

5-6-1942

A Statistical Analysis of the New Mexico State Merit System 1941 Oral Examination Program

Anchard Zeller

Follow this and additional works at: https://digitalrepository.unm.edu/psy_etds



Part of the [Psychology Commons](#)

Recommended Citation

Zeller, Anchard. "A Statistical Analysis of the New Mexico State Merit System 1941 Oral Examination Program." (1942).
https://digitalrepository.unm.edu/psy_etds/250

This Thesis is brought to you for free and open access by the Electronic Theses and Dissertations at UNM Digital Repository. It has been accepted for inclusion in Psychology ETDs by an authorized administrator of UNM Digital Repository. For more information, please contact disc@unm.edu.

UNIVERSITY OF NEW MEXICO-UNIVERSITY LIBRARIES



A14429 095499



ZEPHYRUS

ERIT

SYS

378

789

Un30

2

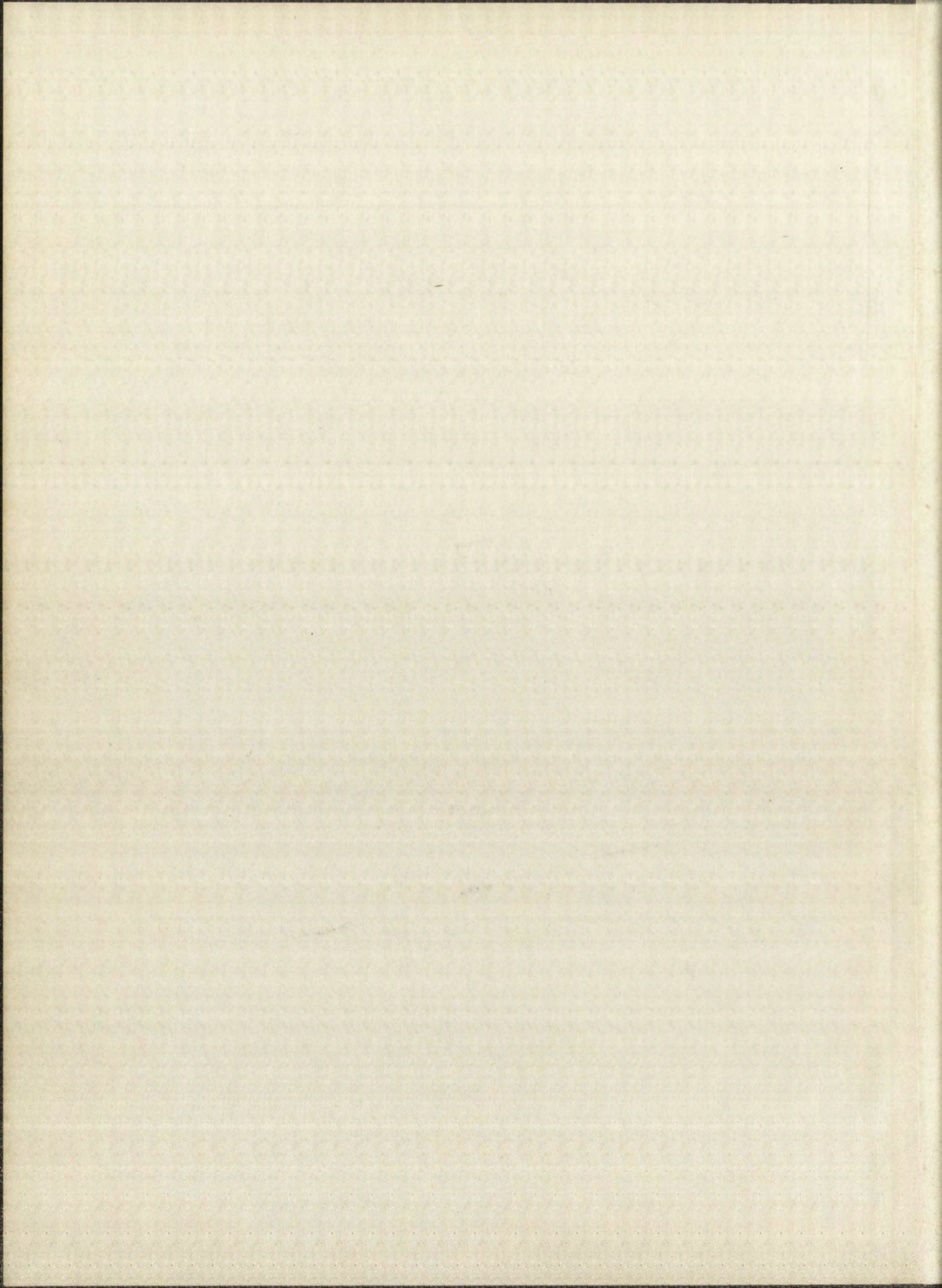
LIBRARY
of
THE UNIVERSITY OF
NEW MEXICO



75985

CLASS
378.789

BOOK
Un30z
1942, cop. 2



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Unpublished theses submitted for the Master's and Doctor's degrees and deposited in the University of New Mexico Library are open for inspection, but are to be used only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 Ancharad Zeller.....
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

A STATISTICAL ANALYSIS
OF THE
NEW MEXICO STATE MERIT SYSTEM
1941 ORAL EXAMINATION PROGRAM

UNIVERSITY OF
NEW MEXICO LIBRARY
ALBUQUERQUE, N.M.

By
Anchard Zeller

A Thesis
Submitted for Partial Fulfillment of the
Requirements for the Degree of
Master of Science in Psychology

University of New Mexico

1942

UNIVERSITY OF
NEW MEXICO LIBRARY
ALBUQUERQUE, N.M.

THE
UNIVERSITY OF
NEW MEXICO
LIBRARY
ALBUQUERQUE, N.M.
1970-1971
302

378.789

Un 302

1942, cop. 2

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 SCIENCE

F. Hammond

DEAN

May 6, 1942

DATE

Thesis committee

Sworn Peterson

CHAIRMAN

B. F. Haught

C. E. Hutchinson

8/5/42 Winter 125

PLATE 10

[Faint handwritten signature]

1884

32

ACKNOWLEDGMENT

The writer wishes to express his sincere thanks to the Merit System Council for the unlimited use of their private files and to the faculty of the Department of Psychology for their interest and cooperation in the face of unforeseen difficulties.

The writer wishes to express his appreciation to the
the staff of the hospital for their kind and efficient
private life and to the family for their support and
hospitality for his stay in the hospital.
of his family and friends.

TABLE OF CONTENTS

CHAPTER	PAGE
I. INTRODUCTION	1
Statement of the Problem	1
Brief Review of Related Literature	2
Procedure and Organization	6
Source of Data	6
Procedure	6
Coding	7
Organization	9
II. PRELIMINARY STATISTICAL TREATMENT	11
Examiners	11
Examination	11
Conduct	13
Initial Statistical Treatment	14
III. STATISTICAL ANALYSIS OF THE EXAMINATIONS	16
Intercorrelation of Traits by Examiners	16
Intercorrelation of Traits	17
Reliability and Validity	27
IV. EXTERNAL FACTORS	29
Factors in the Situation	29
Factors in the Applicant	30
V. CONCLUSIONS AND SUGGESTIONS FOR FUTURE STUDY	44
BIBLIOGRAPHY	47

CHAPTER

I. INTRODUCTION

Statement of the Problem

General Review of Literature

Procedures and Data Analysis

Summary of Results

Conclusions

References

Organization

II. THEORETICAL BACKGROUND

Concepts

Measurement

Control

Initial Statistical Treatment

III. STATISTICAL ANALYSIS

Interpretation of Results

Interpretation of Results

Reliability and Validity

IV. EXPERIMENTAL DESIGN

Factors in the Situation

Factors in the Situation

V. CONCLUSION AND DISCUSSION

Summary

LIST OF TABLES

TABLE	PAGE
I. Exhibit G	12
II. Intercorrelation of Traits by Examiners for Interviewee	18
III. Intercorrelation of Traits by Examiners for Visitor	21
IV. Average of Correlations for Each Trait	24
V. Correlation of Traits with Examiner Evaluation for Visitor	24
VI. Correlation of Traits with Examiner Evaluation for Interviewer	25
VII. Combined Evaluation Correlated with Individual Evaluations for Both Visitor and Interviewer.	25
VIII. Intercorrelation of Traits	26
IX. Average Oral Score by Center	31
X. Average Oral Score by Linguistic Group	33
XI. Average Oral Score	34
XII. Average Oral Score by Education	36
XIII. Average Oral Score by Weight	38
XIV. Average Oral Score by Height	40

TABLE

I.	General Introduction
II.	Historical Development of the Theory
III.	Mathematical Foundations
IV.	Physical Interpretation of the Results
V.	Comparison with Experimental Data
VI.	Conclusions and Outlook
VII.	References
VIII.	Appendix A: Mathematical Derivations
IX.	Appendix B: Numerical Calculations
X.	Appendix C: Bibliography
XI.	Appendix D: Glossary of Terms
XII.	Appendix E: Index
XIII.	Appendix F: List of Figures
XIV.	Appendix G: Summary of Results

LIST OF FIGURES

FIGURE	PAGE
I. Average Oral Score by Race and Sex	32
II. Average Oral Score by Age	35
III. Average Oral Score by Education	37
IV. Average Oral Score by Weight	39
V. Average Oral Score by Height	41

LIST OF FIGURES

FIGURE	
I.	Average Oral Score by Age and Sex 32
II.	Average Oral Score by Age 32
III.	Average Oral Score by Education 33
IV.	Average Oral Score by Weight 33
V.	Average Oral Score by Height 34

CHAPTER I

INTRODUCTION

In accordance with Article V of the Classification Plan of the New Mexico Department of Public Welfare¹ and Regulation 8, Article V, of the Employment Security Commission of New Mexico,² an oral examining board composed of three members toured the State of New Mexico from June 9 to June 26, 1941, interviewing applicants for the position of Visitor, Interviewer, Case Worker in Training, and Receptionist. Only those persons who had successfully passed a written examination were invited to appear before the board.

Statement of the Problem

The purpose of the present study is to analyze the oral examination mentioned above and the results obtained. This analysis will consist of three parts: first, a description of the statistical methods by which the oral ratings were converted into quantitative results; second, an internal

¹New Mexico Department of Public Welfare, "Classification and Compensation Plan of the New Mexico Department of Public Welfare," Rule for a Merit System of Personnel Administration in the New Mexico Department of Public Welfare, Article V, Section 1, Paragraph 1, Part C (Effective March 10, 1941).

²Employment Security Commission of New Mexico, Employment Security Commission of New Mexico Regulation No. 8, Article V, Section 1, Paragraph 1, Part C.

IN-SECRET-2 was received from the
Bureau of the New Mexico State Police
on 10/10/41, Article 1, Section 1, of the
New Mexico Constitution, which provides
that the Governor shall have the honor
of the State.

1941, Inspector General, New Mexico
State Police, Article 1, Section 1, of the
New Mexico Constitution, which provides
that the Governor shall have the honor
of the State. This is a copy of the
original document, which is on file in
the New Mexico State Police, Santa Fe,
New Mexico.

The purpose of this document is to
provide a copy of the original document
to the New Mexico State Police, Santa Fe,
New Mexico. This is a copy of the
original document, which is on file in
the New Mexico State Police, Santa Fe,
New Mexico.

New Mexico State Police, Santa Fe,
New Mexico, Article 1, Section 1, of the
New Mexico Constitution, which provides
that the Governor shall have the honor
of the State. This is a copy of the
original document, which is on file in
the New Mexico State Police, Santa Fe,
New Mexico.

analysis of the examination itself with a brief discussion of the validity and reliability of the examination; and third, a survey of the external factors inherent in the examinees and situations which influenced the ratings of the judges.

A Brief Survey of the Literature

What then is an interview? There are many answers to this question, almost as many as there are fields of endeavor. The sales manager, the psychiatrist, the police officer would all have a different concept, yet each uses a process which could legitimately be called an interview. The definition of Symonds,³ which defines an interview as the process by which one person, called the interviewer, directly addresses another person, called the interviewee, in order to learn from him, seemed to be one of the most adequate. The same thought is expressed more succinctly by Bingham, who defines an interview as a conversation with a purpose.⁴

Since the art of communication has been in existence, one of the most popular pastimes has been the "sizing up" of others. Many laymen with little technical training feel quite capable of arriving at a correct judgment based on very brief

³Perceval M. Symonds, "Research on the Interviewing Process," Journal of Educational Psychology, 30: 346, 1938.

⁴Walter Van Dyke Bingham and Bruce Victor Moore, How to Interview. New York: Harper & Brothers, 1931, 3.

analysis of the examination itself with a view to determining the validity and reliability of the examination; and third, a survey of the external factors inherent in the examination and situations which influenced the ratings of the judges.

A brief survey of the literature

Just then in an interview I made the following statement to this effect, almost as many as there are fields of endeavor. The same manager, the same people, the same officer would all have a different concept, yet each has a process which could legitimately be called an interview. The definition of Synovels, which defines an interview as the process by which one person, called the interviewee, directly communicates with another person, called the interviewer, in order to learn from him, seemed to be one of the best available. The same definition is expressed more succinctly by Synovels who defines an interview as a conversation with a purpose.

Since the art of communication has been in existence, one of the most popular practices has been the asking of questions. Many persons with little technical training feel quite capable of arriving at a correct judgment based on very little

1. Synovels, "The Interview," *Journal of Management Psychology*, Vol. 5, No. 1, 1930.
2. Synovels and John Synovels, "The Interview," *Journal of Management Psychology*, Vol. 5, No. 1, 1930.

acquaintance.⁵ The more scientifically trained, however, realize more and more that judgments based on brief observation are quite often falacious.

Experiments conducted by Hollingworth,⁶ Corey,⁷ and Moss⁸ have all yielded results which would seem to point to definite limitations of the usefulness of oral examinations. Their unanimous conclusion is that judgment by experts is unreliable. Moss, for instance, found that if the lower twenty-five per cent of the candidates seeking admittance to medical school had been refused admittance on the strength of an oral interview that thirty-three per cent of the failures in medical school would have been eliminated, but at the same time twenty-three per cent of those averaging above eighty-five per cent in medical school subjects also would have been eliminated.⁹

On the other hand Bingham feels that, although the interview method is subject to a certain amount of just criticism, it is, nevertheless, a means to an end which can not

⁵Edward N. Hay, "Sizing Up Job Applicants," Personnel Journal, 18-7: 258, 1940.

⁶S. M. Corey, "The Interview in Teacher Selection," Journal of Educational Research, 26: 526, 1933.

⁷Loc. cit.

⁸E. I. Hoveland and E. F. Wonderlie, "Prediction of Industrial Success from a Standardized Interview," Journal applied Psychology, 23: 544, 1939.

⁹Loc. cit.

be approached in any other way. He feels that with careful instruction of interviewers and control of conditions the interview may be made to yield valuable results.¹⁰ This is also the opinion of others who, while admitting the inadequacy of the interview as a measuring instrument, point out that a poor estimate is after all better than no estimate at all.¹¹

One of the most forceful criticisms of the interview, which tries to judge ability and character from a very brief appearance before a group of judges, has been made by Ordway and O'Brien. They point out, quite justifiably, that the habitually sloven fellow may on occasion change his appearance by shaving and putting on a tie, and that the surly person may assume a charming manner. In their opinions such changes merely indicate that at the time of the interview the person was well dressed and agreeable, but that from this evidence no conclusions can be drawn concerning the interviewee's normal dress and manner.¹² They suggest as a remedy that what the person has done in the past should be taken as a criterion of what he will do in the future. This would

¹⁰Bingham, op. cit., p. 204.

¹¹Milton Mandell, "Civil Service Oral Interviews," Personnel Journal, 18-10: 380, 1940.

¹²S. H. Ordway and James C. O'Brien, "New Methods of Interview Hiring Review and Extract from Report," Personnel Journal, 18-4: 128, 1939.

be approached in any way...
investigation of the...
interview...
also the opinion of...
quality of the interview...
that a poor quality...
all.

One of the most...
which tries to judge...
appearance before...
and O'Brien...
habitually...
and by...
person may...
chances...
person was...
evidence...
viewer's...
that when the...
a collection of...

to...
...
...
...
...
...
...
...
...
...

necessarily involve a thorough investigation, which would be both expensive, time consuming, and, consequently, impractical.

This approach has also been attacked, not only because of the time and expense involved; but also because, as the exponents of the character trait rating method point out, there is no evidence to indicate that "legal analysis" of functional evidence is more valid than a psychological analysis of behavioristic evidence.¹³

Some interviewers favor a direct approach in which cross examining is permissible;¹⁴ others favor an indirect approach.¹⁵ Some examiners think that no writing should be done during the interview;¹⁶ others disagree. The only thing on which there seems to be some agreement is that there is and can be no such thing as a compulsory interview.¹⁷ In this connection it might be pointed out that the police would seriously object to that statement.

It can then be seen that interviewing is not a well-developed instrument of precision, but is rather an

¹³Mandell, op. cit., p. 374.

¹⁴Ordway and O'Brien, op. cit., p. 133.

¹⁵Bingham, op. cit., p. 47.

W. White, "Interview-Atmosphere," Employment Service News, 5, No. 3: 12, 1928.

¹⁶W. Williams and L. Kennedy, "Psychology in the Interview," Educational Trends, 9: 23, 1940.

¹⁷Bingham, op. cit., p. 198.

approximating instrument which serves to give indications, not facts. The present status of the interview is summed up by Symonds, who points out that by its very nature an interview must be free and unsystematized, but that interpretation demands a certain amount of standardization.¹⁸

Interviews are used, however, for employment purposes. If this use is continued, more definite knowledge must be obtained.

Procedure and Organization

Source of Data

The information for the present study was obtained from the private files of the New Mexico State Merit System Council. This information consisted of the oral rating forms, application blanks, training and experience records, and final rating sheets for each candidate examined in the 1941 testing program.

Procedure

Due to the fact that the material involved was in some cases very awkward to handle in its original state, the data from all sources were collected and punched into Hollerith cards. The statistical computations were then performed, using the Hollerith Sorter, following the method outlined by DuBois.¹⁹

¹⁸Symonds, op. cit., p. 353.

¹⁹P. H. DuBois, "Some Statistical Operations on the Counting Sorter," Psychometrika, 6: 383, 1941.

As only ten persons were interviewed for the position of Receptionist and less than thirty for the position of Case Worker in Training, these persons were not included in the study. Analysis was then made of two hundred applicants for the position of Visitor and sixty-two applicants for the position of Interviewer. Due to the fact that eighteen persons took both examinations, the total number of cases dealt with was two hundred and forty-four rather than two hundred and sixty-two. All sixty-two of the Interviewers were males as well as seventy-two of the Visitors. The remaining one hundred and twenty-eight Visitors were females. For the Visitor examination the males and females were treated separately.

There were four interviewers used in the program: Kern, Crawford, Wivel, and Gibson. For purposes of this study, Wivel and Gibson were treated as one examiner. This was justified as there was only one point difference in their average score, and as of the two hundred and forty-four cases, Wivel contributed only thirty-five. The other two hundred and nine were all rated by Gibson.

The code system used in punching the cards in order to render them adaptable for statistical treatment was as follows:

Space

1-4 Identification Number

5 Examination Center

6	Kern Trait 1	
7	Kern Trait 2	
8	Kern Trait 3	
9	Kern Trait 4	
10	Kern Trait 5	
11	Kern Trait 6	
12	Kern Trait 7	
13-14	Kern Sum of Traits	
15	Kern Visitor Exceleration	X
16	Kern Interviewer Exceleration	X
17	Crawford Trait 1	
18	Crawford Trait 2	
19	Crawford Trait 3	
20	Crawford Trait 4	
21	Crawford Trait 5	
22	Crawford Trait 6	
23	Crawford Trait 7	
24-25	Crawford Sum of Traits	
26	Crawford Visitor Evaluation	
27	Crawford Interviewer Evaluation	
28	Wivel or Gibson	
29	Wivel-Gibson Trait 1	
30	Wivel-Gibson Trait 2	
31	Wivel-Gibson Trait 3	
32	Wivel-Gibson Trait 4	
33	Wivel-Gibson Trait 5	

1. Introduction	3
2. General Principles	7
3. Theoretical Foundations	8
4. Experimental Methods	9
5. Results and Discussion	10
6. Conclusions	11
7. Acknowledgments	12
8. References	13
9. Appendix A	16
10. Appendix B	17
11. Appendix C	18
12. Appendix D	19
13. Appendix E	20
14. Appendix F	21
15. Appendix G	22
16. Appendix H	23
17. Appendix I	24
18. Appendix J	25
19. Appendix K	26
20. Appendix L	27
21. Appendix M	28
22. Appendix N	29
23. Appendix O	30
24. Appendix P	31
25. Appendix Q	32
26. Appendix R	33
27. Appendix S	34
28. Appendix T	35
29. Appendix U	36
30. Appendix V	37
31. Appendix W	38
32. Appendix X	39
33. Appendix Y	40
34. Appendix Z	41

34	Wivel-Gibson Trait 6
35	Wivel-Gibson Trait 7
36-37	Wivel-Gibson Sum of Traits
38	Wivel-Gibson Visitor Evaluation
39	Wivel-Gibson Interviewer Evaluation
40	Total Visitor Evaluation
41	Total Interviewer Evaluation
42	Spanish Test Evaluation
43	Race
44-45	Age
46	Sex and Marital Status
47-48	Height
49-51	Weight
52	Degree of claimed Spanish proficiency
53	Education
54-56	Visitor Written Score
57-59	Visitor Oral Score
60-62	Visitor Training and Experience
63-66	Visitor Final Rating
67-69	Interviewer Written Score
70-72	Interviewer Oral Score
73-75	Interviewer Training and Experience
76-79	Interviewer Final Rating
80	Veterans Preference

Organization

In the following chapters the material will be

1	...
2	...
3	...
4	...
5	...
6	...
7	...
8	...
9	...
10	...
11	...
12	...
13	...
14	...
15	...
16	...
17	...
18	...
19	...
20	...
21	...
22	...
23	...
24	...
25	...
26	...
27	...
28	...
29	...
30	...
31	...
32	...
33	...
34	...
35	...
36	...
37	...
38	...
39	...
40	...
41	...
42	...
43	...
44	...
45	...
46	...
47	...
48	...
49	...
50	...
51	...
52	...
53	...
54	...
55	...
56	...
57	...
58	...
59	...
60	...
61	...
62	...
63	...
64	...
65	...
66	...
67	...
68	...
69	...
70	...
71	...
72	...
73	...
74	...
75	...
76	...
77	...
78	...
79	...
80	...

Original copy

In the ...

presented in order as outlined in the statement of the problem. Chapter II deals with the preliminary statistical manipulation by which the ratings were converted into quantitative results subject to mathematical treatment. Chapter III presents the results obtained from intercorrelation of various combinations of traits together with a brief discussion of the reliability and validity of the oral test. Chapter IV deals with those characteristics of the examinees which might influence the ratings of the judges, and Chapter V presents the conclusions of the study.

CHAPTER II

PRELIMINARY STEPS

The Examiners

Four oral examiners were used in the 1941 examination program, although the board consisted of only three members at any one time. The chairman was James S. Kern of the Department of Social Work of the University of Denver. Other members were W. L. Crawford, administrative assistant of the Texas State Employment Service; Professor C. B. Wivel of Eastern New Mexico College; and E. Dana Gibson of New Mexico Highlands University. Both Kern and Crawford served during the entire period. Wivel served for two days and Gibson for the remainder of the program.

The examinations were scheduled in various parts of the State on specific days (Table I). As each candidate had previously received a time allotment, usually twenty minutes, there was no delay or confusion.

The Examinations

The Oral Interview Rating Form used in the program was made up of a list of seven traits. These were Voice, Appearance, Command of Language, Alertness, Ability to Present Ideas, Poise and Bearing, and Purpose. The term "Purpose" referred to the candidates purpose in applying for the job and his apparent interest in the work. Under each trait were

Four other members of the
Department, including the
at any one time.
Department of Health,
members were in the
Texas State
Eastern New Mexico
Highlands and
the entire party
the remainder of
The entire
the State on
previously received
there was no
The Oral
made up of
ence, General
these, Jones
returned to
and his

TABLE I

Exhibit G²⁰

Centers	Number Scheduled to Appear	Number Appearing	...	Secretary	...
Albuquerque	109	97	...	Mrs. Virginia Edwin	...
Las Cruces	31	31	...	Mrs. Concha Johnson	...
Las Vegas	68	62	...	Miss Margaret Long	...
Roswell	30	26	...	Capt. T. V. Price	...
Santa Fe	55	46	...	Mrs. Alice Greiner	...

²⁰New Mexico State Merit System Council, Annual Report of New Mexico State Merit System Activities, March 13, 1940, to December 31, 1941. Albuquerque: New Mexico State Merit System Council, 1942-45.

five descriptive choices, each of which expressed a greater degree of the trait in question than the previous choice.

At the end of the list of traits was a five-degree descriptive scale of endorsement which ranged from "Do Not Endorse" to "Endorse Very Strongly." An exact duplication of this last scale was printed on an auxiliary rating sheet, which was used to indicate the final degree of endorsement arrived at by the judges jointly.

Conduct

The candidate first presented himself to the secretary, who checked his admittance slip and identification number. She then introduced him to the examiners, using only his identification number. At the same time the secretary also gave each examiner a blank rating scale. During the course of the examination each of the examiners, working independently, gave the candidate a rating on each of the seven traits and on the endorsement scale. The candidate then retired, leaving the examiners to discuss his case and to arrive jointly at a recommendation, which was then checked on the auxiliary rating scale. This process was then repeated for the next candidate.

Due to the fact that acquaintance with the candidate might lead to prejudiced judgments, the examiners were instructed to refrain from rating anyone with whom they were previously acquainted. They were also instructed to rate each

five descriptive choices, each of which expressed a greater degree of the trait in question than the previous choice.

At the end of the list of traits was a five-point descriptive scale of endorsement which ranged from "no endorsement" to "strongly endorse". An early indication of this scale was obtained in an earlier study (1931) which was used to indicate the final degree of endorsement arrived at by the judges themselves.

Procedure

The candidates first presented themselves to the experimenter who explained the purpose of the study and the identification number. He then introduced him to the experimenter, giving him the identification number. At the same time the experimenter gave each candidate a blank rating scale. During the course of the examination each of the candidates working independently gave the candidate a rating on each of the seven traits and on the endorsement scale. The candidates then left leaving the experimenter to assign his case and to arrive later at a recommendation, which was then passed on to the next rating scale. This process was then repeated for the next candidate.

Due to the fact that acquaintance with the candidates might lead to prejudiced judgments, the examiners were instructed to refrain from rating anyone with whom they were previously acquainted. They were also instructed to rate each

trait separately without reference to any other trait, and were particularly instructed to fill in all blanks.

Initial Statistical Treatment

As all of the scales were in five steps, each step was assigned a quantitative rating from zero to four in ascending order, that is, the greater the quantity of the trait, the higher the numerical value. These trait values were then added in order to find a sum of all of the traits. The sums of the traits for each examiner were then changed into comparable scores by converting each examiner's score to a standard score with a mean of eighty and a standard deviation of fifteen. The individual ratings of each examiner for each position were similarly treated, being converted to scores having a mean of twenty and a standard deviation of six. Similarly the combined evaluation was converted to a scale having a mean of eighty and a standard deviation of twenty-seven. As the means were constant, it is apparent that half the weight of the oral score was the combined individual endorsements.

For each candidate there were then seven scores, that is, three trait scores, three individual endorsement scores, and one joint endorsement score. The seven scores were then added to give a total oral score for each candidate for each position.

This sum was then treated as a raw score and converted

trial separately without reference to any other trial, and
were particularly interested in this in all cases.

Initial Positional Treatment

An analysis of the material was made in five or six cases, and
assigned a descriptive rating from zero to four in ascending
order, that is, the greater the quantity of the trial, the
higher the numerical value. These trials were then
added in order to find a sum of all of the trials. The sum
of the trials for each examiner were then changed into a
percentage score by converting each examiner's total to a
standard score with a mean of eighty and a standard deviation
of fifteen. The individual ratings of each examiner for each
position were similarly treated, being converted to a score
having a mean of twenty and a standard deviation of five.
Similarly the combined examination was converted to a score
having a mean of eighty and a standard deviation of fifteen.
In the same way, if the means were combined, it is apparent that the
weight of the trial score was the combined individual
examinations.

For each position there were three scores, that
is, three trial scores, three individual examiner scores,
and one joint examination score. The scores were then
added to give a total score for each position for each
position.

into the desired range by means of the formula $70 + \frac{90\text{ile W} - 70}{90\text{ile O} - \text{XL}} (X - \text{XL})$, in which 90ile written is the ninetieth percentile of the converted written score; 70 is the lowest passing written score; 90ile oral is the ninetieth percentile of the oral raw scores; XL is the lowest passing oral score; and X is the score to be converted.

This formula converted the oral and written scores into a common range. The ninetieth percentile is used rather than the highest score in order to avoid any error which might arise due to a disproportionally high score.

The oral examination was then given a rating of twenty-five per cent in computing the final Visitor score and twenty-five per cent in computing the final Interviewer score.

Various criticisms have been leveled at the conversion method. The one which carries the most weight is that which points out that the public is very sceptical of grade manipulations, which it does not understand.²¹ This is very true, but due to the large number of unrelated values to be combined, the conversion method seemed to be the most appropriate statistically.

²¹ Mandell, op. cit., p. 382.

late the first part
of the first part
of the first part

percentage of the total population
passing within the limits of the
life of the first part of the
score; and the first part of the
This is the first part of the
into a second part of the
then the highest of the first part
also the first part of the
The first part of the
five per cent of the total
five per cent of the total
method. The first part of the
points of the first part of the
pulsations, and the first part of the
but the first part of the
the conversion of the first part of the
statistical data.

21

CHAPTER III

STATISTICAL ANALYSIS OF THE EXAMINATION

Intercorrelations of Traits by Examiners

For the purpose of analyzing the results of the oral examination, the raw scores, that is, the sum of the seven converted scores before re-conversion, were used as the totals, and the original numerical evaluations of the traits were used rather than the converted scores. At this point the scores of Wivel and Gibson were combined.

From the various intercorrelations of traits by examiners (Table II, Table III) it was found that there was not a very great degree of agreement. These intercorrelations range from .125 to .64 with a combined average of .49 (Table IV). In practical terms of prediction this indicated that given one judge's average score, another judge's average score could be predicted with only thirteen per cent accuracy. For the average sum (Table IV) this degree of prediction was raised to twenty-seven per cent. This was still unsatisfactory from the standpoint of accurate prediction.

When the correlations for each trait were averaged and arranged in order, it was found that there was the least agreement concerning the applicants appearance and the most agreement concerning his alertness. Alertness was closely followed by Purpose with Ability to Present Ideas, Poise and Bearing, Command of Language, Voice, and last, Appearance, following

in descending order (Table IV).

This seemed to indicate that those personality traits with no external criteria received a greater degree of agreement than more objective traits, such as Appearance.

It was also noted that the sum correlations were higher than the individual evaluation correlations or the average of the trait correlation. This indicated that a composit sum leads to a more reliable answer than a single overall judgment.

An examination of Tables V and VI shows that the average individual endorsements by the judges correlated higher with Purpose than with any other trait. Purpose was followed by Ability to Present Ideas, while Command of Language was lowest. This order was the same for both Visitor and Interviewer.

In Table VII will be found the correlation between the individual endorsement and the joint endorsement of the three judges. It will be noted that Crawford's individual endorsement correlated higher with the final endorsement than did those of either Kern or Wivel-Gibson. This might be due to the fact that Crawford, as an employee of the Employment Service, was in a better position to judge potential employees of that service, or that due to his prestige as such an employee, he carried more weight in the joint discussion.

Intercorrelation of Traits

On the basis of this high agreement between Crawford's

In accordance with the 1911

the same is maintained and

with no external change

more than were the

It was a very common

then the individual

the first of these

last of a number of

an example of

are individual

with various

by which is

lowest. This

Interview

In fact, it

individual

judges. It

ment correlated

those of other

the fact that

Service, which

of that service

ployees, the

On the basis of

TABLE II

Intercorrelation of Traits by Examiners
for Interviewee

Crawford

N = .63

	1	2	3	4	5	6	7	S	E
1	.54								
2		.57							
3			.64						
4				.56					
5					.46				
6						.51			
7							.60		
									Average = .55
S								.78	
E									.50

Kern

TABLE II (Continued)

Intercorrelation of Traits by Examiners
for Interviewee
Wivel-Gibson

	1	2	3	4	5	6	7	8	E
1	.29								
2		.39							
3			.46						
4				.62					
5					.62				
6						.43			
7							.70		
									Average = .50
8								.71	
E									.60

Kern

TABLE II (Continued)

Intercorrelation of Traits by Examiners
for Interviewee
Wivel-Gibson

	1	2	3	4	5	6	7	S	E
1	.35								
2		.32							
3			.46						
4				.49					
5					.40				
6						.54			
7							.40		
								Average = .42	
S								.60	
E									.41

Crawford

TABLE II (Continued)

Interpretation of Results by Examination
for Interference
Wavelength

	1	2	3	4	5	6	7	8	9
1	1.35								
2		2.22							
3			3.46						
4				4.42					
5					5.40				
6						6.34			
7							7.40		
8								8.30	
9									9.41

Oxwford

TABLE III

Intercorrelation of Traits by Examiners
for Visitor

Crawford

N = .199

	1	2	3	4	5	6	7	S	E
1	.48								
2		.36							
3			.50						
4				.58					
5					.58				
6						.49			
7							.52		
								Average = .50	
S								.69	
E									.57

Kern

TABLE III (Continued)

Intercorrelation of Traits by Examiner

for Visitor

Wivel-Gordon

	1	2	3	4	5	6	7	S	E
1	.35								
2		.25							
3			.34						
4				.53					
5					.52				
6						.43			
7							.58		
								Average = .43	
S								.63	
E									.56

Kern

TABLE III (Continued)

Intercorrelation of Traits by Examiner
for Visitor
Wivel-Gibson

	1	2	3	4	5	6	7	S	E
1	.34								
2		.44							
3			.46						
4				.59					
5					.61				
6						.51			
7							.55		Average = .50
S								.67	
E									.66

Crawford

TABLE IV

Average of Correlations for Each Trait

[Both positions]

Trait	1	2	3	4	5	6	7	S	E
Average of Six intercorrelations for each trait	.39	.39	.48	.56	.53	.49	.56	.68	.55
Trait Average				.49				.68	.55
Traits in descending order	4	7	5	6	3	1	2		

TABLE V

Correlation of Traits with Examiner Evaluation
for Visitor

				Trait			
Evaluation	1	2	3	4	5	6	7
Kern	.51	.50	.59	.75	.76	.70	.75
Crawford	.61	.46	.60	.66	.72	.65	.74
Wivel-Gibson	.31	.21	.34	.57	.56	.46	.75
Average	.48	.38	.51	.66	.68	.60	.75

TABLE VI

Correlation of Traits with Examiner Evaluation
for Interviewer

Evaluation	Trait						
	1	2	3	4	5	6	7
Kern	.51	.41	.54	.60	.78	.79	.76
Crawford	.60	.64	.57	.59	.68	.60	.71
Wivel-Gibson	.55	.44	.55	.55	.68	.64	.80
Average	.55	.50	.55	.58	.71	.68	.76

TABLE VII

Combined Evaluation Correlated with Individual Evaluations
for Both Visitor and Interviewer

Evaluation	Total Evaluation	
	Visitor	Interviewer
Kern	.77	.76
Crawford	.83	.84
Wivel-Gibson	.78	.64
Average	.79	.75



1915									
1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	100

evaluation and the joint rating, Crawford's cases were chosen for an inter-trait analysis. The results are presented in Table VIII. The least relationship seemed to be between Appearance and Purpose, the highest between Alertness and Ability to Present Ideas.

Validity and Reliability

The question naturally arises as to how valid and reliable was the oral examination. This is rather difficult to answer. According to the basic theory of oral examining the oral examination should be measuring something which is not measured by either a written examination or a training and experience evaluation. If this is the criterion used, the oral examination was in this case valid as the correlation between the oral and written examination, for Visitor was only .29 and between the oral and a training and experience evaluation .00. For Interviewer the oral examination correlated .34 with the written and .14 with training and experience. This, however, does not show what the oral examination was measuring, although the positive agreement among the judges indicates that it was measuring something. Regardless of what that something is, it can be of little use in determining the employment success of an individual until it has been studied in connection with employment success. As no such external criterion was available, no definite statement can be made concerning the

validity of the oral examination.

The question of reliability can be answered a little more exactly. The fact that the sum correlations were higher than the individual endorsement correlations indicated a greater reliability in a seven-fold observation of a single person than in one single overall observation. The ideal method of determining the reliability would be to again have the same individuals rated by the same judges, at the same places, for the same traits. If the examinations are considered as a three-fold observation of a single person, the reliability is as great as the intercorrelation between judges. This is .49 for trait averages, .68 for sum averages, and .55 for average individual evaluations.

A24

Validity of the oral examination.

The question of reliability can be answered a little more exactly. The fact that the same conclusions were drawn from the individual independent observations indicates a greater reliability in a seven-fold observation of a single person than in one single overall observation. The ideal method of determining the reliability would be to again have the same individuals rated by the same judges, at the same place, for the same trials. If the examination was considered as a three-fold observation of a single person, the reliability is as great as the inter-observer reliability. This is .45 for credit averages, .48 for non-credit averages, and .55 for average individual evaluations.

CHAPTER IV

EXTERNAL FACTORS

According to Symonds there are four factor variables in any oral examination. These are variables within the subject, variables in examiners, variables in general situations, and variables in the form and content of the interview.²²

As only one board of three examiners was used in the Merit System oral program, the variability of the board was negligible. The variability then was between examiners. This variability has been pointed out in the previous chapter.

The variability in the form and content of the interview could not be investigated due to lack of available information. As a standardized procedure was followed, this variability should not have been great.

Factors Inherent in the Situation

The variability in general situations was likewise subject to a very cursory examination due to lack of information. As can be seen by a glance at Table II, the examinations were conducted in five centers in the State. Table IX is a summary of the results of the examination by center. The average oral score for the Las Cruces applicants was

²²Symonds, op. cit., p. 349.

...in any case, the ...
...and ...

...As ...
...negligible ...
...variation ...

...The ...
...view ...
...formation ...
...variability ...

...The ...
...subject ...

...section ...
...there ...
...is a ...
...The ...

slightly higher than that for any other center. This was true for both position of Visitor and Interviewer. This may have been due to the errors of sampling or to some unmeasured variable.

Factors Inherent in the Applicant

The last variable on Symond's list is that of personal differences or those variables within the applicant. As was seen by an examination of the average scores by center (Table IX), one possible explanation of the higher average score was unmeasured factors. Further investigation of this possibility showed that non-Spanish men were on the average slightly higher than Spanish men, but that Spanish women were on the average higher than either non-Spanish women or men of either category. A further survey of Table X and Figure I also showed that women of either category were higher than men. The bi-serial correlation between sex and oral score for Visitor was .23, which was not considered high. The bi-serial correlations between race and oral score were lower, being .12 for Interviewer and .13 for Visitor.

In Table XI is presented the average oral score by age. In the case of Interviewers the trend was slightly downward, that is, the older the person, the lower the score (Figure 2). This was shown by a correlation of $-.30$ between age and score for Interviewer. For Visitors the picture was not so clear.

variable.

at 10:00 AM on 10/10/1944, the following was observed:

seen by an observer at 10:00 AM on 10/10/1944.

(1) one bird was seen at 10:00 AM on 10/10/1944.

was observed at 10:00 AM on 10/10/1944.

bird was seen at 10:00 AM on 10/10/1944.

bird was seen at 10:00 AM on 10/10/1944.

on the average at 10:00 AM on 10/10/1944.

either at 10:00 AM on 10/10/1944.

also shown at 10:00 AM on 10/10/1944.

was. The bird was seen at 10:00 AM on 10/10/1944.

for 10:00 AM on 10/10/1944.

seen at 10:00 AM on 10/10/1944.

being 10:00 AM on 10/10/1944.

in 10:00 AM on 10/10/1944.

in the case of 10:00 AM on 10/10/1944.

that the bird was seen at 10:00 AM on 10/10/1944.

This was seen at 10:00 AM on 10/10/1944.

for 10:00 AM on 10/10/1944.

TABLE IX

Average Oral Score by Center

		Interviewer		Visitor					
			N	Male	N	Female	N	Total	N
Albuquerque	1	390.5	25	357.2	23	373.8	52	368.7	75 *
Santa Fe	2	339.5	11	348.9	23	393.2	13	364.9	36 *
Las Vegas	3	397.2	12	354.5	19	401.8	28	382.7	47
Roswell	4	359.3	7	360.0	4	392.3	16	385.9	20 *
Las Cruces	5	422.6	7	405.7	3	411.3	19	410.5	22
		382.82	62	352.9	72	389.7	128	376.5	200

Interviewer

 $\bar{M} = 382.82$ $\bar{O} = 76.53$

Visitor

 $\bar{M} = 376.5$ $\bar{O} = 77.88$

*The total numbers for these centers is larger than the number appearing in Table II due to the fact that one person might take both examinations.

FIGURE I

Average Oral Score by Race and Sex

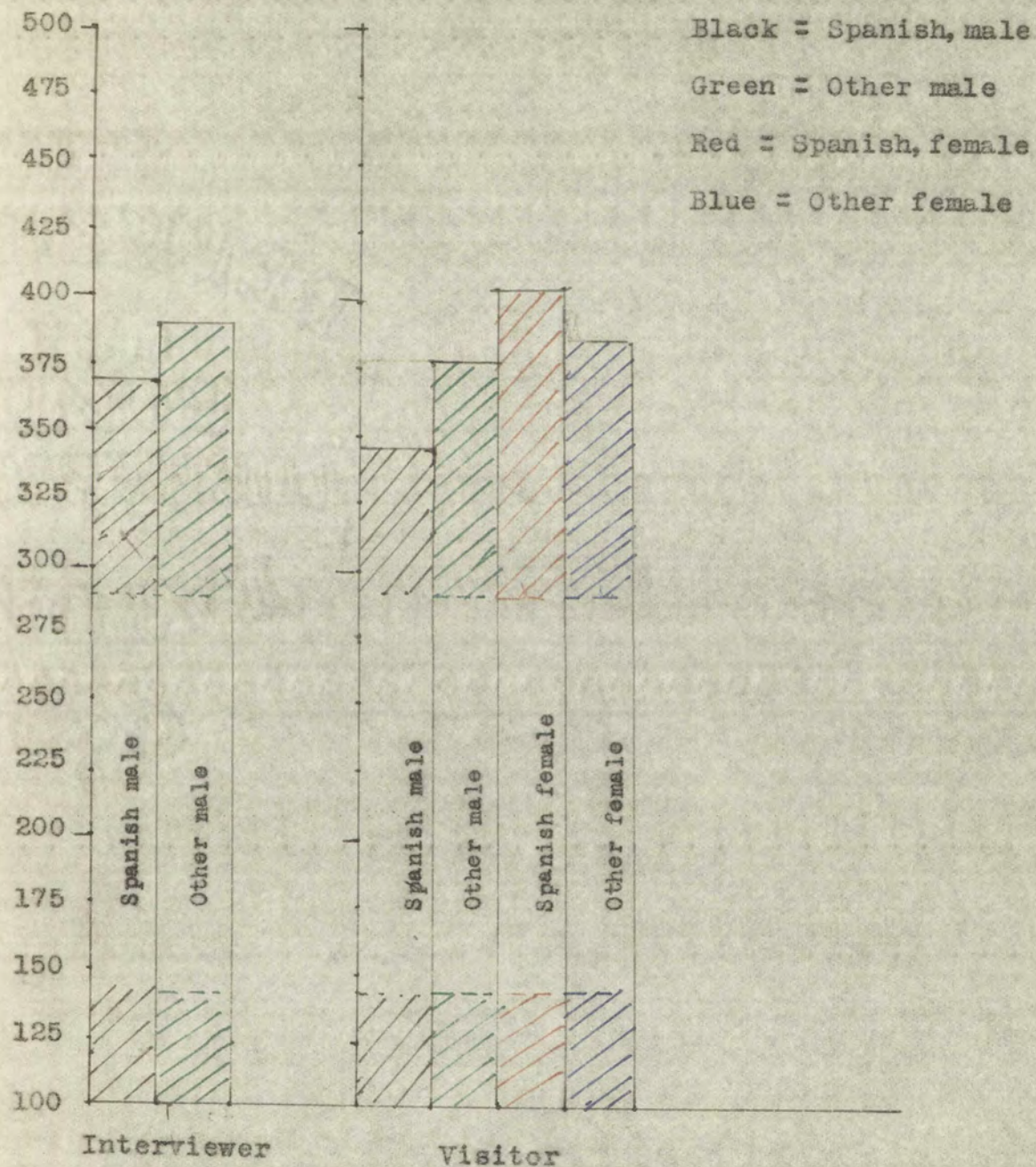


FIGURE 1

Average Oral Score by Race and Sex

Black = Spanish male
Green = Other male
Red = Spanish female
Blue = Other female

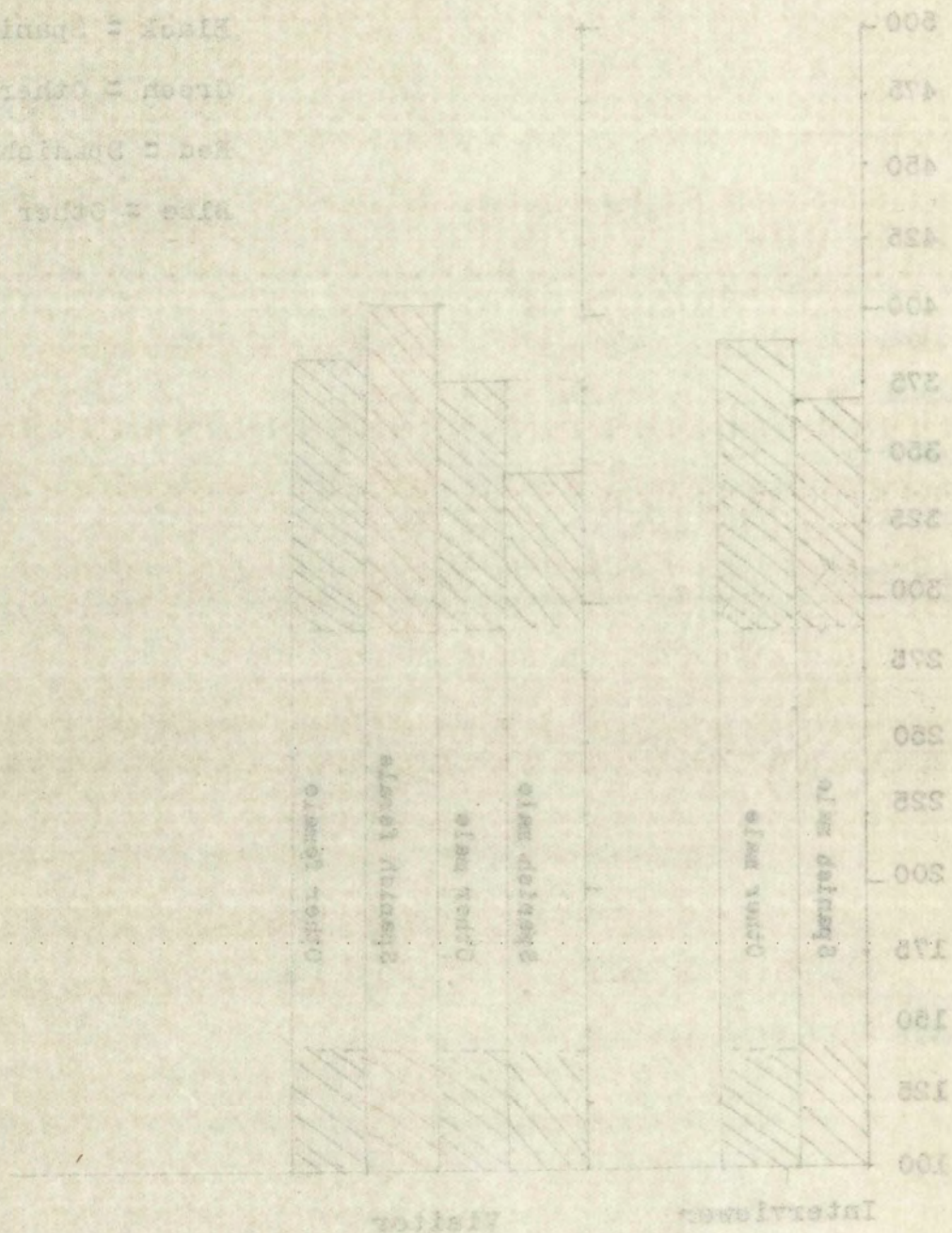


TABLE X

Average Oral Score
by Linguistic Group

Linguistic Group	Interviewer		Visitor					
		N	Male	N	Female	N	Total	N
Spanish-speaking	368.3	17	342.3	49	406.4	24	363.3	73
Spanish-speaking and other	388.3	45	375.7	23	385.9	104	384.1	127
	382.82	62	352.9	72	389.7	128	376.5	200

Visitor r -bis_{sex-oral} = .23

Visitor r -bis_{race-oral} = .13

Interviewer r -bis_{race-oral} = .12

TABLE XI
Average Oral Score

Age	Interviewer		Visitor					
	All Male	N	Male	N	Female	N	Total	N
20-24	426.1	10	387.3	13	361.9	23	372.8	36
25-29	383.3	15	355.9	28	387.8	30	372.4	58
30-34	402.5	11	345.9	19	410.8	26	383.4	45
35-39	363.1	8	325.4	5	390.0	15	373.8	20
40-44	365.5	6	322.3	3	394.7	16	383.3	19
45-49	368.0	3		0	429.4	8	429.4	8
50-54	350.6	5	273.5	2	380.3	7	356.6	9
55-59	344.0	2	327.0	1	337.0	1	332.0	2
60-64	326.5	2	354.0	1	306.0	2	322.0	3
N =	382.82	62	352.9	72	389.7	128	376.5	200

Interviewer $r_{\text{age-oral}}$ = .30

Visitor $r_{\text{age-oral}}$ = .04

Visitor $r_{\text{age-oral}}$ = .18

TABLE IV

Estimated Gross Income

Age	All	Visitor			
		Male	Female	Male	Female
20-24	432.1	10	367.3	15	344.9
25-29	502.3	12	400.2	20	367.4
30-34	405.3	11	345.9	19	300.4
35-39	362.1	9	305.4	8	260.0
40-44	362.3	8	322.2	2	302.7
45-49	362.0	1		0	252.4
50-54	361.8	2	275.3	2	270.5
55-59	364.0	2	321.0	1	207.0
60-64	368.2	3	324.0	1	206.0
H =	362.82	65	322.2	57	252.7

Interviewer = 30
 Age-crea
 Visitor = 40
 Age-crea
 Visitor = 18
 Age-crea

FIGURE II
Average Oral Score
by Age

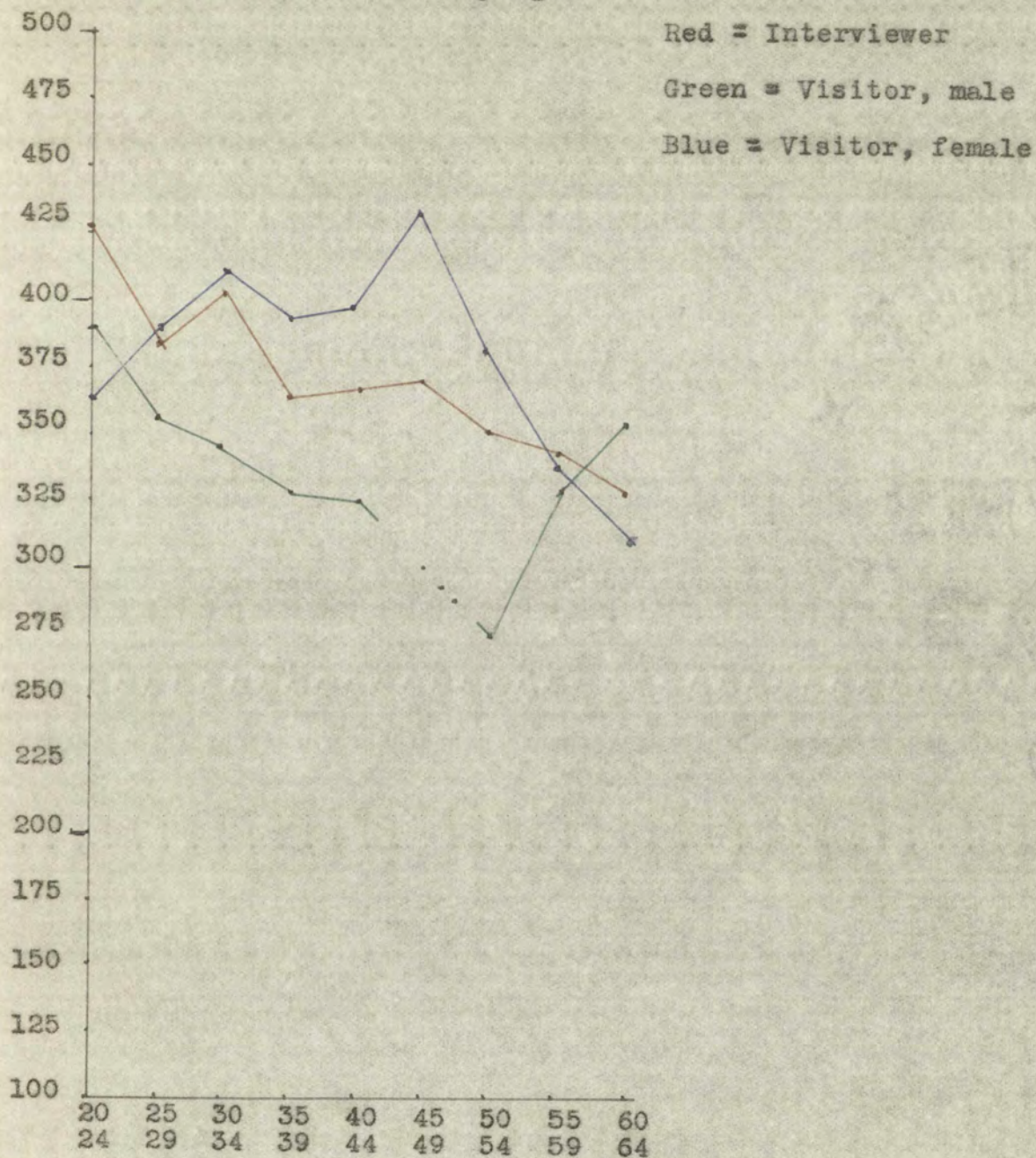


FIGURE II

Average Civil Service

by Age

Red = Incumbents
 Green = Veterans, new
 Blue = Veterans, former

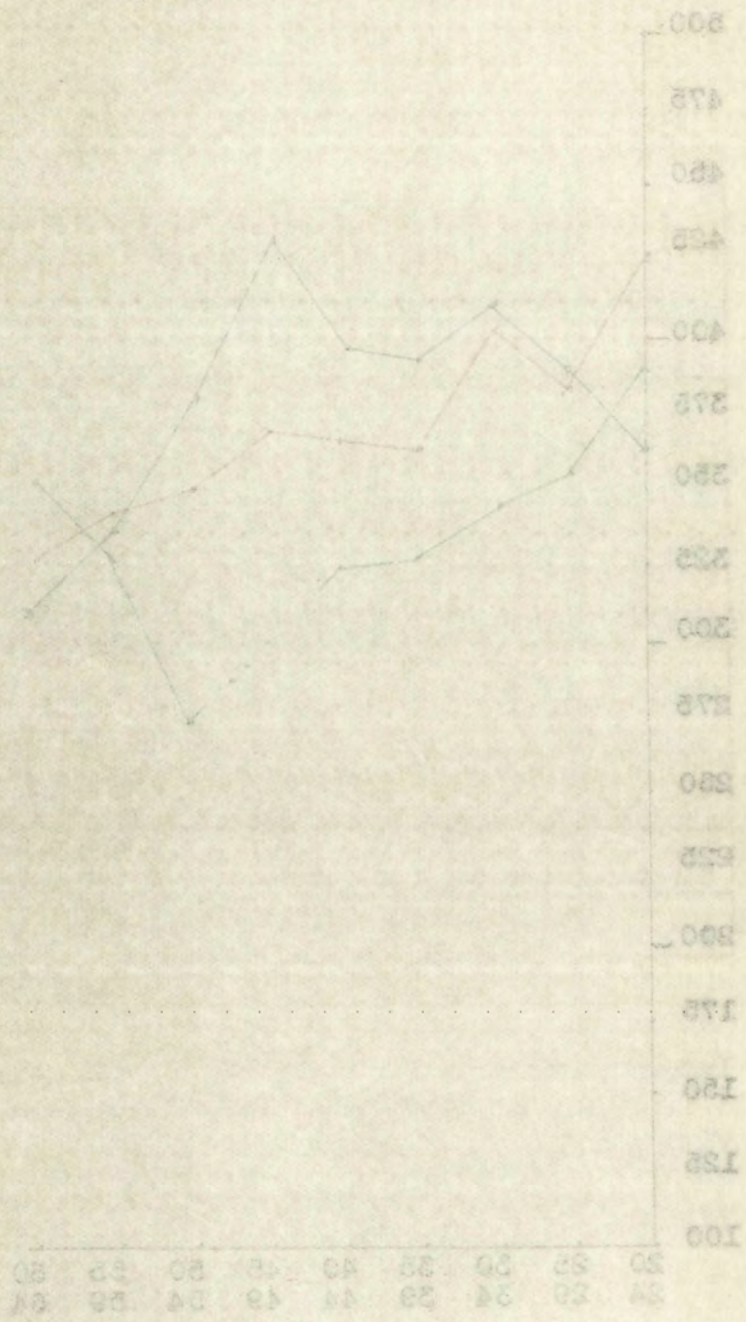


TABLE XII

Average Oral Score
by Education

Education in Years Above High School

Education	Interviewer		Visitor					
		N	Male	N	Female	N	Total	N
0	352.5	15	344.9	10	422.6	16	392.7	26
1	381.0	5	371.6	9	416.8	8	392.8	17
2	403.6	7	336.4	22	370.6	26	354.9	48
3	305.0	1	348.1	12	371.6	13	360.3	25
4	387.8	24	369.1	18	386.6	53	383.2	71
5	408.8	8	479.0	1	424.4	11	428.9	12
6	458.0	1			199.0	1	199.0	1
7	377.0	1						
	382.82	62	352.9	72	389.7	128	376.5	200

r Interviewer-Education = .25

r Visitor-Education = .03

TABLE III

Average Oral Counts

by Location

Station in Room above Main Floor

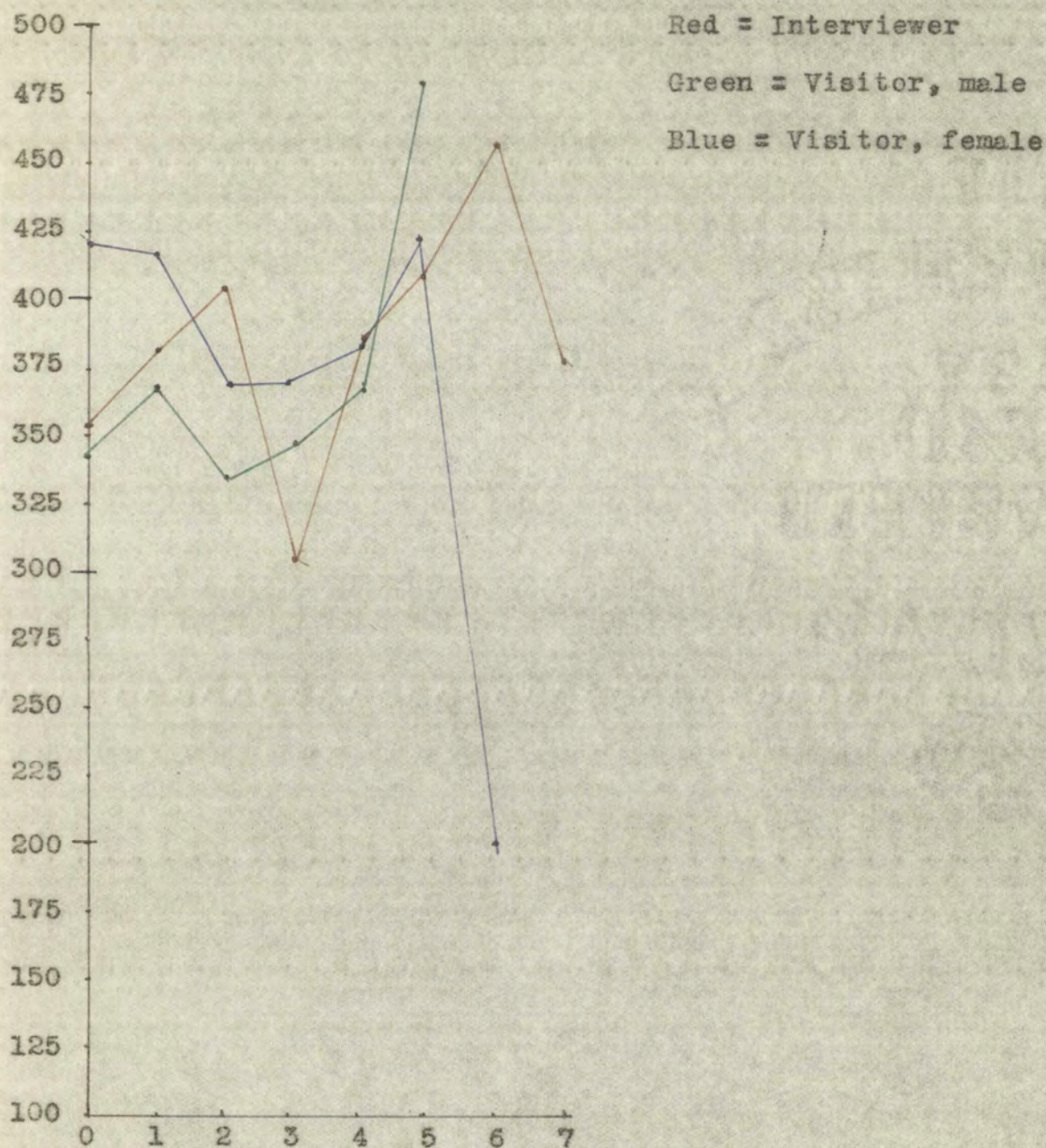
Station	Interviewer		Visitor			
	N	Ratio	1	2	3	4
0	10	14.9	10	4.5	10	10.1
1	3	17.5	3	4.5	3	17.5
2	7	13.4	12	11.3	10	14.7
3	1	14.1	12	11.3	12	14.0
4	24	10.1	15	10.0	22	10.1
5	8	13.0	1	14.1	11	14.7
6	1	14.0		17.5	1	14.0
7	1	17.0				
	22	10.1	18	10.1	18	10.1

r Interview-Station = .35

r Visitor-Station = .01

FIGURE III

Average Oral Score by Education



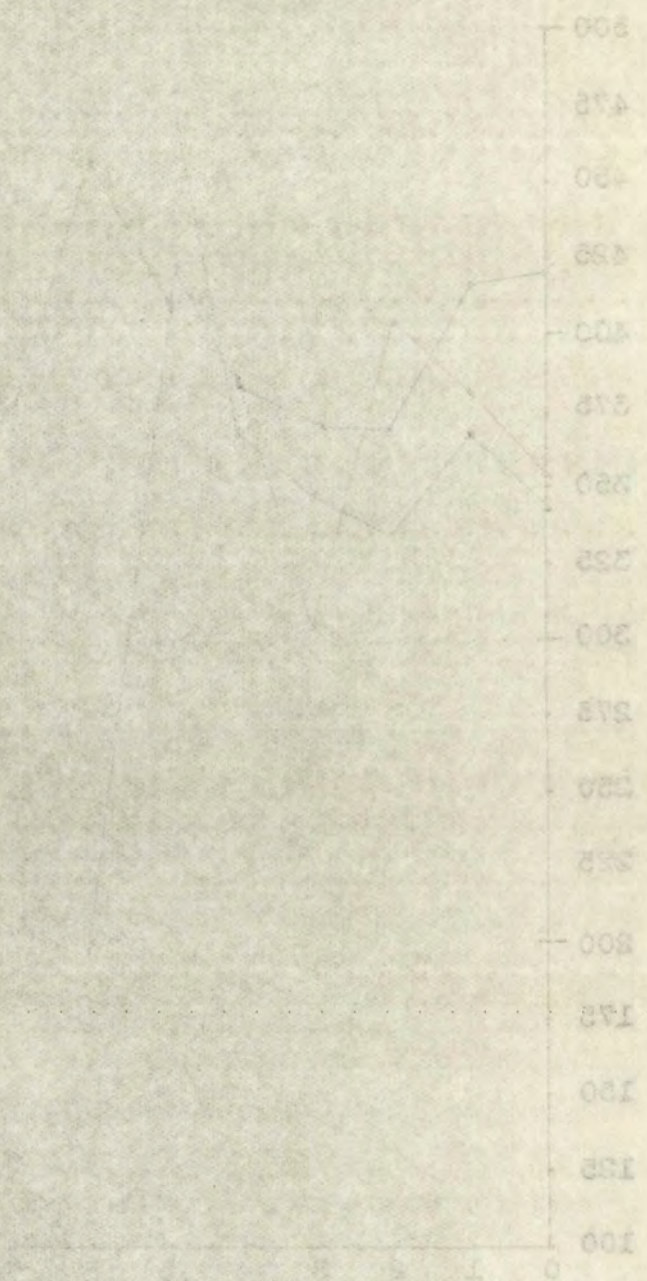


TABLE XIII

Average Oral Score by Weight

Weight	Interviewer		Visitor					
	All Male	N	Male	N	Female	N	Total	N
90-199					365.2	18	365.2	18
100-109					382.0	35	382.0	35
110-119	475.5	2	470.0	1	402.0	23	404.8	24
120-129	410.2	6	377.2	10	414.2	14	398.8	25
130-139	410.3	3	322.1	7	375.4	14	357.6	21
140-149	315.7	14	227.3	23	413.0	10	353.3	33
150-159	406.9	19	370.5	15	368.2	5	370.0	20
160-169	371.2	9	355.5	8	442.0	3	379.1	11
170-179	409.8	4	384.8	5	411.7	3	394.9	8
180-189	382.0	2	341.5	2	367.7	3	357.2	5
190-199	444.5	2		0				0
200-209	350	1	378.0	1			378.0	1
	382.82	62	352.9	72	389.7	128	376.5	200

Interviewer $r_{wt.-oral} = .11$

Visitor, male, $r_{wt.-oral} = .00$

Visitor, female, $r_{wt.-oral} = .20$

JANUARY 1951

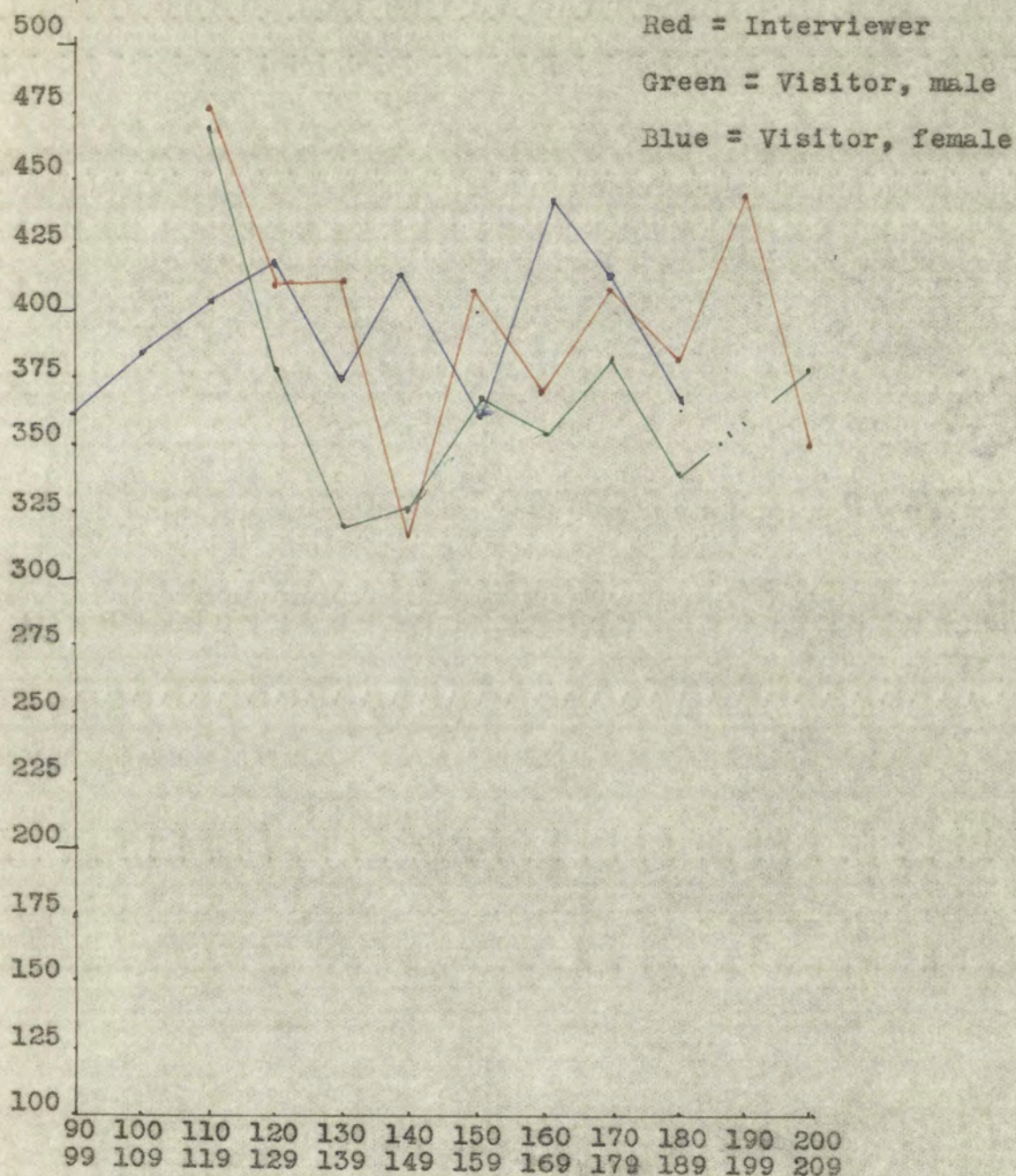
Average Oral Score by Weight

Weight	Interviewer		Visitor			
	All	Male	Male	Female	Total	N
20-100						22
100-150						22
150-175	416.5	2	416.0	1	408.0	22
175-180	410.2	2	371.2	10	344.0	22
180-185	410.1	2	382.1	1	373.4	21
185-190	378.7	14	327.7	22	312.0	22
190-195	408.8	12	370.2	12	308.2	22
195-200	371.2	2	322.2	2	242.0	21
200-205	402.8	4	324.2	2	271.7	22
205-210	322.0	2	201.2	2	227.7	22
210-215	444.2	2		0		2
215-220	320	1	272.2	1	270.0	2
	322.22	22	222.2	72	222.7	222

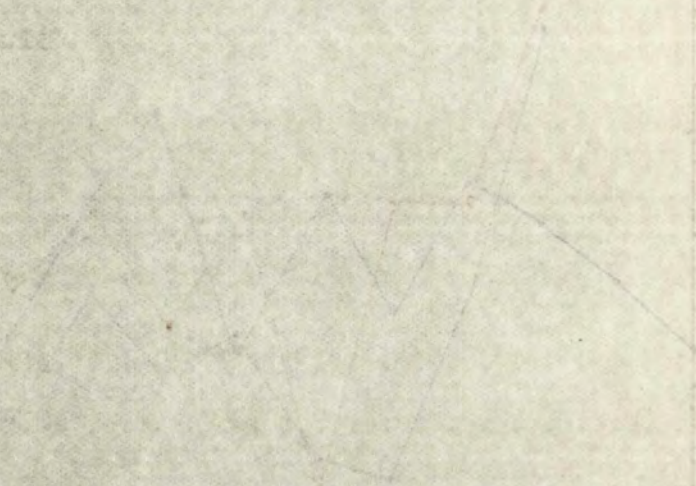
Interviewer = 11
wt.-oralVisitor, male = 20
wt.-oralVisitor, female = 120
wt.-oral

FIGURE IV

Average Oral Score by Weight



300
275
250
225
200
175
150
125
100



90 100 110 120 130 140 150 160 170 180 190 200 210 220 230 240 250 260 270 280 290 300

TABLE XIV

Average Oral Scores

Height	Interviewer		Visitor					
	All Male	N	Male	N	Female	N	Total	N
57-58					367.0	1	367.0	1
59-60					371.4	7	371.4	7
61-62			328.0	2	396.3	29	391.8	31
63-64	388.0	3	339.5	8	379.6	36	372.3	44
65-66	385.7	9	330.8	17	402.1	33	377.9	50
67-68	400.9	19	358.0	21	363.5	20	360.7	41
69-70	387.3	12	377.2	12	284.0	2	361.4	14
71-72	352.2	13	353.9	11		1	353.9	11
73-74	381.0	5	479.0	1			479.0	1
75-76	351.0	1						
	382.82	62	352.9	72	389.7	128	376.5	200

Interviewer $r_{ht.-oral} = .11$

Visitor, male, $r_{ht.-oral} = .21$

Visitor, female, $r_{ht.-oral} = .03$

TABLE XIV

Average Owl Scores

Height	Interviewer		Visitor	
	All	Male	Female	All
57-58			387.0	1 387.0
58-60			371.4	7 371.4
61-62			386.3	22 386.3
63-64	388.0	3 388.0	379.8	23 379.8
65-66	383.7	2 383.7	400.1	13 400.1
67-68	400.9	19 400.9	383.0	20 383.0
69-70	387.8	13 387.8	388.0	5 388.0
71-72	386.8	12 386.8		1 386.8
73-74	381.0	3 381.0		1 381.0
75-76	381.0	1 381.0		
	388.32	62	389.7	122 389.7

Interviewer = II.
nt.-totalVisitor, male = SI.
nt.-totalVisitor, female = S2.
nt.-total

For women there was a slight increase until the age of forty-five to forty-nine. At this point the oral scores were at a maximum. From this point they descended. Due to the fact that the relationship did not follow a straight line, the correlation between oral score and female age should have been slightly higher. For Visitor males the trend was again toward decreasing oral scores with advancing age. $r = -.18$.

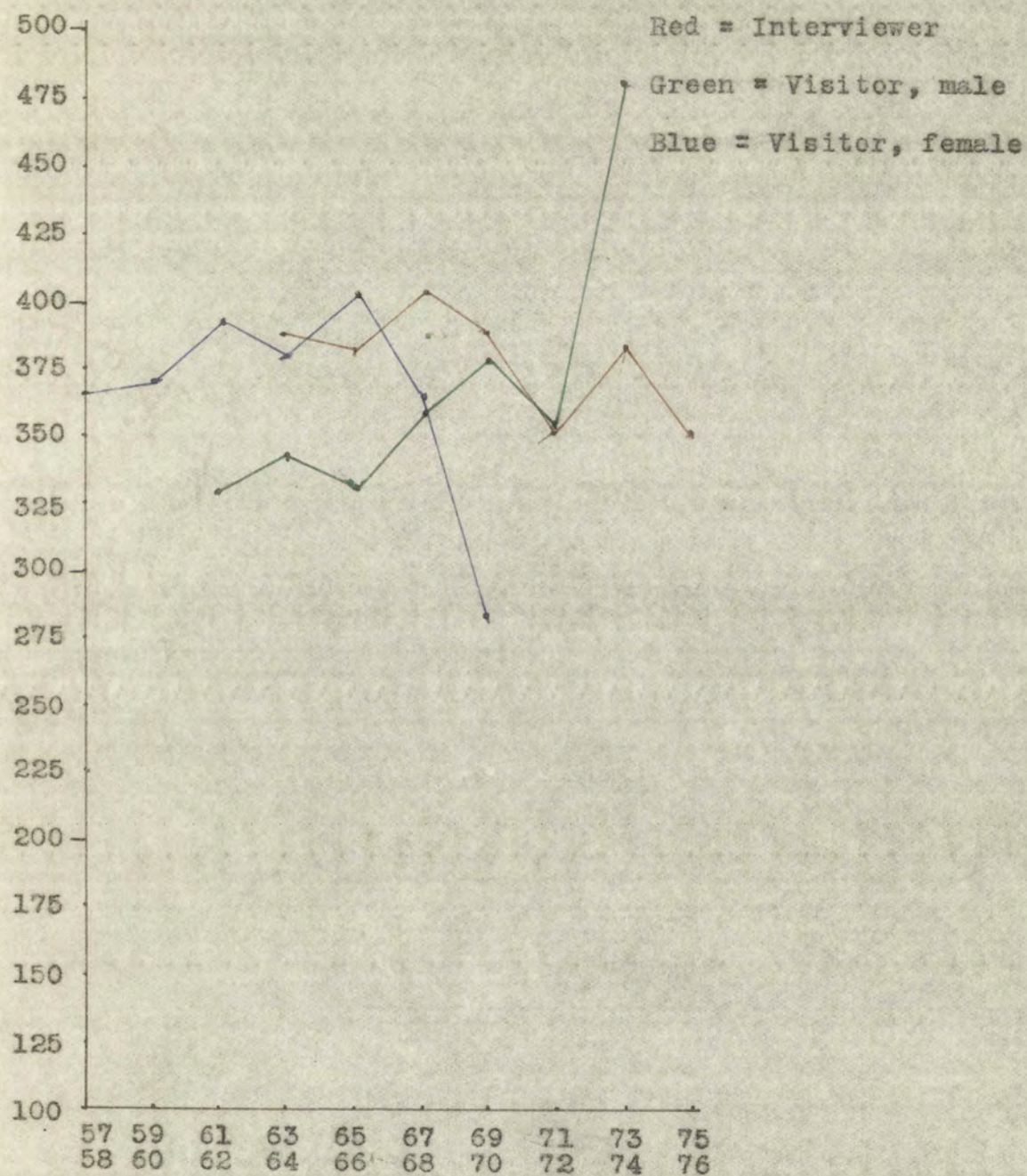
Table XI and Figure III show the results obtained by averaging the oral scores for grades completed above high school. Education apparently had little effect upon the rating given the applicant by the oral examiners. For the position of Interviewer there was a slight trend toward higher scores with greater education, but in the case of Visitor applicants this trend was negligible.

Similarly there seemed to be no significant trend where height and weight were concerned as can be seen by an examination of Table XIII, Table XIV, Figure IV, and Figure V.

From the evidence available it was concluded that the variables studied were of limited significance in determining the rating which the candidate received. If any statement were to be made, it would be that sex seems to be a small influencing factor as females were higher than males. Spanish women were higher than non-Spanish and higher than either male group. Spanish men had the lowest average score of any group. Education showed a slight positive trend, while height and

FIGURE V

Average Oral Score by Height



500
475
450
425
400
375
350
325
300
275
250
225
200
175
150
125
100

57 52 47 42 37 32 27 22 17 12 7 2

weight seemingly were unimportant factors. Applicants for Interviewer had a higher average score than those for Visitor.

weight is usually the only one of importance.

Investigation of a number of cases has shown that

Victims.

111

112

113

114

115

116

117

118

119

120

CHAPTER V

CONCLUSIONS AND SUGGESTIONS

FOR FUTURE STUDY

The oral examination is at best an approximating instrument. There is at present no common agreement as to what this instrument is, how to apply it, or what it measures. From the evidence presented it is clear that an oral examination does not measure the same thing that is measured by either a written examination or a training and experience evaluation. What this something is is not at present known. Traits such as Interest and Alertness seem to be rated more nearly alike by various judges than do more objective traits, such as Appearance. This may be due to several things. One possibility is that there is more actual agreement concerning such traits. Another is that certain common pre-conceptions are used. These may or may not be valid.

Due to the fact that the summary correlations were higher than the individual endorsement evaluation correlations, it can be stated that seven traits, or a seven-fold appraisal of an individual gives a more reliable picture than one overall evaluation. For this reason seven traits are better than one, and three judges are better than a single judge.

The validity of the oral examination could not be determined from the data in hand as no accurate external criterion was available. If such a criterion is considered as

The first thing I noticed when I stepped out of the car was the cold, crisp air. It was a relief after the warm, stuffy interior. I looked around, trying to get my bearings. The street was wide and empty, with a few distant lights visible on the horizon. I felt a sense of isolation, as if I had been dropped into a new world. The silence was broken by the occasional sound of a car engine or a distant siren. I took a deep breath, feeling the cool air fill my lungs. The night sky was dark, with a few stars visible through the light pollution. I felt a mix of emotions, from excitement to apprehension. I knew that this was my chance to start over, to begin a new chapter in my life. I took a step forward, feeling the pavement beneath my feet. The night was young, and I was ready to face whatever came my way.

something not measured by either a written examination or a training and experience evaluation, then the oral is valid. If a more exact standard is set up, no definite statement is possible.

If the ratings by the three judges are considered as a three-fold measure of the same thing, then the examination was as reliable as the intercorrelations between judges, that is, .49 for the trait averages, .68 for average trait sums, and .55 for average individual evaluation.

From an external standpoint the judges were not greatly influenced by any of the factors measured. Sex, race, education, height, and weight made little consistent difference.

After a period of time has elapsed, a rating based on employment success might be obtained and correlated with the oral scores to determine the validity of the examination. Such a study would give a much more complete answer to some of the questions raised. Such a study might further be extended to determine the relative value of the oral examination in employment procedure.

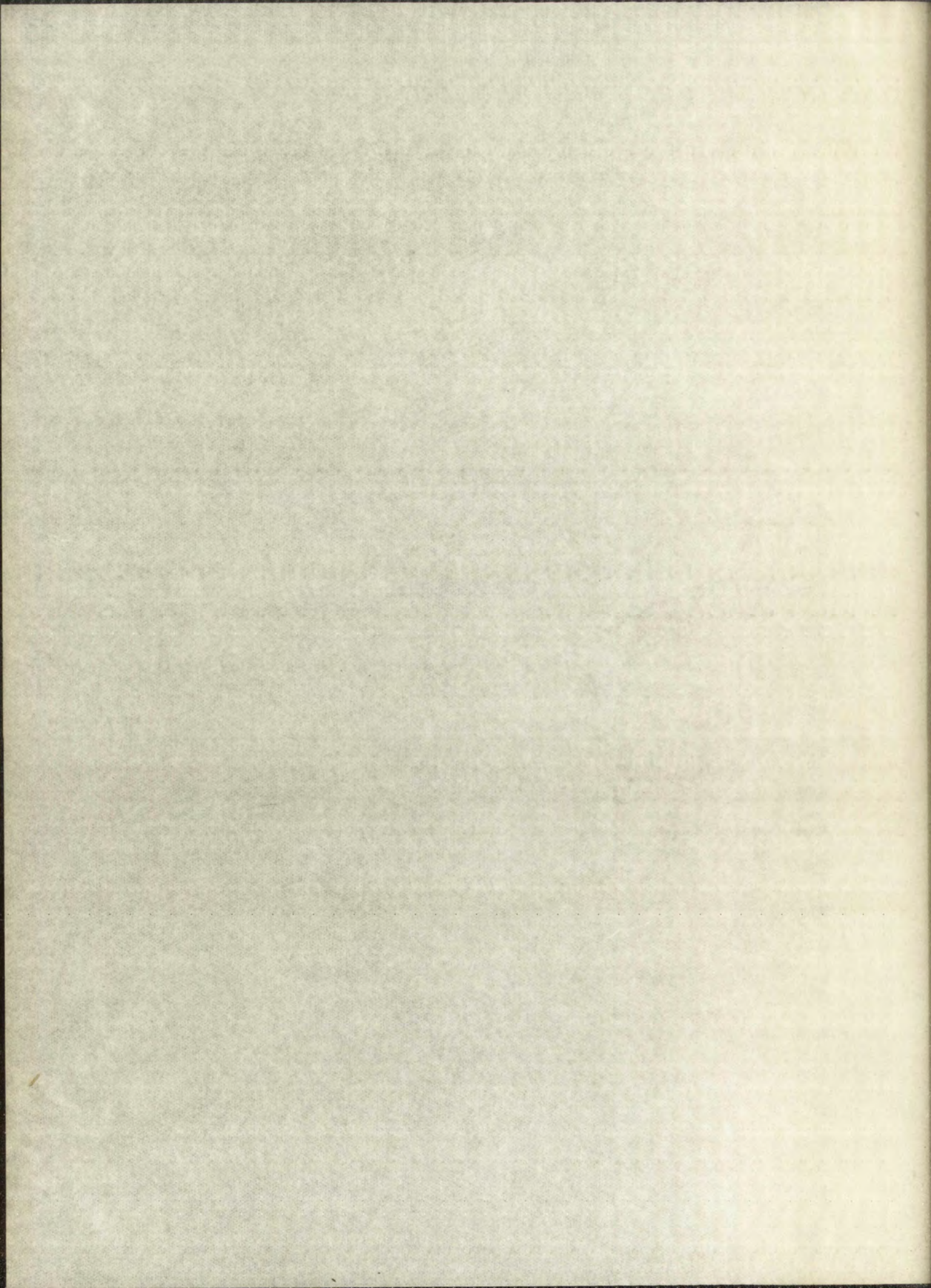
something was measured by either a written examination or a
training and experience evaluation; then the goal is valid.
If a more exact standard is set up, no definite statement is
possible.

If the results by the three judges are considered as
a three-fold measure of the same thing, then the examination
was as reliable as the intercorrelations between judges, that
is, .45 for the three averages, .65 for average first judge
and .55 for average individual evaluation.

From an external standpoint the judges were not usually
influenced by any of the factors mentioned. Sex, race, education,
time, height, and weight were fairly consistent differences.

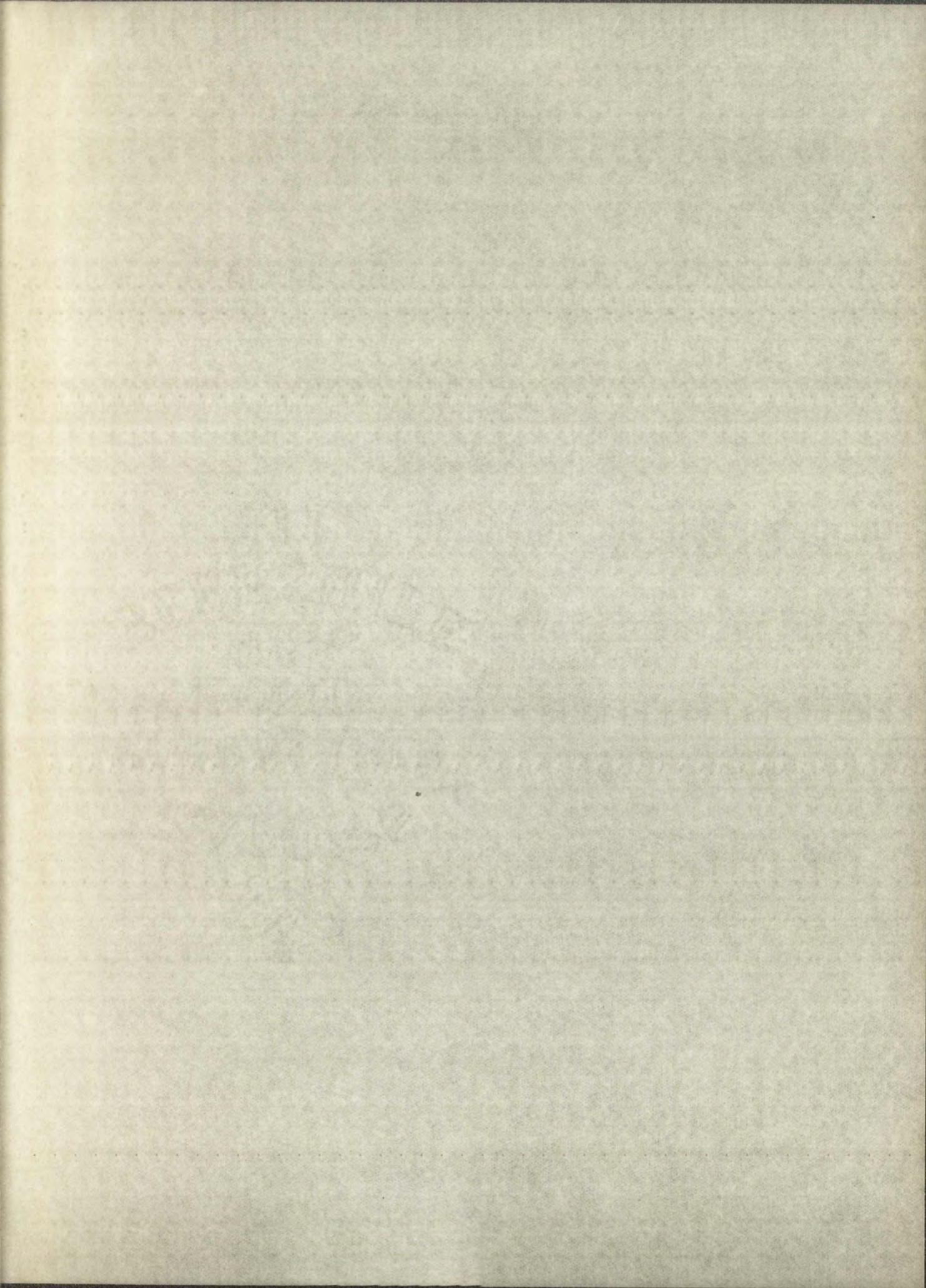
After a period of time has elapsed, a rating based on
employment success might be obtained and correlated with the
oral scores to determine the validity of the examination.
Such a study would give a more complete answer to some of
the questions raised. Such a study might further be extended
to determine the relative value of the oral examination in
employment procedures.

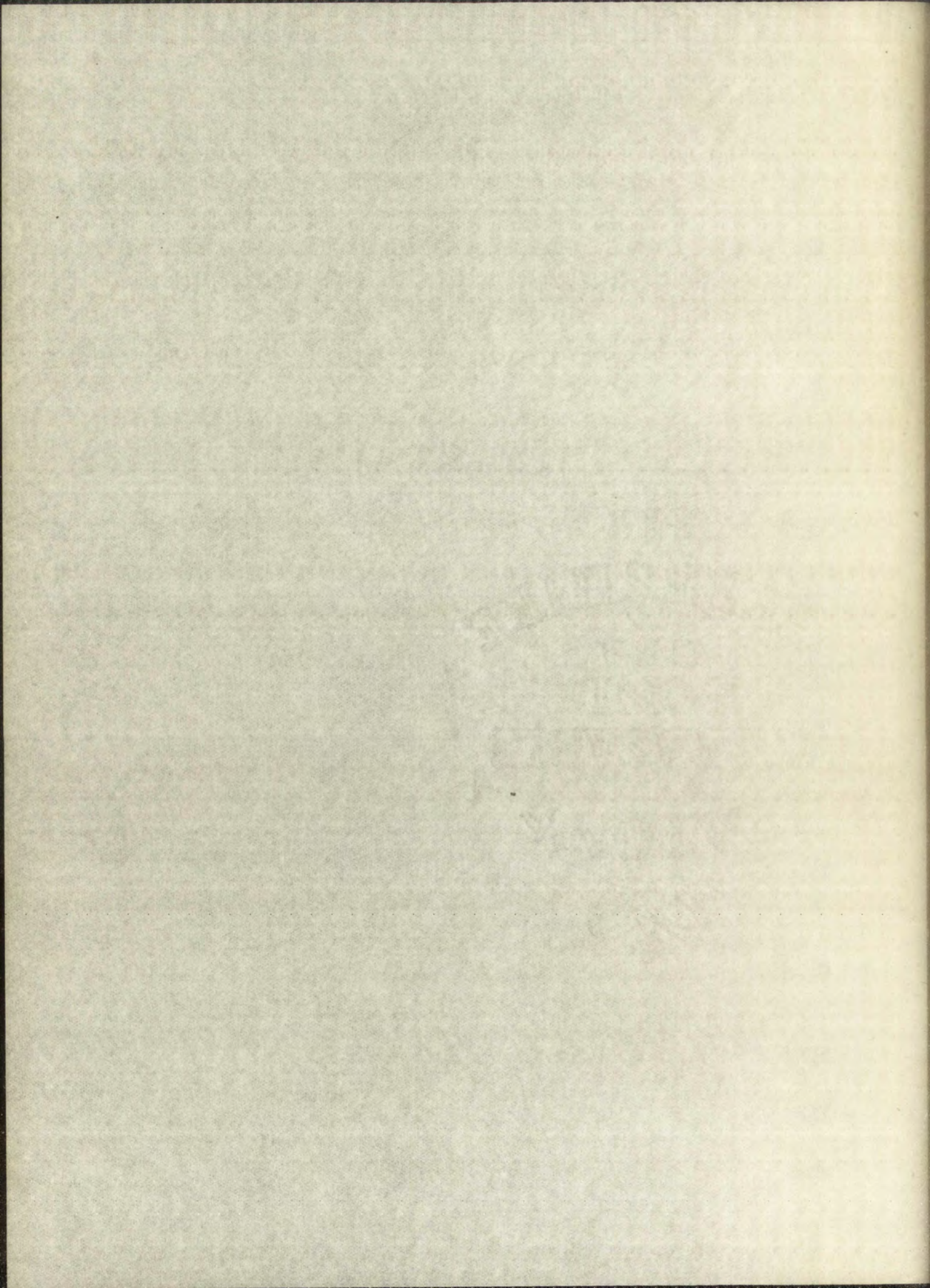
BIBLIOGRAPHY

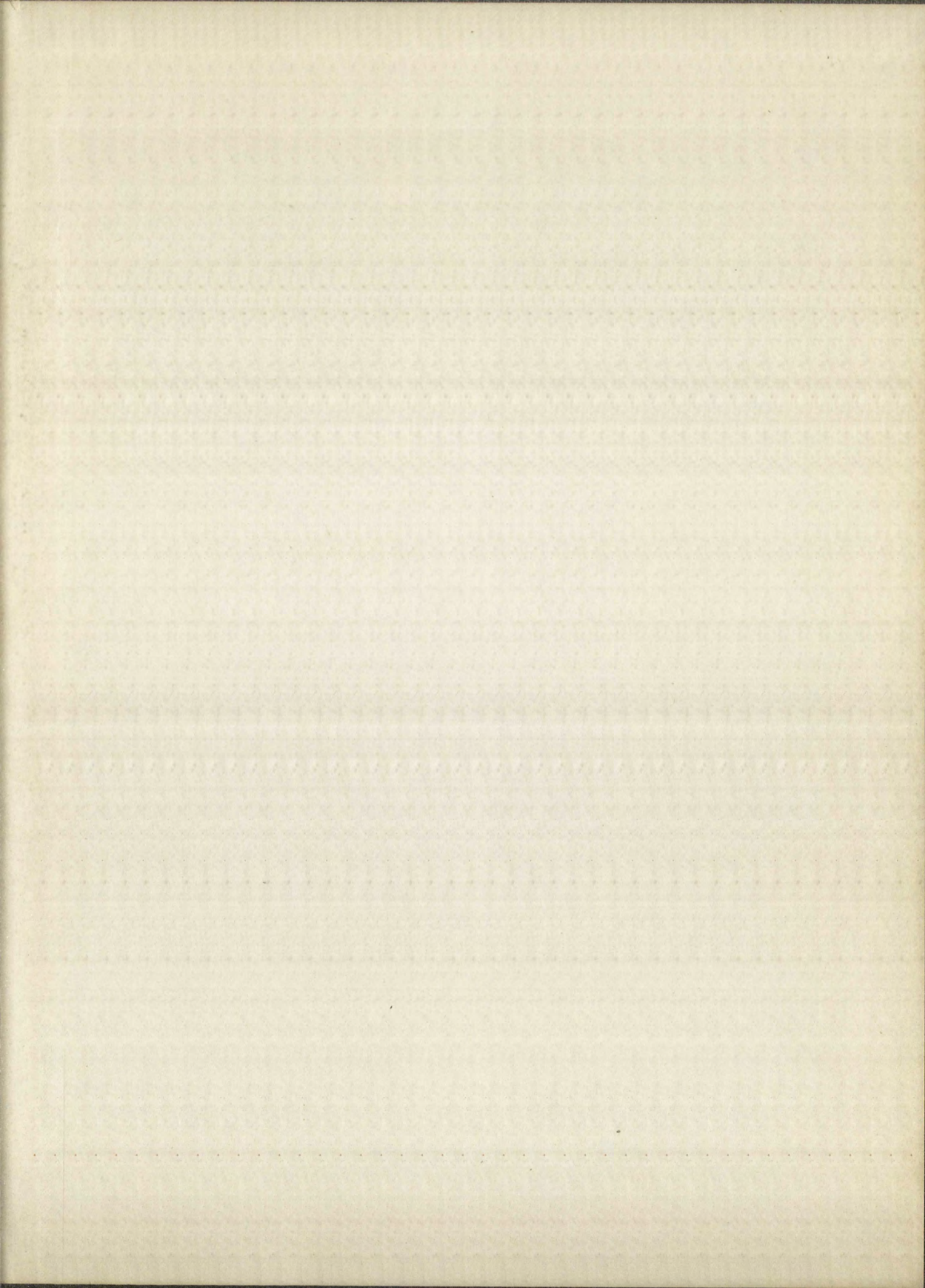


BIBLIOGRAPHY

- ✓ Bingham, Walter Van Dyke, and Bruce Victor Moore, How to Interview. New York: Harper & Brothers, 1931, 320 pp. -- xiv.
- ✓ Corey, S. M., "The Interview in Teacher Selection," Journal of Educational Research, 26: 525-531, 1933.
- DuBois, P. H., "Some Statistical Operations on the Counting Sorter," Psychometrika, 6: 383-390, 1941.
- Employment Security Commission of New Mexico, Employment Security Commission of New Mexico Regulation No. 8, Article V, Section I, Paragraph I, Part C.
- ✓ Hay, Edward N., "Sizing Up Job Applicants," Personnel Journal, 18-7: 258-261, 1940.
- ✓ Hoveland, E. I., and E. F. Wonderlie, "Prediction of Industrial Success from a Standardized Interview," Journal of Applied Psychology, 23: 537-546, 1939.
- ✓ Mandell, Milton, "Civil Service Oral Interviews," Personnel Journal, 18-10: 128-135, 1939.
- New Mexico Department of Public Welfare, "Classification and Compensation Plan of the New Mexico Department of Public Welfare," Rule for a Merit System of Personnel Administration in the New Mexico Department of Public Welfare, Article V, Section I, Paragraph I, Part C (Effective March 10, 1941).
- New Mexico State Merit System Council, Report of Merit System Activities, March 13, 1940, to December 31, 1941. Albuquerque: New Mexico State Merit System Council, 1942, 45 pp.
- ✓ Ordway, S. H., and James C. O'Brien, "New Methods of Interview Hiring Review and Extract from Report," Personnel Journal, 18-4: 128-135, 1939.
- ✓ Symonds, Perceval M., "Research on the Interviewing Process," Journal of Educational Psychology, 30: 346-353, 1938.
- ✓ Williams, W., and L. Kennedy, "Psychology in the Interview," Educational Trends, 9: 18-23, 1940.
- ✓ White, W., "Interview-Atmosphere," Employment Service News 5, No. 3: 12, 1928.







IMPORTANT!

[illegible]

