

The H-T-P was administered to the first group in the standard order. In the second group the order of tests was reversed. The results indicated that females scored higher on both sets of drawings. All subjects scored significantly higher on the first set of drawings presented (significant at .01 level) whether achromatic or chromatic. However, the sample was of a limited age range (sixteen to eighteen years) which restricts the generalizations considerably. The fact that the tests were group administered and that a time limit of thirty minutes was placed on the drawing time may also restrict the conclusions.

## 2. Reliability of the Quantitative Scoring.

One of the few studies in the literature reporting quantitative scoring reliability is reported by Bieliauskas.<sup>7</sup> The sample consisted of forty-three graduate students in psychology who ranged in age from seventeen to thirty-five years with a mean age of 21.7 and a standard deviation of 5.19. Three experienced examiners independently scored the drawings. The correlations between the three scorers on the four raw scores ranged from .68 to .90. All the correlations were significant for group prediction.

---

<sup>7</sup> V. J. Bieliauskas, "Scorer's Reliability in the Quantitative Scoring of the H-T-P Technique", in Journal of Clinical Psychology, Vol. 12, 1956, p. 366-369.

Digianno,<sup>8</sup> using a sample of 102 student nurses, reports test-retest reliabilities as follows: .70 for the Raw G Score; .49 for the Net Weighted Score; .76 for the Good Score; and .50 for the Flaw Score. The drawings were group administered and the test-retest completed within a period of three weeks. All the reliabilities were significant beyond the .01 level. Digianno,<sup>9</sup> using two experienced scorers also reported scorer reliability as follows: .96 for the Raw G Score; .94 for the Net Weighted Score; .91 for the Good Score; and .96 for the Flaw Score.

### 3. Summary and Basic Hypothesis.

There is an extreme paucity of reported research of comparative studies of the achromatic-chromatic quantitative scores. The study using a non-psychiatric sample investigating this relationship does not support or contradict the postulate that emotionally disturbed individuals experience difficulty in shifting from an achromatic to a chromatic method of graphic expression. The results of this study do, however, suggest that other factors such as

---

<sup>8</sup> John J. Digianno, A Factor Analytic Study of the Quantitative H-T-P, unpublished doctoral dissertation presented to the University of Ottawa, 1963, p. 21.

<sup>9</sup> Ibid., p. 22.

fatigue and loss of interest may be partially responsible for a poorer production on the second set of drawings.

Studies investigating the reliability of the H-T-P suggest that this is an adequate tool for use as a research instrument. In terms of scorer agreement, the separate scores are sufficient for group prediction.

This investigation is concerned with the differences in performance on the achromatic and chromatic drawings in two groups of adults differing in degree of emotional adjustment. The differences in performance are in terms of the quantitative discrepancies. The quantitative scores are assumed to be a measure of the present level of functioning and affected by emotional factors. An attempt was made to investigate, through statistical means, relationships either within or between the groups. The hypothesis may be stated as follows: There are no significant differences in the discrepancies between the achromatic and chromatic quantitative H-T-P Scores of psychiatric patients and non-patients.

## CHAPTER II

### EXPERIMENTAL DESIGN

This chapter presents the procedures which were undertaken in order to explore the hypothesis proposed in the preceding chapter. It begins with a description of the psychometric battery followed by a description of the test administration. The sample population is then described, with emphasis given to the explanation of the variables used in pairing the subjects. Finally, an explanation and description is given of the statistical procedures utilized in the investigation.

#### 1. The Psychometric Battery.

The test battery consisted of two tests. The order of presentation of the tests was the same for each individual.

The first test administered was the Shipley-Institute of Living Scale for Measuring Intellectual Impairment. The test consists of two sub-tests, a forty-item multiple choice vocabulary test and a twenty-item completion type abstraction test. Each sub-test has a time limit of ten minutes. The test was included in the belief that it would provide a satisfactory measure of general intelligence.

Shipley,<sup>1</sup> using a sample of 328 army recruits, reports the reliabilities to be: .87 for the vocabulary subtest; .89 for the abstract thinking subtests; and .92 for the two combined. A recent study by Sines and Simmons<sup>2</sup> revealed a product moment correlation of .90 and a standard error of estimate of 6.33, between the Shipley-Hartford and Full Scale WAIS I.Q.'s.

The second test included achromatic H-T-P drawings Post Drawing Interrogation (henceforth referred to as P-D-I), and Chromatic H-T-P drawings. The achromatic H-T-P test consisted of free hand drawings on the standardized House-Tree-Person Drawing Form, using one of several lead pencils with erasers of grade number two, as proposed by Buck.<sup>3</sup> A short form of Buck's sixty-eight question P-D-I<sup>4</sup> was constructed for this study. Twenty questions were selected including at least one question of each type (reality, association, and pressure) from each category. All

---

1 Walter C. Shipley, "A Self-Administering Scale for Measuring Intellectual Impairment and Deterioration", in Journal of Psychology, Vol. 9, 1940, p. 371-377.

2 Lloyd K. Sines and Helen Simmons, "The Shipley-Hartford Scale and the Doppelt Short Form as Estimators of WAIS IQ in a State Hospital Population", in Journal of Clinical Psychology, Vol. 15, 1959, p. 452-453.

3 John N. Buck, "The H-T-P Technique, A Qualitative and Quantitative Manual", in Journal of Clinical Psychology, Monograph Supplement No. 5, 1948, p. 11.

4 Ibid., p. 12-24.

questions that aid in scoring were included. The chromatic drawings were made on the standardized form using a number eight set of "Crayola" crayons as requested by Payne.<sup>5</sup>

Four scores, the Per Cent Raw G, the Net Weighted, the Good Score, and Flaw Score, are obtained from each set of drawings. Buck has proposed the following interpretative hypotheses for the scores:

The Raw G IQ, which is derived from raw points only, without consideration of more than their gross differentiation as Good and Flaw respectively, seems largely to represent the quantity of a subject's creativity; to emphasize how much he does. The Net Weighted score which is derived by subtracting the weighted Flaw score from the weighted Good score and thus accords each factor point a more refined differentiation, seems to emphasize the quality of a subject's concepts; to stress how well he does something. The Good IQ score, which is estimated on the basis of the weighted Good points solely, seems to stress the subject's productivity, since it is a measure of his expression of details, and their size and spatial relationships. The Flaw IQ, which is based solely on the measurable errors a subject commits in producing his H-T-P, stresses his power of criticality, seems largely to emphasize his ability to appraise reality aloofly, objectively and analytically.<sup>6</sup>

These interpretive hypotheses are discussed further in the chapter on discussion of results.

---

<sup>5</sup> John T. Payne, "Comments on the Analysis of Chromatic Drawings", in Appendix of "The H-T-P Technique, A Qualitative and Quantitative Scoring Manual", in Journal of Clinical Psychology, Monograph Supplement No. 5, 1948, p. 119.

<sup>6</sup> John N. Buck, "The Quality of the Quantity of the H-T-P", in Journal of Clinical Psychology, Vol. 7, 1951, p. 352.

## 2. Administration of the Test Battery.

The battery of tests was administered to each subject individually. Time limits were strictly enforced on the Shipley-Hartford. The same order of test administration was followed in each case. The first test administered was the Shipley-Hartford, then the achromatic H-T-P, the P-D-I, and then the chromatic H-T-P. The complete battery was administered in one sitting with each individual. All retests were within seven days. The directions as proposed by Buck<sup>7</sup> were followed carefully. The drawings were scored quantitatively according to the criteria set up by Buck<sup>8</sup> in his manual.

## 3. The Sample.<sup>9</sup>

The experimental group consisted of thirty patients, each tested within seven days of admission to the Psychiatric Department of the Ottawa Civic Hospital, Ottawa, Ontario. The subjects ranged in age from twenty to forty-eight years with a mean age of 28.9 years and standard deviation of 7.64.

---

<sup>7</sup> John N. Buck, "The H-T-P Technique, A Qualitative and Quantitative Manual", Op. Cit., p. 11.

<sup>8</sup> Ibid., p. 28-41.

<sup>9</sup> In this study, the terms patient and experimental, and the terms non-patient and control, are used interchangeably.

There were ten males and twenty females in the patient sample, which is comparable to the yearly departmental statistics of twenty-eight per cent male and seventy-two per cent female admissions. The mean Shipley-Hartford I.Q. Score was 103.23 and the standard deviation 10.18.

A non-patient was paired with each patient on the variables age, sex, and I.Q. Bieliauskas and Heffron<sup>10</sup> found that sex and I.Q. influenced the achromatic-chromatic quantitative scores. Female subjects scored significantly higher than males of a comparable level of intelligence on both sets of drawings. The investigators found that I.Q. was a significant variable. The age variable was controlled by the fact that all subjects were high school seniors.

A total of seventy-three non-patients were tested in order to obtain thirty who paired with the experimental subjects on the variables age, sex and intelligence. The control subjects ranged in age from twenty to fifty years with a mean age of 29.1 and a standard deviation of 7.98. There were ten males and twenty females. The mean Shipley-Hartford I.Q. Score was 103.97 with a standard deviation of 8.61.

---

10 V. Bieliauskas and A. Heffron, "Differences in Performance on the Chromatic vs. Achromatic H-T-P Drawings", in Journal of Clinical Psychology, Vol. 16, 1960, p. 334-335.

The experimental subjects were informed that they were taking part in a research project and that their physician approved. Most of the patients were cooperative and none refused to participate. The control subjects were also informed that they were participating in a research project and most were cooperative at the beginning. The majority, however, once presented with the drawing task, became defensive and required considerable support.

#### 4. Description of Statistical Procedures.

A. Test-Retest Reliability.- It was deemed important to assess the reliability of the quantitative H-T-P scores. A test-retest method was used. In the experimental sample, patients who were not transferred or discharged within a period of seven to ten days were retested. Seventeen patients were retested. In the control group, fifteen of the subjects were selected at random and retested within ten days of the initial administration. Product moment correlations were computed for the combined sample between the raw scores of both tests.

B. Scorer Reliability.- Two scorers were used in this part of the study. The scorers jointly reviewed all of the scoring categories previous to scoring the tests. Fifteen sets of drawings from each group were rescored.

Product moment correlations were then computed for the combined sample between the raw scores obtained by the two scorers.

C. The t Tests.- Discrepancy scores between the four achromatic and the four corresponding chromatic raw scores were obtained. Thirty-six t tests, 24 for correlated means and 12 for uncorrelated means, were computed in an attempt to test the significance of all possible discrepancies between the experimental and control groups as well as within each group. A breakdown by groups of the t tests computed is presented in Chapter Three.

D. Chi Square.- The frequency distributions of discrepancy scores demonstrated an apparent tendency of the experimental subjects to group in the upper and lower intervals while the control subjects were grouped in the middle. To test the significance of this observation, four Chi Squares were computed using the extreme control group frequencies as arbitrary cut-off points. Yates' Correction for continuity was used because of the small cell frequencies.

E. Variance Ratios.- The significance of the difference in variance of the discrepancy score distributions of the experimental and control subjects was tested. The difference in variance of the corresponding male and female sub-groups was also tested. This statistic was applied to the data in an attempt to identify differences between the two groups in the variability of scores.

F. Correlations.- An attempt was made to determine if the size of the discrepancy score is related to the size of the achromatic raw scores. Product moment correlations between the four raw scores and the corresponding discrepancy scores were obtained for the experimental group, the control group, and the two groups combined. These results are presented in Chapter Three.

## CHAPTER III

### PRESENTATION OF RESULTS

The results of the statistical analysis of the collected data are presented in this chapter. The presentation of results is made under the following six headings:

1. Reliability of the Quantitative H-T-P Scores;
2. Scorer Reliability of the Quantitative H-T-P Scores;
3. The t Tests;
4. The Chi Square Tests;
5. The Variance Ratios;
6. The Correlations.

#### 1. Reliability of the Quantitative H-T-P Scores.

The product moment correlations obtained from the two sets of thirty-two achromatic protocols representing the test-retest scores obtained within a ten day interval were: .67 for the Per Cent Raw G Score; .59 for the Net Weighted Score; .82 for the Good Score; and .51 for the Flaw Score. The reliability coefficients for the two sets of thirty-two chromatic protocols were as follows: .72 for the Per Cent Raw G Score; .60 for the Net Weighted Score; .77 for the Good Score; and .53 for the Flaw Score. While the coefficients are low, they are all significant at beyond the .01 level.

## 2. Scorer Reliability of the Quantitative H-T-P Scores.

The product moment correlations between the two sets of scores obtained from the two scorers for the thirty achromatic protocols were as follows: .86 for the Per Cent Raw G Score; .88 for the Net Weighted Score; .85 for the Good Score; and .92 for the Flaw Score. The reliability coefficients for the two sets of scores for the thirty chromatic protocols were as follows: .89 for the Per Cent Raw G Score, .93 for the Net Weighted Score; .86 for the Good Score; and .90 for the Flaw Score. The coefficients were all significant at beyond the .01 level.

## 3. The t Tests.

A. t Tests Between Correlated Means.- Twenty-four t tests were computed to test the significance of the mean discrepancy score drops within the groups. More specifically, the significance of the quantitative discrepancies between the achromatic and chromatic drawings for the male experimental, male control, female experimental, female control, total experimental, and total control groups were all investigated separately. The data obtained appears in summary in Table I on the following page. The Flaw Score consistently showed the highest discrepancy score with all but the male experimental group discrepancy significant

Table I.-

t Tests between the Achromatic and Chromatic H-T-P Quantitative Mean Raw Scores for the Patient and Non-Patient Groups.

Sex	Group	N	H-T-P Score			
			Per Cent Raw G	Net Weighted	Good	Flaw
Male	Non-patient	10	2.89	2.91	1.94	3.29 <sup>a</sup>
	Patient	10	2.21	3.16	2.39	3.06
Female	Non-patient	20	2.94 <sup>a</sup>	2.08	1.11	4.38 <sup>a</sup>
	Patient	20	3.04 <sup>a</sup>	2.80	1.68	3.39 <sup>a</sup>
Total	Non-patient	30	4.11 <sup>a</sup>	5.40 <sup>a</sup>	2.13	4.81 <sup>a</sup>
	Patient	30	3.66 <sup>a</sup>	4.01 <sup>a</sup>	2.75 <sup>a</sup>	4.34 <sup>a</sup>

<sup>a</sup> Significant at the .01 level of confidence.

at beyond the .01 level. The discrepancy scores on the Per Cent Raw G Score for the female groups were also significant at beyond the .01 level.

When the groups were combined, all the discrepancy scores with the exception of the control group Good Score were significant at beyond the .01 level.

The pattern of significant discrepancies indicated that the Flaw Scores were particularly sensitive to the difficulties encountered by all subjects in the chromatic drawings. The Good Score was least sensitive to the quantitative discrepancies between the achromatic and chromatic drawings. However, there was a significant drop in the quantitative scores of both groups on the chromatic drawings. These results are discussed in Chapter Four.

B. t Tests Between Uncorrelated Means.- Once it was observed that the discrepancy scores for both the experimental and the control groups were significant, the next step was to see if one group had a greater discrepancy score than the other. For this reason, twelve t tests for uncorrelated means were computed. The data obtained appears in summary in Table II. The tests included male experimental versus male control, female experimental versus female control, and the experimental group versus the control group, on all four scores.

Table II.-

t Tests Between the Patient and Non-Patient, Male, Female,  
and Combined Groups for the Significance of Differences  
on the H-T-P Quantitative Discrepancy Scores.

Patient vs. Non-Patient Groups	N	H-T-P Discrepancy Scores			
		Per Cent Net Raw G	Weighted	Good	Flaw
Males	10	.42	.55	.19	1.03
Females	20	1.47	.68	.40	1.59
Total	30	1.10	.56	.19	.70

All of the discrepancies were in the expected direction; that is, the discrepancies in the experimental group were greater than in the control group. None of the differences were significant at the .01 or even 0.1 level of confidence.

#### 4. The Chi Square Tests.

On the basis of the observed tendency for the discrepancies of the experimental group to be more widely dispersed than the discrepancies of the control group, four Chi Squares using Yates' Correction for small cell entries were computed. It was felt that this was an important statistic in this case because the t tests based on differences between mean scores may not have picked up this apparent difference in distributions. A four cell, 2 x 2 table was employed. The data obtained appears in summary in Table III.

The observed trend was not sufficient to be statistically significant at the .01 level of confidence.

#### 5. The Variance Ratios.

The variance ratios between the standard deviation for the male experimental versus male control, female experimental versus female control, and combined experimental versus control groups, on the four quantitative raw score

Table III.-

Chi Square Tests for the Relationship between the H-T-P  
Quantitative Discrepancy Score Distributions and  
Patient-NonPatient Category.

H-T-P Discrepancy Score	N	Cut-off Point	Chi Square
Per Cent Raw G	30	16	3.61
Net Weighted	30	34	1.57
Good Score	30	23	1.68
Flaw Score	30	13	1.78

distributions led to no consistent results. This data appears in summary in Table IV. The ratio of female experimental to female control was significant at beyond the .01 level on all but the Good Score. Although the variance ratio was not significant, the trend was for the male control group to have a greater variability of scores than the male experimental group. The combined variance ratio of the experimental group to the combined control group on the Per Cent Raw G and Net Weighted Scores was significant at beyond the .01 level.

#### 6. The Correlations.

The calculation of the product moment correlations between the four quantitative achromatic raw scores and the corresponding discrepancy scores for the control, experimental, and combined groups, produced significant results for all groups. The data appears in summary in Table V. With the exception of the Flaw Scores and the Per Cent Raw G Score for the combined groups, the relationships between the size of the quantitative raw scores and the size of the quantitative achromatic-chromatic discrepancy scores were all significant at beyond the .01 level. The achromatic Flaw Score with the corresponding discrepancy score for

Table IV.-

Variance Ratios Between the Male, Female, and Total Group Quantitative  
H-T-P Discrepancy Score Distributions for  
 Patients and Non-Patients.

Sex	Group	N	H-T-P Score							
			Per Cent		Net		Good Score		Flaw Score	
			Raw G	F	Weighted	F	$\sigma$	F	$\sigma$	F
Male	Patient	10	7.63	1.07	17.18	1.99	13.96	1.95	6.66	2.25
	Non-patient	10	7.37		24.25		19.52		9.92	
Female	Patient	20	13.14	4.31 <sup>a</sup>	39.17	5.27 <sup>a</sup>	14.53	1.01	11.52	7.00 <sup>a</sup>
	Non-patient	20	6.23		17.06		14.50		4.35	
Total	Patient	30	11.70	2.96 <sup>a</sup>	21.35	2.61 <sup>a</sup>	14.57	1.34	10.21	1.69
	Non-patient	30	6.81		13.24		16.87		7.39	

<sup>a</sup> Significant at the .01 level of confidence.

Table V.-

Correlation Coefficients between the Quantitative H-T-P  
Achromatic Raw Scores and the Corresponding  
Discrepancy Scores for the Non-Patient,  
Patient, and Total Groups.

Sample	N	H-T-P Score			
		Per Cent Raw G	Net Weighted	Good	Flaw
Non-patient	30	.72 <sup>a</sup>	.47 <sup>a</sup>	.60 <sup>a</sup>	.14
Patient	30	.46 <sup>a</sup>	.51 <sup>a</sup>	.56 <sup>a</sup>	-.40
Total	60	.25	.40 <sup>a</sup>	.51 <sup>a</sup>	-.25

<sup>a</sup> Significant at the .01 level of confidence.

the experimental and combined groups were negative although not significant at the .01 level of confidence.

The results obtained are discussed in Chapter Four.

## CHAPTER IV

### DISCUSSION OF RESULTS

This chapter is concerned with the discussion of results and is presented as follows:

1. Reliability of the Quantitative H-T-P Scores;
2. The t Tests;
3. The Chi Square Tests;
4. The Variance Ratios;
5. The Correlations.

#### 1. Reliability of the Quantitative H-T-P Scores.

The test-retest reliability coefficients are low but sufficient for group comparisons. The test-retest reliability of the chromatic quantitative scores compares favourably with that of the achromatic quantitative scores.

The correlation coefficients between the two scorers indicate a high degree of consistency on both the achromatic and chromatic scores. Although the correlations reported in this study are somewhat lower than those reported by Digiammo,<sup>1</sup> they compare favourably with those reported by Bieliauskas.<sup>2</sup>

---

<sup>1</sup> John J. Digiammo, A Factor Analytic Study of the Quantitative H-T-P, doctoral thesis presented to the University of Ottawa, 1963, p. 22.

<sup>2</sup> Vytautas J. Bieliauskas, "Scorer's Reliability in the Quantitative Scoring of the H-T-P Technique", in Journal of Clinical Psychology, Vol. 12, 1956, p. 366-369.

## 2. The t Tests.

A. t Tests Between Correlated Means.- The Flaw Score in this study was the most sensitive to discrepancies between achromatic and chromatic quantitative H-T-P scores. The significant Flaw Score achromatic-chromatic discrepancy for all groups reflects the difficulty subjects have in drawing with crayons. Mistakes in the crayon drawings cannot be corrected. However, the Flaw Score discrepancies can be interpreted in the light of Buck's<sup>3</sup> hypotheses. If the Flaw Score is interpreted as a measure of criticality, aloofness, and objectivity, then the large Flaw Scores on the chromatic drawings would support Payne's<sup>4</sup> hypothesis that the usage of color promotes a regressive, crude response. However, all scores proved adequate as a gross indicator of the relatively poorer quality of the chromatic drawings. This could again be interpreted in the light of H-T-P quantitative theory as a poorer quality of creativity and lower productivity for all subjects in this sample when producing chromatic drawings. These results support the

---

<sup>3</sup> John N. Buck, "The Quality of the Quantity of the H-T-P", in Journal of Clinical Psychology, Vol. 7, 1951, p. 352.

<sup>4</sup> John T. Payne, "Comments on the Analysis of Chromatic Drawings", in Appendix of "The H-T-P Technique, A Qualitative and Quantitative Scoring Manual", in Journal of Clinical Psychology, Monograph Supplement No. 5, 1945, p. 119.

findings of Bieliauskas and Heffron<sup>5</sup> that the second drawn series is significantly poorer for all groups than the first drawn series.

B. t Tests Between Uncorrelated Means.- No significant differences were found in the comparisons of the discrepancy scores between the experimental and control groups. In each instance, the mean of the discrepancy score for the experimental group was greater than the mean discrepancy score of the control group but the difference was well below an acceptable level of significance. Factors, other than ability to shift, seem to be at play in the significantly lower chromatic quantitative scores. Possibly, the observed fatigue and loss of interest were the important factors in accounting for the significant but equal discrepancy scores for the experimental and control groups.

### 3. The Chi Square Tests.

Arbitrary cut-off points were used to test the observed tendency of greater dispersions in the discrepancy scores of the experimental group. However, through the use of Chi Square, it was found that the number of experimental subjects achieving high discrepancy scores was not sufficient to be significant.

---

<sup>5</sup> V. Bieliauskas and A. Heffron, "Differences in Performance on the Chromatic versus Achromatic H-T-P Drawings", in Journal of Clinical Psychology, Vol. 16, 1960, p. 334-335.

#### 4. The Variance Ratios.

The significant variability of the female experimental Per Cent Raw G, Net Weighted and Flaw Scores, as compared to the corresponding female control scores, suggests that the chromatic scores of clinically abnormal females are less predictable than clinically normal females. A possible reason for the greater variability in chromatic scores could be the heightened emotional lability of some members of the female experimental group. However, this interpretation is speculative at this point because degree of emotional adjustment within the groups was not controlled for in this study.

The variability of the experimental group was significantly greater than the variability of the control group on the Per Cent Raw G and Net Weighted Scores. However, the greater experimental group variability seems to be largely due to the large female experimental group variance and this is discussed above.

#### 5. The Correlations.

There was a significant relationship between the achromatic Per Cent Raw G, Net Weighted and Good Scores and the corresponding discrepancy scores in all groups. This relationship can possibly be accounted for by the fact that there are ceilings on the scores. An individual with a high achromatic score can only go down on the chromatic score.

This relationship is magnified by the fact that if the achromatic score was high because of the use of good but minute details, the individual cannot possibly reproduce all the details with crayons. However, the relationship between the quantitative H-T-P scores and the achromatic-chromatic discrepancy scores should not affect the results of this study because the individuals were matched for levels of intelligence.

The results obtained from this study tend to cast serious doubt upon the hypothesis that the discrepancy scores are of value in differentiating between all clinically normal and all clinically abnormal subjects. Variables, uncontrolled for in this study, seem to be important in accounting for the wide variance of discrepancy scores within each group. Until further studies using more specific criteria are made, the validity of the discrepancy score as a differential diagnostic sign remains suspect.

## SUMMARY AND CONCLUSIONS

In the analysis of quantitative H-T-P discrepancy score comparisons between psychiatric patients and non-patients, it was found that the hypothesis of no significant differences between the discrepancy scores could not be rejected. It was also found that cut-off scores that would differentiate between groups could not be determined.

In view of the findings that the quantitative differences between the achromatic and chromatic drawings of both groups were significant but that none of the differences between groups were significant, the value of the discrepancy score as a gross differential diagnostic sign remains suspect. On this basis, it seems unlikely that anything could be said of the degree of emotional adjustment purely on the basis of discrepancy scores. The assumption here is that the findings are reversible; that is, no differences were found between groups so it is unlikely that the discrepancy score is diagnostic of level of emotional adjustment. However, all hypotheses concerning the relationship between quality of chromatic drawings and level of emotional adjustment cannot be rejected on the basis of this study. There do appear to be relationships between the quantitative achromatic-chromatic discrepancy scores of individuals of comparative levels of adjustment.

There are several possibilities for future research. Attempts should be made to correlate level of emotional adjustment with the quantitative discrepancy scores.

A second suggestion for further study would be to repeat this study with the addition of an adjustment inventory as one of the matching variables. The adjustment inventory would enable the investigator to study patterns of discrepancy scores as related to the emotional adjustment of the individuals within each of the samples.

Other studies should use a combination of quantitative and qualitative criteria in an attempt to validate hypotheses concerning the relationship between level of emotional adjustment and achromatic-chromatic drawing production.

## BIBLIOGRAPHY

Bieliauskas, V.J., and A. Heffron, "Differences in Performance on the Chromatic versus Achromatic H-T-P Drawings", in Journal of Clinical Psychology, Vol. 18, 1960, p. 334-335.

An important contribution providing quantitative information basic to this study.

Buck, John N., "The H-T-P Technique, A Qualitative and Quantitative Manual", in Journal of Clinical Psychology, Monograph Supplement No. 5, 1948, ii-120 p.

An essential source of information for anyone who wishes to understand the quantitative and qualitative aspects of this technique. Contains detailed instructions for quantitative scoring.

-----, "The Quality of the Quantity of the H-T-P", in Journal of Clinical Psychology, Vol. 7, 1951, p. 352-356.

The author's first attempt to spell out in detail qualitative hypotheses based upon objective quantitative scores.

Digianno, John J., A Factor Analytic Study of the Quantitative H-T-P, unpublished doctoral dissertation, University of Ottawa, August, 1962, viii-54 p.

One of the very few attempts to test experimentally an hypothesis based on the quantitative scoring system. Contains test-retest and scorer reliability data.

Payne, John T., "Comments on the Analysis of Chromatic Drawings", in Appendix of "The H-T-P Technique, A Qualitative and Quantitative Scoring Manual", in Journal of Clinical Psychology, Monograph Supplement No. 5, 1948, p. 119-120.

An important contribution providing the theory largely responsible for this research.

APPENDIX 1

ABSTRACT OF

A Comparison Between Psychiatric Patients and  
Non-Patients on the Achromatic-Chromatic  
H-T-P Discrepancy Scores

## APPENDIX 1

### ABSTRACT OF

#### A Comparison Between Psychiatric Patients and Non-Patients on the Achromatic-Chromatic H-T-P Discrepancy Scores.<sup>1</sup>

Hypotheses have been advanced concerning the relationship between level of emotional adjustment and the ability to shift from achromatic pencil to chromatic crayon drawings. There have been no studies reported in the literature that contradict or support these claims.

Level of adjustment being the variable, thirty non-patients were paired with thirty psychiatric patients on the variables age, sex, and I.Q. The achromatic drawings, an interrogation, and the chromatic drawings were administered individually to the thirty subjects in each group.

The hypothesis of no significant difference between the discrepancy scores of the patient and non-patient groups could not be rejected. The discrepancy score does not seem to be of value in the gross differentiation of all clinically normal from all clinically abnormal types.

---

<sup>1</sup> Ronald L. Trites, Master's thesis presented to the School of Psychology and Education of the University of Ottawa, Ontario, May, 1963, vii-34 p.

