8-9-1961

The Judgment of Intelligence from Photographs

Charlene Diver Fredenburgh

UNIVERSITY OF NEW MEXICO LIBRARY

MANUSCRIPT THESES

Unpublished theses submitted for the Master's and Doctor's degrees and deposited in the University of New Mexico Library are open for inspection, but are to be used only with due regard to the rights of the authors. Bibliographical references may be noted, but passages may be copied only with the permission of the authors, and proper credit must be given in subsequent written or published work. Extensive copying or publication of the thesis in whole or in part requires also the consent of the Dean of the Graduate School of the University of New Mexico.

This thesis by [illegible] has been used by the following persons, whose signatures attest their acceptance of the above restrictions.

A Library which borrows this thesis for use by its patrons is expected to secure the signature of each user.

NAME AND ADDRESS	DATE

THE JUDGMENT OF INTELLIGENCE FROM PHOTOGRAPHS

By

Charlene Diver Fredenburgh

A Thesis

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Arts in Psychology

The University of New Mexico

1961

This thesis, directed and approved by the candidate's committee, has been accepted by the Graduate Committee of the University of New Mexico in partial fulfillment of the requirements for the degree of

MASTER OF ARTS

E. L. Castetter

DEAN

DATE August 9, 1961

Thesis committee

David T. Bradbury

CHAIRMAN

Charles E. Woodhouse

Thom Peterson

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION AND PROBLEM

Between 1918 and 1939 several studies appeared in which the ability of persons to make judgments of intelligence from photographs was assessed. The topic has not been pursued since then, despite the inadequacies of these early studies.

These investigations found that, on the average, the ability of individuals to judge the intelligence of pictured persons is either nonexistant or low. For example the following investigators have reported these average correlation coefficients, with judged intelligence against various criteria: Cook, .05 (3); Pintner, .10 (13); Cogan, .19 (4,5); Moriwaki, .29 (11); Gaskill, .31 (6); Schmidt, .27 (15); Knight, .11 (9); Anderson, .13 (1).

In a few studies, pooled judgments were checked against some "objective" criterion of intelligence. Some of these found at least moderate correlations: Moriwaki, .43 (11); Cogan, .51 (4,5); Gaskill, .43 (6); Schmidt, .70 (15). Others, however, reported low correlations: Cook, .07 (3); Pintner, .16 (13); Knight, .17 (9); Anderson, .27 (1).

All of these studies contain two major weaknesses.

First, the photographs that were used were not standardized, i.e., there were uncontrolled differences in size of image, lighting, background, and in the age, sex, dress, pose and expression of the pictured subjects. Second, the criteria of intelligence have consisted either of the opinions of persons acquainted with the subjects, or scores on intelligence tests which are, now at least, obsolete.

It may be, of course, that despite these limitations the findings of the early studies are essentially correct: that intelligence cannot, on the average, be judged from photographs with very impressive accuracy. It would be highly desirable, however, to ascertain the facts in a study which obviates the flaws of the earlier work; such is the general goal of the present investigation.

The immediate purpose of this project was to produce a set of photographs, standardized both as to subjects and photography, which would replace the old, inadequate photographs in instructional and research work.

The study itself was designed with three objectives in mind: first, to determine how accurately intelligence, as measured by the 1937 Stanford-Binet, can be judged from standardized photographs of homogeneous subjects; second, to determine if there is a difference between college and junior high school students in the accuracy of such judgments; and third, to determine how closely the two groups of judges agree.

CHAPTER II

PROCEDURE

Selection of subjects. Uniformity was the basic criterion for the selection of subjects to be photographed. They all met the following requirements: same sex, boys; same age range, 14 to 16 years; same grade placement, 9.5; and approximately the same size head and shoulders. The subjects had no noticeable facial defects, and had even features; in short, they were "nice looking American boys".

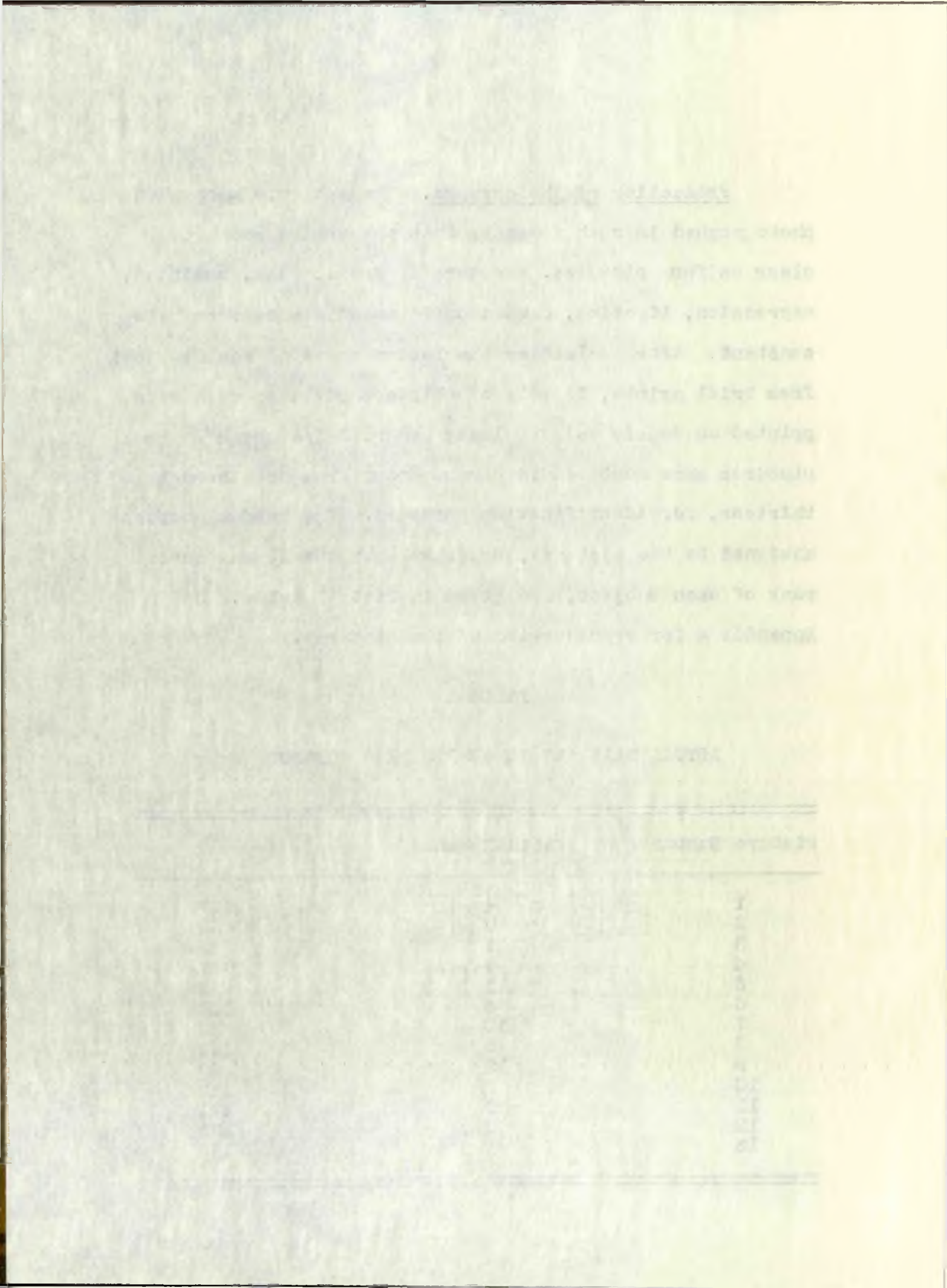
To make screening easier, 40 prospective subjects were selected from 400 high school freshmen boys whose records were available. Those chosen were individuals with whom this writer was acquainted through teaching and other school activities making for some knowledge of IQ, personal appearance, and rapport with the boys and their parents. School records were then checked for IQ, age, achievement, and grade placement. From these records eight individuals were selected whose Otis scores ranged from 70 to 145. After contacting the boys and receiving their parents' approval, the prospective subjects were tested with the Stanford Binet Scale, Form L. The selection and screening procedure, becoming more difficult as more specific IQ's were needed, was continued until thirteen subjects, ranging in IQ from 76 to 146 with intervals of five to seven points, were obtained.

Production of photographs. The subjects were then photographed in such a manner that the prints would be clear uniform pictures. Factors of dress, pose, position, expression, lighting, camera angle and distance were held constant. After selecting the best picture of each subject from trial prints, 33 sets of thirteen pictures each were printed on double weight glossy paper, 3 3/4" by 5". The pictures were numbered in random order from one through thirteen, for identification purposes. The random numbers assigned to the pictures, together with the IQ and actual rank of each subject, are given in Table I below. See Appendix A for reproduction of the pictures.

TABLE I

ACTUAL RANK AND IQ OF PICTURED SUBJECTS

Picture Number	Actual Rank	IQ
1	4	127
2	9	99
3	11	85
4	6	117
5	13	74
6	3	134
7	10	92
8	5	122
9	12	80
10	7	112
11	2	141
12	8	105
13	1	146



The testing procedure. Ninety-four boys and girls from a typical Albuquerque junior high school and 117 University of New Mexico General Psychology students were used as judges. Each judge in each group was given a test of the paired comparison type. In the college group the tests were administered in laboratory sessions devoted to the judging of intelligence from photographs. See Appendix B for reproduction of the test.

The actual procedure began with an introduction concerning the appearance of a person's face as a source of information upon which an estimate of the perceived person's over-all mental ability is commonly made. Materials were enumerated: standard photographs of thirteen junior high school boys, answer sheet for recording 78 paired comparisons, tabulation sheet for determining the rank of the subject's intelligence, and a work sheet for calculation of the Spearman rho between the judged and actual ranks.

Instructions given in the "procedure" section of the test were as follows. The judge was instructed to stack photographs 2 through 13 with picture 1 at the left. He was told to compare 1 with 2 and circle on the answer sheet the one he thought more intelligent, turn 2 face down tc the right of the stack, and continue with 1 and 3, 1 and 4, and so on until 1 had been compared with each of

the remaining 12. The same procedure was continued with picture 2 compared with each of the remaining 11, picture 3 compared with the remaining 10, and so on, until each photograph had been compared with every other one, a total of 78 comparisons.

Data were accumulated by tallying and recording on the tabulation sheet the number of times each photograph was chosen. The judged rank was entered in the appropriate column. After the actual rankings of the subjects were given to the college students, they were told to correlate their ranking with the actual rank order using Spearman's formula: $\rho = 1 - \frac{6\Sigma d^2}{n(n^2-1)}$

When this test was administered to the junior high group, the instructions were explained more carefully and the judges were not required to calculate the correlation coefficient, but only to circle the choices in the paired comparisons.

After checking the judges' work and making the necessary corrections, this investigator calculated the rho for each individual in each group.

CHAPTER III

RESULTS AND DISCUSSION

The judgments secured from the college and junior high groups were compared in two general ways: (1) the two distributions of individual rhos were graphed, and a *t* test of the difference between the two mean rhos was calculated; (2) the mean ranking of each of the thirteen photographs was determined separately for the two groups.

The polygons describing the two distributions of individual rhos are presented in Figure 1. The abscissa of this graph represents the rhos, ranging from -1.0 to +1.0 in intervals of .10, and the frequencies in these intervals are indicated on the ordinate.

A number of facts are conveyed by this graph. The two distributions are similar, both being roughly normal, with a slight negative skew. The ranges are comparable: -.53 to +.96 in the junior high group and -.63 to +.90 in the college group. The individual differences in ability to judge intelligence from photographs, as represented in this graph, are striking.

The central tendencies of the two distributions are somewhat above zero, the point at which the means would fall if the subjects' judgments were merely of a chance nature. The mean rhos were found to be .28 for the college group and .17 for the junior high group. The

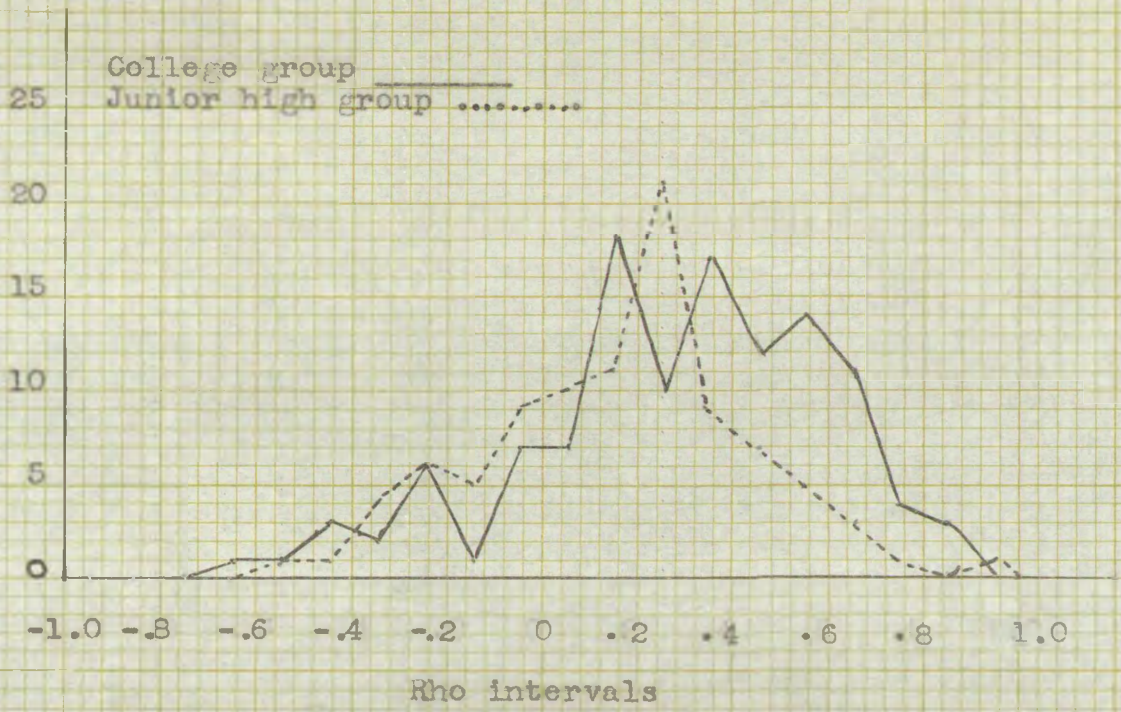
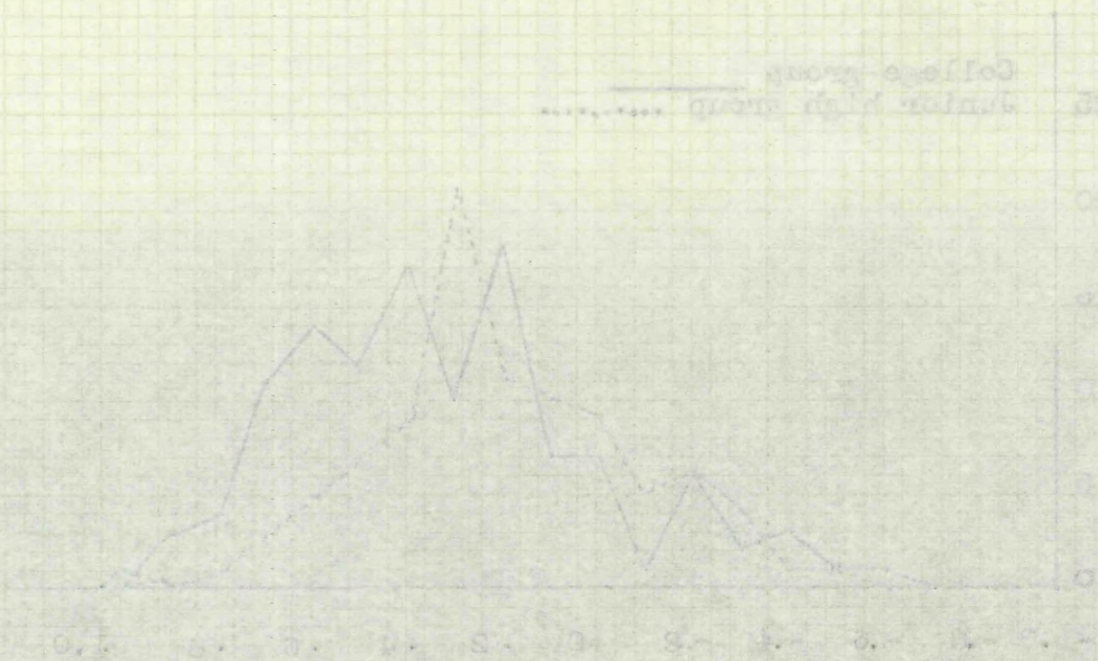


FIGURE 1

DISTRIBUTION OF RHOS OF COLLEGE AND JUNIOR HIGH SCHOOL GROUPS



median rhos, .32 and .21 respectively, differ in the same direction as the means and suggest that the amount of skew is not great.

The two distributions when tested for homogeneity of variance, yielded an F ratio of 1.25, which is not significant. A t test was then run, as shown in Table II.

TABLE II

MEAN RHOS FOR THE TWO GROUPS

Groups	Mean	SD	t	P
Junior High N 94	.165	.288	2.78	.01
College N 117	.284	.323		

Although the absolute difference between the two means is not great, the significant t supports the conclusion that the college group was more accurate than the junior high group in their judgments; further, this significant t indicates that the mean rho of the college students, at least, was significantly different from zero.

The difference between the two groups cannot be attributed wholly to age, inasmuch as they were not equated in intelligence. However, the class sections which made

up the junior high group were at least average in their academic achievement, and probably all of these students had an IQ above 100. In view of this it is probable that the IQ's of the two groups did not differ greatly. Thus there is no reason to believe that the superior performance of the college students is attributable to higher intelligence, although a conclusive statement about this cannot be made because the mental ages of the two groups were not compared.

The survey article by Taft was not mentioned in Chapter One because it concerned studies of the judgment of emotional, personality, and behaviorial characteristics, rather than intelligence (18). However, his general conclusions concerning the factors of intelligence and age as they affect the ability to judge people seem sufficiently relevant to the present discussion to be noted here. Taft concluded that there seems to be a positive relationship between intelligence and the ability to judge others analytically. "The superiority of more intelligent judges is probably the most pronounced in the ability to rate the intelligence of others". Taft also generalized that the ability to judge emotional, personality, and behaviorial characteristics of others does not increase with age after childhood unless the subjects being judged are close to the age group of the judges. Presumably

older judges, less similar to their subjects in age than younger judges, are under a handicap, being less able to use appropriate norms for making their judgments than those judging subjects closer to their own age. If, in general, people are better judges of their age peers, one would expect the junior high group to be better judges of the intelligence of subjects near their own age. The results of the present study, however, do not confirm such an expectation. Further experimentation is necessary to unravel the roles of age and intelligence in the judging of intelligence from photographs.

The second general comparison, as mentioned at the outset of this chapter, concerned the mean ranking of each of the thirteen photographs. Table III presents these mean rankings for the two groups. These mean rankings were in turn ranked from 1 to 13, and are presented in the table along with the actual ranks (based upon the intelligence test results). Figure 2 presents curves representing the mean rankings for the two groups.

TABLE III

MEAN RANKINGS OF EACH PICTURE BY EACH GROUP

Picture Number	Jr. High Mean Rank	College Mean Rank	Jr. High Ranking	College Ranking	Actual Rank
1	6.223	6.098	5	4	4
2	7.654	7.620	9	9	9
3	5.261	5.756	2	3	11
4	7.707	7.607	11	8	6
5	9.601	9.098	12	12	13
6	6.777	6.376	7	5	3
7	10.580	10.671	13	13	10
8	4.266	3.598	1	1	5
9	5.814	6.718	3	7	12
10	7.202	8.355	8	11	7
11	7.660	6.671	10	6	2
12	6.606	7.910	6	10	8
13	5.872	4.291	4	2	1

On the horizontal axis of the graph the picture numbers are arranged from most to least intelligent, left to right. On the vertical axis mean rankings are arranged from high to low (top to bottom), the lowest mean rank being the picture rated the most intelligent, the highest mean rank being the picture ranked the least intelligent.

The similarity of the curves indicates considerable agreement between the two groups as to the judged intelligence of the pictured subjects. This agreement, which is expressed by a rank order correlation of .79, is especially prominent on pictures one through eight. The

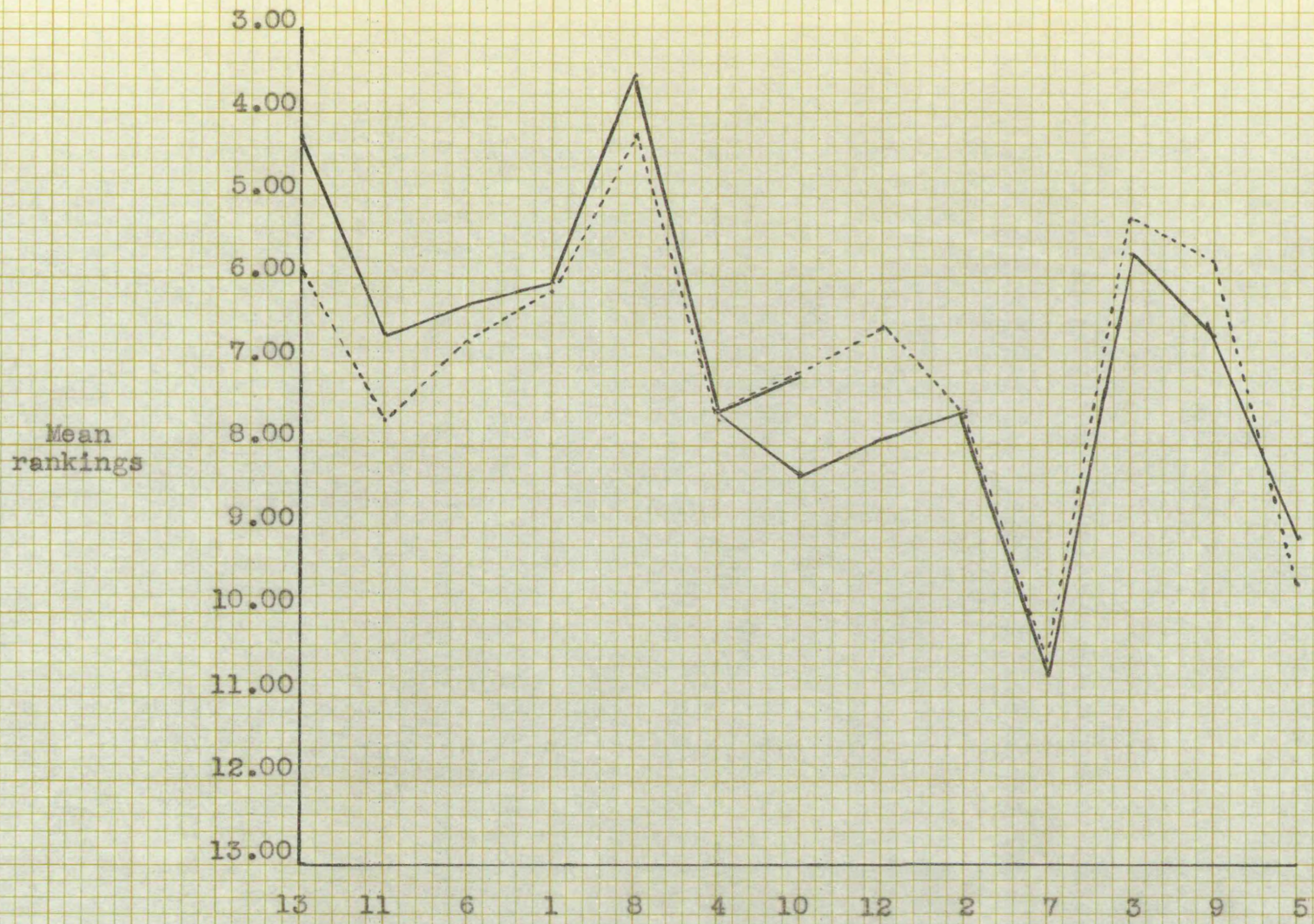
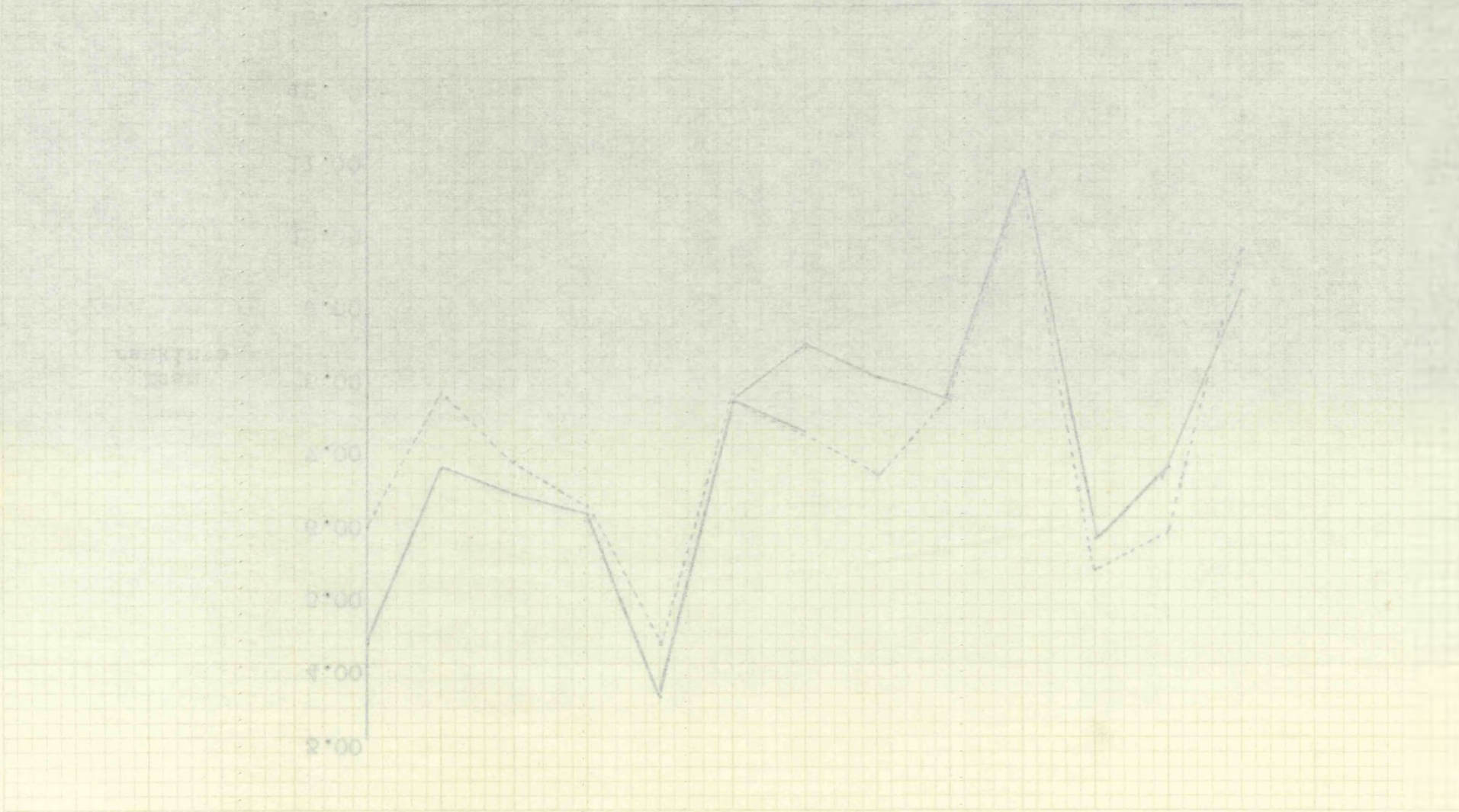


FIGURE 2

MEAN RANKING OF EACH PICTURE BY COLLEGE AND JUNIOR HIGH SCHOOL GROUPS



general downward slope of the curves indicates positive correlation of the over-all rhos, even though the rho for the college group, .56, is appreciably higher than for the junior high group, .16. It is interesting that pooling the judgments raises the rho of the college group so markedly while leaving the rho of the junior high group unaffected. This relationship is apparent from visual inspection of Figure 2 and Table III which show the college group rankings as closer than the junior high group to the actual order. In general, the results suggest that the students are judging <u>something</u>, but not simply intelligence. Their judgments were not empirically accurate, but the two groups tended to be in agreement.

A very interesting feature which may be noted from the analysis of Figure 2 and Table III is that the judgment errors of both groups occurred on essentially the same pictures—3, 9, 8, and 7—in order of most to least error. Both groups ranked picture 3 quite high (second and third), when the subject's actual rank was third from the bottom. This incidentally is congruent with the writer's experience with the subject, in that he did "look" much more intelligent than he actually was. Picture 9, whose actual rank was next to last, was ranked third by the junior high group and seventh by the college group. The error was greater here for the junior high

group. Picture 8, whose actual rank was fifth, was judged most intelligent by both groups. Picture 7, whose actual rank was tenth, was judged least intelligent by both groups.

CHAPTER IV

SUMMARY AND CONCLUSIONS

This study was designed primarily for the purpose of developing a set of photographs, superior to those which have been used in previous investigations, to be made available for further research and instructional purposes. The basic questions of the present study were: (1) How accurately can intelligence, as measured by the 1937 Stanford-Binet, be judged from well-standardized photographs of subjects who are relatively homogeneous in appearance but not in I.Q. (2) Is there a difference between college and junior high school students in the accuracy of such judgments? (3) How closely do the two groups of judges agree in their judgments?

The procedure involved selecting and photographing thirteen homogeneous subjects, 15 year old males, ranging in IQ from 76 to 145 with intervals of five to seven points. The subjects were photographed in such a manner that the prints would be clear uniform pictures. Factors of dress, pose, position, expression, lighting, camera angle and distance were held constant. Two groups of judges were selected: 94 boys and girls from a typical Albuquerque junior high school and 117 University of New Mexico General Psychology students. A test of the paired

comparison type, resulting in a ranking of the thirteen photographs, was constructed and administered to these two groups.

The rank order correlation between the judged and actual ranks was calculated for each judge. The frequency distributions of the individual rhos for both groups were graphed. The mean rho for each group was determined, and the difference tested for significance. The mean ranking of each picture was determined separately for the two groups, and this information was pictured graphically. An over-all rho for each group was computed from these pooled judgments. Also, the correlation between the pooled judgments of the two groups was calculated, and the judgment errors in the two groups were compared.

From the findings which resulted the following conclusions were drawn:

1. Individuals differed widely in their ability to judge intelligence from photographs of fifteen year old boys. The two distributions of individual rhos were comparable, both being roughly normal with a slight negative skew, and ranged from -.53 to +.96 in the junior high group and -.63 to +.90 in the college group.
2. The two groups were not very accurate in their judgments of intelligence from photographs.

The average correlation between the judged and actual rank was very low positive: .28 for the college group and .17 for the junior high group.

3. The older group was more accurate in their judgments of intelligence from photographs than the younger group, the difference between the mean rhos being statistically significant ($t = 2.78$, $p < .01$). Pooled judgments, from the mean rankings of each picture by each group, yielded a rho of .56 for the college group; this was appreciably higher than the rho of .16 for the junior high group. Pooling of judgments increased accuracy only in the older group.
4. There was considerable agreement between the two groups in their pooled judgments of the intelligence of the pictured subjects. This agreement, expressed by a rho of .79, was especially prominent on pictures one through eight. As in previous studies, judges showed more similarity than accuracy in their judgments of intelligence.
5. Departures from accuracy in the judgments of intelligence from photographs were not distributed randomly among the series of

photographs. The judgment errors of both the junior high group and the college group were concentrated on the same four photographs—3, 9, 8, and 7,—in order of most to least error.

REFERENCES

REFERENCES

1. Anderson, L. D. "Estimating Intelligence by Means of Printed Photographs," J. Applied Psychol., June, 1921, 152-155.

2. Camden, C. "Minds Construction in the Face," Bibliog. of Phil., v 20, 400-412, 1941.

3. Cook, S. W. "Judgments of Intelligence from Photographs," J. of Abnormal Psychol., 384-389, 1939.

4. Cogan, L. C. "Judgment of Character on the Basis of Photographs," as reported by Hollingworth, H. L., Vocational Psychology: New York, D. Appleton, 1919, 48-51.

5. Cogan, L. C., Conklin, R. M., & Hollingworth, H. L. "An experimental study of self-analysis, estimates of associates and the results of tests," Sch. & Soc., 1915, 2, 171-179.

6. Gaskill, Fenton, and Porter, "Judging Intelligence of Boys From Their Photographs," J. Applied Psychol., 13, 630-636, 1929.

7. Havells, T. H. "Study of Ability to Recognize Faces," J. Abnormal Psychol., 33: 24-127, 1938.

8. Hollingworth, H. L. "Experimental Studies in Judgment," Arch. of Psychol., Columbia University, No. 29.

9. Knight, H. C. "A Comparison of the Reliability of Group and Individual Judgments". Master's Essay in Columbia U. Lib., 1921. As reported by Murphy, G. and Murphy, L. B., Experimental Social Psych., New York: Harpers 1931, 504-505, 544-545.

10. Kohs and Irle, "Prophesying Army Promotion," J. Applied Psychol., March, 1920.

11. Moriwaki, E. "Notes on the Comparative Validities of Judgments of Intelligence based on Photographs and Interviews," J. of Applied Psychol., 1939.

12. Muenzinger, K. F. and Christoffel, Hans. "Physiognomy and Psychoanalysis," Int. J. Psycho-Anal., 30, 207-208, 1949. Abstract.

13. Pinter, R. "Intelligence as Estimated from Photographs," Psychol. Rev., July, 1918.

14. Pope, D. V. "Interpretation of the Human Face from their Photographs," Psychological Bull., July-Sept., 1922, 3-17.

15. Schmidt, H. "Aussere Erscheinung und Personlichkeitsbeurteilung," (Judging personality from appearances), Industr. Psychotech., 1934, 11, 311-315. As reported in Psychol Abstr., 1935, 788.

16. Simpson, B. R. "Reliability of Judgments of Intelligence," J. of Educational Psychol., April 1915.

17. Stryher, P. "Is there an Executive Face?" Merton System, Fortune 48, 145-147, 1953.

18. Taft, R. "The Ability to Judge People," Psychological Bull., January 1955.

19. Thorndike, E. L. "A Constant Error in Psychological Ratings," J. Applied Psychol., March, 1920.

20. Thorton, G. R. "Ability to Judge Crimes from Photos of Criminals: A Contribution to Technique," J. Abnormal Psychol., 34, 378-383, 1938.

APPENDIX A

APPENDIX A

APPENDIX A

PHOTOGRAPH OF SUBJECTS

APPENDIX A

PHOTOGRAPH OF SUBJECTS

APPENDIX B

APPENDIX B

Judging Intelligence from Photographs

Introduction. In their interactions with each other, people commonly make judgments of others' abilities, traits, attitudes and other psychological characteristics. A variety of cues are used in making such judgments; some of these cues are behavioral (especially verbal) and others are physical ("appearance" cues).

Probably one of the more common interpersonal judgments concerns intelligence: whether or not such a judgment is verbalized, an estimate or "sizing up" of a perceived person's over-all mental ability is commonly made.

It is also probable that the appearance of the person's face is one of the sources of information upon which such a judgment is made. Some persons, it is said, "look bright", while others "look dull". In any event, many people believe that they can get a pretty good idea of another's intelligence by "giving him the once-over".

Today you are going to test, systematically, your own ability to make judgments of the relative brightness of a group of persons when the only information you have is the appearance of the persons' faces, as recorded by photographs.

Before doing this, check the appropriate box below to indicate how well you think you can judge intelligence from photographs;

not at all well ☐ not very well ☐ fairly well ☐ quite well ☐

Materials. Standard photographs of thirteen junior high school boys, numbered 1-13; answer sheet for recording 78 paired comparisons; tabulation sheet to determine your ranking of the boys' intelligence; work sheet for calculation of rank order correlation.

Procedure.

1. The photographs are stacked on the table face up in serial order from 2 to 13 with picture 1 at the left. Compare 1 with 2, circle on the answer sheet the one that you think is more intelligent. For example, 1 2 indicated that 1 and 2 are compared and 2 was judged the more intelligent. Turn 2 face down to the right of the stack and continue with 1 and 3, 1 and 4, 1 and 5, etc., until photograph 1 has been compared with each of the remaining 12.
2. Now turn 1 face down on the left and turn stack face up again. Place picture 2 face up on left of stack and compare it with each of the remaining 11 (3 through 13) on its right. Continue using the same procedure until each photograph has been compared with every

other one. For each comparison circle on the answer sheet the one you judge more intelligent. You will make 78 judgments altogether.

Data.

1. Count the number of times each photograph was chosen and record on tabulation sheet. Indicate your ranking of each picture in the column headed "Judged Rank".
2. Your instructor will give you the actual ranking of the boys, as determined by intelligence testing. Correlate your rank order with the actual rank order, using the following formula:

$$p = 1 - \frac{6 \Sigma d^2}{n\ (n^2-1)}$$

Questions.

1. Describe, as best you can, the methods you used in making the judgments. What characteristics did you use in judging? What difficulties did you experience?
2. Under what conditions would judgments of intelligence be more accurate? Why?
3. Interpret the correlation you obtained, and make the best possible generalized statement you can about the possibility of judging intelligence from photographs.
4. What are the implications of this study for the common practice of requiring applicants for positions to submit portraits? Discuss.

5. Now, after going through this lab session, how well do you think you can judge intelligence from photographs?

not at all well ☐ not very well ☐ fairly well ☐ quite well ☐

Name _____

Answer Sheet

Indicate by circling the number of the boy that you think is more intelligent in each comparison.

1 2	2 3	3 4	4 5	5 6
1 3	2 4	3 5	4 6	5 7
1 4	2 5	3 6	4 7	5 8
1 5	2 6	3 7	4 8	5 9
1 6	2 7	3 8	4 9	5 10
1 7	2 8	3 9	4 10	5 11
1 8	2 9	3 10	4 11	5 12
1 9	2 10	3 11	4 12	5 13
1 10	2 11	3 12	4 13	
1 11	2 12	3 13		
1 12	2 13			
1 13				

6 7	7 8	8 9	9 10	10 11
6 8	7 9	8 10	9 11	10 12
6 9	7 10	8 11	9 12	10 13
6 10	7 11	8 12	9 13	
6 11	7 12	8 13		
6 12	7 13			
6 13				

11 12	12 13
11 13	

Name _____

Age _____

Sex _____

Lab Section _____

Tabulation Sheet

Photograph Number	Number of Times Chosen	Judged Rank (Place a "1" after the largest number, etc.)*
1	____	____
2	____	____
3	____	____
4	____	____
5	____	____
6	____	____
7	____	____
8	____	____
9	____	____
10	____	____
11	____	____
12	____	____
13	____	____
	$\Sigma = 78$	$\Sigma = 91$

*In case of tied ranks: assign each case the mean rank. (For example, if the pictures that would be ranked 3 and 4 both received the same number of choices, assign both of them a rank of 3.5.)

Name _____

Work Sheet

Photograph Number	Actual I.Q. Rank	Your Judged Rank	Diff.	Diff.
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				

d^2 = _____

$$\text{rho} = 1 - \frac{6\Sigma d^2}{n(n^2-1)} = 1 - \frac{6\Sigma d^2}{2184}$$

IMPORTANT!

Special care should be taken to prevent loss or damage of this volume. If lost or damaged, it must be paid for at the current rate of typing.