THE DEVELOPMENT OF A PROJECTIVE DRAWING TECHNIQUE
TO ASSESS ID, EGO AND SUPEREGO INTERACTION

APPROVED:

Audrey Hamilton
Major Professor

Earl W. Koster
Minor Professor

Dwane Kingery
Dean of the School of Education

Robert Torlouse
Dean of the Graduate School
THE DEVELOPMENT OF A PROJECTIVE DRAWING TECHNIQUE
TO ASSESS ID, EGO, AND SUPEREGO INTERACTION

THESIS

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

For the Degree of

MASTER OF SCIENCE

By

Mark Wall, B. A.
Denton, Texas
January, 1970
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>LIST OF TABLES</th>
<th>iv</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chapter</td>
<td></td>
</tr>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>II. SURVEY OF LITERATURE</td>
<td>16</td>
</tr>
<tr>
<td>III. METHOD</td>
<td>28</td>
</tr>
<tr>
<td>IV. RESULTS</td>
<td>32</td>
</tr>
<tr>
<td>V. DISCUSSION</td>
<td>36</td>
</tr>
<tr>
<td>VI. SUMMARY</td>
<td>42</td>
</tr>
<tr>
<td>APPENDIX</td>
<td>44</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>58</td>
</tr>
</tbody>
</table>
### LIST OF TABLES

<table>
<thead>
<tr>
<th>Tables</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Analysis of Variance Between Judges</td>
<td>32</td>
</tr>
<tr>
<td>II. t Test for Means of Three Predictive Scoring Categories</td>
<td>34</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

During the first two decades of the twentieth century psychoanalysis focused upon the exploration of the unconscious mind via dream analysis and free association. Subsequently, with the publication of Freud's *Beyond the Pleasure Principle* (8), *The Ego and the Id* (9), and *Symptoms, Inhibitions, and Anxiety* (10), psychoanalysis left the study of isolated mental phenomena for the study of the dynamic interplay between intrapsychic processes. This was made possible because of the introduction of the concepts of id, ego, and superego. It was Freud's basic tenet that every neurotic symptom represents a compromise between the repressive forces of these intrapsychic components. This means of delineating intrapsychic processes has proved of such value that the concepts are used by disciplines other than psychoanalysis.

Diagnosticians, counselors, and psychotherapists with a variety of orientations use such concepts as "ego strength" with considerable frequency and acceptance. The goals of therapy are often described as a redirection, control, or expression of id impulses in the service of the cognitive processes, the ego. Some disciplines have changed the
labels but retained the concepts. When Rogerians, for instance, speak of "enhancement of self," it is appropriate to consider that the word "ego" was in the original German "self" so that the Rogerian concept is merely a rephrasing of the concept of ego strength. In the behavioristic approaches, desensitization or conditioning processes are often concerned with a diminution of superego strength. Despite the clinical utility of these concepts and despite Freud's statement that they characterized "a new field of research" (6, p. 254), very little research has been done, and it is often assumed that such concepts are inappropriate subject matter for scientific inquiry because they seem extant only hypothetically (13).

However, as Dombrose and Solbin (6) point out, a great many concepts, such as intelligence, cannot be validated but have proved useful after being operationally defined. Our concretely oriented scientific community may sometimes hold such descriptive hypothetical concepts in less esteem than concepts which can be directly measured in millimeters, grams, and spacial displacement, but not since Berkley (4) have philosophers found evidence to support a contention that one might be more real than the other. The lingering assumption that objects are real but concepts are spiritual is a residual effect of dualism. The most central statement of dualism, that body is separate from spirit, has virtually
vanished from psychology, yet it is peripherally revived in such terms as "hypothetical constructs" versus "intervening variables" despite the fact that such a dichotomy between matter and ideation is not scientifically demonstrable. As long as conceptual groupings of data under such metaphysical headings as phylum names or ego defenses are so operationally defined as to become communicable, they are indistinguishable from any other scientific data. Accordingly, the measurement of the interactive forces and relative strengths of id, ego, and superego are no less reputable than the measurement of thrust in ion propulsion.

One of the means of measuring personality variables has been the corollary to free association—projective techniques. Projective testing in general began with sudden popularity and persistence in 1921 with the advent of the Rorschach test, which consists of a series of ambiguous or unstructured ink blots (15). By 1931, Frank (7) pointed out that a class of such tests had developed which he described as "projective techniques." According to Swensen (18), the Rorschach remains the most frequently used projective instrument, with the **Draw-a-Person** (D-A-P) second in popularity among projective techniques. Such projective methods of assessment emanate from the psychoanalytic concept of projection pointed out by Sargent (16). Projection is held to be a mechanism by which the conscious ego defends itself from inadmissible ideas and impulses by ascribing
them to persons and forces outside the organism. As Bell (3) pointed out, the term as used in testing is a general process, not necessarily pathological, by which people unknowingly endow unstructured or incomplete stimuli with meaningful gestalts emanating from their personalities. It is the process recognized since Aristophanes (2) by which men perceive idiosyncratic objects in cloud formations.

Although a projective technique would seem appropriate for assessing the relative strengths and interactions of id, ego, and superego, there are none now available that do so.

Problem

The problem of the present study was threefold: 1) to devise a projective drawing analysis technique which would assess the operationally defined psychoanalytic concepts of id, ego, and superego, 2) to devise a preliminary scoring technique, and 3) to investigate the relationship between the drawings and the original clinical scales of the Minnesota Multiphasic Personality Inventory.

The problem arose because these three Freudian intra-psychic components constitute a frequent means of grouping and communicating personality data in psychiatric hospitals and other facilities despite the fact that assessments pertaining to them are derived from behavioral observations, social history, or test batteries in which no single test
is designed to elicit them. The assumptions clinicians make about their patients in these terms are logical derivatives from data gathered from diverse media and are only as valid as the logical facility of the clinicians making them. The utility of the concepts, coupled with the lack of personality theory underlying the construction of current projective techniques, makes this problem one of theoretical as well as practical significance.

Assumptions

It was assumed that the stimulus "draw something evil" would elicit graphic illustrations of id impulses disavowed by the superego. It was assumed that the stimulus "draw a picture of yourself" would yield a physiological body image of the ego. The stimulus word "self" seemed appropriate because Freud's word for "ego" in the original German was not "ego" but, unpretentiously, "I." Additionally, it seemed appropriate to specify "yourself" rather than Goodenough's term "a person" because it remains in doubt that "a person" will inevitably yield a projection of the self. Lastly, it was assumed that the stimulus "draw a picture of God" would elicit a personification or abstraction of the superego, i.e., the internalized parent or deity which Freud called the "uber Ich," the "over I." Summarily, it was assumed that graphic responses to these stimuli would yield valid and reliable measurements of the dynamic forces
which in interaction generate intrapsychic harmony or malfunction.

Because the key words in the stimulus directions were EVIL-SELF-GOD, the technique was identified by the initials E-S-G.

The problem then necessitated a method by means of which the E-S-G data could best be assessed. In assessing data from any projective technique for purposes of prediction of behavior, there are two avenues of approach—content analysis and sign approach. Utilizing content analysis with the E-S-G, raters might assume that a "self" figure drawn holding a knife is indicative of overt hostility. A "self" drawing dismissed by the subject with the written words "Self—Nothing" might imply low strength. A "God" drawing in which God appears as a crutch might be indicative of sociopathy. While such productions are capable of good inter-rater agreement, they occur with such diversity that extremely large samples are necessary in order to have statistically useful quantities in each category. When interpreters are asked to fit projective productions into a few very general categories in order to obtain sizable categories, the broadness of each category, such as neurotic, psychotic, psychopathic, leaves room for idiosyncratic opinionation by the raters. The ratio between American diagnoses of schizophrenia and psychotic depression is inversely proportional to that found in England (14). This
is assumed to be a function of the diagnosticians rather than the patients.

The second avenue of approach, the sign approach, is more readily assessable with small samples and is less sensitive to rater sophistication and projection. The sign approach additionally has the advantage of being scorable by clerical personnel. Therefore, a sign approach was used, based upon two variables, hardness and size. Hardness, i.e., degree of pencil pressure, was assumed to be reflective of affect. Size was assumed to be related to the subject's perceptions of importance and power. A logical examination of the possible combinations of these variables yielded certain predictions about the subject's behavior on the Minnesota Multiphasic Personality Inventory (MMPI).

It was assumed that inexperienced laymen would produce reliable judgments regarding relative hardness of line in the three drawings on a three-point scale in which "hard" lines were to be rated 3, "secure" lines were to be rated 2, and "light" lines were to be rated 1. It was assumed that these raters would produce reliable judgments regarding relative size of the drawings on a three-point scale in which a rating of 3 represented large drawings and a rating of 1 represented small drawings. It was also assumed that inexperienced laymen would produce reliable judgments on three-point scales for the variables of hardness and size on the E-S-G drawings of a clinical population.
The degree of inter-rater agreement between laymen would indicate the least reliability of the sign approach in actual clinical practice.

Derivation of Hypotheses

Hardness of pencil pressure has long been associated with expressions of hostile impulses in drawing analysis techniques. It was reasoned that the degree of pencil pressure on each of the three drawings would be indicative of the degree to which each intrapsychic component incorporated hostile impulses. An "evil" drawing produced in hard line pressure, exemplary of aggressive id drives, would be expected in normal subjects. Because the normally developed ego does not incorporate this aggression but exerts a rationally directing influence, the "self" (ego) drawing was expected to be less hard, i.e., less aggressive. Because the superego is the "censor" of aggressive impulses, the "God" (superego) drawing would be expected to yield a very light pressure. More succinctly, the aggressive id and the non-aggressive superego are mediated by the moderately aggressive ego. The E-S-G drawings reflective of normal id, ego, superego interaction could be rated for relative hardness as 3-2-1, respectively.

The second dimension, size, was assumed to relate to perceptions of power. In an attempt to deny the power of the forbidden impulses of the id, it was assumed that persons
exhibiting normal id, ego, superego interaction would produce constricted, small, "evil" (id) drawings. In an attempt to express the relative importance and power of the superego, normals would be expected to produce large "God" (superego) drawings. Our cultural milieu espouses the belief that God, the personification of virtue and power, is larger than man and that man is in a constant battle to suppress, control, and constrict his baser impulses. Accordingly, persons with normal id, ego, superego interaction would be expected to project E-S-G drawings with small ids, large egos, and larger superegos, which would be rated 1-2-3.

Psychopathic personalities of the criminal variety would be expected to incorporate hostile impulses into the ego structure so that the ego is in the service of the id. The id and ego could be expected in these instances to be rated 3 in hardness because they are both aggressive. However, problem children, professional criminals, sexual deviates, and others described in these terms are often manifestly religious or tenaciously espouse a value system divergent only in kind from that of normals. Some of these could be expected to project normal superego productions.

Persons who perceive superego functions as small in importance might be expected to produce small "God" (superego) drawings. Those who produced "God" drawings with hard line pressure might be assumed to be expressing hostility toward superego restraints or toward authority in
general. While such attitudes are not necessarily indicative of true psychopathy, they are attitudes incorporated into the Psychopathic Deviate Scale of the MMPI.

While normals, neurotics, and others may be expected to respond to the emotionally charged stimulus words, GOD-SELF-EVIL, with varying emotional responses, emotionally apathetic persons would not be expected to do so. Those schizophrenic personalities marked by apathy and flattened affect could be expected to yield flattened indices of affect for the three drawings. This condition is marked by an incapacity to recognize the relative importance of things. A charged emotional stimulus is likely to elicit the same degree of affective response as a mundane stimulus. Accordingly, such subjects could be expected to yield drawings of invariable size and invariable hardness. However, they would not be expected to yield aggressive drawings or large, important ones.

Hypotheses

The following hypotheses were investigated:

(1) Inexperienced laymen utilizing three-point scales for the dimensions of hardness and size in rating E-S-G productions would produce ratings, which, when summed across both dimensions, would yield no significant interrater difference.

(2) Subjects whose E-S-G drawings were rated 3-2-1 in hardness and 1-2-3 in size would yield MMPI profiles in
which the highpoint clinical scale score mean would be less than a T-score of 70.

(3) Subjects whose E-S-G drawings included E-S drawings both rated 3 for hardness, G drawings rated 3 for hardness, or G drawings rated 1 in size would yield MMPI profiles in which the mean Psychopathic Deviate Scale scores would be greater than a T-score of 70.

(4) Subjects whose E-S-G drawings were rated 1-1-1 for hardness and 1-1-1 for size, or 2-2-2 for hardness and 2-2-2 for size (or deviant from these patterns by only one digit out of the total six, that digit not exceeding the mode by more than a single rating point) would yield MMPI profiles in which the mean Schizophrenic Scale scores would be greater than a T-score of 70.

Description of Measuring Instrument

The E-S-G technique is a procedural method and therefore is appropriately described in Chapter III. The MMPI, which was used as a validating instrument, is a true-false inventory designed for application in routine psychiatric settings. The original eight clinical scales of the MMPI were used in this study. These were constructed in accordance with the frequency of responses given by known diagnostic groups. The scales bear the labels of the corresponding diagnostic groups such as Psychopathic Deviate, Schizophrenic, etc.
These scales are accumulations of correlated response probabilities. Therefore, differences in scores on any scale can be interpreted in terms of differences in probability that the subjects who achieved them are members of the appertaining pathological group, and reduce the probability that they are members of the normal group. A T-score of 70, two standard deviations above the mean, is generally considered the upper limits of the normal range. Accordingly, those who score below a T-score of 70 on their highest clinical scale score are more likely to be members of the normal population than if they scored above a T-score of 70.

A score above a T-score of 70 on the Psychopathic Deviate Scale increases the probability of group membership with those patients diagnosed "psychopathic personality" upon whom the scale was originally validated (12). Similarly, a score above a T-score of 70 on the Schizophrenic Scale increases the probability of group membership with those patients diagnosed a "schizophrenic" upon whom this scale was cross-validated (12).

While the validity of the MMPI for diagnosing is considered limited, Chronbach asserts that the scales "have been given meaning by all manner of research" (5, p. 492). As an example, he points out that any experienced psychologist might correctly predict that certain occupational groups, for instance, actors, will score significantly higher than
other occupational groups on the Psychopathic Deviate Scale. The confirmation of such predictions is evidence of construct validity. Chronbach asserts that the "body of evidence indicates that the Psychopathic Deviate Scale measures some underlying personality structure" (5, p. 493). He characterized this structure as "vague" and broader than the diagnostic group upon whom the test was validated but nevertheless extant.

More certainly, the scales measure the behavior elicited by certain groups of statements. Whether the statements are actually groups, as Chronbach believes, or merely aggregates, is only tangentially relevant to this thesis. The MMPI behavior in this instance is analogous to any other behavior, which if predicted correctly, is evidence of construct validity. Precedent for this method of construct validation was found in the work of Dombrose and Solbin (6).


CHAPTER II

SURVEY OF LITERATURE

An appropriate survey of related research and test development must encompass two areas, namely, drawing analysis as an effectively appurtenant projective technique, and previous attempts to measure specific id, ego, superego interactions by other techniques. Because the two have not been combined prior to the introduction of the E-S-G technique, they will be discussed separately.

Some crude research and experimentation as well as actual clinical use of drawing analysis as a means of assessing personality began in the last century but occurred among isolated workers rather than as an accepted movement (25). The American, Samuel Calverton, elucidated the concept and procedures of such analysis in his Children's Expression Through Drawing (5), published in 1894. The actual movement emanated virtually by accident from F. Goodenough's 1926 Draw-a-Man technique (14), which was intended to yield a culture free estimate of intelligence. Goodenough, however, recognized that the test also yielded personality data in the form of aspirations and anxieties, and that the body image was "A natural figure for such projections" (18, p. 344). A few years later, Machover developed the "Draw-a-Person"
test as a projective instrument, after which she published her 1948 work, Personality Projection in the Drawing of the Human Figure (23), for which she is accepted as the leader of the evolving drawing analysis movement. Quite simultaneously however, Buck introduced the "House-Tree-Person" technique in the form of a scoring manual (4). Utilization and acceptance of the drawing analysis techniques immediately followed, as well as more rigorous modern research.

Although the American acceptance of projective techniques in general has been considered by some European authorities as so preposterous as to be labeled "scandalous" (9), projective drawing techniques have long been established in this country when used in concert with other projective devices. In fact, drawing analysis tests are considered by American projective technique experts to be second only to the Rorschach for diagnosis (16) despite the fact that these researchers are quite aware of the numerous studies which negate the predictive value of drawing techniques for this purpose. This contradiction is justified by the widely accepted American view that a series of tests with low validity individually can and do yield valid judgments when examined in interrelation.

In 1950 and 1952, Fisher (10), later Fisher and Fisher (11), found the D-A-P technique unable to distinguish between diagnostic groups. First the sign approach was examined, using eleven of Machover's indices. Later,
Fisher and Fisher found the same negative results by using the sign approach in conjunction with the global judgments. This was to prove typical of subsequent findings, and was corroborated quite shortly by Blum (3), who found no significant relationship between psychiatric ratings and the D-A-P technique. Recently, Lewinsohn (21) found OVQ ratings of figure drawings not significantly predictive of psychologists' judgments based upon clinical observation and battery test data. Although these have been typical findings, they have been occasionally contradicted by such workers as Holzberg and Wexler (18), who found that the D-A-P could distinguish between diagnostic groups.

To reconcile this disparity in findings, it is seldom considered in this context that psychiatric and other expert judgments are themselves suspect. Studies such as those by Ash (1), Hunt (19), and others indicate that such judgments are sufficiently unreliable even with only three or four pathological categories as to make the use of such judgments questionable as validating criteria against which to measure individual instruments. Taking this into consideration, some drawing analysis studies such as the one by Watson (27) have examined the relative validity, and often invalidity, of scorers with sharply varying degrees of experience. These typically find the scorer's experience irrelevant to the scorer's diagnostic acuity with drawing analysis techniques. It is, however,
quite possible that definitions of a scorer's "level of sophistication," as Watson called it, when based upon academic achievement and/or years of experience, do not designate an actual level of sophistication. Studies such as Holzberg and Wexler's may show better diagnostic prediction because these scorers were actually more expert.

This reconciliation of the mutually exclusive findings as to whether drawing analysis tests are predictive of diagnostic categories only assumes what Ash has already demonstrated, that professional accreditation and experience do not necessarily generate diagnostic agreement.

The weight of evidence shows that current workers usually can not adequately diagnose with drawing analysis techniques alone. The next most specific assumption was that these techniques could distinguish between normal and pathological groups, i.e., greatly broader diagnostic categories.

With this in mind, Whitmyre's 1953 study (28) attempted to determine whether scorers using a global approach could differentiate between psychiatric patients and normals, using groups of twenty-five each. While the results were in the appropriate direction, the correlation was not significant. Hiller and Nesvig (17), similarly contrasting a normal group with a hospitalized group (in this instance adolescents), found that using a broad sign
approach based upon empirically derived criteria, the scorers proved 79 per cent accurate.

The highest validity has been found when the techniques are used for specific trait analysis. Royal (24) found that judgments from eight signs taken as a whole discriminated between anxiety neurotics and normals. Baker, Mathis and Powers (2) found that two of several signs examined were predictive in determining homosexuality. Witkin (29) found a high relationship between personality characteristics and attitudes toward the body in male subjects as opposed to female subjects, thereby concluding that women are more concerned with social relationships and emotional externalization, whereas men are more concerned with body image. Griffith and Peyman (15) found eye and ear emphasis in figure drawing significantly correlated ($X^2$, 7.754 for 1 df; $P < .01$) with ideas of reference. Hoyt and Baron (19) found two sign indicators significantly correlated with manifest anxiety on the Taylor Manifest Anxiety Scale ($P < .05$).

Summarily, the research done in this area indicates that these techniques are not suited for diagnosis, mildly differentiating between psychotic and normal groups but best suited for trait assessment. Research continues in trait assessment which may eventually provide valid diagnoses, but this eventuality does not seem imminent.
A new direction in these techniques began with Lawton and Sechrest (20), who drew attention to the clinical use of the Draw-a-Family test as a means of assessing familial interaction. The Draw-and-Tell-a-Story test, by Levy (22), in an attempt to measure social interaction, asks the subject to draw three figures, then to give them names, and to tell a story about them which will presumably yield data concerning the subject's relationships with significant figures. The trend to assess group interaction via such drawings has culminated in Cookerly's 1965 (7) Draw-a-Group test for measuring interpersonal responsiveness. The rationale for this development is simply that the human figure is more suited for expressing extrapsychic social interaction and perceived relationships than for presenting a cohesive intrapsychic portrait of the personality.

Based upon this related research, it seemed reasonable to assume that drawing analysis techniques with other stimuli than the human figure, or additional to the human figure, might prove more suitable for diagnosis. The present study evolved from this assumption.

The dynamic interpsychic components which the E-S-G aims to assess are the concepts of id, ego, and superego which emanate from Freud's presentation of a structural point of view in 1923 (12). Since that time they have retained a prominent and ever more specifically defined place in psychoanalytic literature. Freud characterized
the id as a "seething cauldron" of alien impulses. He thought of the ego or "self" as the largely conscious cognitive processes, and he characterized the superego as diverse identifications yielding a value system ideal quite similar to the theological concept of conscience. The latter is experienced predominantly in terms of mood, ranging from free self esteem to depression and guilt. This intrapsychic "structuralism" did not actually imply such compartmentalization as many of the minutia-oriented discussions concerning their precise dispositions imply. Freud's "structures" were conceived as logical constructs under which intrapsychic forces could be grouped for convenience in dealing with the data, and, in accordance with Freud's emphasis upon their dynamic aspects, are perhaps best considered as vectorial energies rather than as discrete physiological phenomena.

Attempted measurement of these forces is not without precedent. Dombrose and Solbin (8) in 1958 devised the IES Test (Id-Ego-Superego) as a clinical aid and research tool for measuring the relative strengths of these three intrapsychic components. The first of four subtests consists of titles projected by the subjects to twelve drawings which depict impulse activities. The second elicits conclusions to incomplete cartoons, while the third elicits projections by the subject concerning emotions in nine photographic portraits. Lastly, the subject performs a
perceptual-motor task requiring a solution which may be idiosyncratic and revealing of personality dynamics. The scoring technique yields scores of varying weights, which are held to be proportional to the relative strengths of the three components, id, ego, and superego.

Several evaluative research studies have been conducted using subjects of varying age, sex and diagnoses. The originators of the test made predictions regarding the variations in scores to be expected among normals, constricted neurotics and paranoid schizophrenics, finding 23 of 36 rank order predictions to be correct (8). Charnes (6) found that adults and ten-year-olds react similarly, while adolescents react differently from other groups, from which he concluded:

The test behavior of the different groups indicates that the test taps a basic personality balance which is formed by the age of ten, which is changed by the pressures of adolescence, and which is restored in adulthood to its early equilibrium only somewhat altered by intervening growth, education and socioeconomic status (2, p. 107).

Golden (13) administered the test to eleven-year-old normal children and compared the test results with teacher ratings regarding the categories of impulsivity, constriction, and good adjustment, finding 39 correct rank predictions out of 45.

Verrill and Costanza (26) examined the test's ability to predict ward behavior in thirteen patients as measured
by the Multidimensional Scale for Rating Psychiatric Patients (MSRPP) by the psychiatric nurse. The I score correlated significantly with impulsive behavior.

This preliminary research with the IES indicates that operationally defined measurements of psychoanalytic intrapsychic components can yield valid and reliable measurements of some aspects of behavior relating to diagnostic symptoms. Yet in the decade since the test was introduced, it has not become widely accepted. This may be somewhat due to its complexity in administration and scoring. A simpler, less time-consuming instrument yielding similar results may prove of more practical value in the clinical setting.
CHAPTER BIBLIOGRAPHY


20. Lawton, M. J. and L. Sechrest, "Figure Drawings by Young Boys from Father-Absent Homes," Journal of Clinical Psychology, XVIII (July, 1962), 304-305.

21. Lewinsohn, P. M., "Psychological Correlates of Overall Quality of Figure Drawings," Journal of Consulting Psychology, XXIX (December, 1965), 504-512.

23. Machover, K., Personality Projection in the Drawing of the Human Figure, Springfield, Illinois, Charles C. Thomas, 1948.


CHAPTER III

METHOD

Subjects

The investigation consisted of forty subjects with an age range from ten to sixty-seven. There were sixteen males and twenty-four females. There were four children, twelve teenagers, and twenty-four adults. Thirty-four of these subjects were psychiatric patients, predominantly first admissions, with a wide range of socioeconomic backgrounds.

Procedures for Collecting Data

Each subject was presented an ordinary sheet of blank white paper, approximately eight and one-half by eleven inches, and a sharpened pencil, usually of number two hardness, with an eraser. Each subject was then asked to "draw something evil." Provided with a new sheet of paper, each subject was then asked to draw himself. Finally, on a third sheet of paper, each was asked to "draw a picture of God." Inquiries for more specific instructions were met with the reply, "Draw whatever it represents to you," or "Your idea of yourself." Questions relating to artistic quality were answered by the statement,
"How good an artist you are doesn't matter at all."

All statements and questions were handled in such a way as to be encouraging without guiding in the drawing content.

The E-S-G was administered in the foregoing manner on an individual basis during the regular test battery. The battery included the MMPI.

**Procedures for Analyzing Data**

In order to assess inter-rater agreement between inexperienced laymen, automobile mechanics were chosen as raters. The choice of such an occupational group was expected to emphasize the lack of necessity for psychological sophistication in rating a clearly defined sign approach on a small rating scale. Four mechanics were asked individually to rate each of the drawings on a three-point scale for hardness of line, and subsequently for relative size of the drawings. Each rater was shown an example E-S-G with clearly defined gradations in hardness and size. Each was asked to notice that the "evil" drawing was second hardest in line of the three, and that the "God" figure was lightest of the three. They were advised that such an E-S-G should be rated 3-2-1 for hardness. Drawings to be rated 3 were described as "hard." Those to be rated 2 were described as "secure," and those to be rated 1 were described as "light."
Overall outline rather than specific shading of certain areas was to be the basis of the raters' judgments. The raters were advised that emphasized details such as items of clothing in an otherwise lighter drawing were to be ignored. Lastly, they were asked what rating they would give an E-S-G in which all the drawings were quite light. When each responded with the rating 1-1-1, he was allowed to proceed. When each rater concluded his ratings of the productions of the forty subjects, he was asked to review his judgments and to make any corrections he felt necessary. The review of the judgments was felt necessary in order to equalize any improvement in judgmental acuity which might have occurred as the raters gained experience.

An equivalent procedure was subsequently used regarding the relative size of the drawings in each E-S-G. Large drawings were to be rated 3, average drawings 2, and small drawings were to be rated 1. An example E-S-G rated 1-2-3 was presented in which the drawings increased in relative size. When multiple human figures were presented in one drawing, the raters were instructed to rate the drawing on the basis of the largest figure present. Each rater was asked what rating should be given a drawing in which all the figures were small. When each responded 1-1-1, he was allowed to proceed in rating the productions of the forty subjects.
An analysis of variance was chosen as a means of assessing the variance between judges' ratings. The analysis of variance technique was applied to the summed ratings over both dimensions for each of the forty subjects. The predictive hypotheses were investigated by use of the $t$ statistic as a test of significance.
CHAPTER IV

RESULTS

Presented in this chapter are the results obtained and the statistical analyses of those results. To assess the effects of inter-rater disagreement upon the validity of the three predictive hypotheses, an analysis of variance was conducted on the ratings obtained by the four judges. The ratings of each judge were summed over both dimensions for each subject's E-S-Q productions. The analysis of variance was conducted on these sums. This procedure constituted a test of the hypothesis that no significant difference existed among the mean ratings of the judges. Table I presents the results of the

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Raters</td>
<td>1.7</td>
<td>3</td>
<td>.57</td>
<td>.24</td>
</tr>
<tr>
<td>Between Subjects</td>
<td>627.9</td>
<td>39</td>
<td>16.07</td>
<td>6.87</td>
</tr>
<tr>
<td>Remainder</td>
<td>273.8</td>
<td>117</td>
<td>2.34</td>
<td></td>
</tr>
</tbody>
</table>
analysis. The difference between raters' judgments was not significant at the .05 level, as indicated by the F value between raters in Table I. As opposed to a required F value of 2.68, the obtained F value of .24 indicates the difference between raters' judgments did not approach significance.

The difference between ratings received by subjects was significant at the .001 level, as indicated by the F value between subjects in Table I. The F value of 6.87 indicates that the difference between subjects' ratings was significant at beyond the .01 level.

The data were tabulated to investigate the hypothesis that E-S-G drawings rated 3-2-1 respectively for hardness of line and 1-2-3 for relative size would have been produced by subjects whose mean high point clinical scale scores on the Minnesota Multiphasic Personality Inventory (MMPI) would be equal to or less than a T-score of 70. Four subjects were rated as meeting the above criteria. The highest T-score of each of these four subjects was tabulated without regard to which clinical scale bore this highest score. The mean of the high point clinical scale scores of these four subjects was 64, with the standard error of the mean being 2.34. Subtracting the hypothetical parameter mean of 70 from the sample mean and dividing the result by the standard error of the mean yielded a t-score of 2.32. Since this t-score did not
exceed the critical value of 3.18, the results were not significant at the .05 level. The data are summarized in Table II, the category being described as "normal."

**TABLE II**

**t-TEST FOR MEANS OF THREE PREDICTIVE SCORING CATEGORIES**

<table>
<thead>
<tr>
<th>Categories</th>
<th>X Scale Score</th>
<th>t</th>
<th>df</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>64</td>
<td>2.34</td>
<td>4</td>
<td>P &gt; .05</td>
</tr>
<tr>
<td>Psychopathic Deviate</td>
<td>82</td>
<td>3.33</td>
<td>3</td>
<td>P &lt; .01</td>
</tr>
<tr>
<td>Schizophrenic</td>
<td>91</td>
<td>3.41</td>
<td>19</td>
<td>P &lt; .01</td>
</tr>
</tbody>
</table>

The data were tabulated to investigate the hypothesis that E-S-G drawings in which the "God" figure was rated 3 in hardness or 1 in size or in which the Evil-Self drawings were rated 3-3, respectively, would have been produced by subjects whose MMPI Psychopathic Deviate Scale score mean would exceed a T-score of 70. Twenty subjects were rated as meeting the above criteria. With the sample mean of 84.32 and a standard error of the mean being 13.45, the corresponding t score was found to be 3. Since this t score exceeds the critical value of 2.86, the results were significant at the .01 level. The data were summarized in Table II under the "psychopathic deviate" category.
The data were tabulated to investigate the hypothesis that E-S-G drawings rated 1-1-1 respectively for hardness and 1-1-1 for size, or rated 2-2-2 for hardness and 2-2-2 for size (or deviant from these patterns by no more than a single numerical value out of the total six) would be yielded by subjects whose mean MMPI Schizophrenic Scale scores exceeded a T-score of 70. Five subjects were rated as meeting this criteria. With the sample mean of 91.6 and a standard error of the mean being 3.41, the corresponding $t$ value was found to be 6.3. Since this $t$ score exceeds the critical value of 4.6, the results were significant at the .01 level. The data are summarized in Table II, under the "schizophrenic" category.
CHAPTER V

DISCUSSION

The analysis of variance detailed in Chapter IV indicated that the difference between raters' judgments did not approach significance at the .05 level. Had the variance between the ratings given to subjects been insignificant, the interpretation of the above finding would have been in doubt. However, the difference between the ratings given to subjects was significant at the .01 level. This aspect of the present study suggested that the E-S-G drawings analysis technique can be reliably scored by inexperienced laymen when the scoring is based upon three-point scales for drawing size and intensity of pencil pressure. Such scoring, although preliminary to the investigation of other variables, can be done by clerical personnel while the psychologist can concentrate on evaluative techniques involving professional training.

The second phase of the experiment tested three hypotheses relating to validity. A certain score configuration was hypothesized to be predictive of MMPI protocols with a mean high point clinical scale less than a T-score of 70. Should the high point clinical scale mean be below a T-score of 70, it could be assumed that all clinical
scale scores would be below this point. A second configuration was hypothesized to be predictive of an MMPI Psychopathic Deviate Scale score mean exceeding 70. The third configuration was hypothesized to be predictive of an MMPI Schizophrenic Scale mean exceeding a T-score of 70.

The first of these three hypotheses was not confirmed. All subjects whose E-S-G drawings were rated 3-2-1 respectively in hardness and 1-2-3 in size had no score exceeding an MMPI T-score of 70. However, the category included only four cases. They could not be considered representative of a normal population because the sample was taken from a largely psychiatric population. The results were not significant at the .05 level. These findings indicate that an investigation among a normal population is appropriate regarding this hypothesis.

The hypothesis that certain E-S-G score configurations would be predictive of a MMPI Psychopathic Deviate Scale score mean greater than a T-score of 70 was confirmed at the P<.01 level. This hypothesis assumed that "God" figures depicted as small would be so depicted by persons perceiving authority as diminished in importance. The hypothesis assumed that those "God" figures depicted in hard lines would be reflective of an aggressive impulses toward superego restraints. It further assumed that the combination of hard "evil" and "self" drawings would be
indicative of aggressive id impulses incorporated into the ego structure with a concomitant rejection of social restraints. These productions were assumed to be predictive of high scores on the MMPI Psychopathic Deviate Scale. The confirmation of the hypothesis that the means of such scores would be above a T-score of 70 is consistent with the contention that the E-S-G assesses those intra-psychic components which it was designed to assess.

The hypothesis that a certain E-S-G score configuration would be indicative of an MMPI Schizophrenic Scale score mean greater than a T-score of 70 was confirmed at the P < .01 level. The E-S-G configuration of evenness in ratings for relative line hardness and size was assumed to be indicative of only those schizophrenically oriented subjects who exhibit flattened affect. Whereas neurotics, normals, and others would be assumed to respond with variations in affect exemplified by hardness of line and drawing size to the different stimulus directions, subjects with more flattened affect would not be expected to respond with variations of these variables in their graphic productions. The small quantity of subjects so judged, five out of forty, remains consistent with the assumption that flattened-affect schizophrenically oriented persons entering testing centers in small quantities. While generally elevated MMPI profiles are to be expected, such general
elevations are quite different from meaningful Schizophrenic Scale elevations as the result of ideation concomitant with flattened affect.

The preliminary scoring technique used in this study was intended to sample the possibilities of that method if applied on a larger basis. The variables of hardness and size were chosen largely because they seemed sufficiently objective as to be capable of yielding good interrater agreement. The confirmation of this assumption suggests that similar variables might prove equally reliable. For example, the speed with which the drawings are produced is a similarly objective variable and may well relate to states of depression and excitement. Variables such as these contrast sharply with such highly subjective judgments as degree of psychosexual difficulty. This study shows additionally that highly specific and reasonably objective variables which exclude the possibility of poor professional judgment in scoring can be predictive in a limited way.

A small increment toward a clinically useful sign approach has been made. With other such increments, the E-S-G may become a useful clinical tool.

A final aspect of the present study relates to construct validity. Theoretical concepts emanating from a personality theory were used to make predictions concerning the expected variations in E-S-G configurations. Certain
scoring configurations were hypothesized to be predictive of specified MMPI behavior. This procedure related to that employed by Dombrose and Solbin, who used "theoretical concepts in order to test predictions" (1, p. 361) in establishing the construct validity of the IES test. They assert the methodological soundness of this approach and say, "... if it can be demonstrated that the number of correct predictions is greater than that to be expected by chance, then the tests will be shown to have validity, and the theoretical concepts underlying them to have validity" (1, p. 361).
CHAPTER VI

SUMMARY

An experimental drawing analysis technique called the E-S-G was devised to assess the psychoanalytic concepts of id, ego, and superego in three separate drawings. Hypotheses were based upon these theoretical formulations to investigate certain E-S-G score configurations as predictors of behavior on the Minnesota Multiphasic Personality Inventory (MMPI). The technique was administered to forty subjects from a largely clinical population.

It was hypothesized that inexperienced laymen could rate the drawings reliably along a three-point scale for hardness of pencil pressure and for size of the drawings. An analysis of variance was performed by summing across both dimensions to ascertain if the difference between the judges' ratings was significant. The difference was not significant at beyond the .05 level, implying homogeneity in the judges' ratings.

It was hypothesized that E-S-G drawings rated 3-2-1 respectively for hardness of line, and rated 1-2-3 for relative size would have been produced by subjects whose MMPI high point clinical scale score mean would be less than a T-score of 70. That score lies two standard
deviations above the mean. Scores above that point are not necessarily indications of pathology, and scores below that point are not necessarily indications of normal behavior. However, scores below a T-score of 70 are usually considered to be within the normal range. The t statistic was used to ascertain the probability of the parameter mean being equal to or less than a T-score of 70. The results were not significant at the .05 level. It is suggested that this hypothesis might best be tested using a normal rather than psychiatric population.

It was hypothesized that certain E-S-G score configurations would be indicative of an MMPI Psychopathic Deviate Scale score mean greater than a T-score of 70. The hypothesis was confirmed at the .01 level of significance.

It was hypothesized that certain E-S-Q score configurations would be predictive of an MMPI Schizophrenic Scale score greater than a T-score of 70. The hypothesis was confirmed at the .01 level of significance.

The experimental E-S-G technique has been shown to be reliable when scored by laymen along a three-point scale for hardness of pencil pressure and along a three-point scale for relative size of the three drawings. The successful prediction of two of the hypotheses is supportive of construct validity.
APPENDIX

EXAMPLES OF THE E-S-G DRAWINGS PRECEDED
BY EXPLANATORY NOTES

Each of the drawings shown received evaluations reflecting substantial agreement among the four raters. They are exemplary of the various predictive categories.

Drawings le, ls and lg were, respectively, the E-S-G productions of a single subject whose MMPI exhibited no score above a T-score of 70. The drawings were rated as increasing in size while decreasing in hardness.

Drawings 2e, 2s and 2g were, respectively, the E-S-G productions of a single subject, rated 3 for hardness in the "evil" and "self" drawings, scoring above a T-score of 70 on the MMPI Psychopathic Deviate Scale. The subject was a ten-year-old hospitalized psychiatric patient with a history of overly aggressive behavior.

Drawings 3e, 3s and 3g were productions of a single subject whose "God" figure was rated 3 in hardness and who scored above a T-score of 70 on the MMPI Psychopathic Deviate Scale. The subject was a sixteen-year-old referred for psychiatric treatment due to delinquent sexual behavior and drug abuse. The subject stated in association with the "God" drawing, "You can't draw God, but you can..."
draw a flower." She was diagnosed as schizophrenic and subsequently became hallucinatory.

Drawings 4e, 4s and 4g were, respectively, the E-S-G productions of a single subject scoring above a T-score of 70 on the Schizophrenic Scale of the MMPI. The drawings were rated 2-2-2 in hardness and 2-3-2 in size in accordance with the schizophrenic criterion.
EVIL
Welcome to my kingdom.
BIBLIOGRAPHY

Books


Calverton, S. B., Children's Expression Through Drawing, Baltimore, Maryland, Williams and Wilkins, Co., 1894.


Machover, K., Personality Projection in the Drawing of the Human Figure, Springfield, Illinois, Charles C. Thomas, 1948.


Articles


Hiller, E. W. and D. Nesvig, "An Evaluation of Criteria Used by Clinicians to Infer Pathology from Figure Drawings," *Journal of Consulting Psychology*, XXIX (December, 1965), 520-529.

Holzberg, J. D. and M. Wexler, "The Validity of Human Form Drawings As a Measure of Personality Deviation," *Journal of Projective Techniques*, XIV (December, 1950), 34, 61.

Lawton, M. J. and L. Sechrest, "Figure Drawings by Young Boys from Father-Absent Homes," *Journal of Clinical Psychology*, XVIII (July, 1962), 304-305.
Lewinsohn, P. M., "Psychological Correlates of Overall Quality of Figure Drawings," Journal of Consulting Psychology, XXIX (December, 1965), 504-512.


Whitmyer, J. W., "The Significance of Artistic Excellence in the Judgment of Adjustment Inferred from Human Figure Drawings," Journal of Consulting Psychology, XVII (December, 1953), 421-424.

Unpublished Materials
