Convergent and Discriminant Validity of Human Figure Drawings of Children as Indicators of Psychological Traits

Timothy Ward Starkey
Loyola University Chicago

Follow this and additional works at: https://ecommons.luc.edu/luc_diss

Part of the Psychology Commons

Recommended Citation
Starkey, Timothy Ward, "Convergent and Discriminant Validity of Human Figure Drawings of Children as Indicators of Psychological Traits" (1970). Dissertations. 1058.
https://ecommons.luc.edu/luc_diss/1058

This Dissertation is brought to you for free and open access by the Theses and Dissertations at Loyola eCommons. It has been accepted for inclusion in Dissertations by an authorized administrator of Loyola eCommons. For more information, please contact ecommons@luc.edu.

This work is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 License.
Copyright © Timothy Ward Starkey
Convergent and Discriminant Validity of Human Figure Drawings of Children As Indicators of Psychological Traits

BY

Timothy Ward Starkey

A Dissertation Submitted to the Faculty of the Graduate School of Loyola University in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

June

1970
VITA

Timothy Ward Starkey was born on November 10th, 1942, in Beardstown, Illinois. Upon graduation from Beardstown High School in June of 1960, he enrolled in the physics program at the University of Notre Dame. He graduated from this university in June of 1964 with an A.B. in philosophy, and at that time enrolled in the Graduate School of Psychology at Loyola University of Chicago.

Mr. Starkey was a Trainee in Clinical Psychology at the Catholic Charities Guidance Center during his first year in Graduate School. He then served an eighteen-month clinical internship at the Hines Veterans Administration Hospital from 1966 through 1968. Mr. Starkey has since this time completed a first-year clinical child internship at the Institute for Juvenile Research, and is presently completing a second-year clinical research internship at this institution.

Mr. Starkey received his M.A. in clinical psychology from Loyola University of Chicago in June of 1967, and expects to receive his Ph.D. from that institution in June of 1970.
Acknowledgements

The author wishes to express his sincere gratitude to the members of his Dissertation committee, chaired by Dr. Patricia Barger, and consisting of Dr. Ann Heilman and Dr. Jeanne Foley. These psychologists have given generously of their knowledge, interest, and time toward the completion of this research. The author is deeply indebted to Dr. Nicholas Ciaccio, of the Institute for Juvenile Research, for the intellectual stimulation and unflagging assistance he has provided throughout this investigation, and to Dr. Charlotte H. Altman, also of the Institute for Juvenile Research, for her generous support and practical advice throughout the duration of the research. Mr. Roland McCleary, Principal of the Mark Twain School, and Mr. Richard Reilly, Principal of the Sunnyside School, both in School District 87, are also deserving of special thanks for making available the time of their teachers and students. Finally, the author wishes to express his very special appreciation to Miss Regina Donahue, his fiancee, for her tireless efforts in his behalf during the scoring and typing of this research.
# Table of Contents

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Draw-A-Person</td>
<td>1</td>
</tr>
<tr>
<td>Children's Personality Questionnaire</td>
<td>11</td>
</tr>
<tr>
<td>Behavior Problem Checklist</td>
<td>12</td>
</tr>
<tr>
<td>II. Method</td>
<td>14</td>
</tr>
<tr>
<td>Subjects</td>
<td>14</td>
</tr>
<tr>
<td>Measures</td>
<td>14</td>
</tr>
<tr>
<td>Procedures</td>
<td>18</td>
</tr>
<tr>
<td>III. Results</td>
<td>19</td>
</tr>
<tr>
<td>IV. Discussion</td>
<td>30</td>
</tr>
<tr>
<td>V. Summary</td>
<td>34</td>
</tr>
<tr>
<td>References</td>
<td>35</td>
</tr>
<tr>
<td>Appendix A</td>
<td>40</td>
</tr>
</tbody>
</table>
List of Tables

Table                                      Page

1. Summary of Subject Variables.................. 15
2. Draw-A-Person Scoring Reliabilities........... 17
3. Univariate Statistics for All Subjects........ 20
4. Correlations for All Subjects.................. 21
5. Correlations for Fourth Grade Children........ 23
6. Correlations for Fifth Grade Children.......... 24
7. Correlations for Sixth Grade Children.......... 25
8. Correlations for Boys Only...................... 27
9. Correlations for Girls Only..................... 28
List of Appendices

Appendix | Page
---|---
A. Test Materials | 40
   - Draw-A-Person Trait Scales | 41
   - Behavior Problem Checklist | 42
   - Children's Personality Questionnaire, Factor E | 44
   - Children's Personality Questionnaire, Factor C | 45
Abstract

This research utilized a multitrait-multimethod design to investigate the convergent and discriminant validity of human figure drawings of children as indicators of aggression, anxiety, and emotional instability. Draw-A-Person (DAP) scales for these traits were constructed of items selected from the Koppitz (1968) EI Scale for children's drawings. Together with the DAP, the Children's Personality Questionnaire (CPQ) and the Behavior Problem Checklist scores were obtained for 300 4th, 5th, and 6th grade children. The results indicated convergent validity for the DAP with both the CPQ (p < .05) and the Behavior Problem Checklist (p < .01) on the trait of aggression, less validity for the DAP on the trait of emotional instability, and none for the trait of anxiety. No discriminant validity was demonstrated for any of the methods. The size of the DAP aggression correlations was too small to warrant its use as a test of preference for the individual evaluation of that trait. Further research is required to clarify the independence of traits, and the relationships among them, in the functioning personality. Such research should precede, or be concomitant with, studies of the measurement of personality.
Convergent and Discriminant Validity of Human Figure Drawings of Children as Indicators of Psychological Traits

Timothy W. Starkey

Loyola University of Chicago

The purpose of the present research is to investigate the validity of human figure drawings (HFDs) as indicators of aggression, anxiety, and emotional instability in children. The design utilized is the multitrait-multimethod matrix (Campbell & Fiske, 1959), using the Draw-A-Person as the projective method, the Children's Personality Questionnaire (CPQ) as the objective method, and the Behavior Problem Checklist (Peterson & Quay, 1967) as the behavior rating method. This design was selected for its methodological rigor, in that it makes use of both convergence and discrimination of independent methods as the conceptual basis of validity, and thereby avoids the pitfall of requiring a single unidimensional "hard criterion" measure of validity. A cursory review of the projective literature regarding the selected variables demonstrates that at present no such "hard criterion" has been systematically employed. Children were used as subjects partly because of their greater accessibility, but primarily because drawing is a more natural means of expression and communication for them than it is for adults, for whom drawing can be all too often a self-conscious, artificial exercise rather than a free medium of communication.

Draw-A-Person. Human figure drawings have been used for decades by
psychologists as both a rough estimate of intelligence (Goodenough, 1926) and an indicator of various personality traits (Machover, 1949). Its ease of administration and economy of time have no doubt largely contributed to its present-day status as the second most frequently administered psychological test (Sundberg, 1961). Its success as a crude screening device for intelligence has been well established. Typical of the studies on its efficacy as an intelligence screening device is the Wells and Perdini study (1967), in which Goodenough Draw-A-Man (DAM) IQs were correlated with Stanford-Binet (1960, Form LM) IQs on over 1200 school children. The results of this study indicated that the IQ correlations for children between the ages of five and fourteen ranged from .44 to .54. The mental age correlations at these age levels ranged from .37 to .56 for boys, and from .39 to .75 for girls.

Other studies involving both the Goodenough DAM and other HFD scoring systems for intelligence have generally supported the findings of the Wells and Perdini study. DeMoreau and Koppitz (1968), in a comparison study of the Goodenough DAM scores and the Koppitz Developmental Scale scores of 335 lower-class Mexican children, found a high correlation between the two measures. The authors concluded that HFDs were useful for screening large numbers of children in situations in which ease and rapidity of scoring outweigh the disadvantage of not obtaining specific IQ or MA scores. Datta (1967) found that DAM IQs were substantially lower for Head Start children than for normative groups. Strumpfer and Mienie (1968) studied the Harris-Goodenough DAM scores of 79 latency age British children, and found the reliabilities to be acceptably high. Quast and Ireton (1966)
found that medical students could be trained to give and score DAMs. The results indicated that its use proved effective as a screening device for overall psychological functioning.

The value of HFDs as a projective device, however, has been seriously challenged by a substantial body of evidence. Machover (1949) has published the interpretive hypotheses for HFDs that are most frequently tested in the literature. Although she has stated that the structural and formal aspects of HFDs were reliable (Machover, 1949), or at least were more reliable than content variables, other investigators have reported conflicting evidence.

In a study on the content reliability of college students' drawings, Bradshaw (1952) found percentages of agreement ranging from 65 (hips, lips) to 84 (global characteristics). In another early study of college students' HFDs, Lehner and Gunderson (1952) used percentage of agreement as a measure of reliability, and their results ranged from 42 (breasts) to 70 (hair). Swensen (1957), in his thorough review of the HFD literature, criticizes the use of percentage agreement as an index of reliability, pointing out that the significance of the percentage of agreement is entirely dependent on the base rate of the particular body part or structural aspect of the drawing that is being investigated. The more frequently a particular sign is found in a particular part of HFDs, the higher the percentage of agreement must be in order to be significant. Starr and Marcuse (1959) assessed the reliability of seven aspects of HFDs, taking into account the base-rates, on the drawings of college students. They found that placement of the figure on the page, sex of first figure drawn for males, the direction the figures are facing on the page, incompletions, height of figure, and the
ratio of head size to figure size are all reliable at the .01 level of significance.

In contrast to the Starr and Marcuse study, Litt and Margolies (1966) found a considerable degree of variability on retesting their subjects (341 school children), and recommended that caution be observed in interpreting sex-drawn-first from single drawings. Hammer and Kaplan (1964), in a study of 1276 children in the 4th, 5th, and 6th grades, report that when a child draws a same-sex figure, it tends to be reliable, but that when he draws an opposite-sex figure first, it is not reliable. Swensen (1957) suggests that the unreliability of the sign "opposite-sex drawn first" would explain why research has frequently failed to relate this particular sign to an specific psychopathology or trait.

In a later study of the drawings of 1305 children in the 4th, 5th, and 6th grades, Hammer and Kaplan (1966) report that shading and erasures were reliable, particularly with girls, and that their reliability increased with age. Teeth, however, were found to be unreliable. Gravitz (1966), in a study of the sex drawing preferences of 200 normal adult job applicants who ranged in age from 17 to 59 years, found that 85% of males drew same-sex figures, as did 67% of women subjects. In another study using adult subjects, Apfeldorf, Randolf, and Whitman (1963) found that height area and centeredness scores correlated significantly from first drawing to second. With the particular exception of sex-drawn-first, then, HFDs are satisfactorily reliable in their global characteristics, and in many of their content features, to which interpretive meaning is typically given.

Machover (1949) stated that certain features of HFDs could reveal
underlying aggression, possibly of a paranoid variety. Griffith and Lemley (1967), in a study of adult drawings, found that teeth and threatening looks, when occurring together in a drawing, significantly indicated verbal aggression. Neither sign alone, however, indicated aggression. Arata (1965), in a study of children's and adolescents' drawings, found a significant relationship between HFD aggression signs and overt aggression as observed by hospital employees. Koppitz (1966b) found that stance of figure drawn differentiated significantly between normal children and children with aggressive behavior problems. Stance was also one of Machover's original signs of aggression (Machover, 1949). Again supporting some of Machover's hypotheses, Koppitz (1966a) reports that gross asymmetry of limbs, teeth, long arms, big hands, and the presence of genitals on HFDs significantly differentiated shy from aggressive children among 31 pairs who were patients in a child guidance clinic. The drawings of shy children, she further found, were distinguished by tiny figures, cut-off hands, and omissions of the nose and mouth.

The relationship of the DAP to anxiety has been studied much more exhaustively than has its relation to aggression. Swensen (1968) writes in his review of the HFD literature that one of the main problems with shading as an indicator of anxiety is that it is a sign usually found in drawings of good quality. Since drawing quality is positively related to adjustment (Sherman, 1958; Marais & Strumpfer, 1965; Feldman & Hunt, 1958), even though shading may indicate anxiety, the fact that its presence is limited to drawings of good quality tends to confound it with the adjustment variable. This difficulty is largely resolved by Koppitz (1968) in her
study of 1856 drawings of school children, in which she made use of age expectancy rates for different kinds of shading to determine the relative importance or unimportance of shading as an indication of anxiety. She found that as children get older, shading on HFDs takes on a greater interpretive importance. The only exceptions to this was shading of the face, which was found to be highly significant at all ages. Shading of the body and limbs was found to be common for girls through age 7 and boys through age 8. Thus, it was only at age 8 for girls and 9 for boys that shading assumed clinical significance on the Koppitz EI Scale.

Wysocki and Whitney (1965) found that crippled children shaded their drawings more than did non-crippled children. Since crippled children were assumed to have more body anxiety than non-crippled children, these results were interpreted as supporting the basic shading hypotheses. Craddick, Leipold, and Cacavas (1962) found a negative relationship between shading on HFDs and criminal psychopaths. Since psychopaths are assumed to be less anxious than normals, these results were also interpreted as supporting the shading hypotheses. Handler and Reyher (1964), in a study of the effects of stress on HFDs of college students, found that 15 of 21 anxiety and conflict indexes significantly differentiated between stress and non-stress conditions for the male drawing, while 11 of 21 indexes significantly differentiated between the stress and non-stress conditions for the female drawing. Doubros and Mascarenhas (1967) administered the DAP to 204 14-year-old students in a study of the effects of test-produced anxiety on HFDs, and found no support for shading, erasures, or omissions as indicators or predictors of classroom tension. The contradictory results of these two studies support Swensen's (1968) observation that
in general, HFD studies using adult subjects produce positive results, and those using children as subjects produce negative results. While there are important exceptions to this general observation, it is certainly true that the great majority of DAP studies which produced negative results used children as subjects.

Frisch and Handler (1967), in a study of racial differences in the use of shading in HFDs, found that Negro children tended to grossly overemphasize and distort the hair. They ruled out the "sexual virility hypothesis" of the use of hair on HFDs in favor of a cultural interpretation which reflected the Negroes desire for assimilation and integration. These results are particularly interesting from the point of view of Machover's (1949) original hypotheses regarding the relation of shading to state anxiety and level of tension. Handler and Reyher (1966), in a study of the relation between GSR and anxiety indexes in HFDs with college students, found that ten indexes correlated significantly with either GSR frequency or mean conductance. Craddick, Leipold & Cacavas (1962) found that while the rating reliability of shading and anxiety were high for their subjects, they did not correlate with Taylor Manifest Anxiety scores. Overall, both Handler and Reyher (1965) and Swensen (1968), in their reviews of the literature on HFDs, conclude that the results support the validity of a number of anxiety indexes.

Numerous DAP variables have been linked to emotional instability through the years. More recent studies have tended to offset some of the earlier negative studies. Koppitz (1967), in her study of the HFDs of 1856 school children, identified a number of global characteristics and specific features which she found to be significantly related to emotional problems.
and instability. Poor integration of parts, gross asymmetry of limbs, slanting figures, tiny or big figures, and various transparencies were the main global characteristics found in her study to be related to emotional instability, and short arms, big hands, hands cut off, and omission of neck were the specific features of the children's HFDs found to be significantly associated with this personality variable. The objectivity and clarity with which such qualifiers as "tiny," "big," "gross asymmetry," "slanting," and "poorly integrated" are defined in the Koppitz EI Scale contribute greatly to the scoring reliability, and make possible satisfactory tests of its validity.

Machover (1949) stated that stance was one of the principal DAP indicators of personality stability and balance. Koppitz (1966) found that stance of figure drawn differentiated at a significant level between normal children and children with behavior problems in school. Kahn and Jones (1965) found stance to be sufficiently related to severity of illness among non-hospitalized psychiatric patients to predict admission to a mental hospital. Bieliauskas and Kirkham (1958), however, reported that stance failed to differentiate between normal and disturbed children. Mabry (1964), in a study of the serial drawings of a patient with a malignant brain tumor, found that the figure drawn became more and more reclining as the tumor progressed. Swensen (1968), in his review of the literature summarizes the evidence by concluding that Machover's stance hypothesis tends to be supported, particularly by studies using adults as subjects.

Particularly when used to predict diagnostic categories, the DAP has met with mixed success in indicating overall personality adjustment. Eisen (1951), in a study of the drawings of primary school boys, using a 9-item
teacher rating scale as a criterion measure, found that several commonly
used DAP signs differentiated between good-fair-poor adjustment rated boys
at a high level of significance. He also found that the DAP did better with
5-year-olds than with 6-year-olds, presumably because omissions became
ever. Wanderer (1966), in a study of drawings by adult subjects, found that
even DAP "experts" were unable to match diagnoses with schizophrenic,
neurotic, homosexual, or normal groups on the basis of their DAPs. No
allowance was made, however, for inaccuracy of admission diagnosis, or for
overlapping of diagnostic categories. Wanderer's "experts" were able to
identify mental defectives from the above four categories at a better-than-
chance expectancy on the basis of their drawings, however.

Whitmyre (1953), in a study of students' drawings, found a significant
relationship between psychologists' ratings of drawings for personality ad-
justment and art teachers' ratings of the same drawings for artistic ability.
This is in line with the findings of Feldman and Hunt (1958), who concluded
that the more difficult a part is to draw, the more likely a subject is to
demonstrate some sign of disturbance in drawing that part. Sherman (1958)
found relationships between psychologists' ratings of adjustment and artists'
rating of drawing quality.

Since personality adjustment and stability are closely related to self-
esteeam, studies dealing with the relationship between DAP variables and self-
esteeam are of special interest. Gray and Pepitone (1964) found that low-
self-esteem subjects used fewer colors, and drew smaller figures with less
positive facial expressions than did high-self-esteem subjects. They found
no significant differences between the two subject groups on vertical
placement, perspective sex, or activity drawn. They also found that lowering self-esteem through experimental manipulation had a more powerful effect on figure drawing than did raising it.

Lewinson (1964) found that the height of figure drawn with psychiatric patients was negatively related to feelings of depression. Since depression is inversely related to self-esteem, this study lends support to the interpretation of height of figure drawn as revealing self-esteem. Salzman and Harway (1967), however, compared the size of human figure drawings of a psychotically depressed group of patients with those on non-depressed control patients, and failed to find a relationship between depression and size of drawing. Further, changes in mood in patients who improved after treatment for depression were not reflected in subsequent figure drawings. Lakin (1956), in a study of the drawings of institutionalized aged and normal school child subjects, whose self-concepts and self-esteem could be expected to differ markedly, found significant differences in the formal characteristics of the two groups' drawings. They concluded that the formal aspects of figure drawings are related to the central variables of self-conceptualization and body image. Feelings of self-devaluation and shrinking body image found graphic expression in diminished figures drawn. The possibility that organic factors related to age resulted in constricted figures was offset by the non-constricted Bender-Gestalt performances of the aged group.

In summarizing the literature from 1949 to the present, Swensen (1968) concluded that DAP research has improved substantially in quality since its earlier days, and has produced increased empirical support for the use of human figure drawings as a clinical tool. The evidence he cites suggests that the validity of a particular aspect of drawings is directly related to the
reliability of that aspect of the drawing. Global ratings have proven to be the more reliable and valid, while individual signs have proven to be less reliable and valid. In general, studies using adults as subjects have yielded more positive results than those using children.

Children's Personality Questionnaire. While there exists a plethora of inventory-type measures for personality traits of children, there are relatively few that offer numerical scores for the traits investigated in the present study. The most widely known group test that meets this criterion is the Children's Personality Questionnaire (CPQ), which yields a general assessment of personality by measuring fourteen distinct traits (Porter, Cattell, & Ford, 1968). The CPQ is a widely standardized test with two forms, A and B, each form divided into two parts. The test is designed for children between eight and twelve years of age. The CPQ consists of fourteen scales, each measuring a dimension whose functionally independent nature has been established by factor-analytic research (Cattell & Coan, 1958; Cattell & Howarth, 1962). These personality factors have been found not to be unique to the CPQ, but have been established as unitary entities in many researches across various life situations (Cattell & Coan, 1957; Cattell & Gruen, 1954).

Werner (1966) administered the CPQ to 87 talented or underachieving school children. The personality profiles of talented children were like those of artists, writers, and research scientists, while those of the underachievers were like those of delinquents. Werner interpreted those results as indicating that the CPQ could reliably discriminate between talented and underachieving children. Lessing and Smouse (1967) administered
Form A of the CPQ to 110 patients of a child psychiatric clinic and 117 normal 5th and 6th graders. They used sex and group as independent variables in a two-way analysis of variance, and found that clinic children were less dominant, less happy-go-lucky, more restrained, more guilt-prone, more introverted, and more neurotic than the normal children were. Porter, Cattell, and Ford (1968) report a second-stratum, or derived, anxiety scale for the CPQ. This anxiety scale was used in the present research as the inventory measure of anxiety, in the interest of theoretical and construct consistency with the CPQ aggression and emotional stability scales.

Behavior Problem Checklist. Behavior rating scales for children are typically of the diagnostic category variety (Spirach & Spolts, 1965), or are of the multiple-trait type (Peterson & Quay, 1967). The Behavior Problem Checklist (Peterson & Quay, 1967) is a factor analytically derived three-point rating scale for 55 relatively frequently occurring problem behavior traits in children and adolescents. Its development has involved factor analytic studies of problem behavior in public school children (Peterson, 1961; Quay & Quay, 1965), students in public school classes for the emotionally disturbed (Quay, 1964; 1966), and children seen in a child guidance clinic (Peterson, Becker, Shoemaker, Luria & Hellmer, 1961).

The problem behavior dimensions measured by the Checklist are those of conduct disorder (unsocialized aggression), neuroticism (anxious-withdrawn), inadequacy-immaturity, and subcultural (socialized) delinquency. Quay, Sprague, Shulman, and Miller (1966) obtained ratings from both parents and teachers on child-patients of a child guidance center. The correlations between parents were .78 for conduct problem scores and .67 for personality
problem scores.

The preceding discussion of some of the literature on human figure drawings, personality questionnaires, and behavior rating scales serves to demonstrate the present need for an integrative study, using the multitrait-multimethod design (Campbell & Fiske, 1959), with large numbers of subjects and uniform conditions. The diversity of methods and subject populations makes impossible a comparison of these three methods of personality assessment; it is for this reason that the present research utilizes the multitrait-multimethod design to investigate the convergent and discriminant validity of human figure drawings as indicators of personality traits.

The purpose of the present research is to investigate the convergent and discriminant validity of the DAP with the CPQ and the Behavior Problem Checklist through the multimethod matrix, in which convergent validity is shown by the heteromethod-monotrait values being significantly greater than zero, and discriminant validity is shown by the heteromethod-monotrait values greater than their row and column values.
Method

Subjects. The subjects for this research were 300 public school children in the 4th, 5th, and 6th grades of two Chicago grammar schools. Both schools are representative of the middle socio-economic class in terms of their student populations. One hundred and one subjects were in the 4th grade, 105 were in the 5th grade, and 94 were in the 6th grade. One hundred and thirty-five subjects were boys, and 165 were girls. All 4th, 5th, and 6th grade classes were tested in both schools, and no children were excluded (Table 1).

Measures. The scoring reliability of figure drawings has been questioned by many authors, and has been seriously investigated by some. Roback (1968), in his extensive review of the DAP literature, emphasized that there was great need for standardized scoring scales for estimating personality adjustment and traits from human figure drawings. The DAP scales for aggression (D1), anxiety (D2), and emotional instability (D3) used in the present research were constructed of items selected from the Koppitz (1968) E1 Scale (Appendix A). The Koppitz E1 Scale was standardized on 1856 school children, and the reported interjudge reliability of agreement was 95%. The psychologists were in perfect agreement on 444 items of the 467 produced by the subjects in the reliability sample. In the present research, all scoring was done by the investigator, and score-rescore stability coefficients were computed at one-week and three-month intervals. The one-week stability coefficients for aggression (D1), anxiety (D2), and emotional instability (D3) are .92, .94, and .94 respectively, and the three-month stability coefficients are .81,
Table 1
Summary of Patient Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Fourth Grade</th>
<th>Fifth Grade</th>
<th>Sixth Grade</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Boys</td>
<td>46</td>
<td>50</td>
<td>39</td>
<td>135</td>
</tr>
<tr>
<td>Number of Girls</td>
<td>55</td>
<td>55</td>
<td>55</td>
<td>165</td>
</tr>
<tr>
<td>Total</td>
<td>101</td>
<td>105</td>
<td>94</td>
<td>300</td>
</tr>
<tr>
<td>Average Age In Years</td>
<td>10.3</td>
<td>11.5</td>
<td>12.3</td>
<td>---</td>
</tr>
</tbody>
</table>
The Children's Personality Questionnaire (CPQ), Form A, was selected as the objective method for this research (Porter & Cattell, 1963). The Immediate test-retest stability coefficients for aggression (Factor E) and emotional instability (Factor C) reported in the CPQ Manual are .80 and .69. No reliability information is reported on the Anxiety Factor, which was computed with a formula supplied in the CPQ Manual (Anxiety = .2(D+O+Q4-Q3)-.1(C+H)+4.4). Emotional instability scores were obtained by inverting Factor C, on which low scores indicate unstable and emotional personalities. This inversion was accomplished by subtracting the Factor C raw score from 10 (maximum raw score on any factor). The items comprising Factors E and C are presented in Appendix A, marked to indicate presence of the trait associated with the factor.

The behavior rating method used in this research was a modification of the Behavior Problem Checklist (Peterson and Quay, 1967). Two items, one dealing with masturbation and the other with bed-wetting, were deleted from the original Checklist. The Checklists were scored for aggression (Factor I), anxiety (Factor II), and emotional instability (Factor III). The items comprising these factors are presented in Appendix A. The raw scores obtained were given the weightings recommended in the Manual. For a sample of 126 kindergarten children, Peterson (1961) reports interteacher reliabilities of .77 for the conduct problem dimension (Factor I) and .75 for the personality problem dimension (Factor II). Quay and Quay (1965) obtained behavior ratings with Checklist from two teachers on a sample of 7th and 8th grade children. The interteacher correlations ranged from .58 to .71 for conduct problem scores, and from .22 to .31 for personality problem scores. The authors explained these
## Table 2

**DAP Scoring Reliabilities**

<table>
<thead>
<tr>
<th>Interval</th>
<th>Number</th>
<th>$D_1$ Aggression</th>
<th>$D_2$ Anxiety</th>
<th>$D_3$ Emotional Instability</th>
</tr>
</thead>
<tbody>
<tr>
<td>One Week</td>
<td>30</td>
<td>.92</td>
<td>.94</td>
<td>.94</td>
</tr>
<tr>
<td>Three Months</td>
<td>30</td>
<td>.81</td>
<td>.85</td>
<td>.82</td>
</tr>
</tbody>
</table>
latter low correlations by pointing out that the teachers involved averaged only one hour per day of contact with the students they rated. Becker (1960) found that teachers' behavior ratings of their pupils produced scores with typically high intertrait correlations.

**Procedure.** All test data were collected in two sessions by the investigator and an assisting Ph.D. psychologist within a one-week period six weeks after the beginning of the school year. The testing was done on a classroom-by-classroom basis in one 60-minute session, with the DAP administered first and the CPQ second. The children were initially told that they were taking part in a large study of children's drawings, and that the tests they were about to take would in no way affect their grades in school. The instructions for the DAP were as follows: "Please draw a whole person. Do not draw a monster or a snowman." Questions were handled by repeating the instructions and giving reassurance that their drawings were acceptable. The CPQ was administered in the manner described in the Manual (Porter, Cattell, & Ford, 1968), and liberal assistance was given to any child who had difficulty in reading or understanding the test items.

The investigator met with the teachers involved in the research one week prior to the actual testing, and at that time discussed with them in general terms the nature of the research, and specifically the rating task he was asking of them. It was stressed that only actual behaviors observed were to be scored, and not feelings they believed the children to have. At the end of the actual testing periods, the teachers were given the modified Checklists and were requested to complete them within one week. All Checklists were completed and collected within five days of the testing.
Results

Tables 3 and 4 provide the means, standard deviations, and correlations for all subjects, Tables 5 through 7 present the correlations by grade, and Tables 8 and 9 present the correlations by sex. Pearson Product-Moment correlations were used in all analyses.

Campbell and Fiske's (1959) first criterion for multitrait-multimethod validity is that validity diagonals be significantly greater than zero, and that they be large enough to encourage further examination of validity. An examination of Table 4 reveals that for all subjects two validity diagonal values were significant at the .01 level, three were significant at the .05 level, and four were not significant at either level. For the trait of aggression, the DAP correlated significantly with both the CPQ (.05 level) and the Checklist (.01 level), while the CPQ/Checklist correlation for aggression was significant at the .05 level. For emotional instability, the DAP correlated significantly with the Checklist (.05 level), but not with the CPQ, while the CPQ/Checklist correlation for this trait was significant at the .01 level.

Contrary to expectations, no method correlated significantly with any other method for the trait of anxiety.

The second criterion for validity with a multitrait-multimethod design is that a validity diagonal value be higher than the values lying in its column and row in the heterotrait-heteromethod triangles. This criterion is generally met for aggression. The DAP/CPQ aggression correlation (.14) exceeds both its row (.12, .06) and column (.00, .01) values, as does the CPQ/Checklist...
Table 3
Univariate Statistics for All Subjects

<table>
<thead>
<tr>
<th>Method</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Draw-A-Person</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(D_1) (Aggression)</td>
<td>.46</td>
<td>.70</td>
</tr>
<tr>
<td>(D_2) (Anxiety)</td>
<td>.49</td>
<td>.66</td>
</tr>
<tr>
<td>(D_3) (Emotional Instability)</td>
<td>.52</td>
<td>.85</td>
</tr>
<tr>
<td><strong>Checklist</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Factor I (Aggression)</td>
<td>3.17</td>
<td>4.55</td>
</tr>
<tr>
<td>Factor II (Anxiety)</td>
<td>3.71</td>
<td>3.81</td>
</tr>
<tr>
<td>Factor III (Emotional Instability)</td>
<td>2.18</td>
<td>2.46</td>
</tr>
<tr>
<td><strong>Children's Personality Questionnaire</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Factor E (Aggression)</td>
<td>4.59</td>
<td>2.04</td>
</tr>
<tr>
<td>Factor Anx (Anxiety)</td>
<td>4.75</td>
<td>1.43</td>
</tr>
<tr>
<td>Factor C' (Emotional Instability)</td>
<td>4.06</td>
<td>1.92</td>
</tr>
</tbody>
</table>
Table 4
Correlations for All Subjects

<table>
<thead>
<tr>
<th>Method</th>
<th>CPQ</th>
<th>DAP</th>
<th>Checklist</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>Anx</td>
<td>C¹</td>
<td>D₁</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>1</td>
<td>.29*</td>
<td>.14*</td>
</tr>
<tr>
<td>Anx</td>
<td>1</td>
<td>.56**</td>
<td>.00</td>
</tr>
<tr>
<td>C¹</td>
<td>1</td>
<td>.01</td>
<td>.04</td>
</tr>
<tr>
<td>D₁</td>
<td></td>
<td>1</td>
<td>.18**</td>
</tr>
<tr>
<td>D₂</td>
<td></td>
<td></td>
<td>.09</td>
</tr>
<tr>
<td>D₃</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ch₁</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ch₁₁</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ch₁₁₁</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* p < .05
** p < .01

a. Aggression measures are E, D₁, and Ch₁; anxiety measures are Anx, D₂, and Ch₁₁; emotional instability measures are C¹, D₃, and Ch₁₁₁.
list aggression correlation (.16). The DAP/Checklist aggression correlation (.16) exceeds all of its column values, and one of its row values. The hetero­
method anxiety correlations (.09, .03, and .05), however, do not meet this requirement in any instance. For emotional instability, only the CPQ/Checklist correlation (.20) meets this requirement completely, although the DAP/Checklist correlation (.13) does exceed its row values, and one of its column values. The rather high DAP aggression/Checklist emotional instability correlation (.18) prevents DAP/Checklist discriminant validity on this criterion.

The third criterion for validity is that a variable correlate higher with an independent effort to measure the same trait than with measures designed to get at different traits which happen to employ the same method. An examination of Table 4 reveals that all three methods fail this criterion on all traits.

The fourth criterion for validity is that the same pattern of trait inter­
relationships be shown in all the heterotrait triangles of both the monomethod and the hetero­method blocks. Inspection of the overall matrix reveals no clear pattern in any of the correlation blocks.

When the independent variable of grade is considered apart from the overall matrix (Tables 5, 6, and 7), the significance of the validity values diminishes, while the actual size of the correlations remains relatively the same in most cases. This decrease of significance is a function of reduced sample size, and hence of higher requirements for significance. An exception to this, though, is seen in the case of the DAP/CPQ aggression correlations. Here, the size of the correlation is greatest with the 4th grade
<table>
<thead>
<tr>
<th>Method</th>
<th>CPQ</th>
<th>DAP</th>
<th>Checklist</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>1</td>
<td>.31**</td>
<td>-.05</td>
</tr>
<tr>
<td>CPQ</td>
<td></td>
<td>.28*</td>
<td>.17</td>
</tr>
<tr>
<td>Anx</td>
<td></td>
<td>.63**</td>
<td>-.04</td>
</tr>
<tr>
<td>C'</td>
<td></td>
<td>-.02</td>
<td>.10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.22*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.18</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.12</td>
</tr>
<tr>
<td>D1</td>
<td>1</td>
<td>.27*</td>
<td>.26*</td>
</tr>
<tr>
<td>D2</td>
<td></td>
<td>.16</td>
<td>-.08</td>
</tr>
<tr>
<td>D3</td>
<td></td>
<td>.19</td>
<td>.01</td>
</tr>
<tr>
<td>Ch1</td>
<td>1</td>
<td>.28*</td>
<td>.49**</td>
</tr>
<tr>
<td>Ch11</td>
<td></td>
<td>1</td>
<td>.73**</td>
</tr>
<tr>
<td>Ch111</td>
<td></td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

* $p < .05$

** $p < .01$

a. Aggression measures are E, D1, and Ch1; anxiety measures are Anx, D2, and Ch11; emotional instability measures are C', D3, and Ch111
Table 6
Correlations for Fifth Grade Children

<table>
<thead>
<tr>
<th>Method</th>
<th>CPQ</th>
<th>DAP</th>
<th>Checklist</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Ch1</td>
</tr>
<tr>
<td>E</td>
<td>1</td>
<td>.28*</td>
<td>.02</td>
</tr>
<tr>
<td>Anx</td>
<td>1</td>
<td>.54**</td>
<td>.09</td>
</tr>
<tr>
<td>C'</td>
<td>1</td>
<td>-.04</td>
<td>.02</td>
</tr>
<tr>
<td>D1</td>
<td>1</td>
<td>.09</td>
<td>.16</td>
</tr>
<tr>
<td>D2</td>
<td>1</td>
<td>.15</td>
<td>.14</td>
</tr>
<tr>
<td>D3</td>
<td>1</td>
<td>.09</td>
<td>-.01</td>
</tr>
<tr>
<td>Ch1</td>
<td>1</td>
<td></td>
<td>.44**</td>
</tr>
<tr>
<td>ChII</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>ChIII</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* p < .05
** p < .01

a. Aggression measures are E, D1, and Ch1; anxiety measures are Anx, D2, and ChII; emotional instability measures are C', D3, and ChIII.
Table 7
Correlations for Sixth Grade Children

<table>
<thead>
<tr>
<th>Method&lt;sup&gt;a&lt;/sup&gt;</th>
<th>CPQ</th>
<th>DAP</th>
<th>Checklist</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>E</td>
<td>Anx</td>
<td>C'</td>
</tr>
<tr>
<td>E</td>
<td>1</td>
<td>.27*</td>
<td>-.05</td>
</tr>
<tr>
<td>Anx</td>
<td>1</td>
<td>.51**</td>
<td>-.04</td>
</tr>
<tr>
<td>C'</td>
<td>1</td>
<td>.06</td>
<td>.11</td>
</tr>
<tr>
<td>D&lt;sub&gt;1&lt;/sub&gt;</td>
<td>1</td>
<td>.07</td>
<td>.22</td>
</tr>
<tr>
<td>D&lt;sub&gt;2&lt;/sub&gt;</td>
<td>1</td>
<td>.15</td>
<td>.02</td>
</tr>
<tr>
<td>D&lt;sub&gt;3&lt;/sub&gt;</td>
<td>1</td>
<td>.07</td>
<td>.09</td>
</tr>
<tr>
<td>ChI</td>
<td>1</td>
<td>.60**</td>
<td>.74**</td>
</tr>
<tr>
<td>ChII</td>
<td>1</td>
<td>.75**</td>
<td></td>
</tr>
<tr>
<td>ChIII</td>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* p < .05
** p < .01

<sup>a</sup> Aggression measures are E, D<sub>1</sub>, and ChI; anxiety measures are Anx, D<sub>2</sub>, and ChII; emotional instability measures are C', D<sub>3</sub>, and ChIII.
children (.28), and least with 6th grade children (-.01). This would seem to be in agreement with Kopplitz's (1968) observation that the DAP has the greater validity with younger children, and loses value with increasing age. The DAP/Checklist correlations show no such pattern for aggression and anxiety, but do show this pattern for emotional instability. The CPQ/Checklist correlations also show no patterns for aggression or anxiety as a function of grade, but do show an increasing correlation for emotional instability with increasing grade level. The importance of these findings is obscured by the large increase in monomethod-heterotrait correlation values of the Checklist as a function of increasing grade level. It would appear that the Checklist steadily loses discriminative power as grade level increases. Further complicating the picture is the lack of any heteromethod-monotrait significance at the 5th grade level.

When the independent variable of sex is considered, the results indicate that the CPQ/Checklist correlations on emotional instability (.20 for boys, .21 for girls) are the only significant heteromethod-monotrait values (Tables 8 and 9). Surprisingly, the CPQ monomethod-heterotrait correlation for aggression and anxiety for girls (.33) is much higher than for boys (.14). Since this does not occur at all for the CPQ across grade level, it may indicate that these traits are more diffuse in girls than in boys. Similarly, the DAP monomethod-heterotrait correlation for aggression and emotional instability is much greater for boys (.32) than for girls (.06). Again, a great difference does not occur across grade level on this variable, possibly suggesting that aggression and emotional instability are less distinct as traits in
Table 8
Correlations for Boys Only

<table>
<thead>
<tr>
<th>Method</th>
<th>CPQ</th>
<th>DAP</th>
<th>Checklist</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>E</td>
<td>Anx</td>
<td>C'</td>
</tr>
<tr>
<td>CPQ</td>
<td>1</td>
<td>.14</td>
<td>-.06</td>
</tr>
<tr>
<td>Anx</td>
<td>1</td>
<td>.52</td>
<td>-.11</td>
</tr>
<tr>
<td>C'</td>
<td>1</td>
<td>-.01</td>
<td>.08</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D_1</td>
<td>1</td>
<td>.16</td>
<td>.32**</td>
</tr>
<tr>
<td>D_2</td>
<td>1</td>
<td>.15</td>
<td>.14</td>
</tr>
<tr>
<td>D_3</td>
<td>1</td>
<td>.19*</td>
<td>.04</td>
</tr>
<tr>
<td>Chl</td>
<td>1</td>
<td>.47**</td>
<td>.53**</td>
</tr>
<tr>
<td>Chll</td>
<td>1</td>
<td>.73**</td>
<td></td>
</tr>
<tr>
<td>Chlll</td>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* p < .05
** p < .01

a. Aggression measures are E, D_1, and Chl; anxiety measures are Anx, D_2, and Chll; emotional instability measures are C', D_3, and Chlll.
Table 9
Correlations for Girls Only

<table>
<thead>
<tr>
<th>Method</th>
<th>CPQ</th>
<th>DAP</th>
<th>Checklist</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>E</td>
<td>Anx</td>
<td>C'</td>
</tr>
<tr>
<td>E</td>
<td>1</td>
<td>.33**</td>
<td>-.02</td>
</tr>
<tr>
<td>Anx</td>
<td>1</td>
<td>.60**</td>
<td>-.01</td>
</tr>
<tr>
<td>C'</td>
<td>1</td>
<td>.03</td>
<td>.00</td>
</tr>
<tr>
<td>D1</td>
<td>1</td>
<td>.12</td>
<td>.06</td>
</tr>
<tr>
<td>D2</td>
<td>1</td>
<td>.01</td>
<td>-.05</td>
</tr>
<tr>
<td>D3</td>
<td>1</td>
<td>-.04</td>
<td>.03</td>
</tr>
<tr>
<td>Ch1</td>
<td>1</td>
<td></td>
<td>.38**</td>
</tr>
<tr>
<td>Ch1I</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Ch1II</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ch1III</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* * * p < .05
** p < .01

a. Aggression measures are E, D1, and Ch1; anxiety measures are Anx, D2, and Ch1I; emotional instability measures are C', D3, and Ch1II.
boys than in girls.
Discussion

The results of this study indicate that the DAP demonstrates convergent validity with questionnaire and teacher-rating measures of behavior for the trait of aggression at a level of significance sufficient to warrant its use as a general screening instrument for this trait. The magnitude of the correlations for the DAP are not sufficient to recommend its use as a test of preference for individual evaluation of aggression, however. Further, the DAP demonstrated only fair convergent validity on emotional instability, and none on anxiety. In light of the original hypotheses, it can be said that the DAP afforded little or no discriminant validity with the Checklist, while the CPQ succeeded in this respect only with emotional instability.

It must be stated at this point that the CPQ did not demonstrate appreciably better convergent validity than did the DAP, and afforded far less discriminative power, in spite of its much greater cost in time and effort of administration.

Of the three instruments, the CPQ proved to be the least efficient in terms of information lost through incomplete or randomly marked data. It was the observation of the writer that several children in each classroom marked their CPQ answer sheets without pausing to read each question carefully. Even when both Form A₁ and Form A₂ are used, as was the case in this research, the total number of items for each factor is only 10. With so few items, the haphazard marking of even one or two items could invalidate this test for any given child. Inspection of the CPQ answer sheets revealed a relationship between difficulty in taking the test and score on the CPQ general intelligence
factor (Factor B). Those children who did not complete the test, or who were observed to be randomly marking the answer sheet, tended to get lower raw scores on the general intelligence factor. Not surprisingly, nearly all of the questions omitted, even with the brighter children, were items belonging to the general intelligence factor.

The difficulties encountered with the Behavior Problem Checklist were of a different nature. In line with Becker's (1961) findings, individual response styles of the teachers emerged from the Checklist data as critically important. The homogeneity of within-class ratings was as striking as the heterogeneity of between-class ratings, both within and across grade levels. Many teachers appeared to have a definite set, and marked either many items for everyone in their class, or few if any items for anyone. One teacher, in fact, gave over two-thirds of her class a zero rating on all items, while another teacher, of the same grade, gave nearly every child in her room a score for several items on all traits. Clearly, there was a tendency for some of the teachers to rate their pupils on the basis of their global feelings toward them, rather than on the basis of behaviors they had actually observed. For future research use of this instrument, teacher subjectivity effects could be lessened by selecting classrooms that rotate through several teachers during the course of the day, and averaging the teachers' ratings of the children. Thus, each child's Checklist scores would be a composite average instead of being dependent upon only one teacher's view.

Another major difficulty with the Checklist was the very large size of the method effect. In the multitrait-multimethod matrix, the presence of method variance is indicated by the difference in the level of correlation between
the parallel values of the monomethod block and the heteromethod block, qualified by an assumption of comparable reliabilities for all methods. While this assumption is not necessarily met in the present research, it is worthwhile to examine these correlation differences to estimate the probable size of the method effects. It is apparent that method effects were considerable for the CPQ. The CPQ aggression/anxiety correlation (.29) is much greater than the CPQ aggression/DAP anxiety correlation (.12) or the CPQ aggression/Checklist anxiety correlation (-.07), and the CPQ anxiety/emotional instability correlation (.56) far exceeds the CPQ anxiety/DAP emotional instability correlation (.10) or the CPQ anxiety/Checklist emotional instability correlation (.16).

The DAP, however, particularly with the anxiety correlation comparisons, shows less method effect through lower heterotrait-monomethod correlations (D₁ to D₂ correlation of .18; D₂ to D₃ correlation of .09; D₁ to D₃ correlation of .21). Certainly the minimal method error of the DAP with anxiety is due to the generally low correlations of this variable with any other variable. The Checklist, through its extremely large heterotrait-monomethod correlations (Ch₁ to Ch₁ correlation of .44; Ch₁-to Ch₁I correlation of .57; Ch₁ to Ch₁I correlation of .73), demonstrates the largest method effect of any of the three instruments, and hence has the least discriminative power for the traits studied.

The power of the DAP to identify children projecting the traits studied was lessened by the large number of zero trait scores, or normal drawings, which the children produced. This resulted in relatively low means for the DAP scales, but relatively large variances. Clearly, the evaluation of validity of the DAP was hampered by the choice of a normal subject population. Future
research could profitably restrict the domain of subjects to those whose drawings are judged to be clinically relevant by one or more judges, and correlate those drawings with the corresponding subjects' CPQs and Checklists. This would much more closely approximate the clinical enterprise as it is actually practiced.

Since the methods of assessment were theoretically independent, and there was no overlapping of items in the actual methods, it must be considered why some of the heterotrait correlations exists. One possibility is that the traits do not exist as independent traits in a functioning personality. Further, on the basis of the results of the present study, it appears possible that the expression of the traits studied is different in boys than in girls. Further research is required to clarify the independence of traits, and the relationships among them, in the functioning personality. Such research should precede, or be concomitant with, studies of the measurement of personality.
Summary

This research utilized a multitrait-multimethod design to investigate the convergent and discriminant validity of human figure drawings of children as indicators of aggression, anxiety, and emotional instability. Draw-A-Person (DAP) scales for these traits were constructed of items selected from the Koppitz (1968) E1 Scale for children's drawings. Together with the DAP, the Children's Personality Questionnaire (CPQ) and the Behavior Problem Checklist scores were obtained for 300 4th, 5th, and 6th grade children. The results indicated convergent validity for the DAP with both the CPQ (p < .05) and the Behavior Problem Checklist (p < .01) on the trait of aggression, less validity for the DAP on the trait of emotional instability, and none for the trait of anxiety. No discriminant validity was demonstrated for any of the methods. The size of the DAP aggression correlations was too small to warrant its use as a test of preference for the individual evaluation of that trait. Further research is required to clarify the independence of traits, and the relationships among them, in the functioning personality. Such research should precede, or be concomitant with, studies of the measurement of personality.
References


Cattell, R. & Gruen, W. Primary personality factors in the questionnaire medium for children eleven to fourteen years old. Educational and Psychological Measurement, 1954, 14, 50-76.


35


Koppitz, E. Emotional indicators on human figure drawings and school achievement of first and second graders. *Journal of Clinical Psychology*, 1966, 22, 481-483. (a)

Koppitz, E. Emotional indicators on human figure drawings of shy and aggressive children. *Journal of Clinical Psychology*, 1966, 22-4, 466-469. (c)


Lewinsohn, P. M. Relationship between height of figure drawings and depression in psychiatric patients. *Journal of Consulting Psychology*, 1964, 28, 380-381.


Quay, H. Dimensions of personality in delinquent boys as inferred from the factor analysis of case history data. *Child Development*, 1964, 35, 479-484.


Appendix A. Test Materials
Draw-A-Person Trait Scales

**D₁ - Aggression**
- Gross Asymmetry
- Big Figure (9" or more)
- General Transparencies
- Teeth
- Long Arms (reaching to knees)
- Big Hands (as big or bigger than face)
- Genitals
- Omission of Arms

**D₂ - Anxiety**
- Partial Shading of Face
- Shading of Body
- Shading of Hands
- Specific Transparencies
- Hands Cut Off
- Legs Pressed Together
- Omission of Eyes
- Omission of Mouth
- Omission of Legs

**D₃ - Emotional Instability**
- Poor Integration of Parts
- Shading of Entire Face
- Shading of Neck
- Slanting Figure (15° or more from perpendicular)
- Tiny Head (height of head less than one-tenth of entire figure)
- Omission of Body
- Clouds
- Omission of Neck
Behavior Problem Checklist

1. Oddness, bizarre behavior
2. Restlessness, inability to sit still (Aggression)
3. Attention-seeking, "show-off" behavior (Aggression)
4. Doesn't know how to have fun; behaves like a little adult (Anxiety)
5. Self-conscious; easily embarrassed (Anxiety)
6. Fixed expression; lack of emotional reactivity
7. Disruptiveness; tendency to annoy and bother others (Aggression)
8. Feelings of inferiority (Anxiety)
9. Steals in company with others
10. Bolsterousness, rowdiness (Aggression)
11. Crying over minor annoyances and hurts (Anxiety)
12. Preoccupation; "in a world of his own" (Emotional Instability)
13. Shyness, bashfulness (Anxiety)
14. Social withdrawal, in preference for solitary activities (Anxiety)
15. Dislike for school (Aggression)
16. Jealousy over attention paid other children (Aggression)
17. Belongs to a gang
18. Repetitive speech
19. Short attention span (Emotional Instability)
20. Lack of self-confidence (Anxiety)
21. Inattentiveness to what others say
22. Easily flustered and confused (Anxiety)
23. Incoherent speech
24. Fighting (Aggression)
25. Loyalty to delinquent friends
26. Temper Tantrums (Aggression)
27. Reticence, secretiveness (Anxiety)
28. Truancy from school
29. Hypersensitivity; feelings easily hurt (Anxiety)
30. Laziness in school and in performance at other tasks (Emotional Instability)
31. Anxiety, chronic general fearfulness (Anxiety)
32. Irresponsibility, undependability (Aggression)
33. Excessive daydreaming (Emotional Instability)
34. Has bad companions (Emotional Instability)
35. Tension, inability to relax (Anxiety)
36. Disobedience, difficulty in disciplinary control (Aggression)
37. Depression, chronic sadness (Anxiety)
38. Uncooperativeness in group situations (Aggression)
39. Aloofness, social reserve (Anxiety)
40. Passivity, suggestibility; easily led by others (Emotional Instability)
41. Clumsiness, awkwardness, poor muscular coordination
42. Hyperactivity; "always on the go" (Aggression)
43. Distractibility
44. Destructiveness in regard to his own and/or other's property (Aggression)
47. Negativism, tendency to do the opposite of what is requested (Aggression)
48. Impertinence, sauciness (Aggression)
49. Sluggishness, lethargy (Emotional Instability)
50. Drowsiness (Emotional Instability)
51. Profane language, swearing, cursing (Aggression)
52. Nervousness, jitteriness, jumpiness, easily startled
53. Irritability; hot-tempered, easily aroused to anger (Aggression)
55. Often has physical complaints, e.g. headaches, stomach-ache
Children's Personality Questionnaire, Factor E
(Items marked to indicate aggression)

A1

13. Does your teacher think you are good at sitting still ___ or ___ that you run around too much

17. Would you rather hunt birds ___ or ___ draw pictures of birds

21. Would you rather talk with your teacher ___ or ___ talk with a good friend

25. Would you like better to have bears here now ___ or ___ to hear stories about bears

29. Is mother's way of doing things always better ___ or ___ is your new way sometimes better

A2

13. Would you rather be a school teacher ___ or ___ a great hunter

17. On a playground do you make a lot of noise ___ or ___ play quietly, without much noise

21. Would you rather write a book ___ or ___ be the main actor in a play

25. When you get a new game as a present, do you like to try it first yourself ___ or ___ have someone show you how to play it

29. If teacher scolded you badly, would you cry when you told mother ___ or ___ just laugh when you told her
Children's Personality Questionnaire, Factor C
(Items marked to indicate emotional instability)

A1

3. Do you think you could do well at almost anything ___ or X just a few things

4. In a game on the playground, do you stand around X or ___ run a lot

6. Do you feel nervous at school X or ___ are you happy

8. In your group is someone else the leader X or ___ are you the leader

10. Do you think you smile a great deal ___ or X do not smile much

A2

3. Can you easily persuade your friends to accept your plans ___ or X is it difficult

4. Do you think many children do better work than you X or ___ are you as good as anyone else

6. Do grown-ups think you are naughty X or ___ well-behaved

8. Do you make a lot of mistakes X or ___ just a few

10. If you got lost, would you know what to do ___ or X would you be scared
APPROVAL SHEET

The thesis submitted by Timothy Ward Starkey has been read and approved by the director of the thesis. Furthermore, the final copies have been examined by the director of the thesis and the signature which appears below verifies the fact that any necessary changes have been incorporated, and that the thesis is now given final approval with reference to content and form.

The thesis is therefore accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

30 May 1970

[Signature of Adviser]