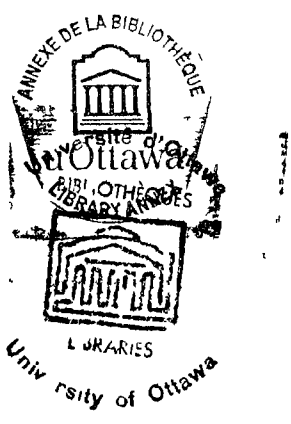
THE FIGURE DRAWING AND THE PHENOMENON OF PROJECTION

by Fern Pickering

Thesis presented to the School of Psychology and Education of the University of Ottawa as partial fulfillment of the requirements for the degree of Doctor of Philosophy in Clinical Psychology



Ottawa, Canada, 1963

ACKNOWLEDGMENTS

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

The author is indebted to W.A. Blair, M.D., for his permission to undertake this research with non-professional staff members of the Ottawa Civic Hospital and to the entire staff of the Department of Psychiatry whose cooperation greatly facilitated the writing of this paper.

CURRICULUM STUDIORUM

Fern Pickering was born November 19, 1935, in Slocan City, B.C. She received a Bachelor of Arts degree from Gonzaga University at Spokane, Washington, in 1958, and a Masters of Science degree in Clinical Psychology from the University of Portland at Portland, Oregon, in 1960. The title of her thesis was Social Desirability and the Lie Scale of the M.M.P.I.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

The figure drawing has in recent years become a much used and abused clinical tool. Its fast rise to popularity appears to have stemmed from its speed of administration, the minimal verbal communication required and its rather dramatic face validity. It is this face validity which can lead the clinician to abuse of the tool. Visual distortions appear to carry more impact than verbal ones, thus encouraging snap judgments which later may prove unwarranted. This phenomenon need not result in condemnation of the tool but it does indicate and, in fact, has touched off a wave of research to provide both qualitative and quantitative norms to counteract the weakness of the clinician's senses.

It is also what theorists claim the figure drawing reveals, and research is beginning to back up, that has added momentum to investigations. A quick look at the patients' unconscious body image and self concept is always sought by the rushed clinician who considers they hold the clue to the nature of the problem. The recent and extensive work of Cleveland and Fisher[1] is one example of a resurgent

---

1 Sidney Cleveland and Seymour Fisher, Body Image and Personality, D. Van Nostrand Co. Inc., Toronto, 1958, xi-420 p.

awareness of the role these images and concepts play in mental illnesses. Wylie's[2] exhaustive survey of self concept research with its many references to the role of the figure drawing is another.

However, in spite of the increasing amount of research seeking a more complete grasp of the meaning of the figure drawing, one receives the impression as much ground has been lost as gained, largely through failure to recognize the complexity of the subject and to design accordingly. It is primarily because of this that Hammer's statement of eight years ago still holds. He said then:

> The analysis of drawing still remains more the intuitive tool of the artisan than the controlled instrument of the scientist.[3]

These problems suggest there are many aspects of the figure drawing where adequately designed research would contribute to our knowledge of this tool.

In view of the design problems inherent in non-phenomenological studies, the present research was directed toward casting some light on the manner in which the figure drawing represents the drawer on the phenomenological level,

---

2 Ruth C. Wylie, The Self Concept, a Critical Survey of Pertinent Research Literature, University of Nebraska Press, Lincoln, Nebraska, 1961, xiii-370 p.

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

while at the same time attempting to cast some light on the manner in which the phenomenon of projection finds expression in the figure drawing.

The first part of this study is concerned with a review of the literature selected to express both the importance of figure drawing research and the design problems this type of research carries with it.

The formulation of the hypothesis is followed by a description of the experimental design which includes a description of the sample and the psychometric battery, its reliability and method of administration. The statistical procedures are also presented.

This is followed by presentation and discussion of the results, weaknesses in the study and, finally, implications for further research.

CHAPTER I

REVIEW OF THE LITERATURE

This chapter reviews the literature pertinent to the development of the problem. It begins with a review of the figure drawing theory, followed by presentation and criticism of studies closely allied to this study, design problems inherent in figure drawing research and, finally, a summary and statement of the working and null hypotheses.

1. Figure Drawing Theory.

That individuals project themselves into any endeavour that requires personal creativity is the basis of psychological diagnostic work. Figure drawings, possibly because of their face validity stemming from frequent obvious visual resemblances to the artists have encouraged much speculation and research by psychologists. From the physical resemblance, speculation and investigation moved to the psychological resemblance of the figure drawing to the drawer, and from there to the more penetrating consideration of the conscious and unconscious psychic resemblances of the drawing to the drawer.

Karen Machover[1] early recognized the psychological value of a person's figure drawings and her 1957 book on the subject, although primarily based on clinical observation rather than research, has become a classic. She explains that:

> A person is impelled to draw from some source when asked to draw a person. External figures are too varied in their body attributes to lend themselves to a spontaneous, composite, objective representation of a person. Some process of selection involving projection and introjection enters at some point. The subject must draw consciously, and no doubt unconsciously upon his whole system of psychic values. The body, or the self, is the most intimate point of reference in any activity. ... consequently, the drawing of a person involving a projection of the body image, provides a natural vehicle for the expression of one's body needs and conflicts. Successful drawing interpretation has proceeded on the hypothesis that the figure drawn is related to the individual who is drawing with the same intimacy characterizing that individual's gait, his handwriting or any other of his expressive movements.[2]

She concludes that the figure drawing is:

> (...) the graphic representation of the subject's self evaluation. The processes of selection and organization of what is to be drawn may take place with varying degrees of awareness and of directness. Most drawings contain elements of self-evaluation in both direct and compensated forms of projection and of both conscious and unconscious phases of self revelation.[3]

---

1 Karen Machover, Personality Projection in the Drawing of the Human Figure, Charles C. Thomas, Springfield, Ill., 4th printing, 1957.

2 Ibid., p. 5.

3 Ibid., p. 9.

Levy,[4] studying the figure drawing about the same time as Machover, outlined essentially the same basic assumptions. However, he was more clear in stating that the nucleus of the figure drawing was determined by the body image. He held other factors such as culture, biochemical makeup, characterological factors and even personal training exerted a peripheral influence on what was produced.

Buck, becoming more inclusive than Machover or Levy, by including the house and tree drawings and at the same time becoming more specific about their meaning, postulates that "each of the drawn wholes in the House-Tree-Person Test is to be regarded as a self-portrait".[5] Hammer, also speaking of the House-Tree-Person Test, says:

> The drawing medium, particularly the drawing of the dynamic and familiar concepts of House, Tree and Person, has been found to enhance the projection of the subject's deepest fantasies, wishes, conflicts and fears on conscious and unconscious levels. The unconscious is bared especially in the language of graphic symbols. In this way the subject often brings to the surface what he does not or cannot say in words.[6]

---

4 Sidney Levy, "Projective Figure Drawing", as quoted by Emanuel F. Hammer in The Clinical Application of Projective Drawings, Charles C. Thomas, Springfield, Ill., 1958, p. 85-86.

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

6 Emanuel F. Hammer, The H-T-P Clinical Research Manual, Western Psychological Services, Beverly Hills, Cal., 1955, Introduction.

Jolles, in his House-Tree-Person manual, initially agrees with Buck, describing the subject's drawing of a person as, "the subject's impression of himself from a psychological and/or physiological standpoint".[7]

It is in Jolles' elaboration of this, without adequate definition of terms, that the confusion possible in interpretation becomes evident. He states that the figure drawing may also represent:

> The subject as he would like to be.
>
> The subject's concept of his sexual role.
>
> The subject's attitude toward inter-personal relationships in general.
>
> The subject's attitude toward a specific inter-personal relationship.
>
> The person in the subject's environment whom the subject most likes.
>
> The person in the subject's environment whom he most dislikes.
>
> A person toward whom the subject has ambivalent feelings.[8]

Jolles' list of possible interpretations of the figure drawing is a clear representation of the confusion that has developed in figure drawing interpretation. Which facet of the self the diagnostician is to choose as the

---

7 Isaac Jolles, A Catalogue for the Quantitative Interpretation of the H-T-P, Western Psychological Services, West Los Angeles, Cal., 1952, p. 93.

8 Ibid., p. 93.

correct interpretation of any particular drawing has become almost a matter of personal preference.

2. Presentation and Criticisms of Studies.

A large quantity of research has followed these initial clinical observations but failure to define operationally exactly what was being investigated and poor designs has increased rather than abated the confusion. Wylie has made the greatest effort to critically evaluate the progress of research in clarifying the role of self concept, both phenomenal and non-phenomenal, in the figure drawing.

She criticizes self concept research, including research into figure drawings and the self concept, as inconclusive because "the theories are vague, incomplete and overlapping and no one theory has received extensive, empirical exploration".[9] She feels

> (...) some of the most crucial difficulties center around the degree to which self-concept theorists wish to be and can fruitfully be, consistently phenomenological.[10]

There have been many studies based on the premise that people draw their body image. In interpreting the results

---

9 Ruth C. Wylie, The Self Concept, A Critical Survey of Pertinent Research Literature, University of Nebraska Press, Lincoln, Nebraska, 1961, p. 317.

10 Ibid., p. 9.

of these studies one is misled by the lack of adequate definition of the phrase 'projection into drawings'. For example: Kotkov and Goodman[11] claim they investigated the premise that 'one's body image is projected into one's drawings'. They compared the drawing of a person by obese women with those of a control group of non-overweight women. They found the drawings of the obese women were significantly larger and wider than those of the control group. In fact, it would appear that what they meant by projection was that the obese subjects actually drew an aspect of their self image.

Berman and Leffel[12] did a similar study. They compared the somatotypes of thirty-nine males with their figure drawings of a man, and a statistically significant correlation emerged in support of the hypothesis of what they termed 'projection of the body image'. By this apparently was meant graphic representation of the body image.

---

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

12 S. Berman and J. Leffel, "Body Type and Figure Drawing", Journal of Clinical Psychology, Vol. 9, No. 4, 1953, p. 368-370.

Cleveland and Fisher[14] defined projection in terms of what the subject verbally attributes to his drawings. In a study comparing the figure drawings of arthritic patients whose disorder involved the musculature of the outer body layer with patients with physical symptoms involving the body interior they found the arthritic patients gave significantly more responses emphasizing body impermeability than the other group. The theory behind the study was that arthritic patients think of their bodies as covered by a hard outer shell as a barrier against psychological threats. Patients with peptic ulcers and other inner disorders, they believed, viewed their body surface as an inadequate defense area easily penetrated. The conclusion was not that the arthritic patients drew heavy body outlines and those with inner illnesses drew weak lines but that the subjects ascribed these characteristics to the drawings when asked to describe them.

More recent research, like that of Cleveland and Fisher's, has recognized that one must distinguish between the phenomenological and non-phenomenological aspects of the figure drawing if valid interpretations of the results are to be made. With this in mind some researchers have attempted to measure only

---

14 S. Cleveland and S. Fisher, "Body Image Boundaries in Various Psychosomatic Illnesses", APA Convention, California, 1955, as quoted by Emanuel F. Hammer, The Clinical Application of Projective Drawings, Charles C. Thomas, Springfield, Ill., 1958, p. 24.

what they feel is unconsciously represented. To do this, they limited themselves to the more easily distinguished body image.

Hunt and Weber,[15] working with the non phenomenlogical body image, devised a Body Image Projective Test and and objective rating scale. The test consisted of various nude figures of women and the subjects were asked to pick the one most like them. The objective rating scale was made up of the parts of the bodies in the projective test. They gave both tests to their subjects and found that both tests were highly reliable but that they did not correlate significantly with each other. They concluded from this that either the two tests were inherently different or that they measure different aspects of the body image. They felt this indicated their projective test was measuing the unconscious body image. The objective rating scale used can be criticized for not controlling set and saturation and also failure to clarify the meaning of their scale which was a rating of: less than normal -- normal -- more than normal. It is unclear what more than normal meant.

However, their greatest error appears to be in the initial assumption that figures devised by the tester for

---

15 Valerie V. Hunt and Mary Ellen Weber, "Body Image Projection Test", Journal of Projective Techniques, Vol. 24, No. 1, 1960, p. 3-10.

evaluation draw forth the same projection phenomenon as when a subject actually draws a figure.

It is exactly this assumption that Silverstein[16] criticizes in a similar study made by Levi.[17] Silverstein holds that the assumption has not yet been demonstrated that the physical body, the body image, and the drawn figure are in, what he terms, "isomorphic relationship".

Levi attempted to show that the physical body and the body image did have a one-to-one relationship as the first step in proving the figure drawing was also similarly related. Taking thirty-eight orthopedically disabled subjects and a matched normal group, he instructed them to pick out which one of each pair of stick figures looked most like the model Levi had devised of three main parts: legs, arms, and a head-spine combined. When the experimental group with leg and arm disabilities showed sensitivity to these disabled parts, Levi concluded a one-to-one relationship between easily perceived physical disabilities and their perception of the human figure. Silverstein's criticism was that demonstrated sensitivity to a limb on a figure by a

---

16 A.B. Silverstein, Comment on Levi's "Orthopedic Disability as a Factor in Human Figure Perception", Journal of Consulting Psychology, Vol. 26, No. 1, 1962, p. 9.

17 Aurelia Levi, "Orthopedic Disability as a Factor in Human Figure Perception", Journal of Consulting Psychology, Vol. 25, No. 3, 1961, p. 253-256.

person with a disabled limb did not constitute a one-to-one relationship, for other parts of his body not included in the model may have been equally sensitive.

3. Problems in Design.

The drawback in this kind of study is to devise an adequate means for tapping unconscious imagery. Freud,[18] Symonds[19] and others consider free association, slips of tongue, dreams and symbolism as the primary expressions of the unconscious. Holt, Havel, et. al.,[20] following the commonly held Rorschach rationale of projection of the unconscious onto ambiguous stimuli, have attempted to score products of the unconscious on the Rorschach. Leary,[21] too, has devised an inter-personal checklist for measuring the various levels of self concept. However, the non-verbal nature of the unconscious does not lend itself to direct verbalization and, consequently, even after the results are

---

18 Sigmund Freud, A General Introduction to Psychoanalysis, Perma Books, Garden City, N.Y., p. 25, 29-30, 105-106.

19 Percival M. Symonds, Diagnosing Personality and Conduct, The Century Company, N.Y., 1931, p. 361-399.

20 Robert Holt, Joan Havel, et.al., Manual for the Scoring of Primary Process Manifestations, Research Center for Mental Health, New York University, 1962.

21 Timothy Leary, "A Theory and Methodology for Measuring Fantasy and Imagination Expression", Journal of Personality, Vol. 25, 1956, p. 159-175.

in and statistically analyzed, one is still forced to make an inference from the results to the degree they represent the unconscious imagery. It is here unwarranted assumptions are most often made.

Another recent study approached the problem from the opposite position. Kamano,[22] with the use of the Semantic Differential, attempted to show that the subject drew himself when he drew a person. This study sidestepped the problem of unconscious and conscious representation simply by assuming that if a subject rated himself on the Semantic Differential Scale similarly to the way he rated his figure drawing on the same scale then he was drawing himself. Kamano concluded:

> For this population at least the subjects tended to draw a figure that approximated their actual self much more than an idealized or unfavourable one.[23]

Kamano did his study with forty-five institutionalized schizophrenic women between the ages eighteen to thirty-five with a mean age of 29.11 years. He had his subjects draw a woman and then he had them rate their drawings of the woman on a Semantic Differential Rating Scale made up of fifteen items representing three of Osgood and Suci's

---

22 Dennis K. Kamano, "An Investigation on the Meaning of Human Figure Drawings", Journal of Clinical Psychology, Vol. 16, No. 4, 1960, p. 429-430.

23 Ibid., p. 430.

factors. He then had them rate, on the scale, their actual, ideal, and least-like selves. The last three ratings were counterbalanced. He found a significant relationship at the .01 level between the way the subjects rated their drawing on the scale and the way they rated their actual selves . From this he concluded that the subjects were, in fact, drawing themselves.

Kamano's study may be criticized for the inference that his subjects were in fact drawing themselves. His data showed a significant relationship between the subjects' ratings of their drawings and their ratings of their actual selves, but these results leave room for other conclusions. It is possible that his subjects may have been projecting some aspect of themselves into their rating of the drawing much in the same manner a subject projects into the Thematic Apperception Test figures. The possibility of this projection occurring may have been heightened by the fact that when a subject makes an object or draws a figure he tends to become somewhat ego involved and may over identify with the object or figure. Vroom[24] found this could happen in a study he did in 1957 using a questionnaire and Q technique to determine conditions under which a person attributes his own

---

24 Victor H. Vroom, "Projection, Negation and the Self Concept", Human Relations, Vol. 12, No. 4, 1959, p. 335-344.

attitudes and opinions to others. He found projection of one's own personal attitudes is towards persons for whom the subject has a positive attitude.

Kamano's study can also be criticized for failure to heed the pitfalls of the Semantic Differential Rating Scale. He allowed close proximity between his subjects' ratings of their drawings and the ratings of the three aspects of self. This was a particular problem because his rating scale had only fifteen scales and responses could possibly have been remembered from one rating to the next. He did not control for set, saturation, or the influence of a preceding scale upon the next scale and he made no attempt to determine test-retest reliability.

4. Summary and Hypotheses.

Studies to date investigating how a figure drawing represents the drawer have created confusion rather than clarification. The research reveals a number of reasons for this. One is a failure to distinguish the phenomenological from the non-phenomenological self in research designs. A second is that conclusions broader than the designs warranted have been drawn. A third is the lack of adequate operational definitions clarifying the meaning of the term 'projection into drawings' for each individual study.

This study, mindful of the above criticism, will attempt to cast some light on the manner in which the figure drawing represents the subject. It will limit itself to a small portion of what Wylie[25] terms the 'phenomenological self'. This study will simply try to determine if the subject when asked to draw a person does draw himself or if the drawing simply becomes a vehicle onto which the subject projects his own conscious view of himself, much in the same manner he would project himself into T.A.T. figures.

For this study, projection was operationally defined as 'unwittingly attributing one's own traits to a figure drawing'.

The working hypotheses for this study were as follows:

1. A subject will be considered to be drawing himself when a reliable judge attributes the same characteristics to the subject's figure drawing as the subject attributes to himself.

2. He will be considered to be projecting himself into his drawings when he attributes the same characteristics to his figure drawing as he attributes to himself but these characteristics are not seen, or are seen to a significantly smaller degree in his drawings by an unbiased reliable rater.

The null hypothesis used to submit the general hypotheses to statistical tests can be summarized as

---

25 Wylie, Op. Cit., p. 6.

follows: There is no significant relationship between the way a subject rates his drawings, the way he rates himself and the way an independent rater rates the subject's drawing as measured by a Semantic Differential Rating Scale.

CHAPTER II

EXPERIMENTAL DESIGN

This chapter presents the procedure followed in order to explore the hypotheses proposed in the preceding chapter. It begins with a description of the sample followed by a description of the psychometric battery, the validity and reliability of the battery and the method of test administration. Finally, a description and explanation is given of the statistical procedures employed in the investigation.

1. The Sample.

The sample consisted of thirty male and thirty female non-professional employees of the Ottawa Civic Hospital. A total of sixty-seven subjects were tested in order to obtain thirty female and thirty male subjects with I.Q.'s over ninety. The subjects ranged in age from twenty to fifty years with a mean age of 30.98 years and a standard deviation of 9.40. I.Q. scores on the Shipley Hartford ranged from eighty-nine to 124. The mean I.Q. score was 108.60 with a standard deviation of 10.08.

2. The Psychometric Battery.

The test battery consisted of two tests and a rating scale.

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 ten minute time limit. The test was included as a screening device to assure the subject had adequate intelligence to understand instructions. Shipley,[1] in validating the test, used a sample of 522 army recruits. He reported a .87 reliability for the vocabulary sub-test, .89 reliability for the abstract thinking sub-test, and .92 for the two combined. A recent study by Sines and Simmons[2] reported a product-moment correlation of .90 and a standard error of estimate of 6.33 between the Shipley Hartford and Full Scale W.A.I.S. I.Q.'s.

The second test administered was the House-Tree-Person Test to be referred to hereafter as the H-T-P. The H-T-P was used rather than the Draw-a-Person Test for two reasons:

1. It was felt that the drawing and rating of the three figures rather than one would conceal from the subject the purpose of the ratings; i.e.,

---

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

2 Lloyd K. Sines and Helen Simmons, "The Shipley Hartford Scale and the Doppelt Short Form as Estimators of WAIS I.Q. in State Hospital Population", Journal of Clinical Psychology, Vol. 15, No. 4, 1959, p. 452-453.

that the subject's rating of his figure drawing was to be compared with his self ratings.

2. To include sufficient ratings that the subject would not be able to remember how he rated his figure drawing as compared to his self ratings.

The third test was a rating scale based on Osgood, Suci and Tannenbaum's Semantic Differential Rating Scale.[3]

The Semantic Scale was composed of eighteen bipolar scales separated by seven rating categories, drawn from the Thesaurus Study.[4] This study and others have determined the usefulness and the factorial loading for specific scales but point out that there are no standard scales which must be administered. The authors suggest that the choice of scales should be directed by the known factorial composition, relevance to the concept to be judged and the semantic stability of the particular scale. They indicate that some connotation scales become denotative depending on what is being judged. Since in this study drawings were being evaluated, such scales as large-small, high-low became denotative and were therefore not used, although their factor loading was high.

---

3 Charles E. Osgood, George J. Suci, and Percy H. Tannenbaum, The Measurement of Meaning, University of Illinois Press, 1957, p. 76-124.

4 Ibid., p. 53-61.

Eighteen scales were selected for the study representing the three factors evaluative, potency, and activity, which were considered most adequate to evaluate various aspects of self concept and figure drawings. However, two scales, strong-weak and hot-cold, were later dropped for reasons to be developed in the discussion of a pilot study.

The initial Semantic Differential Rating Scales chosen were:

The Evaluation Factor
- attractive - repelling
- successful - unsuccessful
- important - unimportant
- healthy - sick
- happy - sad
- sociable - unsociable

The Potency Factor
- hard - soft
- strong - weak
- deep - shallow
- masculine - feminine
- heavy - light
- constrained - free

The Activity Factor
- active - passive
- hot - cold
- tense - relaxed
- aggressive - defensive
- impulsive - deliberate
- motivated - aimless

3. Semantic Differential Reliability.

This study was concerned with three aspects of Semantic Differential reliability:

i) Item reliability
ii) Judge reliability
iii) Subject test-retest reliability.

1) Item Reliability.- The authors report an overall coefficient of correlations of .85 on the test-retest of twenty concepts rated on fifty scales by one hundred subjects.

Osgood, et.al.[5] quote a series of studies on item reliability. They quote a crude study by Solomon to determine if the joint distribution differs significantly from that which would be obtained if subjects assigned check marks purely at random on two testings. They applied the test to fifty scales and found all scales deviated from this change estimate at better than the one per cent level. This indicates that the test-retest data are related on something other than a chance basis, but does not tell how reliable the items are.

Osgood[6] also quotes a study by Wilson who took the actual distributions of both test and retest check marks to determine if the joint distributions of responses were such that complete independency was shown between the two

---

5 L.N. Solomon, "A Factorial Study of Comples Auditory Stimuli (Passive Sonar Sounds)", unpublished doctor's dissertation, University of Illinois, 1954, as quoted by Charles E. Osgood, George J. Suci, and Percy H. Tannenbaum, The Measurement of Meaning, University of Illinois Press, 1957, p. 128-129.

6 Kellogg Wilson, "Multidimensional Scaling of Data Obtained by Method of Triads", Urbana: Control System Lab., University of Illinois, 1954, as quoted by Osgood, et. al., Ibid., p. 128-129.

testings. He found all forty test-retest items to yield a difference significant at the one per cent level or better.

Grigg,[7] using forty-two undergraduates, found significantly greater distance between "ideal self" and "neurotic" than between "self" and "neurotic", concluding from this the validity of using the Semantic Differential for distinguishing self concepts.

On the other hand, Block[8] concluded the Semantic Differential method a poor tool for measuring ideal self, although he found a .73 correlation between Semantic Differential ratings of ideal self and an adjective check list.

ii) Judge Reliability.- Stolz and Coltherp[9] did a study of a judge's ability to predict intellectual, social, and emotional criteria about a subject from the D.A.P. and also to provide data above the ability of individual clinicians to make such predictions. The study found all judges predicted intelligence more accurately than social

---

7 A.E. Grigg, "A Validity Study of the Semantic Differential Technique", Journal of Clinical Psychology, Vol. 15, No. 2, 1959, p. 179-181.

8 J. Block, "An Unprofitable Application of the Semantic Differential", Journal of Consulting Psychology, Vol. 22, No. 3, 1958, p. 235-236.

9 Robert F. Stolz and Frances C. Coltherp, "Clinical Judgments and the D.A.P. Test", Journal of Consulting Psychology, Vol. 25, No. 1, 1961, p. 43-45.

and emotional adjustment. They found also that one judge predicted more accurately than others.

Allport, speaking of the reliability of trained judges in rating others, states:

> It is fairly certain that people who have accurate self knowledge tend to rate another accurately. Most studies show that good judges are socially skillful and emotionally stable and are rated high in leadership and popularity.[10]

Fisher and Fisher,[11] in attempting to differentiate schizophrenics from normals on the Draw-a-Person Test, found trained psychologists showed no greater agreement among themselves than did untrained judges. The study involved evaluating the drawings of thirty-two paranoid schizophrenics by both a detailed atomistic analysis and the total impressions of the drawing.

Because in this study the validity of the design rested upon judge reliability in rating the subject's figure drawing, a pilot study was done to determine the effect of possible variables on the judge ratings. It could be foreseen that results could be confounded if certain factors were operant in the judge's ratings. These were:

---

10 Gordon W. Allport, Pattern and Growth in Personality, Holt, Rinehard and Winston, New York, 1961, p. 508-509.

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

1) projection of personal values or personality traits into the figure drawing ratings; 2) use of different norms; 3) effect of sex on rating; and 4) rating on a deeper psychic level than the subject.

In the pilot study, five judges, two men and one woman trained in the diagnostic use of the figure drawing, and one woman and one man not trained in its use were asked to rate twenty drawings of non-professional subjects employed in the Ottawa Civic Hospital.

The first step was to determine judge reliability for each scale. This was done by finding the correlation between all five judges, i.e., the average correlation when every judge was correlated with every other judge.

The intra-class correlation formula[12] used was:

$$r_{kk} = \frac{V_p - V_e}{V_p}$$

r_{kk} = reliability for mean ratings from k raters.

V_p = variance for persons.

V_e = variance for error.

In contrast to Stolz' and Coltharp's findings, it was found that no one judge, nor any pair of judges were

---

12 J.P. Guilford, Psychometric Methods, McGraw-Hill Book Co. Inc., New York, 1954, p. 395.

significantly reliable. Consequently, the median rating of the five judges was used.

As can be seen in Table I, two scales hot-cold and strong-weak with correlations of .41 and .57 respectively, were considered unreliable and were dropped. The reliabilities for the remaining scales ranged from .71 to .90 and were considered adequate for this study.

The overall reliability of .81 indicated judges were essentially using the same norms and no further investigation of this was necessary.

All the scales were then tested for sex differences using the Mann-Whitney U Test.

$$\text{Critical Ratio } \left(\frac{X}{\sigma}\right) = \frac{U_1 - \frac{N_1 N_2}{2}}{\sqrt{\frac{N_1 N_2 (N_1 + N_2 + 1)}{12}}}$$

$$\text{where } U = N_1 N_2 + \frac{N_1 (N_1 + 1)}{2} - T_1$$

This is a formula[13] for the significance of the difference between two independent groups, when a "distribution - free" type of significance test is required. The description of the terms in this formula and computational details are discussed by McNemar.

---

13 Quinn McNemar, Psychological Statistics, John Wiley and Sons Inc., New York, 1955, p. 359-360.

Table I.-

Intra-Judge Reliabilities on Those Semantic Differential Scales Considered for Use in This Study.

Factor	Scale		r_{kk} [a]
Evaluative	attractive	- repelling	.89
	important	- unimportant	.71
	healthy	- sick	.90
	sociable	- unsociable	.88
	successful	- unsuccessful	.74
	happy	- sad	.85
Potency	hard	- soft	.71
	strong	- weak	.57
	deep	- shallow	.77
	masculine	- feminine	.85
	constrained	- free	.84
	heavy	- light	.71
Activity	active	- passive	.80
	hot	- cold	.41
	tense	- relaxed	.89
	aggressive	- defensive	.80
	impulsive	- deliberate	.80
	motivated	- aimless	.82

[a] r_{kk} = Reliability for mean ratings from k raters.

As indicated in Table II, the scales constrained-free and attractive-repelling received U scores of 2.01 and 2.05 respectively, indicating the sex of the judges was influencing their ratings on these scales at the .05 level of significance. However, in view of their overall judge reliability this was not considered sufficiently high to make them unreliable. However, the scales strong-weak, hot-cold, already dropped because of unreliability, received U scores of 3.70 and 2.82 respectively, indicating sex differences significant at the .01 level. The U test results between the sexes on the remaining scales, ranging from .40 to 1.93 were not significant.

Finally, the relationship between the ratings of trained judges and non-trained judges was obtained, again using the Mann-Whitney U Test. As can be seen in Table III, none of the U test results, ranging from .58 to 1.91 were significant, indicating no significant difference between the ratings of trained and non-trained judges. Subsequently, it was concluded that the judges were rating on the same psychic level as the subjects.

These findings are in agreement with Fisher and Fisher[14] who found trained psychologists showed no greater agreement among themselves in differentiating paranoid schizophrenics from normals than did untrained judges.

---

14 Fisher and Fisher, Op. Cit., p. 727-732.

Table II.-

U Tests of the Significance of Difference Between Male and Female Judge Ratings of the Subject's Figure Drawings For Each Semantic Differential Scale.

Factor	Scales		U Score Between Male and Female Judges[a]
Evaluative	attractive	- repelling	<u>≤.05</u>
	important	- unimportant	1.90
	healthy	- sick	1.93
	sociable	- unsociable	1.66
	successful	- unsuccessful	.69
	happy	- sad	1.07
Potency	constrained	- free	<u>≤.01</u>
	hard	- soft	1.30
	deep	- shallow	1.50
	masculine	- feminine	.40
	heavy	- light	1.73
	strong	- weak	<u>3.70</u>
Activity	active	- passive	1.35
	tense	- relaxed	1.69
	aggressive	- defensive	.98
	impulsive	- deliberate	1.36
	motivated	- aimless	1.81
	hot	- cold	<u>2.82</u>

a Single underline designates significance at .01 (U = 2.57).
Double underline designates significance at .05 (U = 1.96).

Table III.-

U Tests of the Significance of Difference Between Ratings of the Subject's Figure Drawings Made by Trained and Non-Trained Judges for Each Semantic Differential Scale.

Factor	Scales		U Score Between[a] Trained and Non-Trained Judges
Evaluative	attractive	- repelling	1.60
	important	- unimportant	1.34
	healthy	- sick	1.62
	sociable	- unsociable	1.49
	successful	- unsuccessful	1.16
	happy	- sad	1.40
Potency	hard	- soft	1.69
	deep	- shallow	1.71
	masculine	- feminine	1.08
	constrained	- free	1.12
	heavy	- light	1.91
Activity	activity	- passive	1.24
	tense	- relaxed	1.74
	aggressive	- defensive	1.10
	impulsive	- deliberate	.58
	motivated	- aimless	1.13

a U = 1.96 to be significant at .05 level of confidence.

iii) Subject Test-Retest Reliability.- Suinn, Osborne and Winfree[15] investigated the accuracy of recall of the self concept as measured by a 100-adjective check list given to freshmen. Tested again five days later, it was found that items consistent with the subject's self concept were most accurately recalled.

Norman[16] also studied the stability characteristics of the Semantic Differential Scales and found reliability between tests given over a four-week period, as long as there was no systematic intervening treatment.

A pilot study to determine subject test-retest reliability was also done. The retesting was done over a period of one to four weeks. Test-retest reliability for the subject's rating of his figure drawing, rating of his actual self, ideal self and least-like self were found and the mean reliability for the twenty subjects determined.

The formula used was:

$$r = 1 - \frac{\sum d_{ii}^2}{2k}$$

---

15 R. M. Suinn and Osborne, Dorenes and Winfree, "The Self Concept and Accuracy of Recall of Inconsistent Self Related Information", Journal of Clinical Psychology, Vol. 18, No. 4, 1962, p. 473-474.

16 Warren T. Norman, "Stability Characteristics of the Semantic Differential", American Journal of Psychology, Vol. 72, 1959, p. 581-584.

This formula is a derivative of one quoted by Osgood.[17]

$$D_{il} = \sqrt{2k(1-r_{il})}$$

where k = number of scales or rows in the matrix.
r_{il} = the correlation coefficient between the two concepts i and l.

The derivation was obtained through the following steps:

$$\sqrt{2k(1-r_{il})} = D_{il}$$

$$2k(1-r_{il}) = D_{il}^2$$

$$1-r_{il} = \frac{D_{il}^2}{2k}$$

$$r_{il} = 1-\frac{D_{il}^2}{2K}$$

The D-score is central to the Semantic Differential technique. Osgood states that:

$$D_{il} = \sqrt{\sum d_{il}^2}$$

where d_{il} = the difference between the two concepts i and l as rated on one scale of a matrix.

It follows that $D_{il}^2 = \sum d^2_{il}$

If the formula for r_{il} is stated in terms of d, the final formula becomes:

$$r_{il} = 1 - \frac{\sum d_{il}^2}{2k}$$

---

17 Osgood, et.al., Op. Cit., p. 91.

To obtain the overall test-retest correlation for each of the concepts, the sum of the d^2_{11} was obtained and divided by the appropriate 2k.

The formula was:

$$r_T = 1 - \frac{\sum\sum d^2}{N2k}$$

When N = number of subjects.

The overall reliability for all the intra-subject ratings was .72. As indicated in Table IV, this was significant at the .001 level. The mean reliabilities for all subjects on each concept was: P = .51; S = .77, IS = .89, and LL = .71. The mean reliabilities for each subject on the four concepts ranged from .65 to .94.

These findings upheld those of Norman and Quinn, Osborn and Winfree.

More recent studies have investigated the ability of the Semantic Differential Method to measure different aspects of the self concept.

Lochlin[18] questioned whether self ratings were in fact representing the self. He raised the possibility that words can be used differently even in the different ratings of the self. He gave the example that a person when he calls himself solemn may mean serious, but if he rates ideal self

---

18 John E. Lochlin, "Word Meanings and Self Description", Journal of Abnormal and Social Psychology". Vol. 62, No. 1, 1961, p. 28-34.

Table IV.-

Subject Test-Retest Reliability.

Subject	Test-Retest r[b]				
	P[a]	S	IS	LL	r_T
6	.94	.91	.94	.97	.94
7	.45	.96	.97	.82	.80
8	.95	.97	.94	.82	.92
9	.49	.85	.96	.62	.73
11	.54	.59	.97	.51	.65
22	.72	.73	.88	.61	.74
24	.41	.71	.91	.77	.70
26	.55	.86	.93	.86	.81
29	.85	.83	.94	.87	.87
30	.44	.90	.90	.93	.79
31	.48	.63	.88	.63	.66
32	.50	.88	.94	.98	.83
33	.72	.80	.89	.90	.83
34	.49	.78	.89	.88	.76
35	.76	.86	.90	.74	.82
47	.78	.78	.98	.90	.85
50	.85	.79	.87	.97	.87
52	.77	.97	.97	.88	.90
53	.75	.92	.96	.86	.87
57	.47	.75	.97	.96	.79
r_T[c]	.51	.77	.89	.71	.72

a P = subject's test and retest ratings of his figure drawing.
S = subject's test and retest ratings of his actual self.
LS = subject's test and retest ratings of his ideal self.
LL = subject's test and retest ratings of his least-like self.

b Each individual correlation within the table requires a correlation of .63 to be significant at the .01 level.

c Total correlations require a correlation of .20

as solemn he may mean careful, and if rating another person solemn he may mean without humor or obstreperous. This is a fundamental problem of the semantic rating scale. Its effect on this study would be to reduce the chances of correlation rather than increase them.

Some words have different meaning for different people. Lochlin also questioned the possibility that differences in self description may simply represent the difference in the meanings people are giving the words rather than differences in the way they view themselves. He designed a study of college students and found that consistent individual differences in self description amongst the students were no greater than consistent individual differences in the meanings of the words. He suggested that a means of reducing the discrepancy in interpreting words was to define the words for all subjects.

James E. Madden[19] raises the question of forcing a subject to rate concepts which he may consider unimportant in relationships to his picture of himself, thus giving them undue weight as representing his self concept. However, from his study where he compared the subject's self description with his self description on a semantic rating scale, he

---

19 James E. Madden, "Semantic Differential Rating of Self and of Self-Reported Personal Characteristics", Journal of Consulting Psychology, Vol. 25, No. 2, 1961, p. 183.

concluded that if a subject says a concept describes him it does. In this study when a subject felt a concept did not apply he was free to score it in the neutral bracket of the scale, indicating it didn't apply.

Further on this problem Wylie[20] reminds that only a limited structured report of the subject's view of himself is obtained in this type of study. Again, this study was not intended to exhaust all possible views of the self, simply to sample the subject's different views of himself and to determine from the samples any possible phenomenological relationship to his figure drawing.

One of the problems in this study, inherent in all self report methods used to rate more than one aspect of the self at the same time, was the degree to which they overlap and have a common denominator. Wylie considers one portion of this common denominator social desirability. She criticizes a study done by Cowan and Tongas[21] who claimed a method to determine the presence of social desirability loadings by having the items used in a self rating scale rated independently for social desirability loading. When

---

20 Ruth C. Wylie, The Self Concept, A Critical Survey of Pertinent Research Literature, University of Nebraska Press, Lincoln, Nebraska, 1961, p. 154.

21 E.L. Cowan and P.N. Tongas, "The Social Desirability of Trait Descriptive Terms, Applications to a Self-Concept Inventory", Journal of Consulting Psychology, Vol. 23, No. 4, 1959, p. 361.

they found a zero correlation between this scale and rating of ideal self and actual self they concluded no social desirability was present. Wylie considered these results were the product of improper statistical procedures and concluded:

> The problem of the influence of social desirability on the validity of the subjects self report concerning discrepancies between self concept and ideal self, remains unsettled.[22]

Kogan and Fordyce[23] also attempted to devise a check list to estimate social desirability with little success.

Even if the scales do have high social desirability ratings this does not mean they are not part of a subject's conscious self concept.

Wylie[24] has reviewed studies which draw attention to uncontrolled variables when the Semantic Differential Method is used. She lists them as: 1) response sets; 2) halo effect; 3) tendency to check one end of the scale; 4) touchy content areas encourage denial and scoring in the neutral bracket of the scale; 5) saturation causes fatigue and less attention toward the end of the test; 6) a preceding scale can effect both the interpretation of the next

---

22 Wylie, Op. Cit., p. 27-29.

23 William S. Kogan and Wilbert E. Fordyce, "The Control of Social Desirability", Journal of Consulting Psychology, Vol. 26, No. 1, 1962, p. 26-30.

24 Wylie, Op. Cit., p. 32-33.

scale and the way it is scored; and 7) some scales are scored differently by the two sexes.

This study attempted to overcome these weaknesses by randomizing the Semantic Differential Rating Scale four times with each group of fifteen subjects using a different form.

4. Administration.

The test battery was administered to each subject individually.

The procedure followed was:

1. Administration of Shipley-Hartford Scale.
2. Administration of the H-T-P, given according to Buck's instruction.
3. Subjects were then asked to rate their three drawings on the Semantic Differential Rating Scale. The ratings of the drawings were counterbalanced to retain an equal distance between each rating and the self ratings to follow.
4. Subjects were then asked to rate their 'actual selves', their 'ideal selves', and their 'least-like' selves', on the Semantic Differential Rating Scale. These were also counterbalanced to avoid set and saturation, the effect a scale may have on the preceding scale and to keep the various self ratings equal distance from the drawing ratings.

5. Five independent raters, two men and a woman trained in the H-T-P technique, and one man and one woman not trained, rated the subject's drawings of a person. The instructions suggested by Osgood, Suci and Tannenbaum[25] were used in explaining the rating scale to both subjects and judges.

5. Statistical Procedure.

All judge and subject scores obtained on the Semantic Differential Rating Scale were converted into standard scores. These were obtained by summing each subject's scores for each concept across the sixteen scales.

For those subjects who drew figures of the sex opposite their own, the sign of the standard score on the masculine-feminine scale was reversed for each of their three self ratings. This was done to allow for expected disagreement between these self ratings and ratings of a figure drawing of the opposite sex.

Correlations were then obtained between the concepts rated by the subjects, that is:

a) between the way the subject rated his figure drawing and the way the subject rated his actual self;

b) between the way the subject rated his figure drawing and the way the subject rated his ideal self;

c) between the way the subject rated his figure drawing and the way he rated his least-like self.

---

25 Osgood, et.al., Op. Cit., p. 342.

d) between the way the subject rated his actual self and the way he rated his ideal self.

The formula used for this procedure was:[26]

$$r = 1 - \frac{\Sigma d_{il}^2}{2k}$$

The total intra-correlations for all sixty subjects between concepts were then obtained by means of the following formula:[27]

$$r_T = 1 - \frac{\Sigma\Sigma d^2}{N2k}$$

The significance of all the above correlations was determined by means of t tests.

From the five judges it was necessary to obtain one rating for each scale for each subject. This was done by taking the median score. The sixteen resulting scores for each individual were then converted to standard scores by again summing across scales. These sixteen scores were maintained as a profile of scores in all further computations. However, for the sake of clarity, this profile will, in future, be referred to as the judge's rating.

---

26 See p. 29 present study.

27 See p. 31 present study.

The two above formulae were those used for all correlations throughout the study - the former for individuals and the latter for groups.

The judge's rating of each drawing was compared to those made by the corresponding subject. This resulted in correlations between the judge's rating of the figure drawing and: a) the subject's rating of his figure drawing; b) the subject's rating of his actual self; c) the subject's rating of his ideal self; and d) the subject's rating of his least-like self.

The total correlations for these relationships were also computed and the significance of all these correlations were determined by *t* tests.

All the above mentioned correlations were then obtained for the twenty retest subjects and again for the original test scores for the same group. This was done so that it could be determined later if any of the results would have been significantly different had the retest scores been used instead of the test scores. (Although the pilot study indicated subjects were significantly reliable, it was yet possible that familiarity with the tasks during retesting left the subjects less threatened and therefore more objective. In this case, one might expect the retest ratings to be closer to those of the judges.)

For the total group, two relationships were broken down according to sex rather than according to concepts:

a) the relationship between the judge's rating of the figure drawing and the subject's rating of his actual self;

b) between the subject's rating of his figure drawing and his rating of his actual self.

The purpose here was to obtain a rough estimate of the validity of each scale; it was not deemed necessary to repeat this procedure for all possible relationships. Therefore only those relationships including the most independent and fundamental concepts were used.

The total correlations for these relationships were computed by the same method, i.e. r_T method, described above and the significance of these relationships for each scale was determined by *t* tests.

The next statistical procedure was to determine whether any of the total correlations were significantly different from each other. All the important differences were tested for significance by the following formula:[28]

$$t_{r_{12}-r_{13}} = \frac{(r_{12} - r_{13}) \sqrt{(N-3)(1+r_{23})}}{\sqrt{2(1 - r^2_{12} - r^2_{13} - r^2_{23} + 2r_{12}r_{13}r_{23})}}$$

---

28 McNemar, Op. Cit., p. 148.

Also the correlation between the two correlations involved in each test of significance was found by the formula:[29]

$$r_{r_{12}r_{13}} = r_{23} - \frac{r_{12}r_{13}\left(1 - r^2_{23} - r^2_{12} - r^2_{13} + 2r_{12}r_{13}r_{23}\right)}{2(1 - r^2_{12})(1 - r^2_{13})}$$

Both the above formulae took into consideration the presence of a common variable within the two correlations to be compared.

---

29 Lawrence T. Dayhaw, Manuel de Statistique, Editions de l'Universite d'Ottawa, Ottawa, 1958, p. 319.

CHAPTER III

PRESENTATION OF RESULTS

The results and discussion of the statistical analysis of the data are presented in this chapter under the following headings:

1. Total Correlations;
2. t Tests of the Significance of the Difference Between Correlations;
3. Correlated Relationships broken down according to Scales;
4. Discussion.

1. Total Correlations.

The correlations between the various subject ratings for the main sample, the test and retest groups, as presented in the first section of Table V, were between:

a) the subject's rating of his figure drawing and his rating of his actual self (P - S);

b) the subject's rating of his figure drawing and his rating of his ideal self (P - IS);

c) the subject's rating of his figure drawing and his rating of his least-like self (P - LL);

d) the subject's rating of his actual self and his rating of his ideal self (S - IS).

The correlations obtained between the judge's ratings and the various subjects' ratings for the main

sample, the test and retest groups, as presented in the second portion of Table V, were between:

a) judge's rating of the figure drawing and the subject's rating of his actual self (J - S);

b) judge's rating of the figure drawing and the subject's rating of his ideal self (J - IS);

c) judge's rating of the figure drawing and the subject's rating of his least-like self (J - LL);

d) judge's rating of the figure drawing and the subject's rating of the figure drawing (J - P).

All correlations were significant at the .01 level. Closer inspection reveals that the subject's rating of the figure drawing was closer to his various self ratings than was the judge's. The largest correlation of .70 was between the subject's rating of his actual and his ideal self. This was followed by a correlation of .65 between the subject's rating of his drawing and his rating of his ideal self and an equally significant negative relationship of -.52 between the subject's rating of his drawing and his least-like self.

In the correlations involving the judge, the highest relationship was .36 between the judge's rating of the subject's drawing and the subject's rating of his actual self. The judge's and the subject's rating of the drawing revealed a slightly smaller relationship of .34 and the judge's rating of the drawing and the subject's rating of his least-like self received a negative correlation of -.23.

Table V.-

Total Correlations Between the Subjects' Ratings and the Correlations Between the Judge and the Subject Rating for the Main Sample and the Test and Retest Groups.

	Total Correlations (r_T)							
	P - S[a]	P - IS	P - LL	S - IS	J - S	J - IS	J - LL	J - P
Test group N:60[b]	.61[c]	.65	-.52	.70	.36	.29	-.23	.34
Test group N:20[b]	.42[d]	.47	-.45	.67	.24	.25	-.22	.42
Retest group N:20	.49	.52	-.43	.71	.30	.26	-.23	.36

a S = Subject's rating of his actual self.
IS = Subject's rating of his ideal self.
LL = Subject's rating of his least-like self.
P = Subject's rating of his figure drawing.
J = Judge's rating of the figure drawing.

b df for r_T = 16n - 2.

c All correlations significant at .01 (r = .12).

d All correlations significant at .01 (r = .15).

Correlations in the test and retest group retained the same level of significance. Variations in these correlations were in the same direction as the main sample.

It should be noted that the truly independent variables would appear to be the subject's rating of his drawing, his rating of his actual self and the judge's rating of the drawing. The subject's rating of his ideal self had a high component of his rating of his actual self in it and when the actual and the ideal self were rated, the subject's rating of his least-like self was almost totally determined. The only other possible cause of variation would have been unreliability. Consequently, when the relationship between the subject's rating of his drawing and his actual self, his rating of his actual and his ideal self and the judge's rating and the subject's rating of the drawing were known all other relationships could be roughly estimated. Knowing this, certain inner comparisons were not sufficiently independent to warrant determining statistically.

2. *t* Tests of the Significance of the Difference Between Correlations.

The correlations presented in Table V, when tested for significant differences, yielded varied results as indicated in Table VI. In the main sample the correlation between the subject's rating of his figure drawing and his actual self rating was significantly different from the

Table VI.-

t Tests of the Significance of Difference Between Total Correlations for the Main Sample, the Test and the Retest Groups.

r's Compared	t Tests		
	Test Group N:60[a]	Test Group N:20[b]	Retest Group N:20[b]
$(P_1S)-(P_1IS)$	<u>2.76</u>	1.33	.83
$(J_1S)-(J_1IS)$	<u>3.75</u>	.24	1.27
$(P_1S)-(J_1S)$	<u>8.16</u>	<u>3.46</u>	<u>3.63</u>
$(J_1S)-(J_1P)$	.76	<u>3.46</u>	1.20

a Single underline designates significance at .01 (t = 2.66). Double underline designates significance at .001 (t - 3.46).

b Double underline designates significance at .01 (t = 2.84).

correlation between his figure drawing rating and his ideal self rating at the .01 level. There were two sets of correlations significantly different at the .001 level. One was between the judge's rating and the subject's actual self rating and the judge's rating and the subject's ideal self rating. The other was between the subject's rating of his drawing and his actual self rating and the judge's rating and the subject's actual self rating.

In order to draw conclusions from the test-retest groups which would validly apply to the study as a whole, it would have to be demonstrated that the smaller test group was representative of the total sample. In other words, correlations in the test-retest groups significantly different from those in the total test sample would be meaningful only in those cases where the retest group was significantly different in some way from both of the test groups. As the table indicates, the smaller test group was not representative of the total sample since *t* tests significant in one group were not necessarily significant in the other, and in no case was it found that the retest group was significantly different from both tests groups.

The correlations found within the twenty-subject test group were then compared with the corresponding correlations obtained within the retest group. These comparisons, presented in Table VII, are in two forms, correlations

Table VII.-

Comparisons in the Form of Correlations and t Tests Between the 20-Subject Test and Retest Groups for Each Correlation Obtained Between the Judge's Rating and the subject's Ratings.

	$J\text{-}P_1$	$J\text{-}P_2$	$J\text{-}S_1$	$J\text{-}S_2$	$J\text{-}IS_1$	$J\text{-}IS_2$	$J\text{-}LL_1$	$J\text{-}LL_2$
$r_{r_1 r_2}$[a]	.55[b]		.79		.91		.71	
$t_{r_1\text{-}r_2}$	1.41[c]		2.38		.53		.26	

a Arabic number 1 designates the test group and arabic number 2 designates the retest group.

b All correlations between correlations significant at .01 (r = .54).

c All differences between correlations not significant. (t = 2.53 to be significant at .01)

between correlations and t tests of the difference between correlations. Only those relationships involving the judges were computed and it can be seen that the correlations, which ranged from .55 to .91, remained significant at .01 and the t test results, which ranged from .26 to 2.38, were not significant. In effect then, when correlating the judge's rating of the figure drawing with the subject's rating of the figure drawing it made no significant difference whether the subject's rating was obtained during his test session or during his retest session and the same held true for all the subject's ratings.

That no significant difference between the test and retest groups was found for the relationships seen in this table, confirms previous findings that the test-retest groups were generally not significantly different from each other. Therefore these results did not warrant any further attempts to demonstrate that the retest group was unique.

3. Correlated Relationships Broken Down According to Scales.

Correlations obtained when the data was broken down according to the scales indicated all the correlations between the judge's rating of the figure drawing and the subject's rating of his actual self were somewhat lower than the subject's figure drawing rating and his actual self rating.

These results can be seen in Table VIII. The one exception was the scale heavy-light where the judge-actual self relationship was slightly higher. When correlating the subject's rating of his figure drawing with the subject's rating of his actual self, only the scale deliberate-impulsive, with a correlation of .16, was not significant. However, when the judge's rating of the figure drawing was compared with the subject's rating of his actual self, five scales failed to produce a significant relationship. These scales and their correlations were: constrained-free -.17; deep-shallow -.13; deliberate-impulsive -.11; active-passive -.08; and unsociable-sociable -.14.

Correlations for the remaining scales ranged from .54 to .85 between the subject's rating of the figure drawing and his actual self and from .32 to .69 between the judge's rating of the drawing and the subject's actual self.

4. Discussion.

Keeping in mind the terms 'drawing himself' and 'projecting himself' as defined for this study, it can be seen from the statistical analysis that a subject, when asked to draw a person does, in certain respects, draw himself. When asked to go one step further and evaluate his drawing the subject, to some extent, projects himself into his rating. That is, in certain areas he is fairly

Table VIII.-

Correlations Indicating the Relationship Between the Subject's Rating of His Figure Drawing and His Actual Self and Between the Judge's Rating of the Drawing and the Subject's Actual Self for Each Scale

Scale		r between P - S where N:22	r between J - S where N:39
Factor I (Evaluative)			
attractive	- repelling	.73[a]	.54[b]
important	- unimportant	.76	.59
healthy	- sick	.83	.37
sociable	- unsociable	.77	-.14
successful	- unsuccessful	.79	.52
happy	- sad	.54	.32
Factor II (Potency)			
constrained	- free	.57	-.17
hard	- soft	.69	.62
deep	- shallow	.54	-.13
masculine	- feminine	.75	.40
heavy	- light	.62	.69
Factor III (Activity)			
active	- passive	.63	-.08
tense	- relaxed	.57	.43
aggressive	- defensive	.63	.50
impulsive	- deliberate	.16	-.11
motivated	- aimless	.66	.48

a Single underlining designates significance at .01 (r = .52).

b Single underlining designates significance at .01 (r = .39).

Double underlining designates significance at .05 (r = .30).

realistic while in others he becomes less critical and describes his production in idealized terms not warranted by the drawing.

The high level of significance found between the subject's and the judge's rating of the drawing and between both these ratings and the subject's various self ratings indicates a definite relationship between a subject and the figure he draws. This relationship is not strictly representative of one aspect of the person's self concept but includes aspects of his ideal self as well. This conclusion was reached on the basis that both the subject and the judge, to a significant degree, saw traits similar to the subject's actual and ideal self in the subject's drawing; with the least-like self being seen by both as negatively related. Keeping in mind that all relationships were significant, closer analysis was warranted. This revealed that the rating of the drawing by the judge, who could be considered more objective than the subject, corresponded most closely to the subject's actual self rating. This would indicate the subject's drawing revealed more traits descriptive of himself as he felt he really was than of his ideal or least-like self even though the subject did not rate it this way himself.

On one scale, however, when considering either the subject or the judge as being the criteria, this relationship

failed to hold up. This was the scale measuring the trait of impulsiveness.

To distinguish where objectivity ceased on the part of the subject and subjectivity began required a careful evaluation of results. Evidence of projection could be seen in the fact that, in spite of the common factor between judge and subject ratings, the correlations concerning the subject's rating of his figure drawing, his actual, ideal and least-like self were considerably higher than any of those involving the judge. This would tend to indicate that the subject's rating of his drawing was related to his various ratings of himself to a significantly greater degree than would seem to be justified in the light of the judge's objective analysis.

Further, the subject's rating of his figure drawing was significantly more related to his ideal self, that is to what he would like to be, than any of his other self evaluations.

On the other hand, the judge's evaluation of the drawing was more closely related to the subject's actual self, indicating the nature of the projection seemed to be toward idealization of the drawing.

It was interesting to note also that the subject found significant similarity of traits between his drawing and his actual self on four scale where no similarity was

seen by the judge. The traits represented by these scales were: constraint, depth, activity, and sociability. In view of the extremely high correlations between the subject's rating of his drawing, his ideal and actual self, it is likely the subject's analysis of his drawing on these traits reflected his tendency to project the ideal. This raises the possibility that the subject may have rated his drawing as nearer perfect than he rated his actual self, even in the face of the fact that all the judges commented on the poor quality of many of the drawings. This is even more interesting in the light of the fact that the drawing of a person or of any similar visual-motor task would be expected to put a subject in a critical frame of mind. Most subjects did, in fact, make critical comments about their drawings and their own ability during the drawing period and the examiner noted frequent signs of embarrassment.

It was possible to look for verification of projection in one further area. If, as Wylie[1] and other theorists hold, anxiety or threat to self esteem or lack of familiarity with the task can affect a subject's self evaluation, and if subjects when tested initially experienced this (there was some indication they did from their spontaneous comments

---

1 Ruth C. Wylie, The Self Concept, a Critical Survey of Pertinent Research Literature, University of Nebraska Press, Lincoln, Nebraska, p. 198.

and observations made during testing); and if familiarity with the task on retesting reduced this threat, then it could be predicted their self ratings would have varied at retesting. Evidence of this would have supported the above findings that subjective elements were entering into the subject's ratings if not truly projection. It would also have upheld recent findings that projection occurs where there is some emotional need or value of the individual, the need here being reassurance and security. However, retest results did not verify this in that they did not differ significantly from the initial test results when compared to an outside criteria, i.e., the judge. As could have been predicted, the size of the sample appeared to vary the results more than any other factor. This conclusion was supported by the fact that some inexplicable significant differences were found in the twenty-subject test group used for test-retest purposes that were not present in the total sample.

The initial conclusions that the subject can be considered in certain respects to be drawing himself concur to some degree with Kamano's findings. Kamano, using an almost identical Semantic Differential Scale, found a relationship significant at the .01 level between the way his subjects rated their drawings and their actual selves. He found no significant relationships between his subjects' ratings of their drawings and their ideal or least-like

selves. Insofar as the subject's actual self rating was significantly related to the subject's drawing, this study upheld Kamano.

One cannot say with certainty why in this study the subject's rating of his drawing should exhibit a greater relationship with his ideal self than his actual self as Kamano found. However, the obvious explanation, aside from the already noted possibility of uncontrolled variables in Kamano's study, is that the samples were taken from different populations. Kamano's sample of forty-five institutionalized schizophrenic women seemed to recognize less psychic distance between their drawing and their true feelings about themselves while the normals appear to have attributed more idealized traits to them. It is true that the difference between the drawing, the ideal and the actual self in this study, while significant, was not so large as to preclude common elements. In view of this, it is also plausible some explanation may lie in the commonly held assumption that the normal person's actual and ideal self should not be too far apart. Certain Rogers[2] hold this to the degree he uses the reduction of the distance between

---

2 Carl Rogers, Client-Centered Therapy, The Houghton Mifflin Company, Boston, 1951, p. 140.

actual and ideal self concepts as criterion of progress in psychotherapy.

Machover and Levy's conclusions that a person's drawing does represent self evaluation representative of varying degrees of awareness and directness are clearly demonstrated. Buck's conclusion that a figure drawing can be regarded as a self portrait is also supported if one does not go on to define 'self portrait'.

Overall results of this study did not uphold the findings of Spilka and Lewis of disowning projection occurring in a subject's evaluation of his drawing. The relationship between the subject ratings of his drawing and least-like self were consistently and significantly different.

SUMMARY AND CONCLUSIONS

A summary of the design and methodology of the study, the conclusions reached, the limitations placed on them by weaknesses in the study and indications for further research will be presented in this chapter.

The purpose of this research was to cast some light on the manner in which the figure drawing represents the subject who draws it. The method chosen to determine this was to have sixty normals take the H-T-P, then have them rate their drawings, then their actual, ideal, and least-like selves in counterbalanced order on a Semantic Differential Scale drawn from Osgood, Suci and Tannenbaum's Thesaurus Study. Five judges were also asked to rate the subject's figure drawing. Judge and subject ratings of the subjects' figure drawings, their actual, ideal and least-like selves were then tested for significant relationships and differences.

Two pilot studies were undertaken, one to determine intrajudge reliability and one to determine subject test-retest reliability. Both the judges and subjects were found to be reliable.

Statistical treatment of the data involved conversion of all Semantic Differential scores into standard scores, determination of correlations between all important ratings

and determination of the significance of these correlations by means of *t* tests.

Analysis of the statistics revealed a positive but dual relationship between the subject's and the judge's rating of the drawing and between both of these ratings and the subject's various self ratings. Further analysis revealed a positive relationship between the subject's actual self, his ideal self and his drawing and a negative one between his drawing and his least-like self. From this it became evident that a person's figure drawing could be interpreted in terms of both his actual and his ideal self concepts.

That there was at the same time projection into his rating of his drawing on the part of the subject could be seen in the fact that the subject saw a significantly closer relationship between his drawing and his various self concepts than the judges recognized.

Further, it could be seen that the subject was projecting into his rating of his drawing traits not seen by the more objective judges and that these were most likely traits idealizing the drawing and over-looking its weaknesses. It was also important to note that this occurred even when the subject was often embarrassed by the sight of his drawing and verbally critical of it, while at the same time all five judges commented on the poor quality of many of the drawings.

Consequently, it followed that, since a significant relationship was found between the way the subject and independent judges rated the subject's figure drawing, and the way the subject rated himself, the null hypothesis that no significant relationship would be found could be rejected.

Interpretation and practical use of the conclusions are limited by a number of weaknesses in the study and its tools. Firstly, the study purported only to be investigating the drawing - projection phenomenon on the phenomenological level. At best, preconscious aspects of the subject's actual self may have been tapped by the judges' ratings. Secondly, the actual and the ideal self were not as sharply distinguished in the results as might have been desired. Thirdly, although a relationship was demonstrated between the actual and ideal selves and the drawing, this does not exclude the possibility of other concepts not studied, being equally important. Fourthly, the subject was actually seen to be represented in his drawing in this study by only eleven variables. Fifthly, there was no true 'outside' point of reference or comparison; no known, independent zero point for each adjective or trait, i.e., it is not known in what ways or to what extent all of the drawings and the subjects may have had certain of the traits in common. People compared to other people or to drawings

may appear to have differences when only compared among themselves, but compared to some outside criteria, or an outside zero point, they may all be quite alike. All drawings and all people do have some things in common, but how far from this 'common' factor the traits used as measures in this study were, is not known.

Finally, the results were group results with individuals not always agreeing with the group as a whole.

The results from this study suggest several fruitful research areas. Repetition of this study with one design change immediately suggests itself as a means of further clarifying results. The design change would be to have the judges rate the subject by means of his drawing rather than rate the drawing for its own merits. If this resulted in a significantly higher relationship between the judge's rating and the subject's actual self rating while the results of the subject's rating remained unchanged, i.e., his drawing and the ideal self remained most closely related, then projection on the part of the subject would be more clearly demonstrated. Also this would result in more clear-cut evidence that the drawing is revealing at least an unrecognized, likely pre-conscious, level of the subject's self concept.

It would also be of interest to try to determine if the presence of common elements between the subject's drawing, his actual and ideal self was, in fact, a

manifestation of Rogers' theory that the normal's actual and ideal self should be close together.

Further investigation of the phenomenon of projection suggests itself in the denial that seemed to exist in the individual's evaluation of his drawing, often in the face of visual proof of poor quality, verbal recognition of this and signs of embarrassment that included such things as keeping the figure drawing partially hidden from the examiner by an arm or by turning it over on completion or asking the tester not to look at it.

This phenomenon is particularly interesting in that it occurred in a normal sample but not in Kamano's institutionalized schizophrenic sample and gives rise to the speculation that normal individuals possible have more recourse to the ego-defending mechanism of denial than schizophrenics.

The marked difference between the manner in which the normal and schizophrenic subject evaluated their figure drawings also indicates the possibility of perceptual differences in mentally ill persons. This is a question long disputed by theorists who phrase the problem in terms of whether abnormalacy is simply an exaggeration of normalacy or if psychotic disorders have different patterns or traits, among them perceptual ones. Eysenck, one of the leading theorists in this area, has investigated the hypothesis

first put forth by Kretschmer that a normality-abnormality continuum exists. In one factor analytic study done on intercorrelation of tests between one hundred normals, fifty manic-depressives and fifty schizophrenics, he found evidence that:

> Schizophrenic and manic-depressive insanity are not qualitatively different from normal mental states but form a continuum which goes all the way from the perfectly normal to the completely insane, psychotic individual.[1]

However, when he tried to distinguish common patterns of traits between normals and psychotics in this study, he could not.

If such a continuum does not exist, the validity of generalizing knowledge of traits and dynamics obtained from research with normals to abnormal populations would be in question.

This study also suggests possible research into the Semantic Differential Rating Scale method itself. Although the pilot study into the effect of a judge's sex on his rating of the Semantic Differential was significant at the .01 level on only two scales and the .05 level on two others, this does indicate that sex does influence ratings.

---

1 H.J. Eysenck, "Schizothymia-Cyclothymia as a Dimension of Personality", Journal of Personality Vol. 20, No. 3, 1952, p. 345-384.

The explanation of why trained and non-trained judges rated on the same psychic level when trained judges are assumed to be capable of deeper evaluations is likely accounted for by the type of instructions given. Being asked to rate the drawing and not the person doing the drawing possibly caused the judges to take a set on a less intuitive level than usual. Research into this possibility would also seem worthwhile.

SELECTED BIBLIOGRAPHY

Berman, S. and J. Leffel, "Body Type and Figure Drawing", Journal of Clinical Psychology, Vol. 9, No. 4, 1953, p. 368-370.

A study of projection of the body image into figure drawings which exemplifies how failure to operationally define projection confounds results.

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

This manual for the scoring and interpretation of the H-T-P is quoted in this study in support of the basic premise that the figure drawing does, in some manner, represent the drawer.

Cleveland, Sidney and Seymour Fisher, "Body Image Boundaries in Various Psychosomatic Illnesses", A.P.A. Convention, California, 1955, as quoted by Emanuel F. Hammer, The Clinical Application of Projective Drawings, Charles C. Thomas, Springfield, Illinois, 1958, xxii-663 p.

One of the few studies which adequately distinguish between subjects ascribing certain characteristics to their figure drawings and actually drawing them.

Hammer, Emanuel F., The Clinical Application of Projective Drawings, Charles C. Thomas, Springfield, Illinois, 1958, xxii-663 p.

Leaders in the H-T-P and figure drawing technique present their clinical and experimental findings in this text. Hammer concludes with a review of rather highly selected research, presented to support the theory put forth. The text serves best as an introduction to leading theorists' views on drawing.

Hunt, Valerie V. and Mary Allen Weber, "Body Image Projection Test", Journal of Projective Techniques, Vol. 24, No. 1, 1960, p. 3-10.

A recent study attempting to measure projection of the non-phenomenological body image. The assumption that the same projection phenomenon is drawn forth by figures devised by the testers as when subjects draw their own figures is questionable.

Jolles, Isaac, A Catalogue for the Qualitative Interpretation of the H-T-P, Western Psychological Services, West Los Angeles, California, 1952, 97 p.

Strictly a hand-book for interpretation of drawings, including the figure drawing. When used for diagnostic purposes it is of value only in the hands of an experienced clinician. Its value in this research was to serve as an example of the confusion that can arise from unqualified interpretive hypotheses put down in cook-book form.

Kamano, Dennis K., "An Investigation of the Meaning of Human Figure Drawings", Journal of Clinical Psychology, Vol. 16, No. 4, 1960, p. 429-430.

Kamano's research is closely related to this study and, for this reason, served throughout as a guide to help avoid the faulty reasoning, pitfalls in design and inadequate statistical procedures encountered by Kamano. Its similarity also made it an important source of comparison with results obtained in this study.

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

Another study of projection of body image into the figure drawing which fails to operationally define projection.

Levi, Aurelia, "Orthopedic Disability as a Factor in Human Figure Perception", Journal of Consulting Psychology, Vol. 25, No. 3, 1961, p. 253-256.

A study investigating the theory of an isomorphic relationship between the physical body, the body image and the figure drawing.

Machover, Karen, Personality Projection in the Drawing of the Human Figure, Charles C. Thomas, Springfield, Illinois, 1957, ix-181 p.

Although many of the pathological traits considered diagnostic of emotional disturbance in this text remain unproven, Machover's theoretical interpretation of the projection mechanism behind figure drawings is important. Also the text, when taken in its entirety, sharpens one's awareness of the diagnostic possibilities of the figure drawing.

Osgood, Charles E., George J. Suci and Percy H. Tannenbaum, The Measurement of Meaning, Urbana, University of Illinois, 1957, 342 p.

The primary research and early studies on the Semantic Differential are contained in this book.

Silverstein, A.B., "Comments on Levi's Orthopedic Disability as a Factor in Human Figure Perception", Journal of Consulting Psychology, Vol. 25, No. 3, 1961, p. 253-256.

A timely criticism of Levi's study and the theory behind both Levi's study and that of Hunt and Weber.

Vroom, Victor H., "Projection, Negation and the Self Concept", Human Relations, Vol. 12, No. 4, 1959, p. 335-344.

A study investigating projection of personal attitudes, a possibility overlooked by Kamano in evaluating his subjects' ratings of their figure drawings.

Wylie, Ruth C., The Self Concept, A Critical Survey of Pertinent Literature, University of Nebraska Press, Lincoln, Nebraska, 1961, xiii-370 p.

At present this is the leading text surveying and critically analyzing self concept research. This includes considerable figure drawing research. The sections dealing with the phenomenon of projection of self into the figure drawing, design problems arising from failure to distinguish the phenomenological from non-phenomenological aspects of self concept in figure drawing research and the weakness of the tools used, were all important sources of knowledge for this study.

APPENDIX 1

INSTRUCTIONS TO SUBJECTS

APPENDIX 1

INSTRUCTIONS TO SUBJECTS

The purpose of this study is to measure the meaning of various drawings and self concepts. You are to rate these things on this set of scales.[1]

Here is how you are to use these scales:

If you feel that the drawing or concept is very closely related to one end of the scale, you should place your check-mark as follows:

fair--X--:-----:-----:-----:-----:-----:-----:unfair

OR

fair-----:-----:-----:-----:-----:-----:--X--:unfair

If you feel that the drawing or concept is quite closely related to one or the other end of the scale (but not extremely), you should place your check-mark as follows:

strong-----:--X--:-----:-----:-----:-----:-----weak

OR

strong-----:-----:-----:-----:-----:-----X:-----weak

If the drawing or concept seems only slightly related to one side as opposed to the other side (but is not really neutral), then you should check as follows:

active-----:-----:--X--:-----:-----:-----:-----passive

OR

active-----:-----:-----:-----:--X--:-----:-----passive

The direction toward which you check, of course, depends upon which of the two ends of the scale seem most characteristic of the thing you're judging.

If you consider the drawing or concept to be neutral on the scale, both sides of the scale equally associated with the

---

1 Instructions to the judges were identical except that they were asked to rate only the figure drawing.

drawing or concept, or if the scale is completely irrelevant, unrelated to the drawing or concept, then you should place your check-mark in the middle space:

safe-----:-----:-----:--X--:-----:-----:-----dangerous

IMPORTANT: (1) Place your check-marks in the middle of spaces, not on the boundaries:

-----:-----:-----:--X--:-----:-----:-----

(2) Be sure you check every scale for every concept - do not omit any.

(3) Never put more than one check-mark on a single scale.

Do not look back and forth through the pages. Do not try to remember how you checked previous lines. Make each rating a separate and independent judgment. Work quickly and carefully. It is your first impressions, your immediate feelings that are important.

These general instructions were then followed by the specific instructions:

> Now, look at your figure drawing. I want you to rate this drawing on the set of scales before you.

The words 'drawing of a house' and 'drawing of a tree' were substituted for 'figure drawing' in a counterbalanced manner.

Next the subject's drawing was removed and he was instructed as follows:

> Now I want you to rate your actual self, that is rate yourself as you feel you really are (...). Now rate your ideal self, that is rate yourself as you'd most like to be (...). Now rate your least-like self, that is rate yourself as you'd least-like to be.

These instructions were also given in counterbalanced form.

APPENDIX 2

A SAMPLE OF ONE OF THE FOUR SETS OF THE SEMANTIC DIFFERENTIAL RATING SCALES

APPENDIX 2

No._____

A SAMPLE OF ONE OF THE FOUR SETS OF THE SEMANTIC DIFFERENTIAL RATING SCALES

aggressive_____:_____:_____:_____:_____:_____:_____:defensive

unsuccessful_____:_____:_____:_____:_____:_____:_____:successful

deep_____:_____:_____:_____:_____:_____:_____:shallow

hot_____:_____:_____:_____:_____:_____:_____:cold

unimportant_____:_____:_____:_____:_____:_____:_____:important

free_____:_____:_____:_____:_____:_____:_____:constrained

motivated_____:_____:_____:_____:_____:_____:_____:aimless

sad_____:_____:_____:_____:_____:_____:_____:happy

impulsive_____:_____:_____:_____:_____:_____:_____:deliberate

unsociable_____:_____:_____:_____:_____:_____:_____:sociable

soft_____:_____:_____:_____:_____:_____:_____:hard

active_____:_____:_____:_____:_____:_____:_____:passive

healthy_____:_____:_____:_____:_____:_____:_____:sick

repelling_____:_____:_____:_____:_____:_____:_____:attractive

feminine_____:_____:_____:_____:_____:_____:_____:masculine

heavy_____:_____:_____:_____:_____:_____:_____:light

weak_____:_____:_____:_____:_____:_____:_____:strong

tense_____:_____:_____:_____:_____:_____:_____:relaxed

APPENDIX 3

THE SAMPLE, ITS COMPOSITION AND THE INDIVIDUAL CORRELATIONS

APPENDIX 3

THE SAMPLE, ITS COMPOSITION AND THE INDIVIDUAL CORRELATIONS

S	Sex	IQ	r's Between Concepts	r's Judge Vs Subjects	Subject Reliability	S
1	M M	116	98 96 97 -98	86 79 78 -62		1
2	M M	99	66 04 -16 06	-11 28 67 -31		2
3	M M	107	85 48 44 -52	59 49 31 -37	53 53	3
4	M M	118	78 80 92 -82	72 68 60 -46	77 97 97 88 90	4
5	F F	123	02-31 44 23	76-29 79 24	75 68 95 01 60	5
6	F F	108	98 81 92 -75	-18-18-20 22	94 91 94 97 94	6
7	F F	113	43 39-04 -28	29-42-61 53	45 96 97 82 80	7
8	F F	110	92 94 93 -96	58 53 63 -61	95 97 94 82 92	8
9	F F	114	61 57 46 -69	42 82 54 -65	49 85 96 62 73	9
10	F F	119	19 62 24 -34	-31-201.00 72		10
11	F F	101	68 72 59 -64	-12-36-46 25	54 59 97 51 65	11
12	F F	112	74 67 69 -68	44 74 67 -18		12
13	F F	116	38 72 77 -22	70 26 66 -27		13
14	M M	111	55 63 82 -82	30 58 52 -52		14
15	F F	98	88 74 83 -74	64 36 35 -29		15
16	F F	123	39 79 77 -27	50 10 61 -35		16
17	F F	110	35 41-21 -19	60 04-90 -22		17
18	M M	116	94 89 89 -88	90 86 85 -61		18
19	F M	121	87 81 79 -79	55 64 77 -64		19
20	M M	118	87 86 94 -1.	87 92 95 -1.		20
21	M M	108	65 39 57 -28	33 43 51 00		21
22	F M	117	48 86 80 -62	52 44 60 -37	72 73 88 61 74	22
23	F M	119	80 77 79 -87	78 63 79 -69		23
24	F F	107	91 59 40 -47	39 48 56 -41		24

S	Sex of S of P	IQ of S	Rs Between Concepts S-O P-S P-R A-H	Rs Judge Vs Subjects t-P J-S J-R J-H	Subject Reliability P-P S-S R-R H-H M	S
25	F F	103	88 92 97 -94	-13-22-32 33		25
26	F M	118	64 54 63 -50	74 48 63-53	55 86 93 88 81	26
27	F F	115	90 34 37 -51	20 57 61-47		27
28	F F	108	60 74 72 -66	-14 00 -96-19		28
29	M M	110	84 80 77 -66	76 81 92 -98	85 83 94 87 87	29
30	F F	107	93 88 90 -84	34 47 42 -71	44 90 90 93 79	30
31	M M	110	95 69 64 -68	-33-55-72 56	48 63 88 63 66	31
32	F M	110	90 77 89 -49	70 89 81 -88		32
33	F M	110	81 69 80 -84	69 71 82 -49	72 80 89 90 83	33
34	M F	109	86 55 44 -37	49 68 60 -47	49 78 89 88 76	34
35	M F	119	84 66 69 -55	84 70 77 -61	76 86 90 74 82	35
36	F F	97	03 34 48 -54	31 06-42 39		36
37	M M	100	95 92 92 -68	-63 -56-79 22		37
38	M M	120	96 96 96 -78	-34-39-38 43		38
39	M M	110	93 77 78 -74	75 88 91 -96		39
40	M M	118	98 90 92 -89	-22-22-30 32		40
41	M M	119	89 22 16 -30	49 42 42 -24		41
42	M M	100	56 88 94 -87	76 71 71 -62		42
43	M F	117	86 88 88 -86	37 46 38 -29		43
44	M M	120	97 96 96 -96	-21-25-29 12		44
45	M M	98	96 81 77 69	59 61 71 23		45
46	M F	118	63 42 56 -40	44 65 68 -59		46
47	M M	104	75 67 93 -95	01 11-22 -11	78 78 98 90 83	47
48	M M	98	00 78-03 -39	27 49 16 -30		48
49	F F	98	42 20 74 -62	-45-18-99 53		49

S	Sex	IQ	rs Between Concepts	rs Judge Vs Subjects	Subject Reliability	S
50	M F	121	67 76 86 -69	62 60 79 -61	85 79 87 97 87	50
51	F F	114	69 65 71 -78	65 89 77 -1.		51
52	M M	118	84 54 72 -80	66 61 58 -45	77 97 97 88 90	52
53	M M	116	88 52 58 -62	70 70 74 -73	75 92 96 86 87	53
54	F F	110	94 99 94 -48	62 59 60 -57		54
55	M M	99	49-54-37 36	43-54-85 87		55
56	M M	104	51 39 65 62	46 48 81 84		56
57	F F	117	48-50-54 55	-34 48 62 -64	47 75 97 96 79	57
58	F F	113	71 25 23 -13	00-38-60 58		58
59	F F	123	03-35 79 29	93-37 82 -56		59
60	M M	110	92 60 71 -30	25-14-06 -31		60
M			70 61 65 -52	34 36 29 -23	51 77 89 71 72	M

APPENDIX 4

INDIVIDUAL CORRELATIONS OF THE SUBJECTS IN THE TEST-RETEST GROUP

APPENDIX 4

INDIVIDUAL CORRELATIONS OF THE SUBJECTS IN THE TEST-RETEST GROUP

	TEST								RE-TEST								
	r's Between Concepts				r's J - S				r's Between Concepts				r's J - S				
S	S-IS	P-S	P-IS	P-LL	J-P	J-S	J-IS	J-LL	S'-IS'	P'-S'	P'-IS'	P'-LL'	J'-P'	J'-S'	J'-IS'	J'-LL'	S
4	78	80	92	-82	72	68	60	-46	84	54	72	-80	66	61	58	-45	4
5	02	-31	44	23	76	-29	79	24	03	-35	79	29	93	-37	82	-56	5
6	98	81	92	-75	-18	-18	-20	22	87	93	86	-86	-27	-22	-07	15	6
7	43	39	-04	-28	29	-42	-61	53	71	25	23	-13	00	-38	-60	58	7
8	92	94	93	-96	58	53	63	-61	94	99	94	-48	62	59	60	-57	8
9	61	57	46	-69	42	82	54	-65	69	65	71	-78	65	89	77	-1.	9
11	68	72	59	-64	-12	-36	-46	25	86	94	95	-87	-29	-08	-43	17	11
22	48	86	80	-62	52	44	60	-37	86	88	88	-86	37	46	38	-29	22
26	64	54	63	-50	74	48	63	-53	63	42	56	-40	44	65	68	-59	26
29	84	80	77	-66	76	81	92	-98	93	77	78	-74	75	88	91	-96	29
30	93	88	90	-84	34	47	42	-71	91	59	40	-47	39	48	56	-41	30
31	95	69	64	-68	-33	-55	-72	56	75	14	11	-16	16	-37	-71	44	31
33	81	69	80	-84	69	71	82	-49	90	77	86	-49	70	89	81	-88	33
34	86	55	44	-37	49	68	50	-74	89	-05	-13	-09	64	36	35	-29	34
35	84	66	69	-55	85	70	77	-61	80	77	79	-87	78	63	79	-69	35
47	68	-35	-20	-04	08	-47	-36	48	88	01	-20	28	67	-50	-33	51	47
50	67	76	86	-69	62	60	70	-61	87	81	79	-79	55	64	77	-64	50
52	84	54	72	-80	66	61	58	-45	78	80	92	-82	72	68	60	-46	52
53	88	52	58	-62	70	70	74	-73	98	96	97	-98	86	79	78	-62	53
57	48	-50	-54	55	-34	48	62	-64	39	-08	18	-19	40	24	70	-60	57
r_T	67	42	47	-45	42	24	25	-23	71	49	52	-43	36	30	29	-23	

Decimal points omitted on the figures.

APPENDIX 5

COMPARATIVE STUDY OF METHODS OF CORRELATION

APPENDIX 5

COMPARATIVE STUDY OF METHODS OF CORRELATION

A pilot study was done to determine which of several types of correlations best suited the data of this study. The basic measurement throughout was in the form of a profile composed of sixteen scores (each score being a rating on one profile with another by some means which would:

1. be sensitive to the level or height of the profiles, and maintain the individual differences between the corresponding points or scores of the profiles;
2. be the most feasible for doing the rather large number of comparisons needed.

The first point requires that the computation be based upon individual difference scores. Osgood presents a form of intra-class correlation which is computed from difference scores:[1]

$$r1 = \frac{1 - \sum d_{xy}^2}{\sum d_{x\bar{z}}^2 + \sum d_{y\bar{z}}^2}$$

where: d_{xy} = the difference between two scores x and y (i.e. the difference between two ratings on the same scale).

$\bar{z}$ = the mean of all the x and y scores combined.

---

1 Charles E. Osgood, George J. Suci and Percy H. Tannenbaum, The Measurement of Meaning, University of Illinois Press, 1957, p. 92.

In that this formula was suggested for use with Semantic Differential data, it was considered to be the formula best suited for this study. However, as is evidenced by the formula itself, the use of this correlation requires that a new combined mean be obtained for every pair of profiles to be compared. This made it most unfeasible from a practical standpoint.

Cattell[2] has put forth another form of intra-class correlation also based on difference scores:

$$r_p = \frac{2x^2 - \sum d_{xy}^2}{2x^2 + \sum d_{xy}^2}$$

where x^2 = the value of chi-square when P = .50 and n = the number of pairs of scores.
d_{xy} = the difference between two standard scores x and y.

This formula required that each profile of raw scores (ratings) be converted into standard scores. However, after having made this conversion, any two profiles may be correlated quite simply. This seemed to make this the more feasible of the two formulae, from a practical standpoint, but it was not known whether the resulting correlations would be valid when compared to the recommended method.

---

2 R.B. Cattell, "r_p and Other Coefficients of Pattern Similarity", Psychometrika, Vol. 14, No. 4, 1949, p. 297.

Figure 1 shows the results when 120 paired profiles were correlated by both methods and compared. Each point in the graph represents the two values of r obtained by the two methods on the same set of data.

Although there was a relationship, Cattell's correlations were consistently lower than those of Cuci-Tannenbaum, particularly at the extreme of the negative correlations, as can be seen in Figure 1.

The results of this comparison indicated that the use of Cattell's r_p would be statistically questionable, especially in light of the fact that it requires standardizing scores across columns. Osgood warned against such a procedure, considering that information is lost. However, if one is willing to risk the loss of some information, another method of obtaining the correlation between two profiles is available. Osgood[3] states that his D score is related to the product-moment correlation, when it is based on standard scores, as follows:

$$D_i^1 = \sqrt{2k(1-r_{11})} \quad \text{or} \quad r_{11} = 1 - \frac{D_{11}^2}{2k}$$

where: k = number of scales
d_{11} = the sum over k scales of the differences between two standardized scores, i and 1.

---

3 Osgood, et.al., Op. Cit., p. 91.

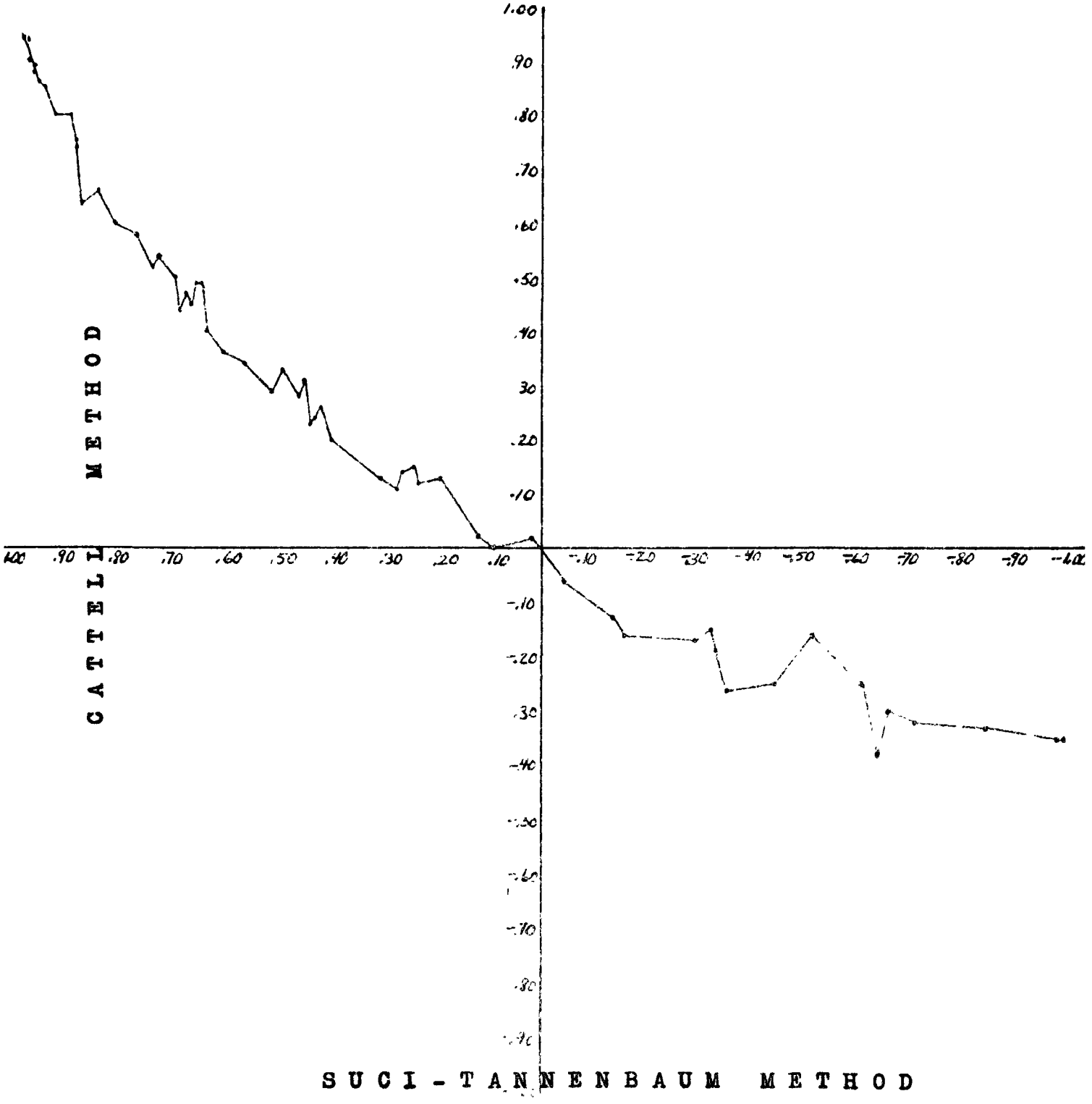


Figure 1.- A graphic representation of the relationship between the Suci-Tannenbaum and the Cattell Methods of intra-class correlation.

This formula will be referred to hereafter as Osgood's formula.

Since the product-moment correlation has the advantage of a known distribution, important in determining significance, and having in mind the advantages of working from standarized scores, a comparison was made between the results obtained with this r and those obtained by means of the Suci-Tannenbaum intra-class r first discussed. The results of these two formulae based on an N of 109 paired profiles can be seen in Figure 2. Although there were minor fluctuations in the relationship, the fluctuations were distributed fairly equally around the line that would indicate a perfect relationship.

From this it was concluded that the correlations obtained by using Osgood's formula were not essentially different from those obtained by the Suci-Tannenbaum formula. Therefore it was concluded that the use of standardized scores on the data in this study would not result in the loss of a significant amount of information.

The reason that a significant loss of information did not occur could be explained by the fact that the dimension of 'height' as it is usually applied in analyzing correlations did not truly exist in the type of profiles obtained in this study. This was because the Semantic Differential Scale is bipolar, has no absolute zero and the

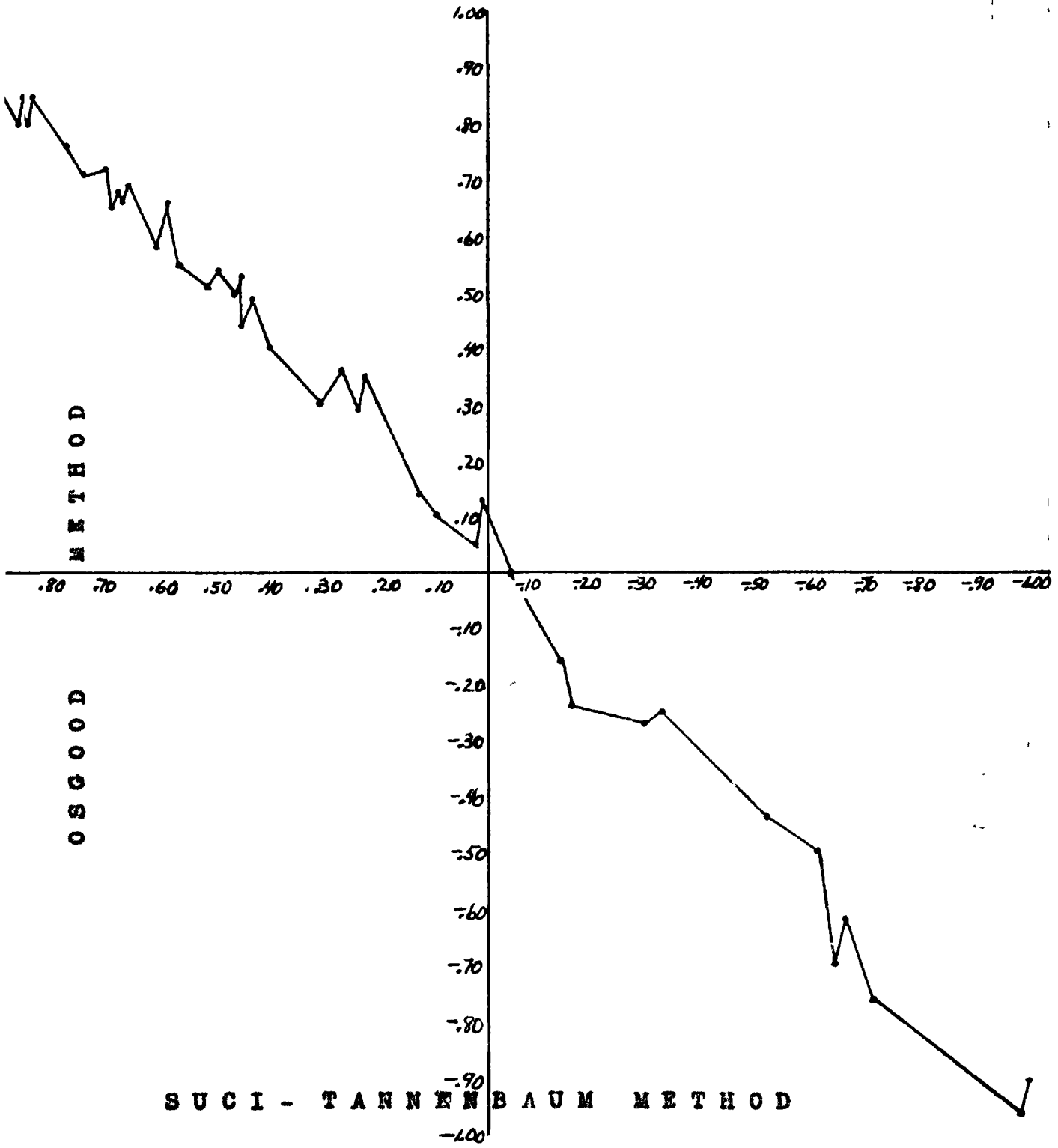


Figure 2.- A graphic representation of the relationship between the Suci-Tannenbaum and Osgood Methods of intra-class correlations.

extreme values of the scale are arbitrarily assigned. Since 'height' is the main dimension affected by standardized scores, this may account for the lack of any loss of information.

Consequently, in view of the fact that it was empirically demonstrated that Osgood's formula did compare favourably with the Suci-Tannenbaum formula which was initially considered to retain the most information, and in view of the fact that Osgood's formula was most feasible from a standpoint of economy and efficiency, it was considered the most suitable formula for this study.

APPENDIX 6

ABSTRACT OF

The Figure Drawing and the Phenomenon of Projection[1]

The present research has been directed toward casting some light on the manner in which the figure drawing represents the drawer. In the review of the literature, it was noted that research into figure drawings has often been misleading, partially because of failure to define terms, partially through failure to distinguish between the phenomenological and non-phenomenological, partially through failure to recognize that projection on the part of subjects when evaluating their drawings was often an uncontrolled variable and, finally, through very real difficulties in design encountered when trying to adequately distinguish the pre-conscious and unconscious psychic levels that drawing theorists and clinicians alike hold to be revealed in the figure drawing.

These problems suggested that more knowledge of the phenomenological aspects of the figure drawing and possible effects of subject projection might cast light on the

---

1 Fern Pickering, doctoral thesis presented to the School of Psychology and Education of the University of Ottawa, Ontario, 1963, ix-83 p.

non-phenomenological and offer a more complete grasp of the manner in which the figure drawing represents the drawer.

The method chosen to investigate this was to have sixty normals take the H-T-P, then rate their drawings, then their actual, ideal and least-like selves in counterbalanced order on a Semantic Differential Rating Scale. These results were then compared to the median rating of five judges' ratings of the figure drawing.

Two pilot studies, one to determine judge reliability and the second to determine subject test-retest reliability, indicated both were reliable.

Statistical analysis indicated a positive relationship between the subject's drawing and his actual and ideal self and a negative one between his drawing and his least-like self. Evidence of projection into the figure drawing was also found. It was of a nature that indicated a tendency to idealize the drawing on the part of the subject. On the basis of this, the null hypothesis was rejected.