This dissertation has been 61–2905
microfilmed exactly as received

SAUNDERS, Mauderie Hancock, 1927–
AN ANALYSIS OF CULTURAL DIFFERENCES
ON CERTAIN PROJECTIVE TECHNIQUES.

The University of Oklahoma, Ph.D., 1961
Education, psychology

University Microfilms, Inc., Ann Arbor, Michigan

Downloaded from SHAREOK. Persistent Handle URL: https://hdl.handle.net/11244/1503

COPYRIGHT BY

MAUDERIE HANCOCK SAUNDERS

1961

THE UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

AN ANALYSIS OF CULTURAL DIFFERENCES ON CERTAIN PROJECTIVE TECHNIQUES

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

DOCTOR OF PHILOSOPHY

BY

MAUDERIE HANCOCK SAUNDERS

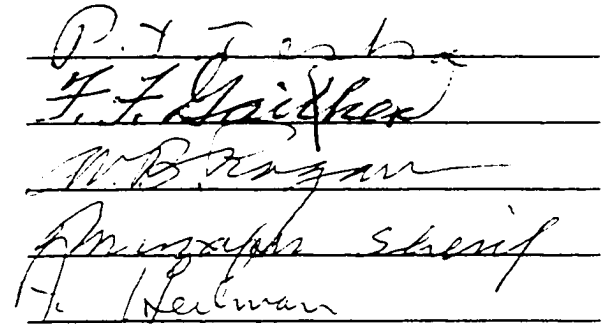
Norman, Oklahoma

1961

AN ANALYSIS OF CULTURAL DIFFERENCES ON

CERTAIN PROJECTIVE TECHNIQUES

APPROVED BY



DISSERTATION COMMITTEE

ACKNOWLEDGMENTS

The writer expresses sincere gratitude to Dr. P. T. Teska. This dissertation would not have been possible without his untiring assistance. Dr. Teska's encouragement, scholarly guidance, and patience made the completion of this study possible.

Sincere appreciation is extended to the judges who gave their time so unstintingly. The writer expresses the same appreciation to the teachers and pupils who participated in this research. Sincere gratitude is expressed to Mesdames Wanda Jo Glidewell, Pauline Krieger, Wanda Sharp Green, Evelyn Williams, and Miss Mary Wells who aided in the collection of the basic data. The same appreciation is extended to all visiting counselors of the Department of Pupil Services, under the direction of Dr. V. T. Hill and Dr. R. R. Phillips, Oklahoma City Public Schools, who rendered services in the study.

The writer expresses sincere appreciation to Mrs. Ora Thompson who typed the first reading copy of this dissertation. The same appreciation is extended to Mrs. Starlin Powell for her cooperation and valuable services rendered in typing the final copy.

Mrs. Barbara Ward Clark and Dr. Valmore Goines completed many tasks valuable to this study. Their enthusiasm, sincere friendship and cooperation will always be remembered. The same appreciation is extended to the administrative staff and to the many staff members of Southern University who arranged class schedules for the writer's convenience and who relieved the writer of classroom duties while this research was being completed.

Sincere gratitude is extended to the members of the dissertation committee who have made many valuable suggestions in this research effort. Their interest and support made the completion of this dissertation possible.

The writer expresses gratitude to her husband and daughter, Leonard and Cheryl, who made many sacrifices during this research. The same appreciation is extended to the writer's brother, A. A. Hancock, who has offered encouragement and rendered many valuable services.

TABLE OF CONTENTS

Page

LIST OF TABLES vii

Chapter

I. THE PROBLEM AREA 1

Introduction 1
Pilot Study 12
Statement of the Problem 16
Significance of the Problem 21
Purpose of the Study 23
Procedure and Method 24
Summary 25

II. REVIEW OF RELATED LITERATURE 27

House-Tree-Person 28
Machover-Draw-A-Person Test 35
Human Figure Drawings 40
Projective Techniques and Reading Achievement 45
Race and Intelligence 47
The Utilization of Clinicians as Judges 55
Summary 57

III. EXPERIMENTAL DESIGN 60

Sample 60
Testing Procedures 67
Hypotheses 68
Experimental Procedure 72
Group I--Ability 74
Group II--Reading 74
Group III--Emotional Adjustment 75
Group IV--Race 76
Statistical Analysis 76
Summary 77

Chapter Page

IV. ACCURACY AND VALIDITY OF JUDGES' RESPONSES 80

Accuracy of Judges' Responses 80
Validity of Judges' Responses 80
Reading Achievement 86
Emotional Adjustment 91
Race 96
Summary 101

V. AGREEMENT AMONG JUDGES 103

Summary 110

VI. COMPARISON OF JUDGES' CORRECT RESPONSES ON NEGRO AND WHITE CHILDREN'S DRAWINGS 112

Summary 122

VII. SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS 124

Conclusions 127
Recommendations 130

BIBLIOGRAPHY 132

APPENDICES 140

I. HYPOTHESES AND CHI-SQUARES FOR PILOT STUDY 141

II. JUDGMENT FORMS IN BASIC DATA FROM THE DICHOTOMIZING PROCESS 147

III. CASE DATA FORMS FOR EXPERIMENTAL AND CONTROL GROUPS 160

LIST OF TABLES

Table		Page
1.	Source of sample drawn from the socio-economic districts of Oklahoma City Public Schools--1960	62
2.	Means and differences of intelligence quotients of white and Negro first-grade children in an integrated school, Oklahoma City, Oklahoma	65
3.	Means and differences of intelligence quotients of Negro and white first-grade children in two racially homologous schools of the above average socio-economic districts, Oklahoma City, Oklahoma	65
4.	Means and differences of intelligence quotients of Negro and white first-grade children in two racially homologous schools of the low socio-economic districts, Oklahoma City, Oklahoma	65
5.	Summary of judges' agreement on H-T-P and DAP results from forty (40) first-grade elementary school children of Oklahoma City, Oklahoma, based on evaluations of mental ability	81
6.	Significance of judges' estimates of Negro and white children's mental ability based on their drawings	83
7.	Significance of judges' estimates of bright Negro and white children's mental ability based on their drawings	83
8.	Significance of judges' estimates on dull Negro and white children's mental ability based on their drawings	84

Table Page

9. Significance of judges' estimates of Negro children's mental ability based on their drawings 85

10. Significance of judges' estimates of white children's mental ability based on their drawings 85

11. Summary of judges' responses on H-T-P and DAP drawings for forty (40) first-grade elementary school children of Oklahoma City, Oklahoma, based on evaluations of reading achievement 87

12. Significance of judges' estimates of Negro and white children's reading achievement based on their drawings 88

13. Significance of judges' estimates of good readers based on Negro and white children's drawings 89

14. Significance of judges' estimates of poor readers based on Negro and white children's drawings 89

15. Significance of judges' estimates of reading achievement based on Negro children's drawings 90

16. Significance of judges' estimates of reading achievement based on white children's drawings 91

17. Summary of judges' responses on H-T-P and DAP drawings for forty (40) first-grade elementary school children of Oklahoma City, Oklahoma, based on evaluations of emotional adjustment 93

18. Significance of judges' estimates of Negro and white children's emotional adjustment based on their drawings 94

19. Significance of judges' estimates of well adjusted Negro and white children's emotional adjustment based on their drawings 94

Table Page

20. Significance of judges' estimates of poorly adjusted Negro and white children's emotional adjustment based on their drawings 94

21. Significance of judges' estimates of Negro children's emotional adjustment based on their drawings 96

22. Significance of judges' estimates of white children's emotional adjustment based on their drawings 96

23. Summary of judges' responses on H-T-P and DAP drawings for forty (40) first-grade elementary school children of Oklahoma City, Oklahoma, based on selections of race 98

24. Significance of judges' selections of race based on Negro and white children's drawings 100

25. Significance of judges' selections of Negro children based on children's drawings 101

26. Significance of judges' selections of white children based on children's drawings 101

27. Chi-square frequencies of judges' correct estimates on ability, reading achievement, emotional adjustment, and race based on children's drawings 105

28. Chi-square frequencies of judges' correct estimates on ability, reading achievement, emotional adjustment, and race based on Negro children's drawings 105

29. Chi-square frequencies of judges' correct estimates on mental ability, reading achievement, emotional adjustment, and race based on white children's drawings 106

30. Chi-square frequencies of judges' correct estimates on mental ability of bright and dull children based on children's drawings of both races 107

Table Page

31. Chi-square frequencies of judges' correct estimates of reading achievement of good and poor readers based on children's drawings of both races 108

32. Chi-square frequencies of judges' correct estimates on emotional adjustment of well adjusted and poorly adjusted children based on children's drawings of both races 109

33. Chi-square frequencies of judges' correct estimates on mental ability based on Negro and white children's drawings 113

34. Chi-square frequencies of judges' correct estimates of mental ability based on bright Negro and bright white children's drawings . . . 114

35. Chi-square frequencies of judges' correct estimates of ability based on dull Negro and dull white children's drawings 115

36. Chi-square frequencies of judges' correct estimates of reading achievement based on Negro and white children's drawings 116

37. Chi-square frequencies of judges' correct estimates of reading achievement based on Negro good readers and white good readers children's drawings 117

38. Chi-square frequencies of judges' correct estimates of reading achievement based on Negro poor readers' and white poor readers' drawings 118

39. Chi-square frequencies of judges' correct estimates on emotional adjustment of Negro and white children based on their drawings . . . 119

40. Chi-square frequencies of judges' correct estimates on emotional adjustment of well adjusted Negro and white children based on their drawings 120

41. Chi-square frequencies of judges' correct estimates on emotional adjustment of poorly adjusted Negro and white children based on their drawings 121

Table | Page

42. Chi-square frequencies of correct responses made on race based on Negro and white children's drawings 122

AN ANALYSIS OF CULTURAL DIFFERENCES ON CERTAIN PROJECTIVE TECHNIQUES

CHAPTER I

THE PROBLEM AREA

Introduction

For over 300 years, American society has provided a conspicuously greater opportunity for whites than for Negroes in cultural, social, and economic areas. In view of the differences that exist, this study is an attempt to determine the diagnostic and predictive value of certain projective techniques in the evaluation of characteristics of Negro children as reflected in their responses to the House-Tree-Person and Draw-A-Person techniques.[1] As a result of these special cultural, social, and economic forces, it is hypothesized that differences will appear when a comparison is made of test responses of whites and Negroes. Such differences in responses when analyzed may be traceable to certain conditions, such as: (1) longer life-span among

---

[1]Hereafter will be abbreviated as H-T-P and DAP.

whites than among Negroes; (2) lower incidence of certain physical diseases--e.g., tuberculosis, gonorrhea, syphilis, rickets, etc., among whites than Negroes; (3) less intense racial self-hate among whites than among Negroes; (4) fewer incidences of juvenile delinquency among white than among Negro youth; (5) greater achievements in political, economic and social status among whites than among Negroes; (6) greater educational achievement among whites than among Negroes in all professions and in all occupations of skill; (7) higher average IQ score achieved by white school age children than by Negro school age children relative to geographical sections of the United States.

There is evidence of a longer life-span among white than among colored persons. Reports prepared by the U. S. Bureau of the Census[2] and various studies made by insurance companies[3] render proof of this. For the period covering 1930, there was a lower mortality rate, as revealed by the U. S. Bureau of Census, for white than for Negroes up to the age of 75 years. Maurice R. Davie reports:

> At the age of ten the life expectancy of white males was 45.6 years in 1911. For white females at the age of ten life expectancy was at 50.6 years. For colored males and females in 1911, life expectancy was 41.3 years. In

---

[2]U. S. Bureau of the Census, Sixteenth Census of the United States, 1940. Vital Statistics Rates in the United States, 1900-1940 (Washington: Government Printing Office, 1943).

[3]U. S. Bureau of the Census, United States Life Tables and Actuarial Tables, 1939-1941 (Washington: Government Printing Office, 1946).

> 1935 for white males life expectancy had increased to 53.7 years and for white females it was 57.6 years. For colored males in the same year it was 47.0 years and for colored females it was 49.4 years.[4]

Again, Davie[5] reveals the difference more vividly. In spite of steady and rapid improvement, amounting to a gain of twenty years in life span since 1900, the average duration of life for Negroes in 1940 was about at the level achieved by the whites before the first World War.

Higher incidence of certain diseases exist among Negroes than among Caucasians. Davie says:

> The National Health survey of 1935-36 found that the amount of disability per person due to illness which incapacitated for a week or longer was 43 per cent higher in the Negro than in the white population. The higher disability rate for Negroes was attributed . . . to the greater prevalence among them of chronic diseases and to their lower economic status. The selective service records over the period 1940 to 1947, according to Smith, showed a rejection rate for military service, because of physical or mental disease or defect, of 27.6 per cent for white and 41.1 per cent for Negro males aged 18 to 45.[6]

By August, 1945, Negro registrants (for military service) accounted for approximately 60 per cent of all those rejected for venereal disease. Peacetime studies had found that the prevalence of venereal disease among Negroes was

---

[4]Maurice R. Davie, Negroes in American Society (New York: McGraw-Hill Book Company, Inc., 1949), p. 237.

[5]Ibid.

[6]Ibid., p. 238.

eleven times as great as among whites.[7]

In 1949, Davie said, "the mortality rate due to tuberculosis for Negroes is three times as great as for Caucasians.[8] Davie quotes Dr. Thomas Parran:

> Poverty and ignorance are the friends of disease in all races, and the Negroes have more than their share of both. The invariably lower income, the higher percentage of relief, and the lower level of living--all contribute to the high mortality rate. In the case of the diseases from which the Negro suffers particularly, social factors are especially prominent.[9]

Some studies have shown the excessive self-hate among Negroes. Clark made a study to determine the development of racial awareness, racial identification, and racial preference. Negro children were shown four dolls identical except for color. Two dolls were brown and two were white. Children as early as three years of age showed awareness of skin color. They were also asked questions which revealed a preference toward the dolls. The majority of the Negro children at each age indicated an unmistakeable preference for the white doll and a rejection of the brown doll.[10] In view of these findings, it is very probable that the H-T-P and DAP would elicit responses that reveal self-hate and personal rejections

---

[7]Eli Ginsberg, The Negro Potential (New York: Columbia University Press, 1956), p. 68.

[8]Davie, op. cit., p. 240.

[9]Ibid.

[10]Kenneth B. Clark, Prejudice and Your Child (Boston: Beacon Press, 1955), p. 23.

such as were found by Clark. An analysis of the questions presented shows that Negro children not only reject themselves but have formed attitudes toward Negroes which are consistent with those of the majority group. This too should be revealed particularly in the results of these two projective devices.

Juvenile delinquency among Negro youth is more prevalent than among Caucasians. Juvenile delinquency in the Negro represents hostile aggression upon society and would probably be reflected in the two projective devices under study, by gross distortions of the self-image. In the cases handled by the juvenile courts in the country, Negro children appear more than three times as often as is to be expected on a population basis.[11]

It seems apparent that there is greater achievement in political, economic, and social status among whites than among Negroes. The American caste system sets limits and barriers for the Negro that are not easy to overcome. In the deep South, the Negro is constantly reminded that he is inferior and is not to enjoy the opportunities available for whites. In the North there are other subtle ways of maintaining barriers in all areas of living. The problem of the Negro is one of a vicious circle. In order to earn more money to improve living conditions, one must be skilled and

---

[11]Davie, op. cit., p. 258.

trained for the higher paid jobs. To obtain education, one must have an income to afford the training. With these social pressures exerted upon the Negro, it is probable that the H-T-P test would show extreme ornamentations indicating a more wishful state of conditions. This social system has limited the Negro's opportunities in the political and economic world. Of the approximately 17 million Negroes in the United States, there are only a few who hold political positions of status in the United States Government. The economic status of the Negro is clearly revealed by the following data:

> In 1954, for the total U. S., the white family's median income was $4,339 while the nonwhite was $2,410. The total urban median money income of United States families for whites was $4,827 while for nonwhites it was $2,876. The total rural-farm median income for whites was $2,157 while for nonwhites it was $763.[12]

Eli Ginsberg makes the following statement which reveals the income of the Negro as compared to that of the Caucasian when related to the amount of education achieved:

> A recent analysis of 1950 Census data on the relation between education and income among the white and Negro population shows that the well-educated Negro has improved his position considerably since 1940. But it also shows that Negroes still earn far less than whites who have the same amount of education. In the North and West as well as the South, Negro college graduates earned less in 1949 than whites who had attended but not graduated from high school. In 1940 Negro college graduates earned less than whites who had not graduated

---

[12]Ginsberg, op. cit., p. 16. Note: Data was taken from Census publications. Separate data for Negroes are not available, therefore, 95 per cent of the non-white population represent a close approximation of data for Negroes alone.

> from grammar school. Both in the South and in other regions, the difference between Negro and white earnings tends to be greater, the more education individuals have acquired.[13]

There is greater educational achievement among whites than there is among Negroes. One out of every five Negro men reaching adulthood in the South in 1950 had not completed the fifth year of school.[14] In the North, one out of five white high school graduates completes college, but only one out of nine Negro high school graduates completes college.[15]

There is evidence revealed by tests with standardized norms based upon a white population that white school children earn a higher average I.Q. score than Negro school children, relative to geographical sections of the United States. In 1944, Otto Klineberg reviewed many studies which have been made concerning the intelligence of Negroes. He found that the median I.Q. for Negroes was 86 while the average I.Q. for the total white population is that of a theoretical value of 100.[16] Robert D. North's[17] review of studies made relative

---

[13] Ibid., p. 38.

[14] Ibid., p. 44.

[15] Ibid., p. 48.

[16] Otto Klineberg, Characteristics of the American Negro (New York: Harper and Brothers, 1944).

[17] Robert D. North, "The Intelligence of American Negroes," Anti Defamation League (November, 1956), p. 28.

to Negro-white intelligence in the United States reveals that the average I.Q. scores of Negroes are below those obtained by whites in the same geographical areas. He found no proof of differences in mental ability levels between whites and Negroes which stemmed from genetically determined "racial traits." However, he did find that differences were traceable to environmental factors. Therefore, with improvement of environmental and educational opportunities, scores made by Negroes on intelligence tests increase which result in smaller differences of I.Q. points earned by Negroes and whites.

D. M. Johnson's study, which was made in 1948, revealed that the average I.Q. score for Negroes being screened for military service during World War I was 83.[18] G. R. Wells made a study in 1923 in which he applied the Binet-Simon tests to groups of white and Negro school children. He found an inverse relationship of Negro children's I.Q. scores and chronological age. The younger Negro children earned comparable I.Q. scores to those obtained by white children.[19]

Obtained I.Q. scores differ among Negroes as to regions parallel with cultural, social, and economic opportunities in

---

[18]D. M. Johnson, "Application of the Standard Score I.Q. to Social Statistics," J. Soc. Psychol., XXVII (1948), 217-227.

[19]G. R. Wells, "The Application of the Binet-Simon Tests to Groups of White and Colored School Children," Psychol. Monogr., IIIXII (1923), 52-58.

the United States. As early as 1935, Klineberg[20] made a study in New York City. The results of the study suggested that I.Q. scores of Negroes in New York City increased with length of residence. Everett S. Lee tested Klineberg's hypothesis in Philadelphia in 1951 and made the following conclusions based upon his findings:

> Klineberg's hypothesis that there is an increase in the intelligence scores of southern Negro migrants to New York with increasing length of residence in New York is, in the main, substantiated by independent evidence in Philadelphia. There is a significant and continuous upward trend in the intelligence test ratings of southern-born Negro children as their length of residence in Philadelphia increases. This increase manifests itself not only on a general intelligence test, but also on each of the sub-tests of the Chicago Tests of primary mental abilities with the single exception of memory. The increase in general score cannot, therefore, be attributed to an increase in any one specific ability. Nor can the increase be laid to increasing familiarity with the tests or the testing situation, or to a trend to be found in all students, since there is no such increase in the Philadelphia-born students. It can further be shown that the migrant children who entered the first grade in Philadelphia are, on the first three tests definitely inferior to the Philadelphia-born on the Philadelphia series of tests, but by the time they have reached the sixth grade, there is no significant difference in their test ratings and those of the Philadelphia-born group, who like them, had not attended Kindergarten.[21]

Evidence of Negroes earning high I.Q.'s when cultural, social, and economic opportunities are present is shown in John B. Miner's report on the work done by Jenkins on high

---

[20]Otto Klineberg, Race Differences (New York: Harper and Brothers, 1935), Chap. IX.

[21]Everett S. Lee, "Negro Intelligence and Selective Migration: A Philadelphia Test of the Klineberg Hypothesis," American Sociological Review, XVI (1951), 227-233.

I.Q.'s among Negro children:

> Sixteen studies report I.Q.'s of above 130; and twelve report cases above 140, almost all of these being from Northern urban communities. Jenkins himself studied eighteen cases with Stanford-Binet I.Q.'s above 160, four of these being above 180, and one above 200. Such high I.Q.'s could not be obtained if native potential, environmental stimulus potential, or motivation were at a low level.[22]

All of these studies cited above seem to reveal a difference in I.Q. scores obtained between whites and Negroes. The differences seem to be more apparent as age increases and as cultural, social, and economic opportunities decrease. It is probable that differences would be revealed in the two projective devices under study just as other devices have revealed differences.

Many tests which did not include Negroes in the standardization population have been made, and we have made interpretations in view of standardized norms based upon white subjects. Two of the most widely used standardized tests of intelligence are the Stanford-Binet Intelligence Scale and the Wechsler Intelligence Scale for Children. Of the large number of studies made on these two tests, a very few have been made on the validity of scores made by Negroes. In the last decade, of numerous studies made on the validity of the Wechsler Intelligence Scale for Children (WISC) (Wechsler, 1949), only two studies have been conducted with

---

[22]John B. Miner, Intelligence in the United States (New York: Springer Publishing Co., 1957), p. 75.

Negro children. Young and Pitts (1951) and Young and Bright (1954) conducted studies applying the WISC to Southern Negro children. They found the test norms to be questionable when applied to these children. Young and Pitts (1951) found the mean WISC Full Scale I.Q. score of a control group representative of their culture to be 69.8.[23] Young and Bright (1954) tested a group of Southern Negro children and found the mean WISC Full Scale I.Q. to be 67.71.[24] As a result of these studies, the authors question the applicability of Wechsler's norms to Southern Negro children.

Studies have been made which reveal differences of earned I.Q. scores between white and Negro children, but the interpretations as revealed are not always sound. In Audrey M. Shuey's[25] analysis of Negro intelligence, it is very vividly suggested that the differences which are found in scores obtained are traceable to racial biological inferiority. This interpretation is made without consideration of environmental factors.

It might be hypothesized that some relationship between

---

[23]Florence M. Young and Virginia A. Pitts, "The Performance of Congenital Syphilitics on the Wechsler Intelligence Scale for Children," Journal of Consulting Psychology, No. 15 (1951), pp. 239-242.

[24]Florence M. Young and H. H. Bright, "Results of Testing 81 Negro Rural Juveniles with the Wechsler Intelligence Scale for Children," Journal of Social Psychology, No. 39 (1954), pp. 219-226.

[25]Audrey M. Shuey, The Testing of Negro Intelligence (Lynchburg, Virginia: Randolph-Macon Women's College, 1958).

the race and cultural influences may be evaluated by means of a valid projective device. It was postulated in the experimental design that structured measuring devices would be contaminated by the nature of the tests, and the socially deprived sub-samples would be quantitatively appraised. Such appraisals would fail to identify salient features of identification which should effectively reflect the peculiar performances of the ethnic element being evaluated.

Since the administration and interpretation of projective devices employed in the analysis of behavior require tremendous time resources, it was deemed appropriate to conduct prior investigations upon a limited sample to determine if the basic hypotheses were reasonably tenable. It was necessary to discover the problems that might emerge in a larger and more sophisticated experimental situation. The pilot study utilized the testing materials and situations of the major design and presented evidence that could be more completely analyzed in the rigid circumstances of the research which followed.

Pilot Study

The pilot study was conducted in which a comparison was made of trained and untrained white and Negro judges with one another in their estimations of the following patterns of behavior of first and fourth-grade Negro and white children on the basis of the H-T-P drawings of these children: mental

ability; reading achievement; socio-economic status; and emotional adjustment. Comparisons were made among six judges: two white judges who were trained in clinical psychology, two white judges who were not trained in clinical psychology, and two Negro judges who were not trained in clinical psychology. The judges were asked to make high and low dichotomized estimations as to mental ability, reading achievement, socio-economic status, and emotional adjustment.

Hypotheses.--The hypotheses to be tested were: (1) if significant relationship existed between the accuracy of the judge and the race of the group of children whose drawings he rated; (2) if significant relationship existed between the accuracy of the judge and his training or lack of it in relation to the race of children whose drawings he rated; (3) if significant relationship existed between the accuracy of the judge and the grade level of the race of children whose drawings he rated; (4) if significant relationship existed between the accuracy of the judge and the race of the children whose drawings he rated; (5) if significant relationship existed between the accuracy of the judge and the kind of drawings made by the two races of children whose drawings he rated such as house, tree, person I, or person II.

Method and Procedure.--Sixty-five first and fourth-grade Negro and white subjects served in the pilot study. However, 12 pupils were randomly selected from each grade

level of each race in order to make the number of subjects in the subgroups equal; hence the drawings of the randomly eliminated pupils were not included for rating. Therefore, a total number of 48 pupils' drawings were actually represented in the pilot study. A final total of 48 pupils attending the public schools of Oklahoma City and Norman, Oklahoma; and 6 judges served in the pilot study.

The testing procedure was very simple. The pupils were merely asked to draw a (1) house, (2) tree, (3) person, and (4) draw another person of the opposite sex of the first person drawn.

The directions for the judges were also very simple. They were merely asked to make estimations, in dichotomized ratings, of mental ability, socio-economic status, reading achievement, and emotional adjustment of the children whose drawings they rated. The last task was for the judges to make judgments as to the race of the child whose drawings he rated.

Results and Conclusions.--Chi-square tables contained frequencies of individual children by grade level and by race categories of correct number of estimates for each of the five variables rated by the judges. Judgments were made on each of the four drawings: house, tree, person I, person II.

Of the ninety-three chi-square values obtained, nine were significant (see Appendix I). This result suggests that

9.7 per cent of the total chi-squares revealed a difference in the ability of trained and/or untrained judges to differentiate in drawings of white and Negro children on the four variables tested. This is to say, the experiment as conducted did not reveal significant differences in trained or untrained judges nor in the ability of the judges to do better on one race of children than they did on the other. However, upon inspection of the total responses of the judges, it was observed that trained judges made more correct responses than did untrained judges. Therefore, the pilot study yielded its most significant result from a subjective examination of the data in addition to the significant differences revealed statistically. It was difficult to handle the statistical data adequately due to the smallness of "N." For example, large differences were apparent before a division was made into categories of both judges correct versus neither judge correct. However, when such division was made, the "N" became too small for the use of X^2.

The pilot study revealed the following trends: (1) trained judges were able to make more correct estimates than untrained judges based on children's drawings; (2) untrained judges differed in their estimates based on children's drawings; (3) untrained judges differed in their estimates based on Negro children's drawings.

The major experimental design was altered in the following manner: (1) judgments were made only by trained

judges; (2) judgments were made on the four drawings (house, tree, person I, person II) as a whole or constellation instead of being made on separate drawing items; (3) judgments were made on a larger sample; (4) judgments were made only on first-grade children; and (5) judgments were made on children who had been equated as to socio-economic status instead of judging socio-economic status of the children whose drawings were being judged.

Statement of the Problem

A fundamental assumption is made in all intelligence testing that differences in obtained I.Q. scores are traceable to differences in native intellective ability where environmental opportunity has been equal. Customarily, when considering the scores made by an entire group of white children from deprived areas, allowances are frequently made. Sherman and Key made a study of white children from the Hollows district of Virginia. The degree of environmental impoverishment was great in this area. They found the I.Q. level decreased as the children's chronological age increased. The study gave impetus to the idea that mental growth takes place only in an environment which demands a certain kind of learning which is measured by intelligence test. These children living in a physically and intelligently deprived environment responded quite differently from the average children on whom the test was standardized.

Wheeler[26] made a study of 3000 or more children living in the mountains of East Tennessee. From the results obtained from intelligence tests, Wheeler compared the I.Q.'s with the I.Q.'s obtained on the same families ten years prior to his study. He learned that the median I.Q. had decreased as the chronological age increased from 102.6 at six years to 81.3 at fifteen years of age.

> However, the later survey yielded a median I.Q. of 93 as compared with the median I.Q. of 82 obtained ten years earlier. This difference in I.Q. level was interpreted as reflecting the general improvement in educational and environmental conditions during the decade separating the two surveys.[27]

Equal care should be taken in the case of Negro children. In the case of the individual Negro or groups of Negroes, the lower score is assigned to Negroes as a whole. Audrey M. Shuey[28] made a study of Negro intelligence. She cited many studies made on Negro children which revealed an I.Q. level lower than that of white children. The environmental factor was stressed as not being important. Dr. Shuey considered Negroes and whites as being exposed to the same culture. For over 300 years Negroes have lived in a discriminatory and isolated environment, however, it is true they have lived in the United States for this period of time. The culture

---

[26]L. R. Wheeler, "A Comparative Study of the Intelligence of East Tennessee Mountain Children," Journal of Educational Psychology, IIIXIII (1942), 321-334.

[27]George G. Thompson, Child Psychology (New York: Houghton Mifflin Co., 1952), p. 421.

[28]Shuey, op. cit.

of the American Negro is best described in the words of Kardiner and Ovesey:

> The culture in which the Negro lives is American. Hence, we expect to find in the Negro an American personality . . . the same basic patterns of adaptation we find in the white man, insofar as they are determined by social conditions common to both. In addition, we expect to find the results of the impact of those conditions that are different to the Negro. These come under the heading of social discrimination. They force the Negro to live within the confines of a caste system which not only interferes seriously with all varieties of social mobility through class lines, but simultaneously tends to stifle effective protest by threat of hostile retaliation from the majority white. Such oppression cannot but leave a permanent impact on the Negro's personality.[29]

While it is well known there has been in America a wide differential in the economic, educational, social, and politcal treatment of Negro and white there is a common practice or tendency to use tests with Negroes which have been standardized on a white population. In view of literature available as to poor validity of test results when tests are standardized on a certain group and used with another, it seems that a careful interpretation of results should be made. In connection with this statement, Thompson has said:

> The conclusion that poor environmental conditions tend to decrease the IQ level of children in a progressive manner has also been drawn by Crissey, Gordon, Skeels, et al. Worbois has further shown that children attending rural one-room schools over a two-year period tended to show some losses in IQ level, while children from the same geographical regions who attended a consolidated (central) school showed IQ gains over the same period.[30]

---

[29]Abram Kardiner and Lionel Ovesey, The Mark of Oppression (New York: W. D. Norton and Co., Inc., 1951), p. 11.

[30]Thompson, op. cit., p. 421.

It is realized that group norms on individual intelligence tests are not available for Negroes, but interpretation should be made in terms of environmental differences for Negroes as they are made for underprivileged whites. Anne Anastasi[31] illustrates the effective role played by standardized norms. She reports that DuBois standardized a Draw-A-Horse Test on Indian children. He found that when this test was administered to white children a lower IQ was earned than was earned by Indian children.

There are three non-verbal tests, the House-Tree-Person, Goodenough-Draw-A-Man, and the Machover-Draw-A-Person tests that are frequently used to get at personality and intelligence. These three tests have presupposed that the tasks required involve only experiences that are common to all persons regardless of cultural differences. However, as Anastasi says, "No existing test is entirely unrestricted in its cultural reference. The difference between 'culture-free' and other tests is one of degree."[32] Therefore, these three non-verbal tests are less culturally influenced than individual verbal tests. Here we have in the House-Tree-Person test an instrument that is (1) easily and quickly administered, and (2) extremely useful in getting at social and emotional

---

[31]Anne Anastasi, Differential Psychology (New York: The Macmillan Company, 1958), Chap. XVI, pp. 542-570, 566.

[32]Ibid., p. 562.

adjustment. This instrument is not as widely used as it could be in view of the above statements and furthermore, the instrument was not standardized on the Negro population. The Goodenough-Draw-A-Man test is another instrument that is (1) easily and quickly administered, (2) extremely useful in getting at social and emotional adjustment and intelligence level, and (3) easy to score. The same advantages may be stated of the Machover-Draw-A-Person test. These tests may be relatively more culture free than any of the individually administered diagnostic intelligence tests. What diagnostic and predictive value do these tests have when they are applied to Negro children?

To answer the above stated question, the writer has made an effort in this research to determine:

1. Are trained clinicians able to estimate the mental ability of first-grade Negro children equally as well as they are able to estimate the mental ability of first-grade white children by their House-Tree-Person and Machover-Draw-A-Person drawing?

2. Are trained clinicians able to estimate the reading achievement of first-grade Negro children equally as well as they are able to estimate the reading achievement of first grade white children by their House-Tree-Person and Machover-Draw-A-Person drawings?

3. Are trained clinicians able to estimate the emotional adjustment of first grade Negro children equally as well as

they are able to estimate the emotional adjustment of first-grade white children by their House-Tree-Person and Machover-Draw-A-Person drawings?

4. Are trained clinicians able to differentiate between Negro and white children's House-Tree-Person and Machover-Draw-A-Person drawings?

Significance of the Problem

Intelligence and personality tests have become of accepted importance among psychologists, educators, sociologists, and other professional groups. These tests have played roles of eminence in predicting, and planning the future of individuals. In our schools, we have used standardized tests without questioning the validity when applying them to individuals of varying socio-economic status and cultural background. It is known that children of high social status tend to earn higher IQ scores than children of low social status. However, we continue to use standardized tests, in many cases, as the sole basis for diagnostic and predictive purposes. Some of the purposes for which tests are used are placement in institutions of various kinds, guidance and vocational planning, and for predicting the future of individuals.

Emanuel F. Hammer made studies which included Virginia children's drawings of a House-Tree, and Person. Hammer[33]

---

[33]Emanuel F. Hammer, "Frustration-Aggression Hypothesis Extended to Socio-Racial Areas: Comparison of Negro and White Children's HTP's," Psychiatric Quarterly, XXVII (1953), 597-707.

found that Negro children, even at early grade levels, tend to be more neurotic than white children, as revealed in comparison of House-Tree-Person drawings by Negro and white children. As the test was not standardized on Negro children, a comparison without the proper interpretation of results is difficult to make. Expressing the idea of Ruth Benedict,[34] behavior is normal relative to the culture. Neurotic behavior in one culture is not necessarily neurotic behavior in another.

The fact that there is a definite and measureable relationship between the scores which pupils obtain on intelligence tests and the social status, or cultural background, of their parents has been known since the time of Binet.[35] As early as 1912, the problem of social class differences as it relates to the results obtained on intelligence was considered.[36] In more recent years, Otto Klineberg and Eli Ginsberg recognized the differences which exist between ethnic groups and made their interpretations in light of environmental factors.

It is important that a way be discovered to detect and to properly evaluate all human resources in terms of

---

[34]Ruth Benedict, Patterns of Culture (New York: Houghton Mifflin Co., 1934).

[35]Kenneth Eells, and Others, Intelligence and Cultural Differences (Chicago: University of Chicago Press, 1950), p. 3.

[36]Ibid., p. 6.

human potentials. It is possible that the Negro one-tenth of the population remains foreign because of the lack of instruments to evaluate the potentials of these people.

Purpose of the Study

The purpose of this study is to test the applicability of the House-Tree-Person Test, Machover-Draw-A-Person Test, and the Goodenough-Draw-A-Man scoring scale to Negro children. The applicability of this test to Negroes may be subject to question. What diagnostic and predictive value to these tests have when applying them to Negro children? In an effort to answer these questions, comparisons were made of the judgments of trained persons who evaluated drawings by white and Negro children. Estimates were made in terms of mental ability, reading achievement, emotional adjustment, and race.

Personal Interest.--It has been observed that many Negro children earn low scores on standardized tests but generally perform in academic areas on a much higher level. In this research effort, the House-Tree-Person, Machover-Draw-A-Person test, and the Goodenough scoring scale were chosen for study. There were two reasons for selecting these tests: (1) They are non-verbal in structure and (2) are diagnostic and predictive instruments for appraising personality and intelligence.

Procedure and Method

This is a research problem utilizing the normative survey technique to evaluate data of a field study and employing small sample probability treatment and chi-square to analyze the data. The normative aspect was appropriate because of the necessity of relating the Negro-white results of the several tests to the established norms. While the detail scoring technique was not applied, the interpretations were based upon clinical standards which have proved to be consistent and valid. However, the normative clinical evidence was derived from Caucasoid samples, and as such, may not have been similar in nature to the experimental Negro groups which were tested.

The white samples of the research were the control units and these were equated with similar Negro samples and the results were pooled for classifying by the judges. The judgments were tests of discrimination involving the capacity of the judges to identify the nature of the samples and testing their accuracy in classifying the productions according to criteria presented by the experimenter.

The projections must be viewed as the matrix effect of the operation of environmental stimuli producing reactions and influencing productions which could furnish clues for judge-teams to succeed in error free classifications.

These data were collected in the schools where the students were regularly enrolled and intensive observations

of the behavior of the children in their usual activities and social milieux were used to assure that the experimentally established classification criteria were as valid as could be in such experimental circumstances. It is assumed that the projections collected in these field circumstances are replete with the cultural loadings, and thus, the differential character of the experimental groups were maximally stimulated.

Summary

In this chapter, the writer has attempted to present the following existing conditions: (1) cultural, educational, economic, and political limitations of Negroes living in an American social system as presently exists; (2) effects of these existing conditions upon Negroes in general; and (3) effects of these existing conditions upon IQ scores earned by underprivileged persons. The writer has further suggested the misuse of, and the improper interpretations of intelligence tests and personality tests applied to Negroes when results are compared with those of whites.

A summary of the pilot study was presented. This summary revealed problems encountered in the pilot study and revealed certain important trends in its results. The major experimental design was altered as a result of these problems and trends.

The important role of intelligence and personality tests have caused many professional groups to apply tests to

various cultural groups without adequate norms for these groups. To conduct a study designed to determine the applicability of two projective devices to Negro children is certainly of great value to psychologists, sociologists, educators and other professional groups.

This research is a normative survey which utilized a field study approach. The data was analyzed by small sample probability treatment and chi-square. Information presented showed that the normative aspect was appropriate in this research. White samples were the control units, and these were equated with similar Negro samples. The results were pooled for classifying by the judges.

Finally, it was indicated, the research was designed to provide for collecting data on children while they were in a familiar and natural environment--the school--and while they were engaged in their usual activities.

CHAPTER II

REVIEW OF RELATED LITERATURE

The purpose of this chapter is to present literature that relates to the present study. This literature will not include all studies that have been conducted in all areas covered by the present study, but a careful selection has been made to present investigations that bear directly upon some aspect of the study. A great deal of literature has been presented in Chapter I and more literature will be presented in following chapters as it relates to the discussion at hand. Many references are listed in the bibliography that will not be reviewed in this chapter. Therefore, studies presented herein represent only a summary of investigations of importance to the subject.

The literature which is to follow was selected on the basis of broad areas involved in this study such as the tests used, the socio-economic factor which has been used as a basis for selecting children's drawings for comparison; teacher ratings, which have been used as a supplementary aid for a basis of selecting children's drawings; reading achievement, which was one of the variables tested; emotional adjustment,

which is another variable tested; and race and intelligence, which was a major aspect of the study. A special effort will be made to mention studies in which the utilization of judges was the major experimental technique employed.

The major concern in this investigation is the applicability of the H-T-P and the DAP techniques to Negro children. As will be observed from the testing procedure, a combination of the H-T-P and the DAP tests has been employed in this study. This testing procedure is one employed by Dr. P. T. Teska in the Special Education Laboratory at the University of Oklahoma, and has been highly recommended by him.

House-Tree-Person

The H-T-P is a two-phased approach to the personality.[1] One phase is unstructured and non-verbal while the other is a structured verbal situation. In the first place, the individual is asked to draw a house, tree, and a person. He is not told what house, what kind of tree, or what person to draw. For this reason, Buck says it is basically an unstructured situation. In the second phase, the situation is more structured in that the individual is asked to define, describe and to interpret verbally his drawings. He is also provided with an opportunity to qualify his drawings so that an interpretation might be made of the individual's interaction with his environment.

---

[1]John Buck, "The H-T-P Technique," Journal of Clinical Psychology, Monograph Supplement, V (October, 1948).

Buck chose the specific items of House, Tree and Person because:

> (1) They were items familiar even to the comparatively young child; (2) they were found to be more willingly accepted as objects for drawings by subjects of all ages than other items suggested, and (3) they appeared to stimulate more frank and free verbalization than did other items.[2]

The H-T-P is recommended as a useful diagnostic projective technique for evaluating the total personality. Buck postulated:

> (1) That each of the drawn wholes (House, Tree, and Person) is to be regarded as a self-portrait, as well as the drawing of a specific or composite House, Tree, or Person, since subjects are believed to draw only those characteristics of a given whole which they someway regard as essential, and often those characteristics are found on objective appraisal to bear little resemblance to that which the subject says he has reproduced; (2) that a subject may indicate that a given detail or combination of details or the method of their presentation, proportionally or spatially individually or as a totality, has special significance for him in two ways: Positive by: (a) overtly exhibiting emotion immediately before, during, or after drawing a given detail or combination of details, or while commenting upon it or them in the P-D-I; (b) by representing the detail or detail complex in sequential order deviant from the average; (c) by exhibiting unusual concern over the presentation as by erasing excessively (and particularly when the erasure is not followed by improvement in form), or as by returning to the detail or detail complex one or more times during the drawing of that or some subsequent whole (and the last is the more pathoformic), or by taking an excessive amount of time in drawing the detail or detail-comples; (d) by presenting the detail or detail-complex in bizarre fashion; (e) by perseveration upon the presentation of a detail; or (f) by his frank comment (spontaneous or induced) concerning a whole or any part or parts thereof. And negatively by (a) presenting a detail or detail-complex incompletely; (b) by omitting altogether one or more so-called essential details; (c) by commenting

---

[2] Ibid., p. 3.

> evasively or refusing to comment at all upon a whole or any part or parts thereof; (3) that interpretation of these "significant" details, detail-complexes, and/or the method of their presentation will provide information concerning the subject's needs, fears, strivings, conflicts, etc.; (4) that it is essential that the subject be afforded every opportunity to aid in interpretation of his productions, because of the clinically observed fact that the so-called universal and absolute meaning of certain symbols may be radically altered in certain configurations (in short, the usually accepted symbolic import of a given item or method of presentation may be almost completely changed by the unique significance ascribed to it by the subject, a significance that may well be overlooked completely if the subject is not given every chance to express it); (5) that adequate interpretation of a specific point can be made only when it is considered in its relationship to the total circumspection and in the light of as complete a knowledge as possible of the subject and his environment (both past and present).[3]

The H-T-P technique is recommended as a valid measure of intelligence as well as other personality factors. In connection with this postulate Buck says:

> It is further postulated: (1) that the H-T-P is a valid measure of adult intelligence, despite its restricted and unconventional intelligence from the standpoints of elemental information (details); spatial relationships (proportion and perspective); and concept formation (as evidenced by the organization and quality of the completed whole and by the subject's spontaneous and/or induced comments concerning it); (3) that the problem presented the subject in the non-verbal phase involves the reproduction as a pencil drawing and in 2-dimensional form of a memory image or a combination of memory images of 3-dimensional form; (4) that because of the relatively primitive method of expression, drawing, subjects who find verbalization difficult may reveal in their H-T-P drawings the presence of an hitherto unsuspected intellectual ability or potential.[4]

---

[3] Ibid., p. 4.

[4] Ibid., p. 5.

From the above paragraph, one can readily see that Buck does recommend the H-T-P technique as a measure of intelligence. His fourth postulate listed above would suggest that the test is an excellent one for culturally deprived subjects. Realizing that Negro children generally are deprived culturally and academically, this technique lends itself as a worthwhile instrument for clinical practices with Negroes.

Before going further into the discussion of the validation of the H-T-P technique, mention should be made of its usefulness as a group test. Buck states that its greatest use in this respect is as a screening device to identify those within a given group who deviate rather sharply from the average in personality maturation and adjustment.[5]

As for the standardization of the H-T-P Technique, John Buck began his experimental studies in 1938, and as a result of its successful employment and as a result of validation studies, it was introduced as a standardized test in 1948. In the standardization process, a selected sample of 120 adults was used. These adults, fifteen years of age and over, represented six intellectual levels: imbecile, moron, borderline, dull average, average, and above average. Most subjects were white patients or employees of the Lynchburg State Colony, Lynchburg, Virginia. The above average subjects were selected from the Universities of

---

[5]Ibid., p. 26.

Nebraska and Virginia. In other studies of validation, it was found that the H-T-P test was reliable in determining personality maladjustment and the intellectual level of adults.

Since Buck's publication of the technique in 1948, many studies of validation have been conducted. Bieliauska[6] conducted a study to determine scorer's reliability in the quantitative scoring of the H-T-P technique. The test was administered to 43 college students who were enrolled in Mental Hygiene, Introductory Psychology, the Experimental Psychology classes at the Richmond Professional Institute of the College of William and Mary. Bieliauska used three judges to score H-T-P drawings. His study suggested that judges agreed or disagreed the same way on the same H-T-P I.Q. scores. As a result of his findings, Bieliauska assumed that judges were having difficulty in the scoring system and suggested a refinement of directions. This study has significance because it reveals the differences in scorer's interpretation of the scoring system as established by Buck. The writer in the present study has avoided the possibility of scoring difficulties by having judges use a dichotomized classification. Buck, even though he protested against using a blind analysis, did so in participating in the study conducted

---

[6]Vytauska Bieliauska, "Scorer's Reliability in the Quantitative Scoring of the H-T-P Technique," Journal of Clinical Psychology, XII (1956), 366-369.

by Smykal and Thorne.[7] He obtained a high percentage of correct estimates of people's personality agreeing with test data, detailed clinical histories and erring only in failure to predict the direction of anti-social impulses of person's whose drawings he rated. Because of the findings of Bieliauska's and Smykal's studies, it was not required of the judges in this present study to use quantitative scoring.

Even though Buck recommends the H-T-P test as a diagnostic tool for determining personality maturation and adjustment, he also suggests that the instrument is useful in determining the intelligence level of adults. Investigations have been made to determine the validity of this instrument as a test of intelligence. Rubin[8] conducted a study in which he made a comparison of the H-T-P and the Wechsler-Bellevue Scale. Rubin's study was prompted by Buck's inconsistent statements on the value of his instrument as a measure of intelligence. The Wechsler-Bellevue was used as the criterion for determining the value of the H-T-P's of 108 patients admitted to the Neuropsychiatric center of a veterans hospital. Rubin says in his summary of the study, "the results of a comparison revealed the H-T-P to be a valid

---

[7]Anthony Smykal and F. C. Thorne, "Etiological Studies of Psycho-pathic Personality: II Associal Type," Journal of Clinical Psychology, VII (1951), 299-316.

[8]Harold Rubin, "A Quantitative Study of the H-T-P and Its Relationship to the W-B Scale," Journal of Clinical Psychology, X (1954), 597-607.

measure of intelligence. The per cent Raw G I.Q. had the greatest value in this respect."[9]

A study in which the H-T-P technique was employed in comparing the performances of white and Negro children was conducted by Hammer.[10] Because of the results obtained in a previous study made by Hammer in which Negro children in a community in Virginia earned an average I.Q. score of 75 on two intelligence tests, he tested the hypothesis that the low earned I.Q. scores were traceable to emotional adjustment. Drawings were selected from a group of Negro and white children who live in a community near Lynchburg Virginia State Colony. It was found by using Buck's criteria, that 10.3 per cent of whites revealed through their drawings a feeling of great frustration produced by a restraining environment with concomitant feelings of hostility and a desire to react aggressively as compared to 28.3 per cent Negroes who revealed these same feelings. A "t" score of these differences was 3.84 which is significant at the .01 level of confidence. He concluded that because of environmental pressures of this society on minority groups, Negro children of that particular community were "relatively more neurotic" than a similar group of white children of the same community.

---

[9] Ibid.

[10] Emanuel F. Hammer, "Frustration-Aggression Hypothesis Extended to Socio-Racial Areas: Comparison of Negro and White Children's HTP's," Psychiatric Quarterly, XXVII (1953), 597-707.

Thus, the preponderant evidence of prior research indicates that the neurotic consequences of the generally traumatic social backgrounds of the Negro experimental group should produce patterns and figures which would identify these subjects as emotional victims of a depressive culture.

Machover-Draw-A-Person Test

The Machover-Draw-A-Person test was an aspect of this study used as an additional part of the testing procedure. Because of the significant role of the Machover test in this study, a section in this chapter has been devoted to a review of literature dealing specifically with this test. It is not the desire of the writer to make a distinction between the Machover test and other human figure drawings except for its immediate role in this research and for clarity in the review of related literature.

The Machover test has shown usefulness as a personality diagnostic tool. The utilization of human figure drawings did not originate with this test, but as Machover states:

> Interest in the projective significance of painting and drawing, which stems from a variety of motivations, is recorded in a wide range of literature and is discussed from both psychological and artistic angles. That all art contains some element of the intimate personality patterns of the artist has long been recognized. The human figure representation as a basic vehicle for projection of these patterns in specific and subtle body language, however, is new to publication, despite the verification and supprt that the method has received in more than 29 years of clinical application to a vast

variety of personalities and clinical problems.[11]

The technique has provided a systemized method of using the human figure drawings as a technique of evaluating personality just as Goodenough introduced human figure drawings as a technique of evaluating intelligence. This statement is substantiated by Machover:

> The writer's technique of personality evaluation through the drawing of a human figure originated in the routine application of the Goodenough test to children for measurement of intelligence. It became increasingly apparent that children received the same IQ on the Goodenough Scale were telling different things. The irrepressible and spontaneous comments of the children, subsequently and elaborated by guided associations to the figures, testified repeatedly to the validity and richness of the language of the body. Most of the years that followed in the development of the drawing technique were devoted to systematization and decoding of this private, ideographic communication. These efforts were rewarded with comprehension of an over-extending horizon of graphic detail and variety of projection. Progress became possible only after the basic key was developed--the projection of the body image in its functional implications.[12]

An attempt was made on the part of this writer to review literature pertaining to the validation of the Machover test. However, it was found that most studies dealt with certain aspects of Machover's theory or certain assumptions made by her. In the writer's opinion, the two studies which

---

[11] Anderson and Anderson, An Introduction to Projective Techniques, "Drawing of the Human Figure: A Method of Personality Investigation," Karon Machover (New York: Prentice-Hall, Inc., 1951), p. 343.

[12] Ibid., p. 344.

are to follow do not discredit the usefulness of the Machover test, because studies will be presented to reveal the usefulness of human figure drawings in determining emotional adjustment.

The validity of the Machover-Draw-A-Person Test was investigated by Richard H. Blum.[13] His study was to determine the validity of this method as a clinical device in understanding adult personality. His sample consisted of 31 patients in an army neuropsychiatric center in Korea. He concluded:

> Ratings based on the Machover interpretation of the DAP were compared with ratings by psychiatrists, a battery of psychological tests and an intuitive DAP scheme on 38 personality characteristics. The Machover ratings were also compared with ratings by chief ward attendants. There were no consistent significant agreements of the Machover with any of these clinical procedures or standards nor were there any significant agreements of the intuitive approach to DAP interpretation with any of these clinical rating scales. It was concluded that the DAP was not clinically valid as here measured.[14]

The most important phase of the study as it applies to the present investigation is the interpretation of the results made by the investigator. Blum indicated that the DAP was comparable to other clinical devices, even though, they likewise yielded low clinical agreement with this particular sample. Therefore, according to Blum's interpretation, the Machover is as valid as the other instruments used for descriptions of personality.

---

[13]Richard H. Blum, "The Validity of the Machover DAP Technique," Journal of Clinical Psychology, X (1954), 120-125.

[14]Ibid., p. 125.

Fisher and Fisher[15] made a study of some of Machover's specific assumptions as to the use of her test. The purpose of their study is stated in the following manner:

> The present study actually has the following goals: (1) To determine if one can objectively detect paranoid trends in figure drawings. (2) To determine how much agreement there is in evaluating figure drawing facial expression and figure drawing stance. (3) To determine if a group of raters with special training in figure drawing analysis show greater agreement with each other in evaluating facial expression and stance than do untrained raters.[16]

The investigators secured thirty-two figure drawings from adult patients who had been diagnosed as paranoid schizophrenics by a psychiatric staff and also by Rorschach and TAT criteria. They employed the services of two groups of raters. The investigators presented the following results:

> 1. The figure drawings of the thirty-two paranoid schizophrenics were evaluated by means of a detailed atomistic analysis and also by means of a total impressionistic analysis for the presence of paranoid trends; and it was found that the majority of these drawings did not fall under the category of paranoid in the terms of the criteria implicit in each method. Furthermore, it was found that the two separate methods of evaluating paranoid trends did not correlate.
> 2. The same thirty-two drawings were also evaluated with regards to figure: facial expression and stance. There was little agreement among various raters in their descriptions of these two figure characteristics.
> 3. Psychologists with training in figure drawing interpretation showed no greater agreement among themselves in their ratings of expression and stance than they did with psychologically untrained raters, such as psychiatrists and stenographers.

---

[15]Seymour Fisher and Rhoda Fisher, "Test of Certain Assumptions Regarding Figure Drawing Analysis," Journal of Abnormal and Social Psychology, XLV (1950), 727-732.

[16]Ibid., p. 727.

> 4. The total results suggest that it is precarious to accept most of the current assumptions regarding figure drawings analysis (particularly as expounded by Machover) without the confirmation of further research.[17]

The two studies presented represent the interest that has been exhibited to determine the usefulness of human figure drawings in personality analysis. Machover[18] has conducted a study, since the initial introduction of her technique, which is of importance to this research. She obtained drawings from children in the kindergarten through sixth-grade. The sample consisted of children who were of Jewish background and children of the Negro race. From an analysis of the drawings, when compared to white children, Machover made the following interpretation:

> The Negro child was pre-occupied with defining himself as a unit in the environment, and in stabilizing himself in it in the manner of a young child. He was concerned less with physical prowess, osdipal competition with a father, prestige factors in pleasing an adult. The question rises as to whether cultural deprivations and the disordered family constellation of the Negro home are chiefly responsible for this retardation in self-development and this casting about for position in the environment, rather than for techniques of growth or self-realization.[19]

From the interpretation of the study mentioned above, it appears that differences in personality development of Negro and white children are revealed through their drawings.

---

[17] Ibid., p. 732.

[18] Karen Machover, "Human Figure Drawings of Children," Journal of Projective Techniques, XVII (1953), 85-91.

[19] Ibid., p. 86.

It may be that these differences between white and Negro children are traceable to differences in cultural patterns of the two ethnic groups.

Human Figure Drawings

Emotional Adjustment.--Clinicians have begun to use figure drawings as a technique for determining emotional adjustment. The applicability of this technique has resulted because of a need for more valid means of evaluating the total personality quickly and accurately. A study which relates to the applicability of figure drawings for determining adjustment and one that offers insight into the reliability and validity of such techniques is the investigation made by Albee and Hamlin.[20] The purpose of this experiment was "limited to an attempt to determine whether global or insightful impressions by clinicians, judging from human figure drawings who are experienced with projective methods, have significant reliability and validity."[21]

The investigators selected ten cases from a file of 200 unselected cases of a mental hygiene clinic which represented a range from normality to extreme abnormality in terms of emotional adjustment. Three judges selected these

---

[20]George W. Albee and Roy M. Hamlin, "An Investigation of the Reliability and Validity of Judgments of Adjustment Inferred from Drawings," Journal of Clinical Psychology, V (1949), 389-391.

[21]Ibid., p. 390.

cases in determining the degrees of adjustment. "Each subject had, during his original testing, been asked to draw a picture of a man and a woman on a plain sheet of 8 1/2 by 11 paper. This variation on the test is in line with rather widespread clinical practice."[22] Separate drawings were made by the ten subjects selected and were presented in pairs to fifteen experienced clinicians. The clinicians were asked to select from pairs of drawings the individuals who were better adjusted.

Applying the Spearman-Brown formula, reliability for the entire group of judges was .977. Using case records as a criterion of the subjects whose drawings were being judged, validity of the clinicians' judgments by the rank-order correlation as to adjustment was .640. This correlation is significant at the .05 level of confidence for eight degrees of freedom. There was no significant correlation found between intelligence and the clinicians' judgments of emotional adjustment. Intelligence was determined by Wechsler-Bellevue Full Scale Intelligence Scale. This indicates that judgments were based on factors other than intelligence. Albee and Hamlin concluded that:

> From the results obtained, it appears that clinical psychologists can make reliable judgments of "global adjustment" from drawings-of-a-man-and-woman under conditions of the present experiment. The reliability coefficient obtained is sufficiently high to warrant rejection of the Null Hypothesis. . . . There is rather convincing evidence with regard to validity that the

---

[22] Ibid., p. 391.

> judgments of adjustment determined from drawings agrees with the adjustment of the subjects determined by the case records. Certainly the results are sufficiently promising to warrant continued use of the technique, and further investigation which would answer some of the questions related to the basis of this validity.[23]

Another study made by Graham[24] suggests the great degree of reliability in human figure drawings. Graham asked 23 "naive graduate students" to draw figure drawings. These drawings were scored by the Steinman Manual. A lecture was given to the students as to the characteristics of drawings that showed various negative aspects of an individual's personality. After the lecture, the students were asked to make a second drawing. The correlation between the paired sets of drawings (.71) suggests there is no difference between the drawings, even though, the students were aware of the kind of characteristics present in the drawings of well adjusted individuals.

Sidney Levy[25] has said, "I regard drawing analysis as so fruitful and economical a source of information about personality that I believe the practice of prefacing other more complicated techniques of personality assessment with the 'drawing-a-person' technique is a defensible clinical practice."

---

[23] Ibid.

[24] Stanley R. Graham, "A Study of Reliability in Human Figure Drawings," Journal of Projective Techniques, XX (1956), 385-386.

[25] Sidney Levy, "Figure Drawing as a Projective Test," Projective Psychology (New York: Alfred A. Knopf, 1950), p. 258.

He further states, "Any observable behavior is appropriate material for the clinical psychologist."[26] In connection with his latter statement, he makes mention of such a great psychologist as Muzafer Sherif who has demonstrated the importance of observable behavior in recent studies.

Murray Wexler and Jules D. Holzberg[27] have contributed to the literature of studies for the purpose of validating human form drawings in personality evaluation. Their study was designed "to determine if objective criteria could be established which would differentiate between human form drawings of a normal and a schizophrenic population. The results revealed that certain structural aspects of human form drawings were isolated which statistically differentiated between normals and schizophrenics."[28]

To reveal the diversity of studies conducted in the area of human figure drawings, De Martino's[29] study with mentally retarded males will be reviewed briefly. He compared drawings of 37 mentally retarded males who were known to be homosexuals with 37 drawings of mentally retarded males who were not known or suspected of being homosexuals. He found

---

[26]Ibid., p. 259.

[27]Murray Wexler and Jules D. Holzberg, "A Further Study of the Validity of Human Form Drawings in Personality Evaluation," Journal of Projective Techniques, XVI (1952), 249-251.

[28]Ibid., 249.

[29]Manfred F. De Martino, "Human Figure Drawings by Mentally Retarded Males," Journal of Clinical Psychology, X (1954), 241-244.

that unique characteristics were exhibited through the drawings by homosexuals. The result of this study suggests that human figure drawings are useful in personality diagnosis of the mentally retarded.

There have been studies conducted with the mentally retarded which suggest human figure drawings can serve a purpose in determining the pattern exhibited through their drawings. De Martino[30] conducted an investigation of human figure drawings by mentally retarded males. This study was made for the purpose of establishing validity as it relates to interpretation of drawings of non-homosexual and homosexual males of low intelligence. This study consisted of two parts, the first dealing with males of 63.4 I.Q. as determined by the Wechsler-Bellevue Forms I and II, the WISC, and the revised Stanford-Binet Forms L and M; and who were known and suspected to be homosexuals. Thirty-seven drawings from the above mentioned sample were compared with 37 drawings from a similar group of males who were not known or suspected homosexuals. The characteristics exhibited through the drawings by homosexuals were: (1) Most drew their own sex first. (2) High heels and eyelashes appeared significantly more in their drawings than in those by non-homosexuals.[31] Thus, personality patterns of mentally retarded males reveal

---

[30] Ibid., 244.

[31] Ibid.

themselves through human figure drawings. This is to say, personality characteristics reveal themselves in spite of low intelligence.

Projective Techniques and Reading Achievement

Have projective techniques been used in detecting reading achievement in children? Even though the literature is limited on this subject, the writer will present three studies in which projective techniques have been employed in the area of reading achievement.

Dr. Pauline G. Vorhaus[32] conducted a study of 309 reading disability cases. For seven years after a previous publication, "the investigator continued to test children suffering from marked reading disability."[33] "This extended investigation has strengthened the earlier belief that emotional factors play an important role in which mentally normal individuals suffer from extreme and long continued reading retardation."[34] In utilizing the Rorschach in her study, Dr. Vorhaus substantiated the above belief. It was found that four Rorschach configurations characterized the non-readers with whom she worked. These four configurations

---

[32]Pauline G. Vorhaus, "Rorschach Configurations Associated with Reading Disability," Journal of Projective Techniques, XVI (1952), 3-19.

[33]Ibid., 3.

[34]Ibid.

seem to . . . "indicate types of adaptations which these individuals make to environmental pressure."[35]

George Meyer[36] has shown in his study that Rorschach records of kindergarten children may be a valid prognostic test of reading achievement in the primary grades and may determine first-grade reading readiness. By using the Rorschach with achieving readers and retarded readers, he found eight differences in Rorschach variables which were significant and distinctive between achieving readers and retarded readers in the primary grades.

Another study made by Robert Bailey[37] has shown the usefulness of projective tests to predict academic reading success at the elementary school level. This study is based on the assumption that academic performance in reading represents an adjustment to an academic and social situation. To test his hypothesis, Bailey employed the Goodenough-Draw-A-Man test, House-Tree-Person test, and the Machover-Draw-A-Person test. He says:

---

[35] Ibid., 19.

[36] George Meyer, "Some Relationships Between Rorschach Scores in Kindergarten Reading in the Primary Grades," Journal of Projective Techniques, XVII (1953), 414-425.

[37] Robert Bain Bailey, A Study of Predicting Academic Success in Elementary School Reading from Projective Tests (a Ph.D. dissertation, University of Oklahoma, 1956).

> Presumably some relationship would exist between academic adjustment and scores on personality tests that assess adjustment status. If projective tests can be interpreted adequately to differentiate levels of intelligence or to identify sufficient maladjustment to impair intellectual functioning, they could be used to ascertain the potentially strong and weak readers in a class.[38]

This study involved second, fourth, and sixth-grade children's drawings of the three projective tests. There were trained and untrained judges who analyzed the drawings to select the readers and non-readers. Bailey concluded that the trained group of judges were able to make more correct selections of readers and non-readers than were the untrained judges. This study and the two cited above suggest the interest that has been shown in utilizing projective tests in determining and predicting reading achievement.

Race and Intelligence

Many studies have made an attempt to determine the relationship of race to intelligence as revealed by intelligence tests. These studies have attempted to show the following: (1) Negroes earn lower I.Q. scores on tests in the South than do whites; (2) Negroes are genetically or inheritedly inferior to whites in intelligence; (3) Negroes earn comparable scores to those earned by whites when equated as to social class or socio-economic status; (4) Negroes earn higher I.Q. points as social, cultural, and educational opportunities

---

[38] Ibid., p. 2.

increase; and (5) Negroes earn IQ scores as age increases, that reveal a greater difference in IQ points between Negro and white children in favor of the latter group.

Studies have revealed that Southern Negro children earn lower IQ scores on tests than do whites. Usually the tests employed are those that have been standardized on a white population, and these tests have a theoretical norm or average of 100. When these tests are employed with a Negro population, the norm obtained is less. Young and Pitts[39] made a study of 40 Southern Negro children in Georgia who were determined as representative of their culture by the judgment of the observers. The investigators found the mean WISC Full Scale IQ score to be 69.8. A follow-up study was made by Young and Bright.[40] These investigators used a larger sample of Southern Negro children in a study which revealed comparable results to that of the former study. The mean WISC Full Scale IQ score was found to be 67.74. The investigators stated, "We must question whether the WISC is a suitable test for the Southern Negro child."[41]

---

[39]Florence M. Young and Virginia A. Pitts, "The Performance of Congenital Syphilitics on the Wechsler Intelligence Scale for Children," Journal of Consulting Psychology, XV (1951), 239-242.

[40]Florence M. Young and H. H. Bright, "Results of Testing 81 Negro Rural Juveniles with the Wechsler Intelligence Scale for Children," Journal of Social Psychology, XXXIX (1954), 219-226.

[41]Ibid., 220.

Another study[42] was conducted which also revealed that Southern Negro children earn a lower average IQ score than the test norm. The Chicago Non-Verbal Examination was administered to 1140 East Tennessee Negro children living in three Eastern Tennessee counties. The ages ranged from six through fourteen years. The CNVE test results revealed: (1) At all age levels, these Negro children scored not less than the equivalent of two years below the respective age norms for this test. At ages 11 through 14, the disparity increased to three or more years. (2) At the six and seven year age levels, the test results lacked completely or largely in discriminability. (3) Limited discriminability between the 13 and 14 year levels is indicated. (4) Taking all age levels combined, no significant sex differences were found.[43]

Klineberg[44] as early as 1935 revealed in a study that there is an increase in the intelligence scores of Southern Negro migrants to New York with increasing length of residence in New York. Lee,[45] in 1951, tested Klineberg's hypothesis

---

[42]T. Ernest Newland and William C. Lawrence, "Chicago Non-Verbal Examination Results on an East Tennessee Negro Population," Journal of Clinical Psychology, IX (1953), 44-47.

[43]Ibid., 47.

[44]Otto Klineberg, Race Differences (New York: Harper and Brothers, 1935), Chap. IX.

[45]Everett S. Lee, "Negro Intelligence and Selective Migration: A Philadelphia Test of the Klineberg Hypothesis," American Sociological Review, XVI (1951), 227-233.

in Philadelphia and confirmed the idea that the longer a southern-born Negro lived in Philadelphia the greater increase in general intelligence as well as greater performance on most of the sub-tests of the Chicago Tests of Primary mental abilities.

Studies have revealed that Northern Negroes earn a higher IQ score than Southerners--both Negro and white. Klineberg's[46] study suggested this, and Ginsberg[47] reports this finding. However, some investigators have attempted to show that Negroes are genetically inferior to whites in intelligence. Shuey[48] has reported many studies revealing the differences in IQ points earned by Negroes and whites with the higher points earned by the latter. There is no doubt that Shuey did an excellent job in reviewing the literature, however, her explanations and interpretations of data did not support her theory of biological inferiority of the Negro. On this point, Dreger and Miller say:

> Aside from the infant studies cited previously, of preschool children in which comparisons are made between whites and Negroes are sparse. Shuey (1958) cites only nine altogether and only five reported in 1944 or after. On the whole, young Negro children score lower than whites. But the differences are very much less than in older groups; and in all of the reports in which average IQs are given, Negroes average well within the normal

---

[46]Klineberg, Race Differences, Chap. IX.

[47]Ginsberg, op. cit.

[48]Shuey, op. cit.

> IQ range for whites. Shuey offers several explanations of the discrepancy between the results for preschool and school children, including inadequate sampling of preschool, the relative invalidity and unreliability of tests at younger ages (although some of the lowest Binet standard errors are found in the late preschool years), more verbal and abstract tests in the school years; she also suggests that mental growth curves may not be the same for both races, and that IQs may be less affected by environment in the preschool years.[49]

Dreger and Miller further suggest that Shuey's explanations and interpretations of some of the studies reviewed presented false impressions. They state:

> Special comment will be made with respect to two investigations. Brown (1944) compared Minneapolis Kindergarten children on the Binet, Form L, discovering that the Negro mean of 100.8 was significantly lower than the white mean of 107.1. The white children in Occupational Classes VI and VII (Minnesota scale) averaged about the same as Negro children. Shuey takes exception to Brown's conclusion that at "nominally similar" socio-economic status Negro children are not inferior to whites. Her observation that Brown has small Ns in Levels VI and VII does not constitute an objection from a statistical standpoint.
>
> . . . . . .
>
> The other study to which special attention must be made is that of Anastasi and D'Angelo (Anastasi & D'Angelo, 1952: D'Angelo, 1950). Five year old children in mixed and unmixed neighborhoods in New York City were administered the Goodenough Draw-A-Man Test and studied for language development in spontaneous conversation recordings. IQs were 101.8 and 101.5 for Negro and white children, respectively. (Negro mean slightly higher.)[50] Language development appeared to be somewhat more advanced for white than for colored children. This important study is dismissed by Shuey as permitting no generalizations, because selection of subjects appeared to be biased. The crux of the matter is whether all

---

[49]Ralph M. Dreger and Kent S. Miller, "Comparative Studies of Negroes and Whites in the United States," Psychological Bulletin, LVII, No. 5 (September, 1960), 365.

[50]The parentheses are the investigator's.

> qualified subjects were utilized or only certain selected ones. Anastasi points out Shuey's error in misjudging the selection procedure, and provides a satisfactory answer to Shuey's objections. With due recognition of the limitations of the Goodenough as a test of intelligence we may yet regard Anastasi and D'Angelo's results as a challenge to nativist theories of intellectual differences between the races.[51]

McGurk[52] is another investigator who is interested in Negro-white comparisons, and is one who seems to favor nativity rather than environment. This investigator compared the performance of 213 Negro high school senior children and 213 white high school senior children on various cultural and non-cultural test items. He concluded that Negro children's performances were more comparable to whites on culturally weighted test items than on non-cultural items. Also, that differences between the two ethnic groups increased in favor of the white children as both groups increased in socio-economic status. It is implied from this study, regardless of an attempt to control various variables, the fact remains that there is a difference traceable to nativity. Dreger and Miller in reviewing McGurk's study make this observation:

> McGurk states that 25% of Negroes overlap whites, from which it might be concluded that only 25% of the specified Negro population have scores in common with the specified white population. Actually, analysis of

---

[51] Ibid.

[52] Frank C. J. McGurk, "On White and Negro Test Performances and Socio-Economic Factors," Journal of Abnormal and Social Psychology, XLVIII (1953), 448-450.

> McGurk's data shows that 91% of the Negroes have scores in common with the whites. Inasmuch as McGurk has addressed himself to the lay public, he should make abundantly clear that "over-lap" is used in the technical sense of exceeding (or for an upper distribution, falling below) the mean or median of another group.[53]

When equating children as to socio-economic status, whites and Negroes earn comparable scores on intelligence tests. Again, attention will be called to Fred Brown's study.[54] This study conducted in the Minneapolis Public Schools, revealed that in spite of the mean white IQ being 107.06 as compared with a mean Negro IQ of 100.78; "a comparison of intelligence at various occupational levels reveals that Negro males are not inferior in tested intelligence to white subjects in Group III (clerical, retail business, and skilled trades.)"[55] He further states, "Negro females score lower than white females at all occupational levels except VI and VII."[56]

Some investigators have shown that social and cultural opportunities play a role in the low IQ scores earned by Negroes. Lee[57] has shown that the longer Negro children, who are migrants from the South, live in Philadelphia the greater

---

[53]Dreger and Miller, op. cit., p. 366.

[54]Fred Brown, "An Experimental and Critical Study of the Intelligence of Negro and White Kindergarten Children," Journal of Genetics Psychology, LXV (1944), 161-175.

[55]Ibid., 175.

[56]Ibid.

[57]Lee, op. cit.

the increase in IQ points. Lee was testing Klineberg's theory, and his results were comparable with those of Klineberg's[58] study on Southern Negro migrants to New York City. It should not be difficult to understand the role of opportunity and culture in the performance on intelligence tests regardless of the race of the testee. Sherman and Key[59] made a study of white children living in the Hollows district of Virginia in which environmental impoverishment was great. They found the IQ norm was below that of the children on whom the test was standardized. Wheeler's study[60] of more than 3000 East Tennessee mountain white children revealed results similar to that of Sherman and Key's study. In both studies, an inverse relationship of chronological age and IQ scores existed. In later years, Wheeler conducted another study in the same community. At this time, he found that the median IQ score had increased. The basic assumption was that in ten years, general improvement in educational and environmental conditions existed; therefore, the environmental factor was stressed as being important.

Another study which reveals cultural influences upon

---

[58]Klineberg, Race Differences.

[59]M. Sherman and C. B. Key, "The Intelligence of Isolated Mountain Children," Child Development, III (1932), 279-290.

[60]L. R. Wheeler, "A Comparative Study of the Intelligence of East Tennessee Mountain Children," Journal of Educational Psychology, XXXIII (1942), 321-334.

test performance is that conducted by Fowler.[61] A comparison was made of intelligence test performance on six tests. The sample consisted of 355 ten year old non-ethnic white, Polish, and Negro children living in Detroit and Hamtramck, Michigan. His general conclusions were: (1) Negro pupils scored lower than white children; (2) Polish children usually scored lower than the white non-ethnic children; (3) white non-ethnic children excelled Negro and Polish children in mean performance. This study highly suggests that children who are representative of the standardization population tend to do better on these tests than children who were not included in the sample. In the above study, Polish as well as Negro children did not perform as well as non-ethnic white children.

The Utilization of Clinicians as Judges

The employment of clinicians as judges may be questioned. A number of studies reviewed in previous pages employed clinicians as judges. Bieliauska[62] employed three judges in his study, Smykal and Thorne[63] employed judges in making analyses of H-T-P drawings. Hammer[64] employed clinical judges

---

[61] W. L. Fowler, "A Comparative Analysis of Pupil Performance on Conventional and Culture-Controlled Mental Tests," Dissertation, University of Michigan, Pub. #18 (1956), 603, 2106. (Microfilm.)

[62] Bieliauska, op. cit.

[63] Smykal, and Thorne, op. cit.

[64] Hammer, op. cit.

in making a comparison of Negro and white children's drawings in order to determine the degree of frustration-aggression of the two racial groups.

Albee and Hamlin[65] made an investigation to determine the reliability and validity of clinicians' ability to make global or insightful impressions from human figure drawings. They concluded, "There is convincing evidence with regard to validity that the judgment of adjustment determined from drawings agrees with the adjustment of the subjects determined by the case records."[66]

A study, which deals directly with the subject at hand, was made by Roesapple and Phelan.[67] This study was made to determine if psychologists possess skill in making analyses of projective tests in order to make accurate diagnosis. In their study, twenty clinicians were asked to match autobiographies, TAT Protocals, Rorschach Protocals and four objective type batteries such as Thurstone's Primary Mental Abilities, and Kuder Preference Record. Raesapple and Phelan concluded:

> Judges as a group performed at a higher level than could have been expected by chance. Individual judges

---

[65]Albee and Hamlin, op. cit.

[66]Ibid., p. 390.

[67]James Quintar Roesapple and Joseph G. Phelan, "The Skills of Clinicians in Analysis of Projective Tests," Journal of Clinical Psychology, X (1954), 307-320.

> performed considerably above chance. Performance of judges was differential, that is, judges who were superior in matching with one test were superior in most. A skill in the analysis of such materials exists.[68]

Thus, the evidence of previous research indicates that the employment of trained psychologists in making accurate diagnoses from human figure drawings and other projective devices is a reliable and valid procedure.

Summary

A survey of studies has been presented in the following order: (1) House-Tree-Person technique; (2) Machover-Draw-A-Person test; (3) emotional adjustment as determined by human figure drawings; (4) reading achievement as determined by projective techniques; (5) race and intelligence as they relate to each other revealed through studies presented; (6) experimental techniques utilizing judges to analyze human figure drawings.

Studies presented involving the H-T-P technique revealed its validity as a personality and intelligence diagnostic tool. A review was given of Buck's validation procedure of the H-T-P. Studies revealed not only its usefulness for appraising personality and intelligence, but also suggested its usefulness in diagnosing the personality and intelligence of Negroes. Hammer's studies were reviewed because they revealed the role culture played when this projective technique was applied to a Virginia Negro population.

---

[68] Ibid., 320.

The Machover Draw-A-Person test has shown usefulness in the area of personality just as the Goodenough has shown usefulness in the area of intelligence. Machover conducted studies which substantiated this statement. Blum's study on the validity of the DAP revealed that it did not have diagnostic and predictive value in appraising personality with the sample used by him. However, the instrument was as valuable as any other tool of personality. Fisher and Fisher's study revealed that the Machover was of little value when applying some of the assumptions made by Machover. It must be kept in mind that the Fishers were testing only a few of Machover's assumptions.

Six studies were reviewed in an effort to determine the reliability and validity of human figure drawings in diagnosing and predicting emotional adjustment. All of the literature reviewed suggests high reliability and validity in predicting emotional adjustment of mentally normal as well as mentally retarded individuals.

The literature was explored in an effort to determine if projective techniques were of diagnostic and predictive value in appraising reading achievement. Even though the literature is limited on this subject, three studies were cited. The three studies showed that projective techniques have diagnostic and predictive value in reading achievement on the elementary school level. Two of the studies employed the Rorschach, and one study employed the H-T-P, Machover

DAP, and the Goodenough Draw-A-Man tests.

The literature on race and intelligence revealed there are many viewpoints as to the relationship of race and intelligence. Some studies made attempts to show that race genetically and biologically determines intelligence. However, most authorities favored the environmental and cultural theories. This is to say that environmental, cultural, and social opportunities effect IQ scores regardless of race.

In an effort to reveal the utilization of psychologists as judges, literature was reviewed which suggested the widespread practice of this technique in conducting studies. Many of the studies reviewed in this chapter employed clinical psychologists as judges. One study was reviewed which tested the reliability and validity in the skill of psychologists to make accurate diagnoses from global analyses of projective tests. It was concluded from the study that psychologists employed as judges have a high degree of reliability and validity with projective tests. Thus, the employment of judges in the present study was within keeping with many studies of this nature.

CHAPTER III

EXPERIMENTAL DESIGN

The purpose of this chapter is to present the experimental design and methods employed in collecting the data of this study. The major aspects of this chapter include: (1) sample, (2) method of collecting the data, (3) method of determining mental ability, (4) method of determining reading achievement, (5) method of determining emotional adjustment, (6) method of determining race, (7) testing procedure, (8) hypotheses, (9) experimental procedure, and (10) statistical analysis.

Sample

The original sample for this investigation included first-grade children from five public schools in Oklahoma City, Oklahoma. The group included Negro, white, and Indian children. There were 303 persons in the original test survey. Of these, 240 were selected for use in the experiment. The 63 subjects were eliminated because of complicating age, grade, academic, or race factors. The drawings of first-grade repeaters and Indian children were eliminated. The

deletion of these children's records made it possible to control experimentally the extraneous and non-pertinent factors of maturity and academic experience.

The five schools were selected on the basis of geographical districts that are basically socio-economic areas of the city. The socio-economic districts were determined by type of houses in the community, sources of income, occupation, place of employment, and position held by the parents. One school, located in an above average district, consisted of all white children. A second school consisted of all Negro children located in a socio-economic district comparable as nearly as possible to that of the first school. A third school was selected to represent one of the lowest socio-economic areas which consisted of white and Indian children. The fourth school represented a low socio-economic area comparable as nearly as possible to the third school but consisted of all Negro children. The fifth school was an integrated school, centrally located, which consisted of approximately a fifty-fifty ratio of Negro and white children. This fifth school was located in a below middle socio-economic district, but cannot be considered as a lower-lower class district. (See Table 1.)

The socio-economic status of the children was determined by the Index of Status Characteristics (I.S.C.),

TABLE 1.--Source of sample drawn from the socio-economic districts of Oklahoma City Public Schools--1960

School	Classification		Number Tested	Number Eliminated	Number Used
1	W	AA	44	1	43
2	N	AA	116	24	92
3	IW	LC	35	8	27
4	N	LC	15	0	15
5	IA	UL	93	30	63
Totals			303	63	240

Legend:

Race	Socio-economic Classification
I-Indian	AA-Above Average Socio-economic
N-Negro	M-Middle Socio-economic
W-White	L-Lower Socio-economic

a Measure of Social Status developed by Warner, Meeker, and Eells.[1]

From the 240 drawings, a random selection was made for four groups to be judged. All groups--one, two, three, and four consisted of an equal number of white and Negro children's drawings. Information necessary to determine socio-economic status of the children was obtained by trained visiting counselors of the Department of Pupil Services of the Oklahoma City Public Schools.

[1]W. Lloyd Warner, Marchia Meeker, and Kenneth Eells, Social Class in America (Chicago: Science Research Association, Inc., 1949).

Mental Ability.--Each child's mental ability was determined by the Goodenough scoring scale. Teacher ratings of each child's mental ability were obtained. The Goodenough Draw-A-Man Test was used because (1) it is a performance test, (2) of its high reliability of .937 for first-grade children,[2] (3) of its high positive correlation with the Stanford-Binet Intelligence Scale, and (4) of its inclusive sample of many ethnic groups in its standardization.

Each child was rated by his teacher as bright, average, or dull. For the purpose of this study, teachers' ratings were only used to provide the researcher with more than one estimate of the child's mental ability.

The mean I.Q. score and the standard deviation of first-graders of each school were determined from the scores earned on the Goodenough test. Three "t" tests were calculated to determine the differences of the I.Q. means of Negro and white children taken from comparable socio-economic districts. A "t" test was also calculated to determine the differences of mean I.Q. scores between white and Negro children of the first-grade in the integrated school. This information is presented in Table 1. From the calculations made, one can observe that there was no significant difference in the mean I.Q. scores between Negro and white children of the above average districts and in the integrated school. (See

---

[2]Florence Goodenough, Measurement of Intelligence by Drawings (Chicago: World Book Co., 1926), p. 48.

Tables 2 and 3.) However, there was a significant difference found in the mean I.Q. scores of the two schools in the lower socio-economic districts. In spite of the two districts being the two lowest socio-economic areas in the city, the Negro district is a socially isolated area and many factors entered into the cultural patterns of this area that were not measurable. Therefore, these two low socio-economic groups were not used in the grouping for ability judgments. From the other districts, forty (40) children--twenty Negro and twenty white children were selected randomly from groups of drawings equated as to socio-economic class and as to ability. An attempt was made to have an equal number of children in the upper, middle, and lower socio-economic classes as well as an equal number of "Bright" and "Dull" children. All children who earned a score on the Goodenough that placed them within a ± 1 standard deviation were not used for this part of the experiment because it was necessary that children of average ability not be used for ability judgments.

Reading Achievement.--A reading achievement score was obtained from the Metropolitan Reading Achievement Test for each child. Teacher ratings were utilized for the purpose of validating the classification of each child. The children's reading achievement scores ranged from levels of non-readers to above third-grade. Forty (40) children--twenty Negro and twenty white children--were randomly

TABLE 2.--Means and differences of intelligence quotients of White and Negro first-grade children in an integrated school, Oklahoma City, Oklahoma

Race	Mean	Standard Deviation	"t"	Number	Probability
Negro	91.06	17.48	.11	32	Non-Significant
White	92.56	15.92		31	

TABLE 3.--Means and differences of intelligence quotients of Negro and White first-grade children in two racially homologous schools of the above average socio-economic districts, Oklahoma City, Oklahoma

Race	Mean	Standard Deviation	"t"	Number	Probability
Negro	104.48	15.26	.846	52	Non-Significant
White	107.23	16.22		44	

TABLE 4.--Means and differences of intelligence quotients of Negro and White first-grade children in two racially homologous schools of the low socio-economic districts, Oklahoma City, Oklahoma

Race	Mean	Standard Deviation	"t"	Number	Probability
Negro	83.2	16.2	2.71	15	Significant .01 level of confidence
White	97.62			16	

selected from groups of drawings. There were ten white good readers, ten Negro good readers, ten white poor readers, and ten Negro poor readers. Good readers were those children who were reading above second-grade. Poor readers were those children reading on pre-primer level and below (non-readers).

Emotional Adjustment.--Emotional adjustment of the children included in this study was determined by (1) trained visiting counselors who interviewed each teacher to obtain a description of each child's behavior in the classroom, (2) psychological summaries of those children who had been referred to a child guidance clinic, and (3) an analysis of the Goodenough-Draw-A-Man test was made in order to serve as another method of determining the emotional adjustment level of the children. Utility of the Goodenough in the above mentioned manner has been suggested by Johnson, Ellerd, and Lahay.[3]

Forty (40) children--twenty Negro and twenty white children--were selected randomly from groups of drawings equated as to "Well" and "Poorly" adjusted. There were ten white poorly adjusted, ten Negro poorly adjusted, ten white well adjusted, and ten Negro well adjusted children's drawings used in this sub-sample.

Three trained judges participated in this study. They

---

[3]A. P. Johnson, A. A. Ellerd, and T. Lahay, "The Goodenough Test as an Aid to Interpretation of Children's School Behavior," American Journal of Mental Deficiency, LIV (1950), 516-520.

were judges trained in the field of psychology. One judge is a clinical psychologist at a state mental hospital and holds a Ph.D. degree. Another is a supervisor of special classes for elementary children on the campus of a state university, instructor of college students, and has completed all course work for the Ph.D. degree. The third judge has completed all course work for the Ph.D. degree and is serving as a psychologist for a special education department at a state university. The judges were selected on the basis of training and availability.

Testing Procedure

The stimulus materials for pupils' drawings consisted of white paper 8 1/2 x 11 and a number 2 pencil with an eraser. Instructions were given orally for each drawing. The directions for the test were given as follows:

First Drawing

1. Each child was given a clean sheet of paper.
2. The children were asked to write his name on the back of the paper.
3. The Examiner said, "Draw a house the best you can."

Second Drawing

1. Each child was given a clean sheet of paper.
2. The children were asked to write their names on the back of the paper.
3. The Examiner said, "Draw a tree--any kind--the best you can."

Third Drawing

1. Each child was given a clean sheet of paper.
2. The children were asked to write their names on the back of the paper.
3. The Examiner said, "Draw a picture of a person, any person you want. It does not matter who it is. Be sure to draw the whole person."
4. "Mark on the back of your paper No. 1."

Fourth Drawing

1. Each child was given a clean sheet of paper.
2. The children were asked to write their names on the back of the paper.
3. The Examiner said, "Draw a picture of a person, any person you want. It does not matter who it is. Be sure to draw the whole person. If you drew a girl, then draw a boy this time. If you drew a boy then draw a girl. If you drew a woman, draw a man this time. If you drew a man, then draw a woman."
4. "Mark on the back No. 2."

Hypotheses

One major concern in this study is to determine whether or not the correct responses rendered by the judges, on mental ability; reading achievement; emotional adjustment; and race, will be significantly greater than by chance. A second major concern is to determine whether or not agreement will exist among judges in their correct responses greater than by chance on ability, reading achievement, emotional adjustment, and race. A third major concern is to determine whether or not judges will be able to determine mental

ability, reading achievement, emotional adjustment, and race of Negro children as well as they will be able to determine mental ability, reading achievement, emotional adjustment, and race of white children based on the children's drawings. These major questions raised have been formulated into hypotheses for testing statistical significance. The hypotheses have been organized as follows:

Group I — The hypotheses consist of validation of judges' estimates.

Group II — The hypotheses consider the amount of agreement among judges in their correct responses.

Group III — The hypotheses deal with comparison of judges' responses made on Negro and white children's drawings.

The major groups have been formulated into three major hypotheses. These hypotheses have been divided into minor hypotheses. They have been stated in the following manner:

I. Judges will make correct responses which will significantly exceed chance in determining mental ability, reading achievement, emotional adjustment, and race of the children whose drawings they judge.

   A. Judges' correct responses to the children's drawings will significantly exceed chance in determining mental ability.

      1. Judgments will significantly exceed chance in determining brightness.

      2. Judgments will significantly exceed chance in determining dullness.

      3. Judgments will significantly exceed chance in determining ability of Negro children.

4. Judgments will significantly exceed chance in determining ability of white children.

B. Judges' correct responses to the children's drawings will significantly exceed chance in determining reading achievement.

1. Judgments will significantly exceed chance in determining reading achievement of good readers.

2. Judgments will significantly exceed chance in determining reading achievement of poor readers.

3. Judgments will significantly exceed chance in determining reading achievement of Negro children.

4. Judgments will significantly exceed chance in determining reading achievement of white children.

C. Judges' correct responses to the children's drawings will significantly exceed chance in estimating emotional adjustment.

1. Judgments will significantly exceed chance in determining emotional adjustment of well adjusted children.

2. Judgments will significantly exceed chance in determining emotional adjustments of poorly adjusted children.

3. Judgments will significantly exceed chance in determining emotional adjustment of Negro children.

4. Judgments will significantly exceed chance in determining emotional adjustment of white children.

D. Judges' correct responses to the children's drawings will significantly exceed chance in determining the race of the children.

II. Judges will agree in their correct responses in determining mental ability reading achievement, emotional adjustment, and race of the children whose drawings they judge.

A. Agreement among judges in their responses to children's drawings will exist beyond chance in determining mental ability.

1. Agreement among judges will exist in their responses to bright children's drawings when compared with their responses to dull children's drawings in determining mental ability.

2. Agreement among judges will exist in their responses to good reader's drawings when compared with their responses to poor reader's drawings in determining reading achievement.

3. Agreement among judges will exist in their responses to well adjusted children's drawings when compared with their correct responses to poorly adjusted children's drawings in determining emotional adjustment.

4. Agreement among judges will exist in their correct responses to children's drawings in determining the race of the children whose drawings they judge.

III. Judges will agree significantly in determining mental ability, reading achievement, emotional adjustment, and race of Negro children when compared to their correct estimates of mental ability, reading achievement, emotional adjustment, and race of white children based on the children's drawings.

A. Judges will agree significantly in determining mental ability of Negro children when compared to their correct estimates of mental ability of white children.

1. Judges will agree significantly in determining mental ability of bright Negro children when compared to their correct estimates of bright white children.

2. Judges will agree significantly in determining mental ability of dull Negro children when compared to their correct estimates of dull white children.

B. Judges will agree significantly in determining reading achievement of Negro children when compared to their correct estimates of reading achievement of white children.

1. Judges will agree significantly in estimating reading achievement of Negro good readers when compared to their correct estimates of white good readers.

2. Judges will agree significantly in estimating reading achievement of Negro poor readers when compared to their correct estimates of white poor readers.

C. Judges will agree significantly in determining the emotional adjustment of Negro children when compared to their correct estimates of white children.

1. Judges will agree significantly in determining the emotional adjustment of well adjusted Negro children when compared to their correct estimates of well adjusted white children.

2. Judges will agree significantly in determining the emotional adjustment of poorly adjusted Negro children when compared to their correct estimates of poorly adjusted white children.

D. Judges will agree significantly in their correct responses in selecting the race of the children whose drawings they judge.

Experimental Procedure

The experiment was divided into four groups of judgments. They were mental ability, reading achievement, emotional adjustment, and judgments requiring decisions to be made about the race of the children whose drawings were being judged.

Preceding the experiment, the procedure was explained to the judges. Terms that would be used in the instructions such as bright, dull, emotional adjustment, and poor and good readers as used in this study were defined. This kind of explanation was necessary so that all judges would be able to use the same meaning of the dichotomized classifications of the variables being used in this study. However, race was not mentioned to the judges until they were faced with

the task of determining the race of the children. This step in the experiment was important because it was necessary that the judges not know they were judging drawings made by two different racial groups.

The drawings were placed on 30 inch conference tables in a sound shielded room free from distractions. The criteria for the classifications were orally presented. Each judge was given an opportunity for questions relating to the procedure. The judges were permitted to move freely in the room and to make their classifications without a time limit. The experimenter was seated at a remote point and offered no assistance or comment while judges were making their selections. The circumstances were designed to assure that all judges would make effective clinical classifications without external clues to bias their choices. Once the experiment was begun no judge could confer with other judges or ask questions of the experimenter. The sequence of sorting was varied so that order of classification would not prejudice the findings.

Drawings were not identifiable to the judges by race or sex and all conclusions were necessarily based upon the details of the figures.

The evaluations for mental ability were made first. Each judge required approximately forty-five minutes to make estimates on the forty drawings. Judges were calm and appeared to enjoy the task. There was no evidence of fatigue by the judges.

Group I--Ability

The stimulus materials for the judges consisted of a rating sheet with an equal number of blank spaces for judgments of bright and dull to correspond with the number of bright and dull children whose drawings were being judged. (See Appendix II.) The instructions were given to the judges orally. They were given as follows:

"Here are six tables of drawings made by first-grade children. These drawings were made by forty children of whom twenty were bright and twenty were dull. Each child made four drawings--house, tree, Person I, and Person II (the opposite sex of Person I). On the front of each drawing a number has been assigned to each child for the purpose of identification in tabulating the results. These numbers have no other significance. Be sure that your number corresponds with the drawings being judged."

Group II--Reading

The stimulus materials for the judges consisted of a rating sheet with an equal number of blank spaces for judgements of "Good" and "Poor" to correspond with the number of good and poor readers whose drawings were being judged. (See Appendix II.) The instructions were given to the judges orally. They were given as follows:

"Here are six tables of drawings made by first-grade children. These drawings were made by forty children of whom

twenty were "good" readers and twenty were poor readers. Each child made four drawings--house, tree, Person I, and Person II (the opposite sex of Person I). On the front of each drawing a number has been assigned to each child for the purpose of identification when tabulating the results. These numbers have no other significance. Be sure that your number corresponds with the drawings being judged.

Group III--Emotional Adjustment

The stimulus materials for the judges consisted of a rating sheet with an equal number of blank spaces for judgements of "Well adjusted" and "Poorly adjusted" to correspond with the number of well adjusted and poorly adjusted children whose drawings were being judged. (See Appendix II.) The instructions were given to the judges orally. They were given as follows:

"Here are six tables of drawings made by first-grade children. These drawings were made by forty children of whom twenty were well adjusted and twenty were poorly adjusted. Each child made four drawings--house, tree, Person I, and Person II (the opposite sex of Person I). On the front of each drawing a number has been assigned to each child for the purpose of identification when tabulating the results. These numbers have no other significance. Be sure that your number corresponds with the drawings being judged.

Group IV--Race

The stimulus materials for the judges consisted of a rating sheet with an equal number of blank spaces for judgments of "White" and "Negro" to correspond with the number of white and Negro children whose drawings were being judged. (See Appendix II.) The instructions were given to the judges orally. They were given as follows:

"Here are six tables of drawings made by first-grade children. These drawings were made by forty children of whom twenty were white and twenty were Negro. Each child made four drawings--house, tree, Person I, and Person II (the opposite sex of Person I). On the front of each drawing a number has been assigned to each child for the purpose of identification for tabulating the results. These numbers have no other significance. Be sure that your number corresponds with the drawings being judged.

Statistical Analysis

The purpose of this research is to examine the applicability of the projective devices and to determine the extent of which the factors related to race tend to influence the evaluations of trained judges. Since it is the task of the judges to categorize the racial sub-groups into appropriate dichotomized levels, it is necessary to test statistically their abilities to make accurate identification generally and to observe specifically if their judgements were

in part or wholly the consequence of chance factors. In this, as in all dichotomizing treatments, the probability of making a proper association is equal to N/2 or .50. Since either classification is equally likely and materially exclusive, the differences between observation and expectancy divided by the square root of the product of N, the number of cases, and pq, the expected proportion of N/2. The assumption here is that as N increases the deviation from expectancy assumes greater significance. It is recognized that an N of 40 may give a biased estimate for the population parameter, but the nature of the sampling technique makes it reasonable to assume that these data are representative for first-graders of Oklahoma City children who live in the socio-economic groupings studied.

Chi-square was employed as an evaluation of independence of judges and attributes. Specifically, the purpose of the chi-square treatment was to determine if a relationship existed between the four qualities being judged and the persons classifying the drawings. The expected proportions would occur as a ratio of the number of classifications being made.

Summary

The original sample consisted of 303 first-grade Negro and white children of the Oklahoma City Public Schools. Of the 303 children tested, 240 participated in the experiment. However, some of the subject's drawings were

utilized in more than one of the sub-samples of the experimental groups. The subjects' drawings were selected for the sub-samples in the following manner:

1. Mental ability was determined by the Goodenough Draw-A-Man test. To be certain that groups were equal in mental ability, means, standard deviations, and "t" tests were calculated. Due to differences found in IQ scores between Negro and white children of the lower socio-economic schools, these two schools were not included in the sub-sample for mental ability.

2. Reading achievement was determined by the Metropolitan Reading Achievement Test scores.

3. Emotional adjustment was determined by the Goodenough Draw-A-Man test and the visiting counselor's evaluation.

4. Race was determined by the school attended and, when necessary, by data available on the permanent records. An effort was made to include an equal number of Negro and white children for the above average, middle, and lower socio-economic groups. Socio-economic status was determined by the geographical economic school districts and from data gathered by visiting counselors which provided for obtaining an Index of Status Characteristics. This was necessary so that judges would be making selections of race rather than selections based on socio-economic status.

A statement was made concerning the qualifications of the judges. All judges were trained in clinical psychology and special education.

The testing procedure was described by which the four drawings were obtained from the subjects. Directions for testing were stated in simple sentences. The same was true for the directions to judges.

Three groups of hypotheses were stated. The first group was concerned with the accuracy and validity of judges' estimates while the second group was concerned with the amount of agreement among judges in their correct responses. The third group of hypotheses dealt with comparisons of judges' responses made on Negro and white children's drawings. Statistical procedures were designated to test the hypotheses. These hypotheses have been accepted when differences obtained or exceeded .05 level of confidence.

CHAPTER IV

ACCURACY AND VALIDITY OF JUDGES' RESPONSES

Accuracy of Judges' Responses

Of the 120 evaluations made by the judges, 98 were correct responses which is an accuracy percentage of 81.66. The accuracy rating for Negro subjects, 81.66, was equal to the accuracy rating of 81.66 for whites. These data which are presented in Table 5 suggest that identification of intelligence of Negro subjects on the H-T-P and DAP is the same as for whites. In fact, under the conditions of this experiment, a comparable degree of accuracy about Negroes was noted. It is recognized that certain clinical inaccuracies might occur under circumstances similar to this experiment, but Negro and white productions were not differentially influenced.

Validity of Judges' Responses

What was the validity of judgments rendered by the trained judges in their analyses of the children's drawings? This chapter has been devoted to an evaluation of some assumptions about the degree of confidence that may be placed

TABLE 5.--Summary of judges' agreement on H-T-P and DAP results from forty (40) first-grade elementary school children of Oklahoma City, Oklahoma, based on evaluations of mental ability

Race and Ability Classifications	Judges I		Judges II		Judges III		Totals		Grand Totals		Percentages of Correct Responses
	Right	Wrong	Right	Wrong	Right	Wrong	Right	Wrong	Right	Wrong	
Caucasian:											
Bright	8	2	7	3	9	1	24	6			
Dull	9	1	7	3	9	1	25	5			
									49	11	81.66
Negro:											
Bright	9	1	8	2	8	2	25	5			
Dull	8	2	8	2	8	2	24	6			
									49	11	81.66
Totals	34	6	30	10	34	6	98	22	98	22	81.66

in the judges' estimates from H-T-P and DAP drawings of mental ability, emotional adjustment, reading achievement, and race. The assumptions concerning the validity of judges' responses are those that have been formulated into testable hypotheses presented in the previous chapter.

Validity of Judges' Responses on Mental Ability

The significant Z scores indicate that the judges' responses in their estimates of ability were valid. As summarized in Table 6, it is indicated that the hypothesis, judges' correct responses to the children's drawings will significantly exceed chance in determining mental ability, was accepted. A probability of .01 indicates rejection of the hypothesis that judgments were made by chance, and further suggests that the judges employed systematic methods in the identification of mental ability of the experimental group.

Table 7 has been summarized to determine if the judges were able to select better than chance bright children's drawings. In every case, the significance of the judgments exceeded chance by better than the .01 point (level of confidence). It is indicated that the hypothesis, judgments will significantly exceed chance in determining brightness, was accepted. This further suggests that the judges employed systematic methods in the identification of bright children of the experimental group.

TABLE 6.--Significance of judges' estimates of Negro and white children's mental ability based on their drawings

Judges	R	W	Probability Differences	Standard Error Differences	Z	P
I	34	6	14	3.16	4.43	.01
II	30	10	10	3.16	3.16	.01
III	34	6	14	3.16	4.43	.01

TABLE 7.--Significance of judges' estimates of bright Negro and white children's mental ability based on their drawings

Judges	R	W	Probability Differences	Standard Error Differences	Z	P
I	17	3	7	2.24	3.13	.01
II	15	5	5	2.24	2.23	.01
III	17	3	7	2.24	3.13	.01

The significant Z scores indicate that the judges' responses in their estimates of dull children were valid. As summarized in Table 8, it is indicated that the hypothesis, judgments will significantly exceed chance in determining dullness, was accepted. A probability of .01 indicates rejection of the hypothesis that judgments were made by chance, and further suggests that the judges employed systematic methods in the identification of the dull children of the experimental group. It may be observed that the

TABLE 8.--Significance of judges' estimates on dull Negro and white children's mental ability based on their drawings

Judges	R	W	Probability Differences	Standard Error Differences	Z	P
I	17	3	7	2.24	3.13	.01
II	15	5	5	2.24	2.23	.01
III	17	3	7	2.24	3.13	.01

judges were as accurate in selecting dull children as they were in selecting bright children.

Table 9 has been summarized to determine if the judges were able to estimate better than chance the mental ability of Negro children based on their drawings. In every case, judgments exceeded chance by better than the .01 point (level of confidence). It is indicated that the hypothesis, judgments will significantly exceed chance in determining mental ability of Negro children, was accepted. The result further suggests that the judges employed systematic methods in the estimation of the ability of Negro children included in the experimental group.

Table 10 has been summarized to determine if the judges were able to estimate better than chance the mental ability of white children based on their drawings. In every case, the significance of the judgments exceeded chance by better than the .01 point (level of confidence). It is

TABLE 9.--Significance of judges' estimates of Negro children's mental ability based on their drawings

Judges	R	W	Probability Differences	Standard Error Differences	Z	P
I	17	3	7	2.24	3.13	.01
II	16	4	6	2.24	2.68	.01
III	16	4	6	2.24	2.68	.01

TABLE 10.--Significance of judges' estimates of white children's mental ability based on their drawings

Judges	R	W	Probability Differences	Standard Error Differences	Z	P
I	17	3	7	2.24	3.13	.01
II	14	6	4	2.24	1.79	.01
III	18	2	8	2.24	3.57	.01

indicated that the hypothesis, judgments will significantly exceed chance in determining mental ability of white children, was accepted. The result further suggests that the judges employed systematic methods in the estimations of the ability of white children included in the experimental group. It may be observed that the judges were able to estimate the mental ability of both racial groups better than by chance.

Reading Achievement

Evaluations of reading achievement followed those for mental ability. However, drawings which included a few of those presented previously were rearranged with different children's drawings added. All drawings were re-examined to be certain that there were no identifying marks. Two of the judges completed the task in twenty minutes while the other judge required thirty minutes for making estimates on the forty sets of drawings ($H\text{-}T\text{-}P_1P_2$). There was no evidence of fatigue by the judges.

Of the 120 evaluations made by the judges, 98 were correct responses which is an accuracy percentage of 81.66. This is consistent with accuracy ratings for mental ability. The accuracy rating for Negro subjects, 75 per cent, was lower than that for whites, 88.33 per cent. The accuracy ratings of judges for Negroes were the same as their ratings for white children on mental ability, but when evaluating reading achievement, accuracy was 12.6 percentage points lower than that for white children. These data which are presented in Table 11 suggest that identification of reading achievement of Negro subjects on H-T-P and DAP is less accurate than for whites.

Validity of Judges' Responses on Reading Achievement

In Table 12, the significant Z scores indicate that the judges' responses in their estimates of reading achievement were valid. A probability of .01 for each of the judges indicates

TABLE 11.--Summary of judges' responses on H-T-P and DAP drawings for forty (40) first-grade elementary school children of Oklahoma City, Oklahoma, based on evaluations of reading achievement

Race and Classification	Judges I		Judges II		Judges III		Totals		Grand Totals		Percentages of Correct Responses
	Right	Wrong	Right	Wrong	Right	Wrong	Right	Wrong	Right	Wrong	
Caucasian:											
Good Readers	8	2	9	1	8	2	25	5			
Poor Readers	8	2	10	0	10	0	28	2	53	7	88.33
Negro:											
Good Readers	7	3	8	2	9	1	24	6			
Poor Readers	7	3	7	3	7	3	21	9	45	15	75.00
Totals	30	10	34	6	34	6	98	22	98	22	81.66

TABLE 12.--Significance of judges' estimates of Negro and white children's reading achievement based on their drawings

Judges	R	W	Probability Differences	Standard Error Differences	Z	P
I	30	10	10	3.16	3.16	.01
II	34	6	14	3.16	4.43	.01
III	34	6	14	3.16	4.43	.01

acceptance of the hypothesis, judges' correct responses to the children's drawings will significantly exceed chance in determining reading achievement, and further suggests that the judges employed systematic methods in their estimates on reading achievement of the experimental group.

To determine if the judges were able to select better than chance the good readers based on their drawings, Table 13 has been summarized. A probability of .01 for each of the judges indicates acceptance of the hypothesis, judgments will significantly exceed chance in determining reading achievement of good readers, and further suggests that the judges employed systematic methods in their selection of good readers of the experimental group.

Table 14 has been summarized to determine if the judges were able to select better than chance poor readers based on children's drawings. A probability of .01 for each of the judges indicates acceptance of the hypothesis,

TABLE 13.--Significance of judges' estimates of good readers based on Negro and white children's drawings

Judges	R	W	Probability Differences	Standard Error Differences	Z	P
I	15	5	5	2.24	2.23	.01
II	17	3	7	2.24	3.13	.01
III	17	3	7	2.24	3.13	.01

TABLE 14.--Significance of judges' estimates of poor readers based on Negro and white children's drawings

Judges	R	W	Probability Differences	Standard Error Differences	Z	P
I	15	5	5	2.24	2.23	.01
II	17	3	7	2.24	3.13	.01
III	17	3	7	2.24	3.13	.01

judgments will significantly exceed chance in determining reading achievement of poor readers, and it further suggests that judges employed systematic methods in their selection of poor readers of the experimental group.

To determine whether or not judges were able to estimate better than chance reading achievement based on Negro children's drawings, Table 15 has been summarized. Probabilities for each of the three judges indicate acceptance of the hypothesis, judgments will significantly exceed chance

TABLE 15.--Significance of judges' estimates of reading achievement based on Negro children's drawings

Judges	R	W	Probability Differences	Standard Error Differences	Z	P
I	14	6	4	2.24	1.79	.04
II	15	5	5	2.24	2.23	.01
III	16	4	6	2.24	2.68	.01

in determining reading achievement of Negro children, and they further suggest that judges employed systematic methods in their estimates of reading achievement based on Negro children's drawings.

Table 16 has been summarized to determine whether or not judges were able to estimate better than chance reading achievement based on white children's drawings. Probabilities for the three judges indicate acceptance on the hypothesis, judgments will significantly exceed chance in determining reading achievement of white children, and they further suggest that judges employed systematic methods in their estimates of reading achievement based on white children's drawings. It may be observed that judges were able to make more correct estimates of reading achievement on white children's drawings than they were able to make on Negro children's drawings, even though, the hypotheses were accepted for judgments on both racial groups.

TABLE 16.--Significance of judges' estimates of reading achievement based on white children's drawings

Judges	R	W	Probability Differences	Standard Error Differences	Z	P
I	16	4	6	2.24	2.68	.01
II	19	1	9	2.24	4.46	.01
III	18	2	8	2.24	3.57	.01

Emotional Adjustment

An evaluation of emotional adjustment was the third group of estimates made. These estimates were made on the second day of the experiment. Instructions were repeated as had been given on the previous day. The task of making estimates on the 40 sets of drawings ($H\text{-}T\text{-}P_1P_2$) was completed in thirty minutes by the three judges. The judges showed no evidence of fatigue.

There were 120 evaluations made by the judges, however, 117 responses were recorded because one drawing was deleted as a result of being placed incorrectly. There were 82 correct responses which is an accuracy percentage of 70.1. This percentage is lower than the accuracy percentage earned by the judges on mental ability and reading achievement. The accuracy rating on Negro children's drawings was 71.6 per cent while for white children it was 70.9 per cent. The accuracy ratings of judges were lower for both ethnic groups on

emotional adjustment than they were on mental ability and reading achievement. These data are presented in Table 17.

Validity of Judges' Responses on Emotional Adjustment

The significant Z scores indicate that the judges' responses in their estimates of emotional adjustment were valid. As summarized in Table 18, it is indicated that the hypotheses were accepted. Significant probabilities indicate acceptance of the hypothesis, judges' correct responses to the children's drawings will significantly exceed chance in estimating emotional adjustment, and further suggest that the judges employed systematic methods in their estimates of emotional adjustment of the experimental group.

Table 19 has been summarized to determine if the judges were able to select better than chance well adjusted children's drawings. In one case the significance of the judgments exceeded chance at the .01 level of confidence. Judge I and II made estimates that placed their responses at the .09 point (level of confidence). It is indicated that the hypothesis was accepted in the one instance mentioned above. The result further suggests that only one judge employed systematic methods in the identification of well adjusted children of the experimental group.

Table 20 has been summarized to determine if the judges were able to select better than chance poorly adjusted children's drawings. The hypothesis, judgments will

TABLE 17.--Summary of judges' responses on H-T-P and DAP drawings for forty (40) first-grade elementary school children of Oklahoma City, Oklahoma, based on evaluations of emotional adjustment

Race and Classification	Judges I Right	Judges I Wrong	Judges II Right	Judges II Wrong	Judges III Right	Judges III Wrong	Totals Right	Totals Wrong	Grand Totals Right	Grand Totals Wrong	Percentages of Correct Responses
Caucasian:											
Well Adjusted	7	3	7	1	6	4	20	8			
Poorly Adjusted	7	2	6	3	6	3	19	8	39	16	70.9
Negro:											
Well Adjusted	6	4	8	3	7	2	21	9			
Poorly Adjusted	6	4	9	2	7	4	22	10	43	19	71.6
Totals	26	13	30	9	26	13	82	35	82	35	70.1

TABLE 18.--Significance of judges' estimates of Negro and white children's emotional adjustment based on their drawings

Judges	R	W	Probability Differences	Standard Error Differences	Z	P
I	26	13	6.5	3.12	2.08	.02
II	30	9	10.5	3.12	3.37	.0003
III	26	13	6.5	3.12	2.08	.02

TABLE 19.--Significance of judges' estimates of well adjusted Negro and white children's emotional adjustment based on their drawings

Judges	R	W	Probability Differences	Standard Error Differences	Z	P
I	13	7	3	2.24	1.34	.09
II	15	5	5	2.24	2.23	.01
III	13	7	3	2.24	1.34	.09

TABLE 20.--Significance of judges' estimates of poorly adjusted Negro and white children's emotional adjustment based on their drawings

Judges	R	W	Probability Differences	Standard Error Differences	Z	P
I	13	6	3.5	2.21	1.58	.57
II	14	5	4.5	2.21	2.04	.02
III	13	6	3.5	2.21	1.58	.57

significantly exceed chance in determining emotional adjustment of poorly adjusted children, was accepted for one judge and rejected for two judges. It may be observed that the hypotheses for Judges I and III were rejected on marginal bases.

The significant Z scores indicate that the judges' responses in their estimates of emotional adjustment based on Negro children's drawings were valid except for one judge. As summarized in Table 21, it is indicated that probabilities of .01 and .04 for Judges II and III, respectively, suggest acceptance of the hypothesis. Judgments significantly exceeded chance in determining emotional adjustment of Negro children. These probabilities suggest that the two judges employed systematic methods in their estimates of emotional adjustment of Negro children of the experimental group.

The significant Z scores indicate that the judges' responses in their estimates of emotional adjustment based on white children's drawings were valid. As summarized in Table 22 it is indicated that the hypotheses were accepted. Probabilities of .02 for Judge I and .05 for Judge II indicate acceptance of the hypothesis, judgments will significantly exceed chance in determining emotional adjustment of white children. These probabilities suggest that the two judges employed systematic methods in their estimates of emotional adjustment based on white children's drawings. It may be observed that judges were able to make more accurate

TABLE 21.--Significance of judges' estimates of Negro children's emotional adjustment based on their drawings

Judges	R	W	Probability Differences	Standard Error Differences	Z	P
I	12	8	2	2.24	0.89	.19
II	17	3	7	2.24	3.13	.01
III	14	6	4	2.24	1.79	.04

TABLE 22.--Significance of judges' estimates of white children's emotional adjustment based on their drawings

Judges	R	W	Probability Differences	Standard Error Differences	Z	P
I	14	5	4.5	2.21	2.04	.02
II	13	6	3.5	2.21	1.60	.05
III	12	8	1.5	2.21	0.68	.24

judgments on Negro children's drawings than they were able to make on white children's drawings determining emotional adjustment.

Race

The selection of race was the fourth group of drawings to be judged. These selections were made on the second day of the experiment. Instructions were repeated for this group of drawings as had been given on previous groups of evaluations.

The task of making selections on the forty sets of drawings (H-T-P_1P_2) was completed at various times by the judges. It was observed that judges erased many times in an attempt to make a final decision as to the race of the child whose drawings he rated. Many spontaneous comments were made such as, "I suppose 'religion' will be next." One judge completed the task in twenty-five minutes. Another judge required forty minutes while another required forty-five minutes to make selections of race.

As may be observed from Table 23, there were 120 selections made by the judges. Of the 120, 61 were correct responses which is an accuracy percentage of 50.83. This accuracy percentage is much lower than selections made on any of the previous judgments. When the ethnic groups were separated, the judges earned on Caucasian children's drawings a rating of 48.33 per cent while for Negro children they earned a rating of 50.00 per cent. This accuracy rating for Negro children was slightly higher than that for Caucasian children. However, ratings on both ethnic groups suggest that judges were unable to select correctly more than half of the drawings on the basis of race.

The selections made on ethnic groups were further divided into socio-economic classes. Judges' accuracy ratings on Caucasian upper class children was 42.86 per cent while for Negro upper-class children it was 66.66 per cent. On middle-class children, the judges' accuracy ratings were

TABLE 23.--Summary of judges' responses on H-T-P and DAP drawings for forty (40) first-grade elementary school children of Oklahoma City, Oklahoma, based on selections of race

Race and Socio-economic Classification	Judges I		II		III		Totals		Grand Totals		Percentages of Correct Responses
	Right	Wrong	Right	Wrong	Right	Wrong	Right	Wrong	Right	Wrong	
Caucasian:											
Upper	2	5	3	4	4	3	9	12			42.86
Middle	2	4	2	4	4	2	8	10			44.44
Lower	3	4	6	1	3	4	12	9			57.14
									29	31	48.33
Negro:											
Upper	3	4	5	2	6	1	14	7			66.66
Middle	4	2	4	2	3	3	11	7			61.11
Lower	0	7	3	4	2	5	5	16			23.81
									30	30	50.00
Totals	14	26	23	17	22	18	59	61	59	61	50.83

44.44 and 61.11 percentages for Caucasian and Negro children, respectively. On lower-class children, judges' accuracy ratings were 57.14 and 23.81 percentages for Caucasian and Negro children, respectively. Judges earned higher accuracy ratings on upper and middle-class Negro children's drawings than they did on white children of the same socio-economic classes. However, they earned a much lower rating on lower-class Negro children's drawings than they did on any of the other sub-divisions. Thus, judges made their highest per cent of correct selections on upper and middle-class Negro children's drawings.

Validity of Judges' Responses on Race

Table 24 has been summarized to determine whether or not judges were able to select better than chance Negro and white children's drawings. Probabilities for each of the judges indicate rejection of the hypothesis, judges' correct responses to the children's drawings will significantly exceed chance in determining the race of the children. It may be observed that Judge I made selections which earned a probability of .98 due to the large number of wrong responses. Judges II and III with probabilities .17 and .25, respectively, made responses which were in the limits of chance. Therefore, the data indicate that judges were unable to determine race based on the children's drawings.

TABLE 24.--Significance of judges' selections of race based on Negro and white children's drawings

Judges	R	W	Probability Differences	Standard Error Differences	Z	P
I	14	26	6	3.16	1.90	.98
II	23	17	3	3.16	.95	.17
III	22	18	2	3.16	.65	.25

Table 25 has been summarized to determine if judges were able to select better than chance Negro children's drawings. Probabilities indicate rejection of the hypothesis, judges' correct responses to the children's drawings will significantly exceed chance in determining the race of the children. Judges I, II, and III with probabilities of .09, .19, and .33, respectively, were within the limits of chance.

Table 26 has been summarized to determine whether or not judges were able to select better than chance white children's drawings. Probabilities indicate the hypothesis, judges' correct responses to the race of the children, was rejected. Judges I, II, and III with probabilities of .09, .33, and .33, respectively, were within the limits of chance. It may be observed that the judges' accuracy on selection of race was comparable when judgments were separately analyzed for both ethnic groups.

TABLE 25.--Significance of judges' selections of Negro children based on children's drawings

Judges	R	W	Probability Differences	Standard Error Differences	Z	P
I	7	13	3	2.24	1.34	.09
II	12	8	2	2.24	0.89	.19
III	11	9	1	2.24	0.45	.33

TABLE 26.--Significance of judges' selections of white children based on children's drawings

Judges	R	W	Probability Differences	Standard Error Differences	Z	P
I	7	13	3	2.24	1.34	.09
II	11	9	1	2.24	0.45	.33
III	11	9	1	2.24	0.45	.33

Summary

The result of the statistical data presented in this chapter indicated that estimates made by judges on children's mental ability, reading achievement, and emotional adjustment were valid. Race was the only aspect of the study in which judges were unable to make accurate selections beyond chance.

It was observed that responses on emotional adjustment was the third group to be analyzed, and the first group of

responses to reveal an inconsistent pattern of acceptance of the hypothesis for each judge. However, this inconsistent pattern only appeared when the responses were analyzed according to sub-groups of the total experimental group. In spite of the inconsistent pattern of accuracy ratings for each judge on various sub-groups, ratings were accurate beyond chance when an analysis was made on the total experimental group for determining emotional adjustment.

As a result of analyses of statistical data gathered from the judges' responses on race, probabilities for each of the judges in all aspects of analyses presented revealed consistent rejection of the hypothesis. The data consistently indicated that judges were unable to determine race based on the children's drawings. Therefore, judges were able to determine mental ability, reading achievement, and emotional adjustment while they were unable to determine race based on children's drawings.

CHAPTER V

AGREEMENT AMONG JUDGES

This chapter has been devoted to an analysis of data pertaining to the agreement among judges in their responses to the H-T-P and DAP productions. The major hypothesis, judges will agree in their responses in determining mental ability, reading achievement, emotional adjustment, and race of the children whose drawings they judge, was statistically tested. Chi-square was chosen for the statistic to be applied. Chi-square is a general-purpose statistic that has many and diverse applications. Its most common use is in connection with data in the form of frequencies, or data that can be reduced to frequencies.[1] Therefore, to determine if the judges agreed among themselves on the four variables being tested, chi-square rendered ratios between a squared discrepancy and an expected frequency for the three judges. As a result, this statistic revealed whether or not a significant difference existed among the judges in their responses.

---

[1] J. P. Guilford, Fundamental Statistics in Psychology and Education (New York: McGraw-Hill Book Company, Inc., 1956), p. 228.

Table 27 has been summarized to determine whether or not judges agreed beyond chance in their responses on mental ability, reading achievement, emotional adjustment, and race based on the children's drawings. A chi-square of 2.37 with 6 degrees of freedom at a probability level greater than .80 and less than .90 requires that the hypothesis, judges will agree in their correct responses in determining mental ability; reading achievement; emotional adjustment; and race of the children whose drawings they judge, be accepted. Therefore, judges agreed beyond chance in their estimates on the four variables tested.

Table 28 has been summarized to determine whether or not judges were in agreement on estimates made of Negro children's ability, adjustment, reading achievement, and race. A chi-square of 2.0083 with 6 degrees of freedom at a probability level greater than .90 and less than .95 requires that the hypothesis, judges will agree in their correct responses in determining mental ability; reading achievement; emotional adjustment; and race of the Negro children whose drawings they judge, be accepted. Therefore, judges will agree more than 90 per cent of the time in their estimates on the four variables based on Negro children's drawings.

Table 29 has been summarized to determine whether or not judges agreed in their estimates made of white children's mental ability, reading achievement, emotional adjustment, and race based on the children's drawings. A chi-square of

TABLE 27.--Chi-square frequencies of judges' correct estimates on ability, reading achievement, emotional adjustment, and race based on children's drawings

Variables	Judges I	Judges II	Judges III	Totals
Ability	34 (30.06)	30 (33.82)	34 (34.11)	98
Reading	30 (30.06)	34 (33.82)	34 (34.11)	98
Adjustment	26 (25.77)	30 (28.99)	28 (29.24)	84
Race	14 (18.10)	23 (20.36)	22 (20.54)	59
Totals	104	117	118	339

Obtained $X^2 = 2.37$ df = 6 $.80 < p < .90$
Hypothesis Accepted.

TABLE 28.--Chi-square frequencies of judges' correct estimates on ability, reading achievement, emotional adjustment, and race based on Negro children's drawings

Variables	Judges I	Judges II	Judges III	Totals
Ability	17 (14.67)	16 (17.60)	16 (16.72)	49
Reading	14 (13.47)	15 (16.16)	16 (15.36)	45

TABLE 28.--Concluded

Variables	Judges I	Judges II	Judges III	Totals
Adjustment	12 (12.87)	17 (15.45)	14 (16.72)	43
Race	7 (9.00)	12 (10.77)	11 (10.24)	30
Totals	50	60	57	167

Obtained $X^2 = 2.008$ df = 6 $.90 < p < .95$
Hypothesis Accepted.

TABLE 29.--Chi-square frequencies of judges' correct estimates on mental ability, reading achievement, emotional adjustment, and race based on white children's drawings

Variables	Judges I	Judges II	Judges III	Totals
Ability	17 (15.40)	14 (16.20)	18 (17.40)	49
Reading	16 (16.60)	19 (17.60)	18 (18.80)	53
Adjustment	14 (12.90)	13 (13.60)	14 (14.50)	41
Race	7 (9.10)	11 (9.60)	11 (10.30)	29
Totals	54	57	61	172

Obtained $X^2 = 1.556$ df = 6 $.95 < p < .98$
Hypothesis Accepted.

1.556 with 6 degrees of freedom at a probability level greater than .95 and less than .98 requires that the hypothesis, judges will agree in their correct responses in determining mental ability; reading achievement; emotional adjustment; and race of the white children whose drawings they judge, be accepted. Therefore, judges will agree more than 95 per cent of the time in their estimates on the four variables based on white children's drawings.

Table 30 has been summarized to determine whether or not judges were able to make more correct estimates of bright

TABLE 30.--Chi-square frequencies of judges' correct estimates on mental ability of bright and dull children based on children's drawings of both races

	Judges			
Ability	I	II	III	Totals
Bright	17 (17.00)	15 (15.00)	17 (17.00)	49
Dull	17 (17.00)	15 (15.00)	17 (17.00)	49
Totals	34	30	34	98

Obtained $X^2 = 0.00$ df = 2 p > .99
Hypothesis Accepted.

children than they were able to make of dull children based on their drawings. A chi-square of 0.00 with 2 degrees of freedom requires that the hypothesis, agreement among judges will exist in their responses to bright children's drawings when compared with their responses to dull children's drawings in determining mental ability, be accepted. Therefore, judges will agree more than 99 per cent of the time in their estimates of mental ability of bright children as well as their estimates of dull children based on children's drawings.

Table 31 has been summarized to determine whether or not judges agreed in their estimates of good readers as well as they agreed in their estimates of poor readers based on

TABLE 31.--Chi-square frequencies of judges' correct estimates of reading achievement of good and poor readers based on children's drawings of both races

	Judges			
Reading	I	II	III	Totals
Good	17 (16.2)	17 (17.3)	17 (17.3)	51
Poor	15 (15.8)	17 (16.7)	17 (16.7)	49
Totals	32	34	34	100

Obtained $X^2 = .1015$ $df = 2$ $.90 < p < .95$
Hypothesis Accepted.

children's drawings. A chi-square of .1015 with 2 degrees of freedom at a probability level greater than .90 and less than .95 requires that the hypothesis, agreement among judges will exist in their responses to good reader's drawings when compared with their responses to poor reader's drawings in determining reading achievement, be accepted. Therefore, judges will agree more than 90 per cent of the time in their estimates of reading achievement of good and poor readers.

Table 32 has been summarized to determine whether or not judges agreed in their estimates of well adjusted children as well as they agreed in their estimates of poorly adjusted children based on children's drawings. A chi-square of

TABLE 32.--Chi-square frequencies of judges' correct estimates on emotional adjustment of well adjusted and poorly adjusted children based on children's drawings of both races

	Judges			
Adjustment	I	II	III	Totals
Well Adjusted	13 (13.31)	15 (15.36)	15 (14.33)	43
Poorly Adjusted	13 (12.69)	15 (14.64)	13 (13.67)	41
Totals	26	30	28	84

Obtained X^2 = .0918 df = 2 $.95 < p < .98$
Hypothesis Accepted.

.0918 with 2 degrees of freedom at a probability level greater than .95 and less than .98 requires that the hypothesis, agreement among judges will exist in their responses to well adjusted children's drawings when compared with their correct responses to poorly adjusted children's drawings in determining emotional adjustment, be accepted. Therefore, judges will agree more than 95 per cent of the time when making estimates of well adjusted and poorly adjusted children.

Summary

The result of statistical data presented in this chapter indicated that in every case of estimates made by judges on the four variables there was a high probability level of agreement. Analyses of statistical data were made on the following: (1) agreement among judges in their responses on mental ability, reading achievement, emotional adjustment, and race based on all children's drawings; (2) agreement among judges in their responses on mental ability, reading achievement, emotional adjustment, and race based on Negro children's drawings; (3) agreement among judges in their responses on mental ability, reading achievement, emotional adjustment, and race based on white children's drawings; (4) agreement among judges in their responses on mental ability of bright and dull children based on all children's drawings; (5) agreement among judges in their responses on reading achievement of good and poor

readers based on all children's drawings; and (6) agreement among judges in their responses on emotional adjustment of well adjusted and poorly adjusted children based on all children's drawings.

The major hypothesis, judges will agree in their correct responses to children's drawings to determine mental ability; reading achievement; emotional adjustment; and race, was accepted on the statistical results of Tables 27, 28, 29, 30, 31, and 32. Data from these tables indicated that the combinations of responses analyzed, the judges will agree in their estimates more than 80 per cent of the time on a like sample utilized in this study. Even though high probabilities of agreement were obtained from the six analyses, the highest probability, more than 99 per cent, was obtained from the responses to drawings on mental ability of bright and dull children.

CHAPTER VI

COMPARISON OF JUDGES' CORRECT RESPONSES ON NEGRO AND WHITE CHILDREN'S DRAWINGS

This chapter has been devoted to an analysis of data pertaining to comparisons of judges' correct responses on white and Negro children's H-T-P and DAP drawings. The major hypothesis, judges will agree in determining mental ability; reading achievement; emotional adjustment; and race of Negro children when compared to their correct estimates made of white children based on the children's drawings, was statistically tested. Chi-square was chosen for the statistic to be applied for the same reasons as presented in Chapter V of this study. This statistic revealed whether or not a significant difference existed in the ability of judges to make estimates of white children when compared with the estimates made of Negro children based on the children's drawings.

Table 33 has been summarized to determine whether or not judges were able to estimate mental ability of Negro children as well as they were able to estimate mental ability of white children based on their drawings. A chi-square of

TABLE 33.--Chi-square frequencies of judges' correct estimates on mental ability based on Negro and white children's drawings

Race	Judges I	Judges II	Judges III	Totals
Negro	17 (17.00)	16 (15.00)	16 (17.00)	49
White	17 (17.00)	14 (15.00)	18 (17.00)	49
Totals	34	30	34	98

Obtained $X^2 = 1.3990$ df = 2 $.30 < p < .50$
Non-significant. Hypothesis Accepted.

1.3990 with 2 degrees of freedom at a probability level greater than .30 and less than .50 requires that the hypothesis, judges will agree in determining mental ability of Negro children when compared to their correct estimates of mental ability of white children, be accepted. Therefore, judges did not significantly differ in their correct responses on mental ability of Negro and white children.

Table 34 has been summarized to determine whether or not judges were able to estimate the mental ability of bright Negro children as well as they were able to estimate mental ability of bright white children based on their drawings. A chi-square of .2127 with 2 degrees of freedom at a probability level greater than .80 and less than .90 requires that the hypothesis, judges will agree in determining

TABLE 34.--Chi-square frequencies of judges' correct estimates of mental ability based on bright Negro and bright white children's drawings

Race	Judges I	Judges II	Judges III	Totals
Negro	9 (8.70)	8 (7.70)	8 (8.70)	25
White	8 (8.30)	7 (7.30)	9 (8.30)	24
Totals	17	15	17	49

Obtained X^2 = .2127 df = 2 .80 < p < .90
Non-significant. Hypothesis Accepted.

mental ability of bright Negro children when compared to their correct estimates of bright white children, be accepted. Therefore, judges did not significantly differ in their correct responses on mental ability of bright Negro and white children.

Table 35 has been summarized to determine whether or not judges were able to estimate mental ability of dull Negro children as well as they were able to estimate mental ability of dull white children based on the children's drawings. A chi-square of .2127 with 2 degrees of freedom at a probability level greater than .80 and less than .90 requires that the hypothesis, judges will agree in determining mental ability of dull Negro children when compared to their correct

TABLE 35.--Chi-square frequencies of judges' correct estimates of ability based on dull Negro and dull white children's drawings

Race	Judges I	Judges II	Judges III	Totals
Negro	8 (8.30)	8 (7.30)	8 (8.30)	24
White	9 (8.70)	7 (7.70)	9 (8.70)	25
Totals	17	15	17	49

Obtained X^2 = .2127 df = 2 .80 < p < .90
Non-significant. Hypothesis Accepted.

estimates of dull white children, be accepted. Therefore, judges did not significantly differ in their correct responses on mental ability of dull Negro and white children. It may be observed that judges were able to estimate brightness and dullness with the same degree of accuracy of the two ethnic groups.

Table 36 has been summarized to determine whether or not judges were able to estimate reading achievement of Negro children as well as they were able to estimate reading achievement of white children based on their drawings. A chi-square of .0908 with 2 degrees of freedom at a probability level greater than .95 and less than .98 requires that the hypothesis, judges will agree in determining reading

TABLE 36.--Chi-square frequencies of judges' correct estimates of reading achievement based on Negro and white children's drawings

Race	Judges I	Judges II	Judges III	Totals
Negro	14 (13.80)	15 (15.60)	16 (15.60)	45
White	16 (16.20)	19 (18.40)	18 (18.40)	53
Totals	30	34	34	98

Obtained $X^2 = .0908$ $df = 2$ $.95 < p < .98$
Non-significant. Hypothesis Accepted.

achievement of Negro children when compared to their correct estimates of reading achievement of white children, be accepted. Therefore, judges did not significantly differ in their correct responses on reading achievement of Negro and white children.

Table 37 has been summarized to determine whether or not judges were able to estimate reading achievement of Negro good readers as well as they were able to estimate reading achievement of white good readers based on the children's drawings. A chi-square of .0688 with 2 degrees of freedom at a probability level greater than .95 and less than .98 requires that the hypothesis, judges will agree in estimating reading achievement of Negro good readers when

TABLE 37.--Chi-square frequencies of judges' correct estimates of reading achievement based on Negro good reader's and white good reader's drawings

Race	Judges I	Judges II	Judges III	Totals
Negro	7 (7.30)	8 (8.30)	9 (8.30)	24
White	8 (7.70)	9 (8.70)	8 (8.70)	25
Totals	15	17	17	49

Obtained X^2 = .0688 df = 2 .95 < p < .98
Non-significant. Hypothesis Accepted.

compared to their correct estimates of white good readers, be accepted. Therefore, judges did not differ significantly in their correct responses on reading achievement of Negro good readers and white good readers.

Table 38 has been summarized to determine whether or not judges were able to estimate reading achievement of Negro poor readers as well as they were able to estimate reading achievement of white poor readers based on the children's drawings. A chi-square of .0623 with 2 degrees of freedom at a probability level greater than .95 and less than .98 requires that the hypothesis, judges will agree in estimating reading achievement of Negro poor readers when compared to their correct estimates of white poor readers, be

TABLE 38.--Chi-square frequencies of judges' correct estimates of reading achievement based on Negro poor reader's and white poor reader's drawings

Race	Judges I	Judges II	Judges III	Totals
Negro	7 (6.40)	7 (7.20)	7 (7.20)	21
White	8 (8.60)	10 (9.80)	10 (9.80)	28
Totals	15	17	17	49

Obtained $X^2 = .0623$ df = 2 $.95 < p < .98$
Non-significant. Hypothesis Accepted.

accepted. Therefore, judges did not differ significantly in their correct responses on reading achievement of Negro and white poor readers. It may be observed that judges were able to estimate good and poor reader's achievement with the same degree of accuracy for the two ethnic groups.

Table 39 has been summarized to determine whether or not judges were able to determine the emotional adjustment of Negro children as well as they were able to determine the emotional adjustment of white children based on their drawings. A chi-square of .5310 with 2 degrees of freedom at a probability level greater than .70 and less than .80 requires that the hypothesis, judges will agree in determining the emotional adjustment of Negro children when compared to their

TABLE 39.--Chi-square frequencies of judges' correct estimates on emotional adjustment of Negro and white children based on their drawings

Race	Judges I	Judges II	Judges III	Totals
Negro	12 (13.31)	17 (15.36)	14 (14.33)	43
White	14 (12.69)	13 (14.64)	14 (13.67)	41
Totals	26	30	28	84

Obtained $X^2 = .5310$ df = 2 $.70 < p < .80$
Non-significant. Hypothesis Accepted.

correct estimates of white children, be accepted. Therefore, judges did not differ significantly in their correct responses of emotional adjustment of Negro and white children.

Table 40 has been summarized to determine whether or not judges were able to estimate the emotional adjustment of well adjusted Negro children as well as they were able to estimate the emotional adjustment of well adjusted white children based on their drawings. A chi-square of .1867 with 2 degrees of freedom at a probability level greater than .90 and less than .95 requires that the hypothesis, judges will agree in determining the emotional adjustment of well adjusted Negro children when compared to their correct estimates of well adjusted white children, be accepted.

TABLE 40.--Chi-square frequencies of judges' correct estimates on emotional adjustment of well adjusted Negro and white children based on their drawings

Race	Judges I	Judges II	Judges III	Totals
Negro	6 (6.35)	8 (7.33)	7 (7.33)	21
White	7 (6.65)	7 (7.67)	8 (7.67)	22
Totals	13	15	15	43

Obtained $X^2 = .1867$ df = 2 $.90 < p < .95$
Non-significant. Hypothesis Accepted.

Therefore, judges did not differ significantly in their correct responses on emotional adjustment of well adjusted Negro and white children.

Table 41 has been summarized to determine whether or not judges were able to estimate the emotional adjustment of poorly adjusted Negro children as well as they were able to determine the emotional adjustment of poorly adjusted white children based on their drawings. A chi-square of .5567 with 2 degrees of freedom at a probability level greater than .70 and less than .80 requires that the hypothesis, judges will agree in determining the emotional adjustment of poorly adjusted Negro children when compared to their correct estimates of poorly adjusted white children, be accepted.

TABLE 41.--Chi-square frequencies of judges' correct estimates on emotional adjustment of poorly adjusted Negro and white children based on their drawings

Race	Judges I	Judges II	Judges III	Totals
Negro	6 (6.98)	9 (8.05)	7 (6.98)	22
White	7 (6.02)	6 (6.95)	6 (6.02)	19
Totals	13	15	13	41

Obtained χ^2 = .5567 df = 2 .70 < p < .80
Non-significant. Hypothesis Accepted.

Therefore, judges did not differ significantly in their correct responses on emotional adjustment of poorly adjusted Negro and white children.

Table 42 has been summarized to determine whether or not judges were able to select Negro children better than they were able to select white children based on the children's drawings. A chi-square of .0349 with 2 degrees of freedom at a probability level greater than .98 and less than .99 requires that the hypothesis, judges will agree in their correct responses in selecting the race of the children whose drawings they judge, be accepted. That is, judges did not differ significantly in their inability to select Negro children as compared to their inability to select white

TABLE 42.--Chi-square frequencies of correct responses made on race based on Negro and white children's drawings

Race	Judges I	Judges II	Judges III	Totals
Negro	7 (7.10)	12 (11.70)	11 (11.20)	30
White	7 (6.90)	11 (11.30)	11 (10.80)	29
Totals	14	23	22	59

Obtained $X^2 = .0349$ df = 2 $.98 < p < .99$
Non-significant. Hypothesis Accepted.

children based on the children's drawings.

Summary

This chapter was devoted to an analysis of data pertaining to comparisons of judges' correct responses on white and Negro children's drawings. The result of an analysis of statistical data indicated that judges were just as accurate when making estimates based on Negro children's drawings as they were when making estimates based on white children's drawings.

The major hypothesis, judges will agree in determining mental ability; reading achievement; emotional adjustment; and race of Negro children when compared to their correct estimates made on white children, was statistically tested.

The hypothesis was accepted based on Tables 33 through 42. Each variable was tested separately. In all cases, there was no significant difference found in judges' correct responses for white and Negro children. Even in the aspect of selecting race, significant differences did not exist in the judges' inability to select drawings made by Negro children when compared to their inability to select drawings made by white children.

CHAPTER VII

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

What diagnostic and predictive value have these projective techniques, H-T-P and Machover DAP, when applied to Negro children? This was the major problem presented in this research. The following inquiries were made:

1. Are trained clinicians able to estimate the mental ability of first-grade Negro children equally as well as they are able to estimate the mental ability of first-grade white children by their H-T-P and DAP drawings?

2. Are trained clinicians able to estimate the reading achievement of first-grade Negro children equally as well as they are able to estimate the reading achievement of first-grade white children by their H-T-P and DAP drawings?

3. Are trained clinicians able to estimate the emotional adjustment of first-grade Negro children equally as well as they are able to estimate the emotional adjustment of first-grade white children by their H-T-P and DAP drawings?

4. Are trained clinicians able to differentiate between Negro and white children's H-T-P and DAP drawings?

An analysis was made of the projective responses of 240 elementary school children enrolled in the Oklahoma City Public Schools. Two projective devices were employed to evoke the reactions of the subjects. The responses were compared with the Goodenough, social histories, and teacher ratings. The evidence of this aspect of the research indicates that there was essential agreement among the several devices. The projective tests agreed in consequence with the Goodenough and both of these devices gave similar results about the mental ability, reading achievement, and emotional adjustment of the experimental groups.

The projective responses made by the judges were analyzed by the experimenter, who had corroborating evidence from social histories, teacher appraisals, and Goodenough results. These categories were submitted to trained judges for the experimental classification. The judges were instructed to identify the mental ability, reading achievement, emotional adjustment, and race as revealed by the subjects' presentations. Three judges were employed in each of the attempts. It was assumed that the training of the judges in the effective use of the particular projective devices would result in non-subjectively determined descriptions of the response variables being categorized. It was further assumed that the judges, being similarly trained, would basically render equivalent evaluations of the projections.

A primary assumption was that the peculiar ethnocultural circumstances of the Negro would produce a patterned personality structure that would be revealed by uniform attributes analyzable in the projective responses evoked by the stimulus effect of the test devices. The clinicians, in their dichotomizing tasks, might be capable of identifying the racial origin of the subject by observation of the elements presumably produced by the individuals influenced by common pressures from similar social forces.

The sample was representative of sex and age factors of first-grade school children. The data collected were treated statistically by utilizing probability evaluation and chi-square. It was possible to determine the extent to which judgments tended to depart from proportions attributable to chance and chance alone. Since it is possible for dichotomized matching to occur by chance, it was necessary to evaluate the accuracy of judges in order to assure that their observed proficiencies were not artifacts of the experimental design.

It was hypothetically assumed that the objective training of the judges would eliminate chance as a factor accounting for their accuracies of classifications. It was further assumed that their appraisals would be correct significantly beyond that of chance. The possible confounding of the chance feature and skill was analyzed to assure that the clinical classifications were significantly higher than

would occur by chance and chance alone.

The clinical sub-samples were selected by the clinicians on the basis of upper and lower classifications in the areas of (1) mental ability, (2) reading achievement, (3) and emotional adjustment. Additionally, judges were requested to attempt the recognition of racial origin from the stimuli of the test productions. Each of the three judges participated in the four ratings of the 240 children's drawings. The experimental design was fashioned to determine the accuracy of ratings made by the judges, agreement among judges in their ratings, and comparisons of judges' ability to use these tests with both races. Therefore, the applicability of these projective devices to Negro children was determined.

Conclusions

Three major hypotheses were statistically tested to determine if cultural differences revealed themselves in projective techniques. More specifically, the major purpose of this study was to test the applicability of the H-T-P and DAP tests to Negro children. The evidence of this research indicates that the H-T-P and DAP tests are valid instruments for measuring many aspects of first-grade Negro children. Currently, no projective device can positively isolate and identify all of the behavioral entities investigated by this study. The H-T-P naturally suffers the limitations that are

associated with unstructured devices which depend upon the capacity of a human being to utilize elements of figures to determine circumstances related to events of a real world. However, from the result of this research, it is evident that the trained judges were able to correctly appraise the Negro and white children for their intellectual, academic, and emotional qualities but failed to discriminate between the experimental Negro sample and the white control group when they attempted to seek clues to ethnic origin in the picture productions. The first-graders who were the subjects of this research might conceivably reflect the consequences of the many oppressive events of their racial history, but the traumatic histories of their forebears could hardly be expected to account for all aspects and dimensions of the qualities measured by these two projective devices.

The first hypothesis, judges will make correct responses which will significantly exceed chance in determining mental ability; reading achievement; emotional adjustment; and race of the children whose drawings they judge, was accepted except for the variable of race. Race was the only aspect of this major hypothesis that revealed the inability of the clinicians to make correct judgments. In view of the general acceptance of the first hypothesis, it can be concluded that these tests are applicable to Negro children. Apparently, cultural differences that exist in the backgrounds of these two ethnic groups do not reveal themselves in the

H-T-P and DAP productions. If cultural background was projected in these drawings, it was of such limited degree that the validity of the tests was not destroyed. However, in the opinion of the writer, it is quite evident that Negro children did not show unique patterns in their productions. If such uniqueness had existed, the judges would have been able to determine the race of the children whose drawings they judged.

After establishing the fact that judges' responses were valid in estimating mental ability, reading achievement, and emotional adjustment; the second major hypothesis was tested. This hypothesis, judges will agree in their correct responses in determining mental ability; reading achievement; emotional adjustment; and race, was accepted. Many combinations of responses were analyzed statistically. In all analyses, judges had a high probability of agreement in their responses. While the judges were generally inaccurate in their attempts at racial identification, they were nevertheless uniform in their opinions. This unanimity indicates that the decisions were based upon a clinical stereotype.

The third hypothesis, judges will agree significantly in determining mental ability; reading achievement; emotional adjustment; and race of Negro children when compared to their correct estimates of white children based on the children's drawings, was accepted. Many combinations of responses were analyzed statistically. In all analyses, differences did not

exist in the correct number of estimates made on Negro children when compared with their estimates made on white children. This was true of all variables including race. That is, judges were just as inefficient in determining the race of white children as they were in determining the race of Negro children.

The basic assumptions posited at the outset are thus largely accepted as tenable, at least in terms of their statistical probabilities. Thus, philosophically and statistically it seems valid to conclude that skills which determine mental ability, reading achievement, and emotional adjustment are accurately evaluated by the H-T-P and DAP techniques. While the selection of race was a method employed to determine whether or not Negro children conveyed their cultural and social limitations through drawings, clinicians' racial judgments also revealed that a system of interpreting the tests to identify race is subject to considerable error.

Recommendations

It is recommended that the study be expanded to determine if the H-T-P and DAP are valid instruments for rural Negroes. This seems a feasible extension since the urban Negroes of the experimental group may show greater similarity to the proximate Caucasoid populace because of the extent of the social interactions which occur in urban existence. It is possible that the socially isolated Negro groups may possibly exhibit some unique racial traits in

their projective productions because of isolation.

While every care was taken to appropriately stratify the experimental and control groups to assure adequate representativeness, all non-parametric data are subject to some distortion of results because of the statistical treatment. Thus, it is possible that larger samples might yield data of greater validity.

It is recommended that further investigations be made under similar experimental conditions in different regions of the United States utilizing Negro judges to observe the effect of these variations upon the outcomes.

BIBLIOGRAPHY

Books

Anastasi, Anne. Differential Psychology. New York: Macmillan Company, 1958.

Anderson, Harold H., and Anderson, Gladys L. An Introduction to Projective Techniques. New York: Prentice-Hall, Inc., 1952.

Clarke, Kenneth B. Prejudice and Your Child. Boston: Beacon Press, 1955.

Davie, Maurice R. Negroes in American Society. New York: McGraw-Hill Book Company, Inc., 1949.

Davis, Allison. Social Class Influence Upon Learning. Cambridge, Mass.: Harvard University Press, 1948.

Davis, Allison, Gardner, Burleigh B., and Gardner, Mary. Deep South. Chicago: University of Chicago Press, 1941.

Eells, Kenneth, and others. Intelligence and Cultural Differences. Chicago: University of Chicago Press, 1950.

Frazier, E. Franklin. The Negro in the United States. New York: Macmillan and Company, 1957.

Garth, T. A. Race Psychology: A Study of Racial Mental Differences. New York: McGraw-Hill Book Company, Inc., 1931.

Ginsberg, Eli. The Negro Potential. New York: Columbia University Press, 1955.

Goodenough, Florence L. Measurement of Intelligence by Drawings. Chicago: World Book Company, 1926.

Greville, Thomas. U. S. Life Tables and Actuarial Tables in U. S. Washington: Government Printing Office, 1946.

Kardiner, A., and Ovesey, L. The Mark of Oppression: A Psychological Study of the American Negro. New York: Norton and Company, 1951.

Karon, B. P. The Negro Personality: A Rigorous Investigation of the Effects of Culture. New York: Springer, 1958.

Klineberg, Otto. Characteristics of the American Negro. New York: Harper and Brothers, 1944.

Machover, Karon. Personality Projection in the Drawings of the Human Figure. Springfield, Ill.: C. C. Thomas, 1950.

Miner, John B. Intelligence in the United States. New York: Springer Publishing Company, 1957.

Myrdal, Gunnar. An American Delemma: The Negro Problem and Modern Democracy. New York: Harper and Brothers, 1944.

Prothro, E. Terry, and Teska, P. T. Psychology--A Biosocial Study of Behavior. New York: Ginn and Company, 1950.

Shuey, Audrey M. The Testing of Negro Intelligence. Lynchburg, Virginia: Randolph-Macon Women's College, 1958.

Sherif, Muzafer, and Sherif, Carolyn W. An Outline of Social Psychology. New York: Harper and Brothers, 1956.

Siegel, Sidney. Nonparametric Statistics for the Behavioral Sciences. New York: McGraw-Hill Book Company, Inc., 1956.

Thompson, George G. Child Psychology. New York: Houghton-Mifflin Company, 1952.

Warner, W. Lloyd, and Lunt, Paul S. The Social Life of a Modern Community. New Haven: Yale University Press, 1941.

Warner, W. Lloyd, Meeker, Marchia, and Eells, Kenneth. Social Class in America. Chicago: Science Research Associates, Inc., 1949.

U. S. Bureau of the Census. Sixteenth Census of the United States: 1940. Vital Statistics Rates in the United States, 1900-1940. Washington: Government Printing Office, 1943.

Articles and Periodicals

Albee, George W., and Hamlin, Roy M. "An Investigation of the Reliability and Validity of Judgments of Adjustment Inferred from Drawings," Journal of Clinical Psychology, V (1949), 389-391.

Alpers, T. G., and Boring, E. G. "Intelligence Test Scores of Northern and Southern White and Negro Recruits," Journal of Abnormal Social Psychology, XXXIX (1944), 471.

Anastasi, Anne, and De Jesus, C. "Language Development and Goodenough Draw-A-Man I.Q. of Puerto Rican Pre-School Children in New York City," Journal of Abnormal Social Psychology, XLVIII (1953), 357-366.

Anderson, C. A. "Inequalities in Schooling in the South," American Journal of Sociology, LX (May, 1955), 547-61.

Ansbacher, H. L. "The Goodenough Draw-A-Man Test and Primary Mental Abilities," Journal of Consulting Psychology.

Bagley, N., and Jones, H. E. "Environmental Correlates of Mental and Motor Development," American Sociological Review, II (1937), 329-341.

Bean, K. L. "Negro Responses to Verbal and Non-Verbal Test Materials," Journal of Psychology, XIII (1942), 343-353.

Bieliauska, Vytauska. "Scorer's Reliability in the Quantitative Scoring of the H-T-P Technique," Journal of Clinical Psychology, XII (October, 1956), 366-69.

Birdie, R. F. "Measurement of Adult Intelligence by Drawings," Journal of Clinical Psychology, I (1945), 288-295.

Blum, Richard H. "The Validity of the Machover DAP Technique," Journal of Clinical Psychology, X (1954), 120-125.

Boas, Franz. "Evidence of the Nature of Intelligence," Educational Abstracts, V (1940), 258.

Brown, Fred. "An Experimental and Critical Study of the Intelligence of Negro and White Kindergarten Children," Journal of Genetic Psychology, LXV (1944), 161-175.

Brown, Fred. "H-T-P and Human Figure Drawing," Progress in Clinical Psychology, I (1948), 181.

Buck, John. "The H-T-P Technique," Journal of Clinical Psychology, Monogram Supplement, V (October, 1948).

Canady, H. G. "The Psychology of the Negro," in P. L. Harriman (ed.), The Encyclopedia of Psychology. New York: Philosophical Library, 1946.

Davis, Allison. "Caste, Economy, and Violence," American Journal of Sociology, LI (July, 1945), 7-16.

De Martino, Manfred F. "Human Figure Drawings by Mentally Retarded Males," Journal of Clinical Psychology, X (1954), 241-244.

Fisher, Seymour, and Fisher, Rhoda. "Test of Certain Assumptions Regarding Figure Drawing Analysis," Journal of Abnormal and Social Psychology, XLV (1950), 727.

Fromm, Eria, and Hartmen, Marschak. "Children's Intelligence Tests as a Measure of Dynamic Personality Functioning," American Journal of Orthopsychiatry, XXVII (1957), 134-144.

Gilliland, A. R. "Socio-economic Status and Race as Factors in Infant Intelligence Test Scores," Child Development, XXII (1951), 271-273.

Goodenough, Florence, and Harris, Dale. "Studies in the Psychology of Children's Drawings," Psychological Bulletin, II (1928-1949), and XLVII (1950), 369-433.

Graham, Stanley R. "A Study of Reliability in Human Figure Drawings," Journal of Projective Techniques, XX (1956), 385-386.

Hammer, Emanuel F. "Frustration-Aggression Hypothesis Extended to Socio-Racial Areas: Comparison of Negro and White Children's HTP's," Psychiatric Quarterly, XXVII (1953), 597-707.

_____. "Negro and White Children's Personality as Revealed by a Comparison of Their Drawings," Journal of Clinical Psychology, IX (1953), 7-10.

_____. "An Investigation of Sexual Symbolism: A Study of H-T-P's of Eugenically Sterilized Subjects," Journal of Projective Techniques, XVII (1953), 401-413.

Hammer, Emanuel. "Study of Hostility Possessed of Clinicians in Their Judgments," Journal of Projective Techniques, VII (1953).

Hanvik, Leo J. "The Goodenough Test As a Measure of Intelligence in Child Psychiatric Patients," Journal of Clinical Psychology, IX (1953), 71-72.

Havighurst, R. J. "Environment and the Draw-A-Man Test: The Performance of Indian Children," Journal of Abnormal and Social Psychology, XLI (1946), 50-83.

Hay, Edward N. "A Note on Small Samples," The Journal of Applied Psychology, XXXVII:6 (1953), 445.

Hollowell, A. I. "The Use of Projective Techniques in the Study of the Socio-Psychological Aspects of Acculturation," Journal of Projective Techniques, XV (1951), 27-44.

Hulse, W. D. "The Emotionally Disturbed Child Draws His Family," Journal of Child Behavior, III (1952), 66-79.

Israelite, Judith. "A Comparison of the Differences of Items for Intelligence of Northern Children and Mental Deficiency in the Goodenough Drawing Test," American Journal of Orthopsychiatry, VI (1936), 494-503.

Johnson, A. P., Ellerd, A. A., and Lahay, T. "The Goodenough Test as an Aid to Interpretation of Children's School Behavior," American Journal of Mental Deficiency, LIX (1950), 516-520.

Johnson, D. M. "Application of the Standard Score I.Q. to Social Statistics," Journal of Social Psychology, XXVII (1948), 217-227.

Jolles, Isaac. "Some Advances in Interpretation of the Chromatic Phase of the H-T-P," Journal of Clinical Psychology, XVIII (1957), 81-83.

King, Francis W. "The Use of Drawings of the Human Figure as an Adjunct in Psychotherapy," Journal of Clinical Psychology, XL (1954).

Klineberg, Otto, et al. "On Race and Intelligence: A Joint Statement," American Journal of Orthopsychiatry, XXVII (1957), 420-422.

Lee, Everett S. "Negro Intelligence and Selective Migration: A Philadelphia Test of the Klineberg Hypothesis," American Sociological Review, XVI (1951), 227-233.

Levy, Sidney. "Figure Drawing as a Projective Test," Projective Psychology. New York: Alfred A. Knopf, 1950.

Machover, Karen. "Human Figure Drawings of Children," Journal of Projective Techniques, XVII (1953), 85-91.

Margolis, M. E. "A Comparative Study of Figure Drawings at Three Points in Therapy," Rorschach Research Exchange and Journal of Projective Techniques, XII (1948), 94-105.

McCurdy, H. G. "Group and Individual Variability on the Goodenough Draw-A-Man Test," Journal of Educational Psychology, XXXVIII (1947), 428-436.

McGurk, Frank C. J. "Socio-Economic Status and Culturally-Weighted Test Scores of Negro Subjects," Journal of Applied Psychology, XXXVII:4 (1953), 276-277.

McHugh, G. "Changes in Goodenough IQ at the Public School Kindergarten Level," Journal of Educational Psychology, XXXVI (1945), 17-30.

Meyer, George. "Some Relationship Between Rorschach Scores in Kindergarten and Reading in the Primary Grades," Journal of Projective Techniques, XVII (1953).

Molish, Herman. "Contributions of Projective Tests to Problems of Psychological Diagnosis in Mental Deficiency," American Journal of Mental Deficiency, LXIII (1958), 282-293.

Needham, N. E. "A Comparative Study of the Performance of Feebleminded Subjects on the Goodenough Drawing, the Goldstein-Scheerer Cube Test, and the Standford-Binet," American Journal of Mental Deficiency, XLIX (1944), 155-161.

Newland, T. Ernest, and Lawrence, William C. "Chicago Non-Verbal Examination Results on an East Tennessee Negro Population," Journal of Clinical Psychology, IX (1953), 44-47.

North, Robert D. "The Intelligence of American Negroes," Anti-defamation League of B'nai B'rith (November, 1956), p. 28.

Reznikoff, Marvin, and Reznikoff, Helga R. "The Family Drawings," Journal of Clinical Psychology, XII (1956), 167.

Roesapple, James Quintar, and Phelan, Joseph G. "The Skills of Clinicians in Analysis of Projective Test," Journal of Clinical Psychology, X (1954), 307-320.

Royal, R. E. "Drawing Characteristics of Neurotic Patients Using a Drawing of a Man and Woman Technique," Journal of Clinical Psychology, V (1949), 392-395.

Rubin, Harold. "A Quantitative Study of the H-T-P and Its Relationship to the W-B Scale," Journal of Clinical Psychology, X (1954), 597-607.

Sarason, Seymour B., and Gladwin T. "Psychological and Cultural Problems in Mental Subnormality," American Journal of Mental Deficiency, LXII (1958), 1115-1307.

Sherman, M., and Key, C. B. "The Intelligence of Isolated Mountain Children," Child Development, III (1932), 279-290.

Sloan, William. "A Critical Review of HTP Validation Studies," Journal of Clinical Psychology, X (1954).

Sloan and Guertin. "A Comparison of HTP and W-B I.Q.'s in Mental Defectives," Journal of Clinical Psychology, IV (1948), 424-426.

Smith, Charles U., and Prothro, James. "Ethnic Differences in Authoritarian Personality," Sociological Forces, XXXV (1957), 334-338.

Smykal, Anthony, and Thorne, F. C. "Etiological Studies of Psycho-pathic Personality: II. Associal Type," Journal of Clinical Psychology (1951), 229-316.

Vorhaus, Pauline G. "Rorschach Configurations Associated with Reading Disability," Journal of Projective Techniques, XVI (1952), 3-19.

Weider, Arthur, and Noller, Paul A. "Objective Studies of Children's Drawings of Human Figures: II. Sex, Age, Intelligence," Journal of Clinical Psychology, IX (1953), 20-23.

Wells, G. R. "The Application of the Binet-Simon Tests to Groups of White and Colored School Children," Psychology Monogram, XXXII (1923), 52-53.

Wexler, Murray, and Holzberg, Jules D. "A Further Study of the Validity of Human Form Drawings in Personality Evaluation," Journal of Projective Techniques, XVI (1952), 249-251.

Wheeler, L. R. "Intelligence of Mountain Children," Educational Psychology, XXIII (1932), 351.

Whitmyre, John W. "The Significance of Artistic Excellence in the Judgment of Adjustment Inferred From Human Figure Drawings," Journal of Consulting Psychology, XVII (1953), 421-424.

Yepsen, Lloyd N. "The Reliability of the Goodenough Drawing Test with Feeble-Minded Subjects," Journal of Educational Psychology, XX (1929), 448-51.

Young, Florence M., and Pitts, Virginia A. "The Performance of Congenital Syphilitics on the Wechsler Intelligence Scale for Children," Journal of Consulting Psychology, XV (1951), 239-242.

Young, Florence M., and Bright, H. H. "Results of Testing 81 Negro Rural Juveniles with the Wechsler Intelligence Scale for Children," Journal of Social Psychology, XXXIX (1954), 219-226.

Unpublished Literature

Bailey, Robert Bain. "A Study of Predicting Academic Success in Elementary School Reading From Projective Tests." A Dissertation submitted to the Graduate Faculty in Partial Fulfillment of the Ed.D., University of Oklahoma, 1956.

Brown, John A. "The Ferndale Drop-Out Study: A Comparative Analysis of Values, Attitudes and School Achievements of Negro and White Youth in a Suburban Area of Detroit." Unpublished Master's Thesis, School of Social Work, Wayne State University.

Fowler, William Louis. "A Comparative Analysis of Pupil Performance on Conventional and Culture-Controlled Mental Tests." Dissertation, University of Michigan, Pub. #18 (1956), 603, 2106.

APPENDICES

APPENDIX I

HYPOTHESES AND CHI-SQUARE TABLES FOR THE PILOT STUDY

Hypotheses

1. Trained, untrained white judges, and untrained Negro judges were able to make more correct estimations on white children's drawings than they were able to make on Negro children's drawings.

2. More correct estimations were made on white children's drawings than on Negro children's drawings when judging for ability.

3. Untrained judges made more correct estimations on white children's drawings than on Negro children's drawings.

4. According to significant chi-square calculations, untrained Negro judges made more correct estimations on Negro children's drawings than did the untrained white judges.

5. Although, more correct estimates were made on white children's drawings on both grade levels, more correct estimates were made on white children's drawings on the fourth-grade level than on the first-grade level.

6. Of the nine significant chi-square tables, the set of trained judges was the only set of judges to make more correct estimations on Negro children's drawings than on white children's drawings.

7. More correct estimations were made on white children's drawings than on Negro children's drawings when judgments were based on "Tree" drawings than when judgments were based on "House" or "Person" drawings.

Chi-Square Tables

TABLE 1.--Chi-square table of frequencies of individual first-grade white and Negro children in categories of correct number of ability estimates made by untrained Negro judges on basis of tree drawings

	Both Judges	Neither Judge	Totals
White	7	4	11
Negro	2	10	12
Totals	9	14	23

Obtained $X^2 = 5.3$. Significant $\angle$ .05 level of confidence.

TABLE 2.--Chi-square table of frequencies of individual first-grade white and Negro children in categories of correct number of ability estimates made by untrained white judges on basis of tree drawings

	Both Judges	Neither Judge	Totals
White	7	5	12
Negro	2	10	12
Totals	9	15	24

Obtained $X^2 = 4.31$. Significant $\angle$ .05 level of confidence.

TABLE 3.--Chi-square table of frequencies of individual first-grade white and Negro children in categories of correct number of estimates of reading achievement made by trained white judges on basis of tree drawings

	Number Both Judges Selected Correctly	Number Neither Judge Selected Correctly	Totals
White	9	3	12
Negro	3	9	12
Totals	12	12	24

Obtained $X^2 = 6.00$. Significant $<$ .02 level of confidence.

TABLE 4.--Chi-square table of frequencies of individual fourth-grade white and Negro children in categories of correct number of estimates of ability made by untrained Negro judges on basis of tree drawings

	Number Both Judges Selected Correctly	Number Neither Judge Selected Correctly	Totals
White	8	4	12
Negro	2	10	12
Totals	10	14	24

Obtained $X^2 = 6.12$. Significant $<$ .02 level of confidence.

TABLE 5.--Chi-square table of frequencies of individual fourth-grade white and Negro children in categories of correct number of estimates of ability made by trained white judges on basis of tree drawings

	Both Judges	Neither Judge	Totals
White	1	11	12
Negro	6	6	12
Totals	7	17	24

Obtained $X^2 = 5.04$. Significant $/$ .05 level of confidence.

TABLE 6.--Chi-square table of frequencies of individual fourth-grade white and Negro children in categories of correct number of estimates of ability made by trained white judges on basis of Person I drawings

	Both Judges	Neither Judge	Totals
White	8	4	12
Negro	3	9	12
Totals	11	13	24

Obtained $X^2 = 4.19$. Significant $/$ .05 level of confidence.

TABLE 7.--Chi-square table of frequencies of individual fourth-grade white and Negro children in categories of correct number of estimates of ability made by untrained white judges on basis of Person I drawings

	Both Judges	Neither Judge	Totals
White	10	2	12
Negro	3	9	12
Totals	13	11	24

Obtained $X^2 = 8.22$. Significant $<$ .01 level of confidence.

TABLE 8.--Chi-square table of frequencies of individual fourth-grade white and Negro children in categories of correct number of estimates of emotional adjustment made by untrained white judges on basis of house drawings

	Both Judges	Neither Judge	Totals
White	8	4	12
Negro	2	10	12
Totals	10	14	24

Obtained $X^2 = 6.17$. Significant $<$ .02 level of confidence.

TABLE 9.--Chi-square table of frequencies of individual fourth-grade white and Negro children in categories of correct number of estimates of emotional adjustment made by untrained white judges on basis of Person I drawings

	Both Judges	Neither Judge	Totals
White	9	3	12
Negro	3	9	12
Totals	12	12	24

Obtained $X^2 = 6.00$. Significant $/$ .02 level of confidence.

APPENDIX II

JUDGES' RATING SHEETS AND THEIR RESPONSES

GROUP I--ABILITY

JUDGE I

Bright		Dull	
Number of drawing		Number of drawing	
157	1	110	0
168	1	5	1
91	1	24	1
124	0	153	1
185	1	170	1
78	1	45	1
148	0	112	1
47	1	163	1
88	1	155	1
115	0	102	1
145	1	12	1
180	1	105	1
14	1	147	1
59	1	187	0

95	1	34	0
39	1	26	1
54	1	31	1
42	1	7	1
184	1	62	1
152	1	103	1

Wrong: 1 White
2 Negro

Wrong: 2 White
1 Negro

GROUP I--ABILITY

JUDGE II

Bright		Dull	
Number of drawing		Number of Drawing	
168	1	24	1
88	1	124	1
170	0	31	1
45	0	5	1
152	1	147	1
148	0	153	1
62	0	7	1
180	1	147	0
105	0	102	1
95	1	155	1
59	1	163	1
42	1	12	1
39	1	103	1
54	1	115	1
145	1	14	0
184	1	187	0
157	1	112	1
91	1	34	0
78	1	110	0
185	1	26	1
Wrong:	3 White 2 Negro	Wrong:	3 White 2 Negro

GROUP I--ABILITY

JUDGE III

Bright		Dull	
Number of drawing		Number of drawing	
124	0	31	1
168	1	5	1
47	1	170	1
59	1	45	1
39	1	112	1
54	1	155	1
187	1	102	1
148	0	147	1
155	1	103	1
185	1	14	0
62	0	115	1
184	1	7	1
180	1	34	0
95	1	110	0
42	1	26	1
78	1	24	1
157	1	153	1
91	1	163	1
152	1	105	1
Wrong:	1 White 2 Negro	Wrong:	1 White 2 Negro

GROUP II--READING

JUDGE I

Good Readers		Poor Readers	
Number	of drawing	Number	of drawing
168	1	26	1
77	1	24	1
180	1	31	1
95	1	124	1
59	1	5	1
42	0	170	1
39	0	45	1
184	1	7	1
187	1	112	1
91	1	47	0
78	1	102	1
185	1	12	1
88	1	14	0
22	0	34	0
49	0	121	0
32	0	6	1
67	1	13	1
85	1	80	0
166	1	46	1
181	1	169	1
Wrong:	2 White 3 Negro	Wrong:	2 White 3 Negro

GROUP II--READING

JUDGE II

Good Readers		Poor Readers	
Number	of drawing	Number	of drawing
168	1	26	1
124	0	24	1
180	1	31	1
95	1	5	1
42	0	7	1
59	1	112	1
39	0	45	1
78	1	170	1
91	1	169	1
185	1	102	1
166	1	12	1
181	1	14	0
80	1	34	0
67	1	22	1
121	1	32	1
187	1	46	1
184	1	49	1
88	1	85	0
47	1	6	1
77	1	13	1
Wrong:	0 White 3 Negro	Wrong:	1 White 2 Negro

GROUP II--READING

JUDGE III

Good Readers		Poor Readers	
Number	of drawing	Number	of drawing
168	1	26	1
124	0	24	1
47	1	31	1
180	1	102	1
95	1	112	1
77	1	7	1
185	1	45	1
78	1	170	1
88	1	12	1
166	1	14	0
181	1	34	0
121	1	6	1
80	1	13	1
187	1	49	1
184	1	32	1
59	1	67	0
42	0	46	1
91	1	5	1
85	1	169	1
39	0	22	1
Wrong:	0 White 3 Negro	Wrong:	2 White 1 Negro

GROUP III--ADJUSTMENT

JUDGE I

Well Adjusted		Poorly Adjusted	
Number of drawing		Number of drawing	
186	1	172	0
160	1	162	1
56	1	5	1
9	0	10	1
52	1	101	1
171	1	50	1
158	0	48	0
140	1	159	0
79	1	179	1
93	1	164	1
39	-	161	1
133	1	141	0
92	1	76	0
67	1	125	1
136	0	104	0
134	0	27	1
185	1	114	1
130	0	100	0
142	0	131	1
61	1	98	1
Wrong:	2 White 4 Negro	Wrong:	3 White 4 Negro

GROUP III--ADJUSTMENT

JUDGE II

Well Adjusted		Poorly Adjusted	
Number of drawing		Number of drawing	
160	1	172	0
52	1	186	0
171	1	162	1
179	0	5	1
158	0	125	1
140	1	131	1
141	1	10	1
76	1	101	1
39	-	50	1
133	1	48	0
92	1	114	1
67	1	164	1
56	1	161	1
9	0	134	1
185	1	98	1
61	1	136	1
93	1	104	0
27	0	100	0
159	1	130	1
79	1	142	1
Wrong:	1 White 3 Negro	Wrong:	3 White 2 Negro

GROUP III--ADJUSTMENT

JUDGE III

Well Adjusted		Poorly Adjusted	
Number of drawing		Number of drawing	
186	1	162	1
160	1	5	1
171	1	136	1
52	1	159	0
9	0	48	0
56	1	10	1
158	0	114	1
179	0	161	1
76	1	164	1
61	1	104	0
92	1	100	0
39	-	50	1
142	0	172	0
79	1	93	0
140	1	27	1
134	0	125	0
141	1	131	1
133	1	130	1
101	0	125	1
185	1	98	1
Wrong:	3 White 3 Negro	Wrong:	3 White 4 Negro

GROUP IV--RACE

JUDGE I

White		Negro	
Number	of drawing	Number	of drawing
182	1	152	0
181	1	145	0
63	1	203	0
23	1	158	0
160	1	227	0
146	1	209	0
77	0	171	0
138	0	37	0
21	0	164	0
206	1	175	0
202	0	172	0
14	0	76	1
92	0	95	1
121	0	161	0
61	0	79	1
137	0	13	1
4	0	130	1
140	0	136	1
141	0	55	1
27	0	166	0
Wrong:	13	Wrong:	13

GROUP IV--RACE

JUDGE II

White		Negro	
Number of drawings		Number of drawings	
63	1	145	0
181	1	209	0
203	1	171	0
182	1	158	0
227	1	152	0
23	1	76	1
37	1	172	0
166	1	164	0
146	1	175	0
202	0	95	1
206	1	21	1
138	0	77	1
14	0	161	0
92	0	136	1
27	0	4	1
55	0	130	1
61	0	13	1
137	0	79	1
141	0	140	1
160	1	121	1
Wrong:	9	Wrong:	9

GROUP IV--RACE

JUDGE III

White		Negro	
Number	of drawing	Number	of drawing
61	0	76	1
4	0	136	1
172	1	27	1
14	0	77	1
13	0	202	1
140	0	92	1
141	0	121	1
138	0	166	0
175	1	95	1
164	1	21	1
182	1	55	1
181	1	160	0
209	1	146	0
171	1	152	0
130	0	145	0
161	1	203	0
206	1	227	0
158	1	37	0
63	1	23	0
137	0	79	1
Wrong:	9	Wrong:	9

APPENDIX III

CASE DATA FORM FOR EXPERIMENTAL AND CONTROL GROUPS

I. ABILITY

Teacher's Estimate: Record I.Q. score from permanent record. Be sure to list the name of test administered and the date.

(1) Bright (2) Average (3) Dull

II. SOCIO-ECONOMIC STATUS

(1) Upper (2) Middle (3) Lower

Comments:

III. ADJUSTMENT OR BEHAVIOR

(1) Is he well adjusted?

Does he get along well with others?

(2) Is he poorly adjusted?

Does he get along badly with other children?

(3) Is he a normally adjusted child?

Comments:

IV. READING LEVEL

Indicate reading level as:

(1) Non-reader_____ (3) Primer_____

(2) Pre-primer_____ (4) First-grade reader_____

(5) Second-grade reader_____ (6) Third-grade reader _____

Reading score_____ Test administered_____Date of Test____

Birthdate: Age: Grade: