DOCUMENT RESUME

ED 068 529 24 TM 001 894

AUTHOR Jacobs, John F.; DeGraaf, Carl A.
TITLE Expectancy and Race: Their Influences upon the Scoring of Individual Intelligence Tests. Final Report.
INSTITUTION Southern Illinois Univ., Carbondale. Dept. of Special Education.
SPONS AGENCY Office of Education (DHEW), Washington, D.C. Bureau of Research.
BUREAU NO BR-1-E-096
PUB DATE 14 Mar 72
GRANT OEG-5-71-0043(509)
NOTE 46p.

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS Attitudes; *Bias; Caucasian Students; Cultural Factors; Elementary Grades; *Examiners; *Intelligence Tests; Minority Group Children; Negroes; *Race Influences; Scoring; Socioeconomic Influences; Statistical Data; Tests; *Test Validity
IDENTIFIERS Wischler Intelligence Scale for Children

ABSTRACT

A research study is reported which investigated the influence of the variables of race (examiner and child race) and expectancy (high and low) upon the scoring of individual IQ tests. Video tapes of two children (one black, one white, 10 years old, in the 85 to 95 IQ range) being administered the Wischler Intelligence Scale for Children were shown to 32 practicing school psychologists for their evaluations. Half of the psychologists were led to believe the children were brighter than they actually were (high expectancy); the other half, that they were duller (low expectancy). The expectancy and order of child presentation (high-black, high-white, low-white, low-black, etc.) were randomized. The results indicate that in the case of expectancy white and black examiners score children (white or black) in a similar manner; i.e., in the direction of expectancy. Expectancy has the most influence upon scores obtained when children and examiners are of the same race. The implications for psychological examiners are that expectancy of student ability has a predictable influence upon the IQ attributed to that student. The question remains, however, as to whether such a small average difference in IQ scores has any educational significance. (Author/LH)

FINAL REPORT

Project No. 1-E-096
Grant No. OEG-5-71-0043(509)

EXPECTANCY AND RACE: THEIR INFLUENCES UPON THE SCORING OF INDIVIDUAL INTELLIGENCE TESTS

John F. Jacobs and Carl A. DeGraaf
Department of Special Education
Southern Illinois University
Carbondale, Illinois 62901

March 14, 1972

U. S. Department of
Health, Education, and Welfare

Office of Education
Bureau of Research

FINAL REPORT

Project No. 1-E-096
Grant No. OEG-5-71-0043(509)

EXPECTANCY AND RACE: THEIR INFLUENCES UPON THE SCORING OF INDIVIDUAL INTELLIGENCE TESTS

John F. Jacobs and Carl A. DeGraaf
Department of Special Education
Southern Illinois University
Carbondale, Illinois 62901

March 14, 1972

The research reported herein was performed pursuant to a grant with the Office of Education, U. S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are enccuraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U. S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

CONTENTS

SUMMARY

Although school psychologists spend a large proportion of their time in the administration of individual intelligence tests and important decisions regarding children are frequently based upon these tests, relatively little is known about contextual variables such as race and examiner expectancy. Tacit admission that such variables may influence test results is common and most school psychologists are aware that obtained and potential IQ scores may differ, particularly with children of minority races and low socioeconomic status. Given the importance of the decisions made regarding children's educational futures and the social-racial ramifications of recent court decisions it would appear that investigation of contextual variables is germaine. This is the report of a study which investigated the influence of the variables of race (examiner and child race) and expectancy (high and low expectancy) upon the scoring of individual IQ tests.

The procedure for this study was to video tape the WISC being administered to a white and a black child. Both of the children were in the 85 to 95 IQ range, were 10 years old, and presented subtest scores with a wide range of variance. The two resulting video tapes, each about an hour in length, were shown to thirty-two practicing school psychologists (16 black and 16 white psychologists). The instructions given to the psychologists were to evaluate the two children in their normal, professional manner. Half of the psychologists were led to believe, through the use of ficticious referral forms, that the children were brighter than they actually were (high expectancy) and the other half that the children were duller than their true IQs would indicate (low expectancy). The expectancy and order of child presentation (high-black, high-white, low-white, low-black, etc.) were randomized. The psychologists were not informed as to the real intent of the investigation until they had produced subtest and full scale scores for each of the two children.

The results of the study indicate that in the case of expectancy white and black examiners score children (white or black) in a similar manner. When a psychologist is given reason to expect a certain level of performance from a child, either high or low, the results are in the predicted direction. In other words when a psychologist evaluates children who are thought to be of low potential, he is influenced to a significant degree in that direction. Expectancy has the most influence upon scores obtained when children and examiners are of the same race. There seems to be a dimunition of the influence of expectancy when child and examiner are of different races.

The implications for psychological examiners, if the results from such an investigatory procedure can be extrapolated to the real world, are that expectancy of student ability has a predictable influence upon the IQ attributed to that student. This influence is lessened when examiner and child are of differing race. However, although statistical significance was present, the question remains as to whether such a small average difference in IQ scores has any educational significance.

INTRODUCTION

PROBLEM: It was the purpose of this project to attempt to answer two questions: First, what is the extent of the effect of experimental expectancy upon the behavior of the examiner when determining intelligence? Secondly, what is the extent of the effect of race of examiner as linked to race of students on the measurement of intelligence?

With recent mandatory special education legislation in many states (e.g., Florida, Indiana, Illinois, etc.), large numbers of children are being given individual intelligence tests to determine their eligibility for special class placement; admission to special classes of various types often being based upon an arbitrary IQ score. The importance of accurate evaluations cannot be overemphasized when they determine to such a large extent the type of educational program that an individual will receive for a period of at least several years which may ultimately affect his adjustment and contributions to society.

Further, with desegration of public schools and busing becoming more than nominal throughout the country, and with the relatively greater interaction and need for evaluation of both black and white children by specialists of both races, the problems of reliable information being developed by these specialists becomes crucial.

A more complete and objective awareness of the part expectation plays in the behavior of psychologists, psychometricians, and educators, in itself, should contribute to more objective techniques in evaluating clients.

Under ideal circumstances it would be possible that relatively direct relationships exist between expectation and subjective distortion on the part of evaluators. If such is the case, simple mathematical corrections could be developed to compensate for discrepancies in evaluations. However, it would seem more likely that techniques will have to be developed, for those portions of evaluative procedures most subject to distortion, that would enforce objectivity in evaluations.

In addition to public school applications, it was expected that data resulting from this study would have implications for higher education, vocational rehabilitation, and industry. Particular usefulness might be anticipated for the numerous Federal projects dealing with race, intelligence testing and programs involving minority groups. Recent court cases reviewed by Ross et al (1971) indicate the relevance of such an investigation to current social questions involving education and racial groups.

REVIEW OF LITERATURE:

Experimental Expectancy Effects - Recently the attention of many educators and psychologists has been attracted to the phenomenon of experimental expectancy as it affects experimental variables. Perhaps most familiar is Rosenthal's Pygmalion In The Classroom. Though recently criticized (Buckley, 1968; Gephart, 1970; etc.) the phenomenon of a "self fulfilling prophecy" has yet to be experimentally refuted or to accrue sufficient support to be put forward in more than the form of a tentative hypothesis.

Rosenthal (1969) traces the history of experimenter expectancy to - and perhaps beyond - the early 1900's with the following illustrations:

> Just about the time Alfred Binet was developing his test in France, there was taking place in Germany one of the longest, most thorough, and most famous intellectual evaluations of the century. That evaluation was of the alleged intellectual abilities of Clever Hans, the horse of Mr. von Osten. Hans' considerable talents in the fields of linguistic, mathematical, and musical analysis, it will be remembered, were contingent upon his examiners' giving him unwitting and subtle cues, cues which seemed to derive from his examiners' expectations that Hans would accurately tap out with his foot the answers to an amazing variety of questions (Pfungst, 1911: 1965).

Although most of the earlier studies involved animal-experimenter problems, investigations with human subjects began appearing in the literature during the late 1950's, with the most often cited studies being conducted in the 1960's.

Ekren (1962), reports a distinction between examiners administering the Block Design subtest of the Wechsler Adult Intelligence Scale (WAIS) to high achievers and low achievers. Although WAIS scores showed, as a result of reported procedure difficulties, no distinction between subjects, the subjects contrasted affective behaviors on the part of the examiners directly related to the expectations held by the examiners. An analysis of the subject's ratings of their examiners' behavior showed that when examiners were contacting subjects alleged to be earning higher grades they behaved in a more friendly, likeable, interested and encouraging manner, showed a more expressive face, etc. Although no significant differences existed in scores in the Ekren study, Gordon and Durea (1948) reported substantial differences in IQ scores related to affect of the examiner.

On the basis of the above studies, and several others, Rosenthal (1968) concludes ". . .it appears that (the WAIS) may not be susceptible to the effects of the examiner's expectancy . . ."

Further studies have been reported using the Rorschach (Masling, 1965; Marwit and Marcia, 1967; Strauss, 1968), the Taylor Manifest Anxiety Scale (Pflugrath, 1962), word association tests (Silverman, 1968) and several other instruments. In an analysis of 35 such experiments, Rosenthal concludes the probability that the results could have occurred at random as one in several million.

The most often cited study, and the one with most interest to educators is Rosenthal's (1966) study which appeared in Psychological Reports and provides the focus for Pygmalion In The Classroom.

Rosenthal, using 18 classrooms in the San Francisco area, suggested to teachers that 20% of the children showed, on the basis of a "test for intellectual blooming," unusual potential for intellectual gains. At the end of the experimental period, the "unusual" children (who had actually been selected at random) showed significantly greater gains in IQ (as measured by Flanagan's (1960) nonverbal test of General Ability) than did the remaining children in the control group. Rosenthal indicated that the effects of teacher expectancy were especially apparent among the younger children.

In a study dealing with the effects of experimental expectancy on latency in word association tests, Silverman (1968) found that when examiners expected longer latencies they obtained longer latencies. Silverman found a significant tendency to commit scoring errors in the direction of the expectations, but holds there was evidence to suggest such scoring errors did not account for all the effect obtained.

Rosenthal (1969) discusses an unpublished study by Bees done in 1967 wherein 60 teachers taught the meaning of a series of symbols to preschool children. Half the teachers had been led to expect good symbol-learning and half had been led to expect poor symbol learning. Assessment of the children was on the basis of an examination by a "blind" examiner who did not know what the child's teacher had been led to expect of the child's learning ability. Rosenthal maintains that the results of this study strongly support the concept of the effectiveness of teacher expectancy, since the probability of the discrepancy being due to chance was less than .000002.

Rosenthal points out that a logical explanation would be that the teacher's expectations of superior learning gave incentive for superior teaching. When an attempt was made to consider the data, holding constant for teaching behavior, the effect of expectancy was diminished, but still significant. The "teaching benefits" for which an attempt was made to control, in this case, seemed to be limited to the number of symbols taught rather than differences in teaching style. It would appear that teaching style and skill would be as important in measurements of this sort as the number of concepts the teachers attempted to teach.

In a similarly designed study treating the effect of teacher expectancy as social acceptance rather than IQ, Jacobs (1970) found no significant difference between experimental and control populations.

Rosenthal (1969) also discusses an unpublished experiment by Larrabee and Kleinsasser (1967) in which five examiners administered the Wechsler Intelligence Scale for Children to 12 sixth-graders of average intelligence. Each subject was tested by two different examiners; one examiner administering the even-numbered items and the other examiner administering the odd-numbered items. For each subject, one of the examiners was told the child was of above average intelligence while the other examiner was told the child was of below average intelligence. When the child's examiner expected superior performance the total IQ earned was 7.5 points higher on the average than when the child's examiner expected inferior performance. When only the performance subtests of the WISC were considered, the advantage to the children of having been expected to do well was less than three IQ points and could easily have occurred by chance. When only the verbal subtests of the WISC were considered, the advantage of having been expected to do well, however, exceeds 10 IQ points. The particular subtest most affected by examiners' expectancies was Information. The results of this study were especially striking in view of the very small sample size (12) of subjects employed.

Again, however, no mention was made of systematic errors which might unintentionally have been made by the examiners. This seems particularly striking when it is noted that the scores reflecting the greatest discrepancy were the verbal portions of the WISC requiring the greatest degree of examiner judgment in scoring.

Race Effects - Despite the controversy precipitated by Jensen (1969) regarding IQ and race, the most plausable response to the problem appears to be that the question cannot be answered in present circumstances (Bodmer and Cavalli-Sforza, 1970). Indeed, Havighurst and Neugarten (1962) said much the same when they suggested that 'most anthropologists and psychologists now

believe that there is no innate difference in intelligence between racial or ethnic or religious groups. There are innate differences between individuals within these groups, but the average intelligence of the groups is the same, it is thought, if the groups have equal opportunity and similar training in solving the ordinary problems of life."

Thorndike and Hagen (1955) state that the most widely used individual tests are tests of intelligence. They point out that "An IQ of 90 has a rather different meaning for a Negro child who spent his early years in a sharecropper's cabin in the rural south from what it has for the son of a local banker."

In a comprehensive review of the literature, Sattler and Theye (1966) point out the weaknesses of previous research, suggesting that most of the studies used only one or no black examiners compared to a great many white examiners. These reviewers also suggested design, sampling, and statistical problems inherent in many of the studies. Levy (1956) suggested designs involving analysis of variance should be attempted in order to facilitate meaningful interpretation: to date, these suggestions have not been implemented.

Of particular relevance to the present project was Goodenough's (1940) early discussion of the possibilities of systematic errors in test administration and in test scoring. Sacks (1952) for example, found differences in test results on the Stanford-Binet related to the relationships established between examiner and subject. In this study, differences were attributed to improved or deteriorated performance on the part of the subjects and differences due to scoring or interpretation by the examiner were not considered.

Within the general field of psychological evaluation (Masland et al, 1958; Valett, 1963; White, 1961) it is common practice for authorities to urge practicing psychological examiners to consider the student variables of race and socio-economic status. The usual rationale for this is that the psychologist may estimate intellectual potential as opposed to the obtained or indicated test scores. Nalven et al (1969) found that "a child's social class background, to a great extent, and his race, to a lesser extent significantly shape psychologists' judgments as to whether his obtained IQ scores are representative." Although this study looked at student race and WISC IQ score relationships, examiner race was not considered.

METHODS

INSTRUMENT: The Wechsler Intelligence Scale for Children was the individual intelligence test used in this study. Each examiner was asked to score, in the manner described below, the Information, Comprehension, Arithmetic, Similarities, Vocabulary, Picture Completion, Picture Arrangement, Block Design, Object Assembly, and Digit Symbol subtests (all but Maze's). Each examiner was required to complete the calculation of the verbal, performance and full scale intelligence quotients, having been given the birth date of the individuals to be scored and the date of the testing.

SUBJECT EXAMINERS: The examiners were thirty-two individuals fully trained, certified and experienced in the administration of the WISC. The examiners were selected from the staffs of public school districts, public and private agencies. Included were school psychologists, clinical psychologists and psychometricians. Sixteen of the examiners were Negroid and sixteen were Caucasoid. Attempts were made to stratify the examiner sample so that examiner characteristics would correlate well with the National Survey of School Psychologists (1971). Factors considered to be relevant here were examiner sex, age, experience, type of employment and level of professional preparation.

STUDENTS: Two male children, one black, one white, were used as the students whose responses on the WISC were scored by the subject examiners. The students were both ten years old and had IQ's, as recorded in their school records, within the 80 to 90 range. Students in this range were selected since according to Jacobs and others (1967) (Cunningham and others, 1969), it is from this group that children are most often erroneously assigned to special education classes and it is with children in this IQ range that teachers most often experience difficulties in such areas as academic success, emotional problems, absenteeism, and attention seeking. Both of the students' WISC profiles were characterized by a good range of variability within the subtest scores.

PROCEDURES: Examiners were invited to participate in an experiment involving the use of TV tape for the scoring of intelligence tests. It was, when necessary, explained further that the study was designed to attempt to eliminate systematic errors in test administration through the use of TV tape and as a potential teaching device for use with psychologists in training. At the conclusion of the testing, the complete purpose was more carefully explained to the participants and they were asked not to discuss the full nature of the study to other potential participants. Each examiner received an honorarium of $100 for his participation in the study.

A production quality TV tape was made of each student as he took the WISC. The WISC was administered by a qualified, experienced psychologist who did not appear on the TV tape but was heard administering the test. The resulting tapes were each less than an hour in length. Average examiner time required to complete the entire testing and scoring task was a little over three hours. As a lead to each tape one of two fictitious case histories (see Appendix A) of the child was read by the subject examiner resulting, in effect, in four different presentations consisting of a case description followed by the test administration. One case description presented the subject as coming from a middle class home, as having a history of doing well in school, liking school and his teachers. This was followed by the administration of the WISC. The second case description presented a case study wherein the subject was depicted as coming from an impoverished, low income area, as doing poorly in school, and as disliking his school and his teachers. This was followed by the segment depicting the administration of the WISC.

Each subject-examiner individually viewed and scored two tapes; one of a black and one of a white student. Whether the particular referral was of a student presented as having done well or poorly in school and whether a black student or a white student was seen first, was randomly determined.

During their viewing of the two test administrations, each examiner was provided with all of the equipment necessary to score the test; scoring protocols, stop watch, manual, etc. Copies of student completed Coding subtest of the WISC were reproduced and available to the examiner. After all test scoring was completed the examiners were asked to respond to a brief questionnaire (see Appendix B) which included demographic questions as well as questions about the usefulness of the video tape procedures.

STATISTICAL DESIGN: Basically, the major analysis of the data is a plot x treatment, 2 x 2 x 2 factorial analysis of variance (Mendenhall, 1968; 1969), on each subtest as well as on the verbal, performance, and overall IQ scores (see Figure I). Significance in this study is defined at the .05 level of confidence.

RACE OF SUBJECT	CAUCASIAN		NEGROID	
ATTITUDE TOWARDS SCHOOL	GOOD	POOR	GOOD	POOR
RACE OF EXAMINER	1		1	
	2			2
Caucasian		3	3	
		4		4
(saw Black Subject First)	5		5	
	6			6
		7	7	
		8		8
	9		9	
	10			10
		11	11	
		12		12
(saw White Subject First)	13		13	
	14			14
		15	15	
		16		16
	17		17	
	18			18
Negroid		19	19	
		20		20
(saw Black Subject First)	21		21	
	22			22
		23	23	
		24		24
	25		25	
	26			26
		27	27	
		28		28
(saw White Subject First)	29		29	
	30			30
		31	31	
		32		32

Figure I: Model of Analysis of Variance used. Numbers indicate examiner numbers.

HYPOTHESES: The hypotheses to be tested were:

1. There is no significant difference in IQ scores of Negro subjects when the test is scored by Negro or Caucasian examiners.

2. There is no significant difference in IQ scores of white subjects when the test is scored by Negro or Caucasian examiners.

3. There is no significant difference in IQ scores when examiners believe a subject does well in school and when examiners believe a subject does poorly in school.

4. There are no significant interactions between race of subject, race of examiner and/or expectancy on the IQ score.

RESULTS

EXAMINER CHARACTERISTICS: Self report data obtained from the examiners after they had finished scoring both students, indicates that the examiners mean age was 37 and their mean number of years experience as a school psychologist was 4.5. Nearly all the examiners were at the masters level in terms of academic preparation (28 of 32) and they averaged 117 individual psychological evaluations per year in their present employment. The examiner subjects were balanced (50-50) in terms of sex as well as race. The sample favored large urban school districts over suburban and county districts by a two to one margin. All but three examiners scored the video tapes as requested without skipping any subtests. In the cases of the three who objected to one or more subtest inclusion in the Full Scale IQ, they were requested to finish the subtest as best they could in spite of what they considered to be an error in test administration. Thus all data cells were completed for purposes of data analysis.

The major analyses in the study consisted of a factorial Analysis of Variance on the intelligence quotients (Full scale, Verbal and Performance) derived by the subject examiners followed by analysis of subtest scores. Factors considered in the analysis of variance were race of subject, race of examiner and expectancy and the first and second order interactions of these factors.

FULL SCALE: The results of the analysis of variance of full scale scores are presented in Table I. Using the data presented in Table I, the following inferences are drawn.

A. There was no significant difference in the scoring of the WISC between the black and white subject examiners (mean Full Scale IQs were 88.8 and 88.6 respectively).

B. There was no significant difference in the Full Scale IQs of the students (white student's = 88.6, black student's = 88.6).

C. There was a significant ($P < .05$) difference in Full Scale IQ between the high expectancy (IQ = 98.6) and low expectancy (IQ = 87.9) conditions, suggesting the positive expectations developed in the student's referral form affected the scoring of the WISC by the subject examiners.

D. None of the first order interactions approached significance.

Table I: Factorial Analysis of Variance of Full Scale Wechsler Intelligence Scale Scores

SOURCE	SS	df	ms	F	p
TOTAL	794	63	-	-	-
Between Subjects	215	15	-	-	-
Race of Examiner	.5626	1	.5626	.0367	ns
$Error_b$	214.4374	14	15.317	-	-
Within Subjects	579	48	-	-	-
Race of Subject	1	1	1	.0539	ns
Expectancy	42.25	1	42.25	5.0909	<.05
Race of Examiner X Race of Subject	5.0624	1	5.0624	.2727	ns
Race of Examiner X Expectancy	3.0624	1	3.0624	.3690	ns
Race of Subject X Expectancy	9.0000	1	9.0000	1.1514	ns
Race of Examiner X Race of Subject X Expectancy	33.0626	1	33.0626	4.2296	<.10
$Error_1$	259.938	14	18.567	-	-
$Error_2$	116.188	14	8.2991	-	-
$Error_3$	109.436	14	7.8169	-	-

E. The second order interaction, race of examiner X race of student X expectancy approached ($P < .10$) but failed to achieve significance. This interaction is shown in Figure II. It should be noted that expectancy seems to have its most dramatic effect, in the predicted direction, when examiners are evaluating students of their own race. There is no apparent effect when examiner and subject are of a different race.

VERBAL SCALE: The analysis of the verbal scale is presented in Table II. As in the analysis of Full Scale Scores, only the factor of expectancy achieved significance ($P < .001$). However, both race of student and the second order interaction of race of examiner X race of student X expectancy (see Figure III) approached significance. It is interesting to note that on the verbal portion black examiners score high expectancy white students lower than low expectancy students.

On the Verbal scale, as on the Full Scale, expectancy differences were much greater in the predicted direction, when examiner and student were of the same race. When examiner and student were of a different race, the magnitude of the effect was greatly dimished and, in fact, was not in the predicted dir direction for black examiner scoring white students.

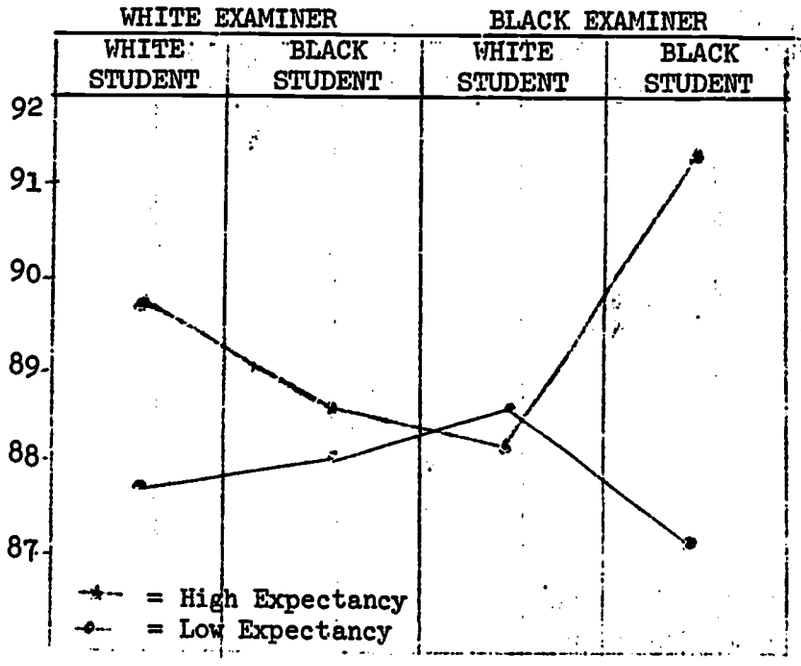

Figure II: Full Scale IQ Showing Race of Examiner X Race of Subject X Expectancy ($P < .10$) Interaction

Table II: Anova = Verbal

SOURCE	SS	df	ms	F	p
TOTAL	1201.7344	63	-	-	-
Between Subjects	331.9844	15	-	-	-
Race of Examiner	.1407	1	.1407	.0059	ns
$Error_b$	331.845	14	23.7032	-	-
Within Subjects	869.75	48	-	-	-
Race of Subject	92.6407	1	92.6407	3.5905	ns
Expectancy	125,7969	1	125,7969	26.5087	<.001
Race of Examiner X Race of Subject	21.3905	1	21.3905	.8290	ns
Race of Examiner X Expectancy	.0158	1	.0158	.0033	ns
Race of Subject X Expectancy	6.8905	1	6.8905	.6112	ns
Race of Examiner X Race of Subject X Expectancy	37.5158	1	37.5158	3.3275	ns
$Error_1$	361.2182	14	25.8013	-	-
$Error_2$	66.4373	14	4.7455	-	-
$Error_3$	157.8443	14	11.2746	-	-

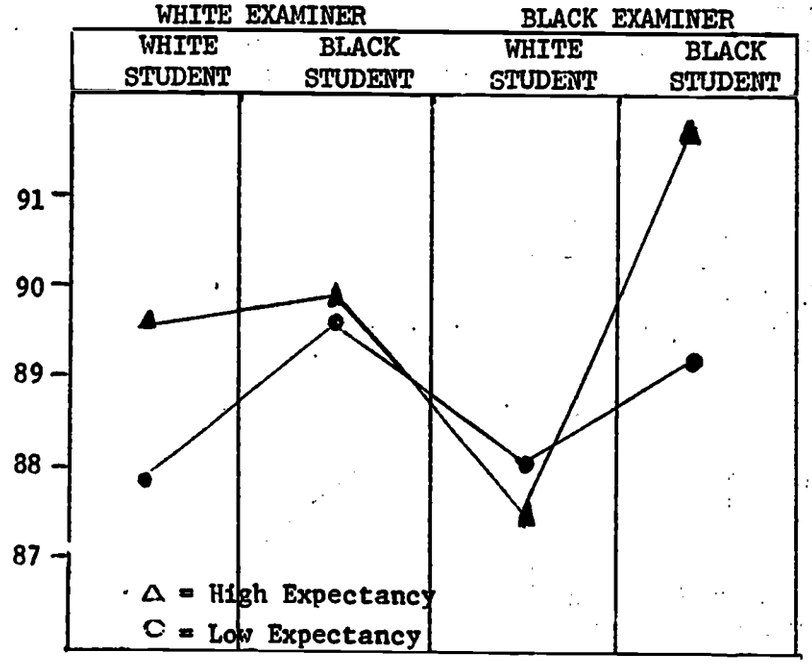

Figure III: Verbal

PERFORMANCE: When scores on the Performance portion of the WISC were analyzed, significant differences were noted for race of student as well as for expectancy (see Table III). Performance IQ for the white student was 90.3 and for the black student 88.1 ($P < .03$). Performance IQ under the condition of high expectancy was 90.0 and low expectancy was 88.3 ($P < .05$).

It can be assumed the difference found for race of student was an actual difference in the performance of the two students. However, the difference in expectancy ($P < .05$) can only be attributed to experimental treatment differences. None of the interactions on Performance approached significance.

It can be seen from Figure IV, the phenomenon noted on the Full Scale and Verbal portions, of magnification of expectancy effect when student and examiner are of the same race, is present on the Performance protion of the WISC, though not of as great a magnitude.

SUBTEST SCORES: The analysis of variance tables for subtest scores are included in the appendix (see Appendix C). Each subtest is discussed below.

Information - The analysis of variance on the Information subtest scores indicates a significant difference ($P < .001$) between the black and white student. Race of examiner and expectancy both approached ($P < .10$) significance.

Table III: Anova = Performance

SOURCE					
TOTAL	845.1094	63	-	-	-
Between Subjects	178.3594	15	-	-	-
Race of Examiner	3.5157	1	3.5157	.2815	ns
$Error_b$	174.8437	14	12.4888	-	-
Within Subjects	666.75	48	-	-	-
Race of Subject	78.7657	1	78.7657	6.1831	<.03
Expectancy	47.2657	1	47.2657	5.9766	<.05
Race of Examiner X Race of Subject	.1405	1	.1405	.0110	ns
Race of Examiner X Expectancy	8.2655	1	8.2655	1.0451	ns
Race of Subject X Expectancy	9.7655	1	9.7655	.6316	ns
Race of Examiner X Race of Subject X Expectancy	17.0158	1	17.0158	1.1005	ns
$Error_1$	178.3438	14	12.7388	-	-
$Error_2$	110.7188	14	7.9085	-	-
$Error_3$	216.4687	14	15.4621	-	-

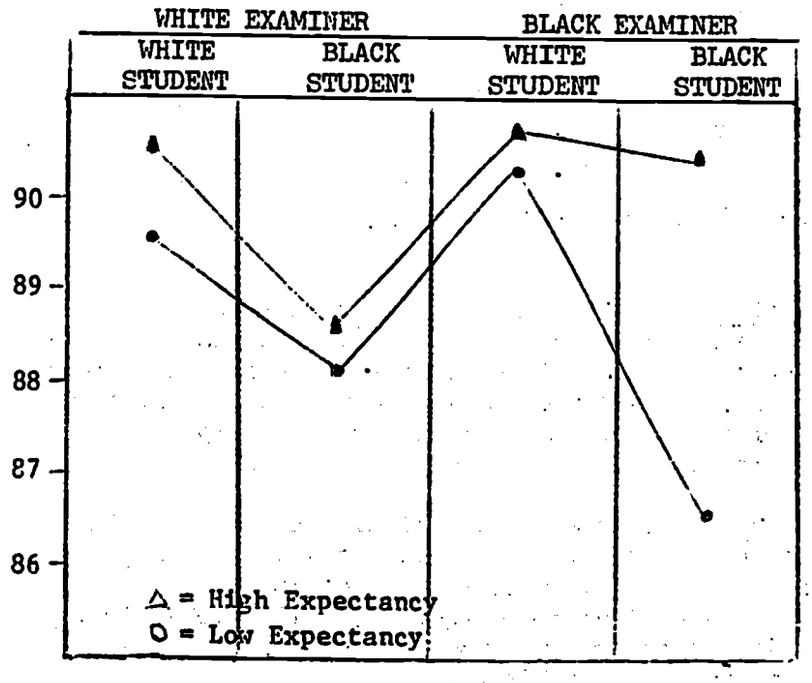

Figure IV: Performance

Comprehension - As on the Information subtest, there was a significant ($P < .01$) difference between the black and white students on Comprehension (white students means scaled score = 69.0; black students = 50.0). The interaction race of student X race of examiner approached ($P < .10$) but failed to achieve significance.

Arithmetic - On the Arithmetic subtest, the second order interaction, race of student X race of examiner X expectancy was highly significant ($P < .001$). Figure V illustrates this interaction. Race of examiner X expectancy approached ($P < .01$) significance.

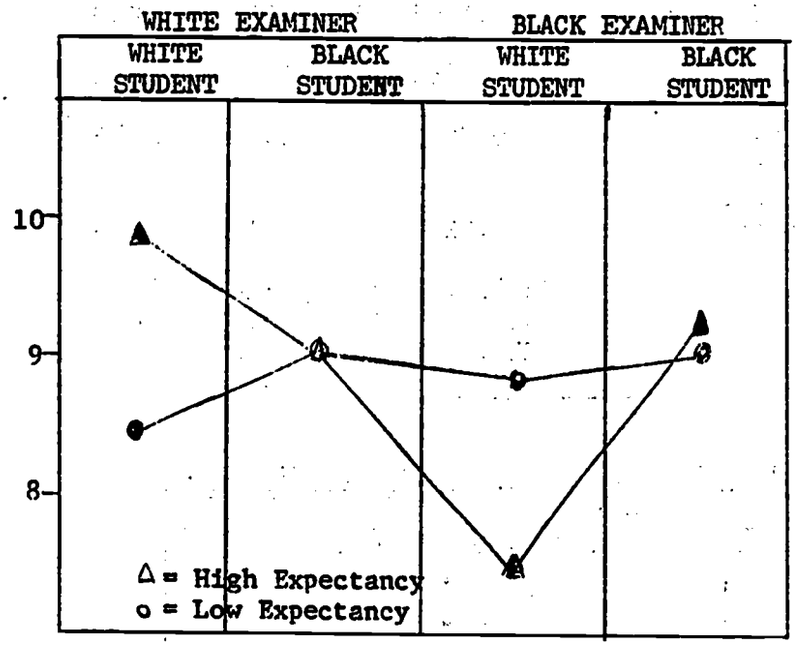

Figure V: Scaled Scores. Arithmetic Subtest

Similarities - On the Similarities subtest, only race of student achieved significance ($P < .0001$). The black student had a mean scaled score of 10.97 and the white student a mean scaled score of 8.0. The race of examiner X race of student ($P < .10$) interaction and the race of student X expectancy ($P < .06$) interaction both approached significance.

Vocabulary - There was again a significant ($P < .001$) difference between black and white student's Vocabulary, the black student having a mean scaled score of 10.0 and the white student a 7.5. Race of examiner approached ($P < .10$) significance, with black examiners scoring the vocabulary items lower and white examiners scoring higher.

Digit Span - There was a significant difference ($P < .05$) in expectancy in the Digit Span subtest: Mean scaled score under the condition of high expectancy was 6.18 and for low expectancy 5.875. The race of examiner X expectancy interaction approached ($P < .10$) significance.

Picture Completion - Only the race of student provided a significant difference ($P < .001$) on the Picture Completion subtest. The white student had a mean scaled score of 6.43 and the black student 10.78.

Picture Arrangement - Again, only the race of student provided a significant ($P < .001$) difference. Mean scaled score on the Picture Arrangement subtest for the white student was 51.0 and for the black student 7.13.

Block Design - The black student had a mean scaled score of 9.2 on the Block Design subtest and the white student a mean scaled score of 10.1. This difference was highly significant ($P < .001$). None of the interactions approached significance.

Object Assembly - As with most of the performance subtests, the only significant factor ($P < .001$) in the analysis of the Object Assembly subtest was the race of student. In this case the mean scaled scores were: white student = 10.53 and black student = 7.56.

Coding - The Coding subtest provided significant differences in race of student ($P < .001$), expectancy ($P < .05$), and race of examiner X expectancy ($P < .05$). Mean scaled score for the white student was 9.6 and for the black student 6.8. Under the condition of high expectancy the mean scaled score was 8.53 and under low expectancy was 7.9. The interaction (race of examiner X expectancy) is shown in Figure VI. Expectancy seemed to exert a significantly greater influence on the scoring of the Coding subtest by black examiners than by white examiners. As has been previously suggested this phenomenon is especially apparent when scoring subjects of different races.

		WHITE EXAMINER	BLACK EXAMINER
High Expectancy	White Student	9.75	10.25
	Black Student	6.75	7.5
	Mean SS	8.25	8.81
Low Expectancy	White Student	9.625	9.125
	Black Student	6.875	6.00
	Mean SS	8.25	7.56

Figure VI: Race of examiner X expectancy on Coding subtest. Numerals represent mean subtest scaled scores.

CONCLUSIONS

On the basis of the data analyzed in the present study, the following conclusions can be drawn:

1. There is no evidence to support the hypothesis that Negro examiners score white students differently than Caucasian examiners.

2. When considering only main effects there is no evidence to support the hypothesis that Caucasian examiners score black students differently than Negro examiners.

Previous research of questionable design has reported students performing differently as a function of race of examiner. Performance of examiners as a function of race of examiner and/or race of student has not been reported. The race of examiner X race of student interaction was not significant in any of the analyses performed: Full Scale, Verbal, Performance nor any of the subtests.

3. Expectancy had a significant effect on the scores derived by the examiners. When the psychologists were led to expect high achievement and ability in the child being tested the scores were significantly higher than when they were led to expect low ability achievement. Expectancy had a significant effect on Full Scale IQ, Performance IQ and Verbal IQ. The difference between high and low expectancy condition scores was especially apparent in the Verbal portion of the WISC (achieving an F of 26.5 with 1 and 14 degrees of freedom; P .001) suggesting that the Verbal scale is more subjective than the Performance scale.

This finding should be considered when, for example, the Larrabee and Kleinsasser (1967) study is evaluated. Larrabee and Kleinsasser found, with individually administered WISCs, an elevation of performance of the students on the Verbal portion under conditions of positive expectancy. However, no attempt would seem to have been made to control for the effect of positive expectancy in the scoring behavior of the examiners.

4. Apparently the most meaningful interaction considered was race of examiner X race of student X expectancy. The data suggest that expectancy has its most noticeable effect when examiners are testing children of their own race. When white examiners scored black students or when black examiners score white students, the effects of expectancy were diminished. This effect, of course, would tend to reduce the main effect of expectancy and would suggest the reexamination of relevant expectancy studies when race was not controlled for.

RECOMMENDATIONS

The results of this investigation seem to suggest that differences between black and white psychologists on the scoring of black and white students is minimal. In view of this, it is suggested that those psychologists (e.g., Williams, 1970) contending that only black psychologists be used in the evaluation of black subjects may want to reconsider their position in light of empirical data. This recommendation would seem especially germane in light of the tendency of the psychologists in the present study to be more influenced by variables other than performance in the examining situation when dealing with children of their own race. It is also recommended that a similar investigation be undertaken with evaluations of personality variables using projective techniques; techniques reportedly more susceptible to subjective unreliability.

Although the magnitude of difference between high and low expectancy conditions would not seem to be educationally important, it must be remembered that the laboratory conditions of the present study may have minimized expectancy difference which may be found under field conditions. It would seem important that consideration be given to the following recommendations when results of intelligence testing may have dramatic effects on the subject: as when the subject is being evaluated for institutionalization or removal from regular academic situations for placement in special classes for the retarded.

1. The use of "blind" evaluations by another psychologist or psychometrician who is not familiar with the reason for referral; with interpretation of the results then being made by the psychologist who can integrate the findings into a knowledge of other relevant variables before making recommendations.

2. The testing only of "other race" students (since the effect of expectancy appeared to be minimized under these conditions).

It is the opinion of the investigators that the most important finding of the study is the remarkable similarity of IQ scores derived for the two students by dissimilar examiners. This would suggest, if the sample studied was representative of the general population and if the evaluation procedure is a valid measure of school psychologists' "in the field" behavior, then students are being reasonably accurately assessed regardless of race.

Furthermore, the actual procedure of video presentation of students for psychological evaluation has much to offer in terms of an examiner training and standardizing technique. It is recommended that college and university programs involved in the preparation of psychologists for the evaluation of children consider the potential of this technique for inclusion in their training programs.

24

REFERENCES

Blackwood, B. A study of mental testing in relation to anthropology. Mental Measurements Monographs, 1927, Serial No. 4

Bodmer, Walter F. and Cavalli - Sforza, Luigi Luca. Intelligence and Race. Scientific American, 1970, 223-4, 19-29.

Buckley, James J. Who Is Pygmalion, Which is Galatea? Phi Delta Kappan, L, 124

Cunningham, Myron A., Pierce, Marnell L., and Jacobs, John F. Prevalence of emotional disturbance as the primary disability in children in classes for the mentally retarded, Paper presented at Council for Exceptional Children, annual meeting, Denver, Colorado, 1969: Also in press.

Ekren, U. W. The effect of experimenter knowledge of a subject's scholastic standing on the performance of a reasoning task. Unpublished master's thesis, Marquette University, 1962

Farling, W. H. and Hoedt, K. C. National, Regional, and State Survey of School Psychologists. U. S. Department of Health, Education and Welfare. (Office of Education) Final Report, Project No. 9-E-150. February 1971

Flanagan, J. C. Tests of General Ability: Technical Report. Chicago: Science Research Associates, 1960

Gephart, William J. Will the real Pygmalion please stand up. American Educational Research Journal, A70; VII, 473-475

Goodenough, F. L. Some problems of nature-nurture research. Yearbook of the National Society for the Study of Education, 1940, 39(1), 367-384.

Gordon, L. V. and Durea, M. A. The effect of discouragement on the revised Stanford-Binet scale. Journal of Genetic Psychology, 1948, 73, 201-207.

Havighurst, Robert J. and Neugarten, B. L. Society and Education (2nd Edition) Allyn and Bacon, Inc., Boston, 1962

Jacobs, John F., Cunningham, Myron, Pierce, Marnell, and Cortazzo, Arnold. Slow Learner: Problem in The Classroom, Fla. Educ. Research and Development Council, Gainesville, Fla. 1967

Jacobs, J. F. Teacher expectancies: Their effect upon poor acceptance, Paper presented before American Educational Research Association, Minneapolis, Minn., 1970. Also in Press.

Jensen, A. R. How much can we boost IQ and Scholastic Achievement? Harvard Educational Review, 1969a, 39, 1-123

Larrabee, L. L. and Kleinsasser, L. D. The effect of experimenter bias on WISC performance. Unpublished paper, Psychological Associates, St. Louis, 1967

Levy, L. H. A note on research methodology used in testing for examiner influence in clinical test performance. Journal of Consulting Psychology 1956, 20, 286

Marwit, S. J. and Marcia, J. E. Tester bias and response to projective instruments. Journal of Consulting Psychology, 1967, 31, 253-258

Masland, R. L., Sarason, S. B. and Gladwin, T. Mental Subnormality, New York: Basic Books, 1958

Masling, J. Differential indoctrination of examiners and Rorschach responses. Journal of Consulting Psychology, 1965, 31, 153-158

Mendenhall, William. The Design and Analysis of Experiments, Wadsworth Publishing Co., Inc., Belmont, Calif., 1968

Mendenhall, William. Personal communication, 1969

Nalven, F. B., Hofmann, L. J. and Bierbryer, B. The effects of subjects age, sex, race, and socioeconomic status on psychologists estimates of "true IQ" from WISC scores. Journal of Clinical Psychology, 1969, 25, 3, 271-274

Pflugrath, J. Examiner influence in a group testing situation with particular reference to examiner bias. Unpublished master's thesis, University of North Dakota, 1962

Pfungst, O. Clever Hans (the horse of Mr. von Osten): A Contribution to Experimental, Animal, and Human Psychology. (Translated by C. L. Rahn) New York: Holt, 1911, New York, Holt, Rinehart and Winston, 1965

Rosenthal, R. Pygmalion in the Classroom, New York: Holt, Rinehart and Winston, 1968

Rosenthal, Robert Personal Communication, 1969

Rosenthal, R. and Jacobson, Lenore Teacher's expectancies; Determinants of pupils' IQ gains, Psychological Reports, 1966, 19, 115-118

Ross, S. L., DeYoung, H. G. and Cohen, J. S. Confrontation: Special Education and the law. Exceptional Children, 1971, 38, 5-12

Sacks, E. L. Intelligence scores as a function of experimentally established social relationships between child and examiner. Journal of Abnormal and Social Psychology, 1952, 47, 354-358

Sattler, J. M. Statistical reanalysis of Canady's "The effect of 'rapport' on the IQ: A new approach to the problem of racial psychology." Psychological Reports, 1966, 19, 1203-1206

Wilverman, I. The effect of experimenter outcome expectancy on latency of word association. Journal of Clinical Psychology, 1968, 24, 60-63

Strauss, M. E. Examiner expectancy: Effects of Rorschach experience balance. Journal of Consulting Psychology, 1968, 24, 60-63

Thorndike, Robert L. and Hagen, Elizabeth, Measurement and Evaluation in Psychology and Education. J. Wiley and Sons, Inc., New York, 1953

Valett, R. E. The Practice of School Psychology. New York, Wiley, 1963

White, M. A. and Harris, M. W. The School Psychologist. New York, Harper & Row, 1961

Williams, R. L. Blackpride, academic relevance and individual achievement. Counseling Psychologist, 1970 2(1) 18-22

APPENDIX A

Referral Forms

XXXXXXXXXX PUBLIC SCHOOLS
PSYCHOLOGICAL SERVICE REFERRAL FORM

Name Lenny XXXXXXX Date October 1, 1971

Address XXXXXXXXXXXXX Phone XXX XXXX

School Marquette Elementary Grade 4th Birthdate 9/4/61

Teacher Mrs. Gloria Simmons Principal Mr. Frank Crowell

Reason for referral (State the problem briefly) Lenny is being considered for the advanced fourth grade room. He is a hard worker and deserves the chance. At a recent staff meeting it was agreed that we need an outside opinion on how to deal with the special sort of problems which he presents for placement.

FAMILY	Name	Address	Occupation
Father	George	XXXXXXXXXXXX	Painter
Mother	Roberta	XXXXXXXXXXXX	Housewife

Number of children 4 Position of child in family group 3

Comments on family (Marital status, others living in house, etc.) Lenny comes from a perfectly wonderful family. The parents have been very cooperative with the school and seem to be supportive with him. The parents report that the entire family works together on school assignments.

HEALTH:

Name of family physician Dr. E. L. Russell Comments on physical problems (vision, hearing, coordination, etc.) Perfect as far as school is concerned. He excels in all sorts of physical activities and is a leader of the other children on the playground.

Has pupil been seen by the Child Guidance Clinic or other agency? No If yes, please attach report to this form.

TESTING INFORMATION

INTELLIGENCE TESTS (Individual or Group)

Date	Test	Grade when given	CA	MA	IQ	Other Results
69	Peabody	2nd	8	9-7	113	

ACHIEVEMENT TESTS

Date	Test	Grade when given	Age	Read. Level	Arith. Level	Other Results
68	Iowa	1st	7	2.3	2.7	
70	Iowa	3rd	9	4.8	4.9	

1. Make out in duplicate. Send original copy to Office of Pupil Personnel Services, Administration Building.
2. Use reverse side for supplying any additional information.
3. Referral assigned to Clinic Date Oct 4, 71
   (To be completed by Office of Pupil Personnel Services)

Tested on October 12, 1971

XXXXXXXXXX PUBLIC SCHOOLS
PSYCHOLOGICAL SERVICE REFERRAL FORM

Name Lenny XXXXXXXXX Date October 1, 1971

Address XXXXXXXXXXXXX Phone XXX XXXX

School Marquette Elementary Grade 4th Birthdate 9/4/61

Teacher Mrs. Gloria Simmons Principal Mr. Frank Crowell

Reason for referral (State the problem briefly) Lenny is so far behind the rest he is a very real problem to himself and the other children. I question the results of the group tests and the wisdom of his continuing in the regular grades. Individual testing, I am sure, will indicate the proper placement for him.

FAMILY	Name	Address	Occupation
Father	George	XXXXXXXXXXXX	Janitor
Mother	Roberta	XXXXXXXXXXXX	Housewife

Number of children 4 Position of child in family group 3
Comments on family (Marital status, others living in house, etc.) Lenny's family seems to care little for him or his problems. Other youngsters from this family have received similar treatment when they experienced academic difficulties. The parents seem to have enough trouble of their own with him and won't listen to us regarding school problems.

HEALTH:
Name of family physician Dr. E. L. Russell Comments on physical problems (vission, hearing, coordination, etc.) There are no records of any major problems. I observe, however, that he is usually very tired and he is one of the most accident prone children I have ever had in my class.

TESTING INFORMATION

INTELLIGENCE TESTS (Individual or Group)

Date	Test	Grade when given	CA	MA	IQ	Other Results
69	Peabody	2nd	8	5-8	72	

ACHIEVEMENT TESTS

Date	Test	Grade when given	Age	Read. Level	Arith. Level	Other Results
68	Iowa	1st	7	Kg. 2	Kg. 3	
70	Iowa	3rd	9	1.6	1.8	

1. Make out in duplicate. Send original copy to Office of Pupil Personnel Services, Administration Building.
2. Use reverse side for supplying any additional information.
3. Referral assigned to Clinic Date Oct, 4, 71
   (To be completed by Office of Pupil Personnel Services)

Tested on October 12, 1971

XXXXXXXXX PUBLIC SCHOOLS
PSYCHOLOGICAL SERVICE REFERRAL FORM

Name Gary XXXXXXXXX Date Spetember 24, 1971

Address XXXXXXXXXXXX Phone XXX XXXX

School Burton El. Grade 4th Birthdate 12/3/60

Teacher Mrs. Priscilla Duncan Principle Mr. David Jarvis

Reason for referral (State the problem briefly) I am convinced that Gary would be better placed in the advanced section of 4th grade. His parents and I agreed to let an outsider evaluate him and to be the judge of such a placement. Gary is a warm and helpful person in the class. I want to do all I can to let him use his high potential.

FAMILY	Name	Address	Occupation
Father	Floyd	XXXXXXXXXX	General telephone
Mother	Margaret	XXXXXXXXXX	Housewife

Number of children 4 Position of child in family group 2
Comments on family (Marital status, others living in house, etc.) Gary's family is well thought of in the community and at school. I had his older sister (a fine student) tow years ago so I have a good relationship with the family.

HEALTH:
Name of family physician Dr. Wm. McCarthy Comments on physical problems (vision, hearing, coordination, etc.) No known physical or emotional problems. Gary is a very hard worker and is very anxious to please the teacher. I sometimes think he is too eager to please.

Has pupil been seen by the Child Guidance Clinic or other agency? No If yes, please attach report to this form.

TESTING INFORMATION

INTELLIGENCE TESTS (Individual or Group)

Date	Test	Grade when given	CA	MA	IQ	Other results
69	Pea-body	2nd	8	9-5	111	

ACHIEVEMENT TESTS

Date	Test	Grade when given	Age	Read. Level	Arith. Level	Other Results
68	Iowa	1st	7	2.4	2.9	
70	Iowa	3rd	9	4.7	5.1	

1. Make out in duplicate. Send original copy to Office of Pupil Personnel Services, Administration Building.
2. Use reverse side for supplying any additional information.
3. Referral assigned to Clinic Date Sept. 27, 71
   (To be completed by Office of Pupil Personnel Services)

Tested on October 4, 1971

XXXXXXXXX PUBLIC SCHOOLS
PSYCHOLOGICAL SERVICE REFERRAL FORM

Name Gary XXXXXXXXX Date September 24, 1971

Address XXXXXXXXXXXXX Phone XXX XXXX

School Burton E. Grade 4th Birthdate 12/3/60

Teacher Mrs. Priscilla Duncan Principal Mr. David Jarvis

Reason for referral (State the problem briefly) Gary is in severe academic difficulty. His level of competence is well below that of others in the classroom. Because of this he is a constant source of disruption which hurts the other children as well as himself. I think he cheats on group tests. He needs help from someone other than myself.

FAMILY	Name	Address	Occupation
Father	Floyd	XXXXXXXXXX	Gas Station
Mother	Margaret	XXXXXXXXXX	Housewife

Number of children 4 Position of child in family group 2

Comments on family (Marital status, others living in house, etc.) Gary was very likely an unwanted child like his younger siblings. The family (if that's what you call it) seems to run in all directions at once as the parents offer little direction over or concern for the children. All of the children have had difficulties in school and some are in special classes.

HEALTH:

Name of family physician Dr. Wm. McCarthy Comments on physical problems (vision, hearing, coordination, etc.) Gary is always tired and has a running nose. Although he is not sick; he sure seems anemic or something. Sometimes I wonder if he hears me when I speak to him.

Has pupil been seen by the Child Guidance Clinic or other agency? No If yes, please attach report to this form.

TESTING INFORMATION

INTELLIGENCE TESTS (Individual or Group)

Date	Test	Grade when given	CA	MA	IQ	Other Results
69	Peabody	2nd	8	5-7	70	

ACHIEVEMENT TESTS

Date	Test	Grade when given	Age	Read. Level	Arith. Level	Other Results
68	Iowa	1st	7	Kg. 1	Kg. 4	
7-	Iowa	3rd	9	1.7	1.8	

1. Make out in duplicate. Send original copy to Office of Pupil Personnel Services, Administration Building.
2. Use reverse side for supplying any additional information.
3. Referral assigned to Clinic Date SEPT, 27/1
   (To be completed by Office of Pupil Personnel Service)

Tested on October 4, 1971

APPENDIX B

Examiner Questionnaire

This data will be used only in the statistical analysis of the study

Age _____ Sex _____ Highest professional degree _____

Years Experience as Psychologist _____

Approximate number of individual Intelligence tests administered

in past year _____ in lifetime _____

What degree of accuracy would you attach to the IQ Score you obtained?

within plus or minus 2.5 points _____
" " " " 5 " _____
" " " " 10 " _____
" " " " 15 " _____
more than " " " 20 " _____

Type of employment you now hold (i.e. school district, clinic, private practice, etc.) _____

How would you approximate the racial proportion of the clients you evaluate (i.e. 60% white, 30% Black, 10% Mexican American, etc.) _____

How would you rate your clinical skills in relation to others in the field in which you are employed?

well above average _____
above average _____
slightly above average _____
average _____
slightly below average _____
below average _____
well below average _____

Briefly relate your opinions regarding the use of video tape equipment for the purpose of training psychologists? _____

in scoring individual IQ tests? _____

Briefly relate your feelings about this study and/or the implications of it.

APPENDIX C

ANOVA Tables For Subtests

Information

SOURCE	SS	df	ms	F	P
TOTAL	27.	63	-	-	-
Between Subjects	2.5	15	-	-	-
Race of Examiner	.5625	1	.5625	4.0643	<.10
Error$_b$	1.9375	14	.1384	-	-
Within Subjects	24.5	48	-	-	-
Race of Subject	14.0625	1	14.0625	89.9712	<.001
Expectancy	.5625	1	.5625	3,5988	<.10
Race of Examiner X Race of Subject	.25	1	.25	1.5995	ns
Race of Examiner X Expectancy	.25	1	.25	.9491	ns
Race of Subject X Expectancy	.0	1	0.	0.	ns
Race of Examiner X Race of Subject X Expectancy	.5625	1	.5625	2.6811	ns
Error$_1$	2.1875	14	.1563	-	-
Error$_2$	3.6875	14	.2634	-	-
Error$_3$	2.9375	14	.2098	-	-

Comprehension

SOURCE	SS	df	ms	F	p
TOTAL	271.75	63	-	-	-
Between Subjects	50.25	15	-	-	-
Race of Examiner	7.5625	1	7.5625	2.4802	-
$Error_b$	42.6875	14	3.0491	-	-
Within Subjects	221.5	48	-	-	-
Race of Subject	90.25	1	90.25	31.0532	<.001
Expectancy	6.25	1	6.25	2.5225	ns
Race of Examiner X Race of Subject	10.5625	1	10.5625	3.6343	<.10
Race of Examiner X Expectancy	7.5625	1	7.5625	3.0522	ns
Race of Subject X Expectancy	1.0000	1	1.000	.46	ns
Race of Examiner X Race of Subject X Expectancy	.0625	1	.0625	.0287	ns
$Error_1$	40.6875	14	2.9063	-	-
$Error_2$	34.6875	14	2.4777	-	-
$Error_3$	30.4375	14	2.1741	-	-

Arithmetic

SOURCE	SS	df	ms	F	p
TOTAL	63.734	63	-	-	-
Between Subjects	15.484	15	-	-	-
Race of Examiner	2.640	1	2.64	2.879	ns
$Error_b$	12.844	14	.917	-	-
Within Subjects	48.25	48	-	-	-
Race of Subject	1.265	1	1.265	1.379	ns
Expectancy	1.40	1	1.4	.974	ns
Race of Examiner X Race of Subject	2.64	1	2.64	2.879	ns
Race of Examiner X Expectancy	4.517	1	4.517	3.141	<.10
Race of Subject X Expectancy	.142	1	.142	.966	ns
Race of Examiner X Race of Subject X Expectancy	4.513	1	4.513	30.101	<.001
$Error_1$	12.843	14	.917	-	-
$Error_2$	20.127	14	1.438	-	-
$Error_3$	2.061	14	.147	-	-

Similarities

SOURCE	SS	df	ms	F	p
TOTAL	145.984	63	-	-	-
Between Subjects	.734	15	-	-	-
Race of Examiner	.015	1	.015	.2918	ns
$Error_b$	.719	14	.0514	-	-
Within Subjects	145.25	48	-	-	-
Race of Subject	141.015	1	141.015	3325.8255	<.00001
Expectancy	.015	1	.015	.1318	ns
Race of Examiner X Race of Subject	.142	1	.142	3.3491	<.10
Race of Examiner X Expectancy	.142	1	.142	1.2478	ns
Race of Subject X Expectancy	.392	1	.392	4.5006	<.10
Race of Examiner X Race of Subject X Expectancy	.138	1	.138	1.5844	ns
$Error_1$	.593	14	.0424	-	-
$Error_2$	1.593	14	.1138	-	-
$Error_3$	1.22	14	.0871	-	-

Vocabulary

SOURCE	SS	df	ms	F	p
TOTAL	226	63	-	-	-
Between Subjects	30	15	-	-	-
Race of Examiner	6.25	1	6.25	3.6843	<.10
Error$_b$	23.75	14	1.6964	-	-
Within Subjects	196	48	-	-	-
Race of Subject	100	1	100.	41.4817	<.001
Expectancy	.0625	1	.0625	.0332	ns
Race of Examiner X Race of Subject	2.25	1	2.25	.9333	ns
Race of Examiner X Expectancy	1.5625	1	1.5625	.8294	ns
Race of Subject X Expectancy	.5625	1	.5625	.2510	ns
Race of Examiner X Race of Subject X Expectancy	.0625	1	.0625	.0279	ns
Error$_1$	33.75	14	2.4107	-	-
Error$_2$	26.375	14	1.8839	-	-
Error$_3$	31.375	14	2.2411	-	-

Digit Span

SOURCE	SS	df	ms	F	p
TOTAL	23.9755	63	-	-	-
Between Subjects	4.9375	15	-	-	
Race of Examiner	.0625	1	.0625	.18667	ns
$Error_b$	4,6875	14	.33482	-	-
Within Subjects	19.038	48	-	-	-
Race of Subject	.0625	1	.0625	.14737	ns
Expectancy	1.5625	1	1.5625	4.92964	<.05
Race of Examiner X Race of Subject	1.00	1	1.00	2.35788	ns
Race of Examiner X Expectancy	1.00	1	1.00	3.15497	<.10
Race of Subject X Expectancy	.5625	1	.5625	1.86369	ns
Race of Examiner X Race of Subject X Expectancy	.25	1	.25	.82831	ns
$Error_1$	5.9375	14	.42411	-	-
$Error_2$	4.4375	14	.31696		
$Error_3$	4.2255	14	.30182		-

Picture Completion

SOURCE	SS	df	ms	F	p
TOTAL	431.2344	63	-	-	-
Between Subjects	22.4844	15	-	-	-
Race of Examiner	.1407	1	.1407	.0882	ns
$Error_b$	22.3436	14	1.5960	-	-
Within Subjects	408.75	48	-	-	-
Race of Subject	301.8907	1	301.8907	144.0318	<.001
Expectancy	.0157	1	.0157	.0056	ns
Race of Examiner X Race of Subject	.0155	1	.0155	.0074	ns
Race of Examiner X Expectancy	.0155	1	.0155	.0055	ns
Race of Subject X Expectancy	2.6405	1	2.6405	1.0424	ns
Race of Examiner X Race of Subject X Expectancy	.1408	1	.1408	.0556	ns
Error	29.3438	14	2.096	-	-
$Error_1$	39.2188	14	2.8013	-	-
$Error_3$	35.4624	14	2.533	-	-

Picture Arrangement

SOURCE	SS	df	ms	F	p
TOTAL	34	63	-	-	-
Between Subjects	5.5	15	-	-	-
Race of Examiner	0.0	1	0.	0.	ns
$Error_b$	5.5	14	.3929	-	-
Within Subjects	28.5	48	-	-	-
Race of Subject	9	1	9	50.005	<.001
Expectancy	.0625	1	.0625	.2	ns
Race of Examiner X Race of Subject	0.	1	0	0.0	ns
Race of Examiner X Expectancy	.0625	1	.0625	.2	ns
Race of Subject X Expectancy	.0625	1	.0625	/-9-5	ns
Race of Examiner X Race of Subject X Expectancy	1.5625	1	1.5625	2.0115	ns
$Error_1$	2.5	14	.1786	-	-
$Error_2$	4.375	14	.3125	-	-
$Error_3$	10.875	14	.7768	-	-

Block Design

SOURCE	SS	df	ms	F	p
TOTAL	30.109	63	-	-	-
Between Subjects	4.359	15	-	-	-
Race of Examiner	.015	1	.015	.048	ns
$Error_b$	4.344	14	.31	-	-
Within Subjects	25.75	48	-	-	-
Race of Subject	13.14	1	13.14	94.5324	$<.001$
Expectancy	.39	1	.39	1.2226	ns
Race of Examiner X Race of Subject	.017	1	.017	.0769	ns
Race of Examiner X Expectancy	.392	1	.392	1.2288	ns
Race of Subject X Expectancy	.142	1	.142	.500	ns
Race of Examiner X Race of Subject X Expectancy	.138	1	.139	.489	ns
$Error_1$	3.093	14	.221	-	-
$Error_2$	4.468	14	.319	-	-
$Error_3$	3.970	14	.284	-	-

Object Assembly

SOURCE	SS	df	ms	F	p
TOTAL	188.859	63	-	-	-
Between Subjects	6.609	15	-	-	-
Race of Examiner	.39	1	.39	.8780	ns
$Error_b$	6.219	14	.4442	-	-
Within Subjects	182.25	48	-	-	-
Race of Subject	141.015	1	141.015	104.0778	<.001
Expectancy	1.89	1	1.89	2.5278	ns
Race of Examiner X Race of Subject	.767	1	.767	.5661	ns
Race of Examiner X Expectancy	.392	1	.392	.5243	ns
Race of Subject X Expectancy	.392	1	.392	.6677	ns
Race of Examiner X Race of Subject X Espectancy	.138	1	.138	.2351	ns
$Error_1$	18.968	14	1.3549	-	-
$Error_2$	10.468	14	.7477	-	-
$Error_3$	8.22	14	.5871	-	-

Coding

TOTAL	SS	df	ms	F	p
TOTAL	236.9375	63	-	-	-
Between Subjects	28.4375	15	-	-	-
Race of Examiner	.0625	1	.0625	.0308	ns
Error$_b$	28.3750	14	2.0268	-	-
Within Subjects	208.5	48	-	-	-
Race of Subject	132.25	1	132.25	76.3524	.001
Expectancy	6.25	1	6.25	5.3437	.05
Race of Examiner X Race of Subject	0.00	1	0.00	0.0	ns
Race of Examiner X Expectancy	5.875	1	5.875	5.0231	.05
Race of Subject X Expectancy	.0625	1	.0625	.0389	ns
Race of Examiner X Race of Subject X Expectancy	.9375	1	.9375	.5833	ns