

THE EFFECTS OF SPECIFIC TRAINING ON
ACCURACY IN JUDGING OTHERS

Thesis for the Degree of Ph. D.
MICHIGAN STATE UNIVERSITY

John H. Wakeley

1961



3 1293 10343 3102

This is to certify that the
thesis entitled
THE EFFECTS OF SPECIFIC TRAINING ON ACCURACY IN
JUDGING OTHERS

presented by
John H. Wakeley

has been accepted towards fulfillment
of the requirements for

28V

Professor

LIBRARY
Michigan State
University

ABSTRACT

THE EFFECTS OF SPECIFIC TRAINING ON ACCURACY IN JUDGING OTHERS

by John H. Wakeley

Research in the area of training to improve accuracy in judging others has been remiss in specifying training and in relating training to a conceptualization of the problem of improving accuracy in judging others. In the present study an analysis of the activities performed by judges in a situation for measuring accuracy in judging others was made to provide a framework for the specification of training.

The analysis identified three aspects of the judging task. The judge gathers information about the stimulus person within the limits of the situation; he makes inferences about the stimulus person; and he records judgments by answering questions posed by the experimenter. From this analysis six training programs were developed and assessed for effectiveness in improving accuracy in judging others.

Two of the programs were concerned with training to improve judges' observations of others. One program, Observing-Self, instructed trainees to focus on their own reactions to the stimulus person when observing, and the second program, Observing-Other, instructed trainees to focus on the other person when observing.

Two of the programs were concerned with training to improve judges' inferences about others. One program, Inferring-Individual Differences, instructed trainees to concentrate on the individual's unique characteristics when making inferences, and the second program, Inferring-Pooling, instructed trainees to concentrate on the characteristics of the individual which made him similar to people who were well-known to the judge when making inferences.

One program, Recording-Rating, instructed trainees on avoiding three common rater errors in the use of rating scales.

The sixth program, Combination, instructed trainees on the three aspects of judging behavior, presented all of the principles of the other five programs and instructed trainees to use their discretion in choosing which principles to use.

All training programs consisted of a brief lecture, two practice periods with knowledge of results given and a discussion session.

The programs were assessed by using seven groups, six training groups and a control group, of introductory psychology students ranging in number from 17 to 22. Each group completed two criterion measures of accuracy in judging others, participated in one of the training programs or the control condition, and again completed the criterion measures. Each group was in session for three consecutive hours with one hour devoted to each of the three phases of the session. Sound-color films were used to present the stimulus persons to the judges.

Results for the criterion instrument, Accuracy in Judging People, did not suggest any change in performance after training. Results for the criterion instrument, Ability to Judge Differences Between People, indicated that only the programs Combination and Inferring-Pooling were effective in improving performance.

A second study used three groups, $N = 10$ or 11 , of engineers and personnel workers from an industrial population. The procedures of the experimental sessions were the same as for the first study but only the criterion instrument, Ability to Judge Differences Between People, was used.

Study Two confirmed the effectiveness of the Inferring-Pooling program and suggested that a modification of the Combination program was effective.

The results of the two studies indicated that:

1. training to improve observation does not appear to improve accuracy in judging others;
2. training to emphasize unique features of an individual when making inferences does not appear to improve accuracy in judging others;
- ✓3. training to emphasize the use of individually determined reference groups, a kind of stereotyping, improves accuracy in judging others;
4. training in the proper use of rating scales does not appear to improve accuracy in judging others; and
5. training with some combinations of principles does appear to improve judging accuracy.

Approved

John H. Wakeley
Major Professor

Date

11/11/11

THE EFFECTS OF SPECIFIC TRAINING
ON ACCURACY IN JUDGING OTHERS

By
John H. Wakeley

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Psychology

1961

G 22828
3/28/62

To Sue who smiled
at Matthew 7: 1-2

ACKNOWLEDGMENT

I wish to acknowledge the valuable guidance and encouragement of Dr. Carl F. Frost who served as Chairman of my Guidance Committee. I also wish to thank Drs. T. M. Allen, A. I. Rabin, and H. C. Smith whose assistance and criticism in ways official and unofficial aided the completion of this research.

My special thanks go to Dr. Harold Gray, Continuing Education Service, Regional Director, Benton Harbor, Michigan, whose interest and cooperation made possible the off-campus portion of this research.

TABLE OF CONTENTS

CHAPTER	Page
I. INTRODUCTION	1
Review of the Literature	2
Appraisal of the Literature	5
Statement of the Problem	6
Rationale for the Specification of Training Programs	6
Training Programs for the Studies	8
II. STUDY ONE	12
General Design	12
Comparisons of the Study	12
Subjects	13
Criterion Instruments	13
Procedures for the Experimental Sessions	15
Results	18
Summary of Results	25
III. STUDY TWO	
Comparisons of the Study	27
Subjects	28
Criterion Instrument	28
Procedures	29
Results	30
Summary of the Results	31
IV. DISCUSSION OF THE STUDIES	32
Training to Emphasize Observation	32
Training to Emphasize Inference	33
Training to Emphasize Recording	37
Training to Emphasize the Analysis of Judging Behavior	38
Further Research	39

TABLE OF CONTENTS - Continued

CHAPTER	Page
V. SUMMARY AND CONCLUSIONS	41
REFERENCES	44
APPENDICES	47

LIST OF TABLES

TABLE	Page
1. Reliabilities for the Two Criterion Measures	19
2. Covariance Analysis of the Test, Accuracy in Judging People	19
3. Means and Standard Deviations for the Test, Accuracy in Judging People, for All Groups of the Study	20
4. Covariance Analysis for the Test, Ability to Judge Differences Between People	21
5. Means and Standard Deviations for the Test, Ability to Judge Differences Between People, for all Groups of the Study	22
6. Comparison of Each Training Condition with the Control Condition for the Second Administration of the Test, Ability to Judge Differences Between People	23
7. t-Values for the Comparison of Differences Between the Differences	24
8. Comparison of the Pretraining Means on the Test, Ability to Judge Differences Between People, for All Subjects, Study One and Study Two	30
9. Matched t-Tests for the Three Groups of the Study . . .	31
10. Comparison of the tests, Accuracy in Judging People and Ability to Judge Differences Between People	86

LIST OF APPENDICES

APPENDIX	Page
A. Training Lectures for Study One	48
1. Observing-Self	49
2. Observing-Other	51
3. Inferring-Individual Differences	53
4. Inferring-Pooling	55
5. Recording-Rating	57
6. Combination	59
B. Criterion Instruments	61
1. Accuracy in Judging People	62
2. Ability to Judge Differences Between People. .	69
C. Instrument Used in Training	74
D. Training Lecture for Study Two (Other-Pooling). . .	80
E. Table 10: Comparison of the Tests, Accuracy in Judging People and Ability to Judge Differences Between People	85

CHAPTER I

INTRODUCTION

In everyday situations, we depend necessarily on our capacity to perceive and predict the behavior, thoughts and feelings of the other person. . . . Our socialization is reared on this foundation of perception of persons in terms of prediction. . . . All the subtle interchanges of love and friendship rest, however insecurely, on this tenuous skill in perception and prediction.

Bender and Hastorf (1950, p. 556)

There is little doubt that improvement in making finer and finer discriminations among other people is part of general social development. One does learn to tell males from females; one does learn to distinguish among people on the basis of age; and one does learn to distinguish among people on the bases of skin color, voice, size and many other variables of appearance, dress and behavior. The socialization process also requires that one make judgments about other people's attitudes, beliefs, personality characteristics and future behaviors. An important practical question concerning the ability to judge these aspects of other people is: Can the ability to judge others more accurately be improved by training?

Many psychologists have the belief that it is possible to train people to be more accurate in judging others. A major purpose of training programs in much of clinical, personnel and counseling psychology is to make the trainee a more accurate judge of others. Training to increase accuracy of interpersonal judgments is not, however, restricted to psychologists. Many training programs for salesmen, many training programs for interviewers and many training programs

which are designed to make individuals better leaders focus on improving accuracy in judging others.

However, most research which has assessed the value of training for making judgments more accurate has failed to substantiate the positive effects of training.

Review of the Literature

Early studies concerning the effects of training typically used photographs (posed or unposed), drawings or actors as stimuli. The judge's task was to determine the emotion being experienced from the facial expression. Jenness (1932) reviewed the research on facial expression and stated the general conclusion that training is not effective in increasing accuracy of judgments.

In a paper read in 1949 Cottrell (1950) discussed possible areas of research for social psychology and urged that more work be devoted to the empathic process. Among other questions which he thought might profitably be investigated was one which asked if training could be effective in improving accuracy in judging others. In the same paper Cottrell outlined a method for measuring accuracy in judging others. Dymond (1949) gave a more complete description of the method. The procedure advocated by these authors, which is presented below, is the one that has been most widely used in this research area.

Judges observe some person or group of people, and subsequently make judgments concerning some aspect of the judged person or group for which an operational measure is available. The operational measure is usually obtained by self report of the judged or through consensus of experts about the judged. The judgments are compared to the criterion measure to determine accuracy. The good judge is one who makes many judgments in accordance with the criterion measure, and the bad judge is one who makes few judgments in accordance with the criterion measure.

The measure obtained by the procedure outlined above has been called "empathy," "sensitivity," "social insight," "understanding of others," and "predictive accuracy." Most research using this measuring procedure has been concerned with determining characteristics of the good judge, characteristics of the judged which make him easy or difficult to judge, the form of and scoring procedures for criterion measures, methods for gathering criterion data, and real and assumed relationships between judge and judged which are related to accurate judging. Bruner and Tagiuri (1954) and Taft (1955) give reviews of this research and indicate the substantive findings. Bronfenbrenner, et al. (1958) also review these areas of research although not as extensively as the other reviewers.

Despite the interest in interpersonal perceptions, the availability of a popular research method and the interest in training psychologists and others to be better judges of others, relatively little work has considered the effects of training. Bruner and Tagiuri (1954) virtually ignore the topic of training in their review. They devote one paragraph to the topic and state, "Experience (training) has generally been assumed to be a correlate of ability to judge others accurately. Little systematic testing is available to prove or disprove the point." (p. 644) Taft (1955) devotes more space to the topic and reviews 11 studies done between 1924 and 1954. Taft concludes

The results on the comparative ability of nonpsychologists, psychology students, and professional psychologists to judge other people are partly obscured by the effect of similarity in age and academic status between J (judge) and S (subject). Attempting to allow for this effect in the reported results, physical scientists, and possibly other nonpsychologists, e.g., personnel workers, appear to be more capable of judging others accurately than either psychology students or clinical psychologists. . . . There is also evidence that suggests that courses in psychology do not improve ability to judge others and there is considerable doubt whether professional psychologists show better ability to judge than do graduate students in psychology.

(p. 12)

Since these reviews of the literature, little additional research has been accomplished; however, three studies are of importance.

Lundy (1956) found that by giving different sets of instructions to subjects he could influence the accuracy with which others were judged. Each of 52 subjects completed 15 items of the Allport-Vernon scale which was used as the criterion measure. Each subject, S_1 , met a second subject, S_2 , and completed the criterion measure as he thought the other had. Before meeting S_2 again each S_1 was given one of two sets of instructions: (1) pay attention to yourself in the interview which you are about to have; (2) pay attention to the other person in the interview which you are about to have. S_1 and S_2 met and talked together for five minutes. S_1 again judged S_2 . Those who had been told to pay attention to themselves judged S_2 to be more like themselves after the interview and were less accurate. Those who had been told to pay attention to the other person made more accurate judgments of S_2 after the interview.

Crow (1957) used movies of ten people for which criterion data were available on seven personality scales. Four groups of medical students ($N = 72$) judged the people. Two groups received training which was identified only as "a course in establishing physician-patient relationships," and two groups served as controls. Crow found that after the training period the trained groups were less accurate than the control groups in judging the standard people. Based on a negative correlation between the judges' variability across people and their accuracy, Crow concluded, "These results indicate that training programs devoted to increasing accuracy of interpersonal perception . . . run the risk of decreasing accuracy when they increase the trainee's responsiveness to individual differences." (1957, p. 358)

Crow and Farson (1960) in a study similar to the Crow study found that a training program which reduced responsiveness to individual differences significantly increased accuracy.

Appraisal of the Literature

Studies relating training and accuracy in judging others have generally employed one of two experimental designs. In both of the general designs, training is treated as the independent variable and is measured by amount of exposure. Procedures for training, content of the training and principles emphasized in the training are seldom specified.

One of the general designs measures training as amount of course work taken in psychology, e.g., Rabin (1950); Luft (1950). Studies which use this method for measuring training compare non-psychologists, (e.g. physical scientists, personnel workers), undergraduate students in psychology, graduate students in psychology and psychologists in various combinations. When training is measured in this way, what has been measured is not clear. An individual receiving training in psychology may study such diverse material as statistics, principles of interviewing, the physiology of rats and techniques of projective testing. The assumption that this training combines additively and is directly related to a criterion of interpersonal judging accuracy is one that is difficult to support.

The second general design employed is to give before and after measures of the criterion instrument with training interpolated (with or without control groups), e.g., Crow, (1957). In this approach training is usually measured by stating the course title and indicating the duration of the training. Again what goes on in training is not clear. The possibility exists that several, perhaps conflicting, principles for judging are presented and discussed. Interpretation of findings when the independent variable has not been specified is an ambiguous undertaking.

Statement of the Problem

The review of literature above gives little support to the common belief that accuracy in judging others can be improved by training. However, training has not been specified sufficiently to permit clear appraisal of training programs. If specification of training is to be accomplished so that more definite information about its effects on judging accuracy can be obtained, it is necessary to provide a framework for the training. A deficiency apparent from the review of the literature concerning accuracy in judging others and the effects of training is that no conceptualization of training in this area has been made.

The present study provides an analysis of judging behavior in the interpersonal situation. The analysis is taken as a framework which permits better specification of training and better definition of the effects of training. From the analysis simple, specific training programs have been formulated. The purpose of this experimental investigation is to examine the effects of the specific training programs on judging others accurately.

Two separate, related experiments were completed to accomplish the purpose of this study. The first experiment used introductory psychology students to test the effects of the training programs. The second experiment used engineers and personnel workers from an industrial population to test particular programs which appeared to be effective in the first experiment.

Rationale for the Specification of Training Programs

In order to provide a rationale which would permit specification of training programs and a systematic manipulation of training programs, an analysis was made of the interpersonal judging situation. The analysis is similar in approach to Johnson's (1955) general analysis of

judgment. The present analysis differs from Johnson's by being more specifically concerned with the behavior of people acting as judges of other people.

The analysis considers the behavior of the person acting as a judge in the usual situation which is used to measure judging accuracy and identifies three general classes of behavior. When a person in the research situation is asked to make a judgment about another person, he does three things. He gathers information about the stimulus person within the bounds established by the experimenter. He uses the information which he has gathered as a basis for inferences about the stimulus person's personality, attitudes, beliefs and behaviors in unobserved situations. Finally, the judge records some portion of his inferences by answering the specific questions posed by the experimenter. The judge who is accurate in gathering information, accurate in inferring and accurate in recording should be the judge who makes accurate judgments of others.

From this analysis it appears that an ideal training program would focus on providing trainees with sound principles to follow in gathering information, in making inferences and in recording judgments. Additional features of a training program to improve accuracy in judging others would be opportunities to practice the principles presented in the program and opportunities for the trainees to get knowledge of the results of applying the principles of the training program. To implement the ideal training program outlined above, it would be necessary to identify the sound principles which should be offered. Previous research has not clearly identified the necessary principles.

From the analysis of the judging situation, six training programs were formulated. The exact procedures of the programs are discussed in a later section of this report. All programs presented simple, specific instruction to trainees, provided for practice, provided for

knowledge of results, and provided for discussion. All programs were approximately one hour in duration. The use of volunteer subjects precluded training programs of any greater duration.

Training Programs for the Studies^{*}

In the experimental sessions specific to this study, subjects gathered information by observing color-sound movies of interviews with the individuals to be judged. Since watching movies is only one of several ways in which judges might gather information about another, e.g., the judge might read test protocols or examine biographical information, the gathering of information aspect of judging behavior was referred to as "observing" in this study. The terms "inferring" and "recording" were used to identify the other two aspects of judging behavior.

Six training programs were formulated. Two training programs differed in content but related to the observing aspect of the judging situation; two programs differed in content but related to the inferring aspect of the judging situation; one program related to the recording aspect of the judging situation; and one program outlined the analysis of the judging situation and presented all of the principles of the other five programs.

The guiding ideas in determining the content of the two programs for observing were (1) each program should contain a single, specific principle for observing and (2) the two programs should be opposed in content. The situation in which judging occurs was considered to have two major foci of activity, the observer and the observed (see Lundy (1956) above). Two programs were formulated so that one focus of activity was emphasized in each program.

^{*} See Appendix A for a more complete account of the training programs.

One program, identified as Observing-Self, presented the principle that the subject should concentrate on his own reactions to the other in the observing situation. The judge was urged to attend to his emotional responses to the other, to the way the other person made him feel and to how he (the observer) would feel if he were behaving as the observed was behaving. The focus of attention was the observer.

The second program, identified as Observing-Other, presented the principle that the judge should concentrate on the other person in the situation. The observer was urged to remain actively engaged in noticing and remembering as much as he could about the other person's appearance and behavior. The focus of attention was the observed.

In formulating the content of the programs for inferring, the guiding ideas were the same as those for the observing programs, namely, (1) each program should present a single, specific principle, (2) the two programs should be opposed in content. The formulation of the content of these programs was aided by the work of Cronbach (1955); Stone, Gage and Leavitt (1957); Bronfenbrenner, Harding and Gallway (1958) and Cline and Richards (1960). All of the above investigators support a position that the ability to judge others is a complex ability constituted of two major specific abilities. The factors of the complex ability are the ability to predict for the generalized other and the ability to predict for the specific other.

The measure of ability to predict for the generalized other (called stereotype accuracy by some investigators) is the accuracy of a judge when he is predicting the typical response of the social category to which the judged belongs. The measure of the ability to predict for the specific other (called differential accuracy by some investigators) is the accuracy of the judge when he predicts the individual's responses in accordance with unique aspects of the judged. The content of the two training programs relating to the inferring aspect of the judging situation emphasized the

generalized other in one case and the specific other in the second case.

One of the programs relating to inferring, identified as Inferring-Individual Differences, emphasized using unique features of the judged as the basis for making inferences. Judges were instructed that when they were making judgments about an individual they should be constantly aware of the ways in which the judged was unique. Further, they were urged to consider how the judged's behavior would be influenced by his unique qualities.

The second program of this pair, identified as Inferring-Pooling, emphasized using features of the judged which were common to his group as a basis for making inferences. In this program information concerning the social class of the judged was not given to judges. Judges were instructed to relate the judged person to other people who were well-known to the judge. Each judge thus used an individually determined reference group as the basis for making inferences about the judged.

A single training program was formulated regarding the recording aspect of the judging situation. This program, identified as Recording-Rating, gave instruction designed to reduce the occurrence of common rater errors. Judges were introduced to the most common rater errors (see Guilford (1954) and Ghiselli and Brown (1955)), and were urged to use care to avoid these errors. The rater errors discussed were the leniency error, the central tendency error, and the halo effect.

The sixth program, identified as Combination, presented the three-part analysis of the judging situation and discussed all of the principles which appeared in the other five programs. The judges were instructed to use the general analysis of the judging situation and to select from the principles which had been presented those principles which would permit them to make the most accurate judgments. No attempt was made to indicate when to use which principle, rather, the selection was entirely at the discretion of the judges.

These six training programs were the ones used to examine the effects of specific training on judging others accurately. Chapter II presents the first study which used these training programs; Chapter III presents the second study which used these training programs; and Chapter IV discusses the results of the two studies and the relationships of the results to the three-part analysis of interpersonal judging behavior which was made for this research.

CHAPTER II

STUDY ONE

General Design

The first study used seven groups with 17 to 22 subjects per group. Each group was pretested on two criterion measures, received one of the six training programs (the seventh group was a control group which received no training) and was posttested with the two criterion measures used in the pretesting. The total time per group was three consecutive hours with two ten-minute breaks.

Comparisons of the Study

Comparisons were made between pretraining and posttraining performance on both criterion measures for each of the seven groups. The control group, $n = 22$, was included to obtain a measure of the effects of practice on the criterion measures. The six experimental groups were used to determine the effects of the training.

Comparisons were made to determine the effects of:

1. training in attending to one's self when gathering information (This comparison used a group of 21 subjects who received the Observing-Self training program.);
2. training in attending to the other person when gathering information (This comparison used a group of 18 subjects who received the Observing-Other training program.);
3. training to emphasize unique features of the individual when making inferences about the other person (This comparison

used a group of 17 subjects who received the Inferring-Individual Differences training program.);

4. training to emphasize features of the other person which make him similar to a particular reference group when making inferences about the other person (This comparison used a group of 18 subjects who received the Inferring-Pooling training program.);
5. training on the use of rating scales in recording judgments about the other person (This comparison used a group of 22 subjects who received the Recording-Rating training program.); and
6. training with several principles presented within the analysis of judging behavior made in Chapter I (This comparison used a group of 21 subjects who received the Combination training program.).

Subjects

Subjects in this study were volunteers from the introductory psychology course. Subjects were recruited by posting seven sign-up sheets in a central location. All those who signed a particular sheet constituted an experimental group. The groups were designated according to the training program which they received.

Criterion Instruments

Both of the criterion instruments used in this study are based on the movie developed and tested by Cline (1955) and Cline and Richards (1958). The movie presents color-sound, filmed interviews with each of six different people. Each interviewee is questioned for four or five minutes about personal values, personality strengths and weaknesses,

reactions to the interview, hobbies and activities, self-conception and temper. Cline and Richards have gathered data on each interviewee and have used the data to develop criterion items and tests. Data were gathered from the interviewees by further interviews and by psychological tests. Data were also gathered from friends, family and close associates of the interviewees by interviews and by psychological tests. A complete list of tests used to gather data about the interviewees is found in Cline and Richards (1958, p. 3).

One criterion measure (See Appendix B, Number 1) used in the present study is a test of ability to judge others which was developed by Cline and Richards and modified by H. C. Smith. The test, Accuracy in Judging People, is based on the first three interviews in the film and consists of 30 questions about the interviewees. The 30 questions are divided into three groups. The first group deals with the interviewee's religious beliefs, e.g., the judge is asked to indicate on a five-point scale how much the interviewee agreed with such statements as, "I have sometimes been very conscious of the presence of God." The second group of items deals with personality traits, e.g., the judge is asked to indicate on a five-point scale ranging from "very unlike" to "very like" how well the interviewee is described by such adjectives as, "Shy." The third group of items deals with predictions of behavior in unobserved situations, e.g., the judge is asked to choose the one best completion for items such as,

"With her children:

- (1) she maintains quite firm and strict discipline.
- (2) she is about average in this regard.
- (3) they usually get their way."

A second criterion measure (See Appendix B, Number 2) used is a test of ability to judge differences among people. This test, Ability to Judge Differences Between People, was developed by Peiper (1960). The test requires that the first three filmed interviews from the Cline-Richards movie be shown to the judge. The judge is then presented with

74 items which comprise the test. The items require the judge to identify which of the three interviewees is best described by the particular items. For example, the judge is asked which of the three interviewees is most practical. Material covered in the test includes statements of beliefs, descriptive adjectives, sentence completion responses, statements about daily behavior, personality traits and brief personality descriptions.

The two criterion measures do not have any demonstrated empirical validity; however, both tests do have rational validity. Both tests purport to measure an individual's ability to judge other people and base their claim to validity on the operation needed to complete the test, namely, making judgments about another person. The extensive background material gathered on the interviewees permits a definition of the right judgment. Thus, individuals can be ranked on either instrument on the basis of the number of correct judgments which are made.

The major difference between the two criterion measures provides the rationale for including both tests in the study. In completing the test, *Accuracy in Judging People*, the judge makes absolute judgments. For example, an item might ask the judge to indicate on a five-point scale the extent to which the interviewee would agree with the statement, "There is a God." In completing the test, *Ability to Judge Differences Between People*, the judge makes comparative judgments. For example, an item might ask the judge to indicate which of the three people would agree most strongly with the statement, "There is a God." By using both of these instruments, the training programs can be assessed with regard to the two different judging processes.

Procedures for the Experimental Sessions

Each of the seven sessions was run according to the same procedure. The only difference from one session to another was the content of the

training programs. Each group was in session for three hours. Five groups were tested during one week and the final two groups were tested the following week.

All subjects received a machine scoring pencil and a file folder containing three packets of testing materials. Packet I contained the two criterion measures, Accuracy in Judging People and Ability to Judge Differences Between People, and appropriate answer sheets. Packet II contained a test to be used during the training session (See Appendix C). The items in this test were the same as those for the test, Accuracy in Judging People, except that they were for interviewees numbered five and six on the Cline-Richards movie. Packet III was identical to Packet I. Subjects completed all tests anonymously.

To begin the session subjects removed Packet I from the file folder and read the directions for both of the tests. The experimenter explained,

You will see an unrehearsed sound-color filmed interview which will last about four or five minutes. When you have seen the interview, the projector will be stopped and the lights will come on. You will then answer the first thirty items on the test, Accuracy in Judging People. After you have followed this same procedure for a second and third interview, you will then complete the first 74 items of the test, Ability to Judge Differences Between People. Are there any questions?

When all subjects in the group had completed the two instruments, total time about one hour, they returned Packet I to the file folder and left the room for a five- to ten-minute break.

After the break the group received a training program. Subjects removed Packet II from the file folder and examined the test. The control group received no training but simply saw interviews five and six of the Cline-Richards film and completed the test in Packet II. The control group took approximately a 30 minute break before returning to complete the last hour of work. All other groups followed the procedure indicated below.

To begin the training session the experimenter gave a five- to ten-minute lecture concerning one of the six training programs formulated for this study. For five of the training sessions, those given to the groups, Observing-Self, Observing-Other, Inferring-Individual Differences, Inferring-Pooling, and Recording-Rating, the experimenter stated that the purpose of the lecture was to help the subjects to be better judges of people. He told subjects that the lecture did not exhaust the topic of improving accuracy in judging people but that it did present some important information about the topic. The lecture then mentioned one general aspect of judging behavior, i. e., observing, inferring, or recording, and stated a specific principle which could be used to improve performance related to the general aspect. The remainder of the lecture gave examples and further elaboration of the principle.

For the training session given to the Combination group, the experimenter stated the same purpose as for the other five groups. However, the lecture for this group presented all three general aspects of judging behavior and presented all of the principles of the other five programs. The choice of which principle or principles to use when judging others was left to the discretion of the subjects.

The experimenter answered all questions at the end of each lecture. See Appendix A for the virtually verbatim accounts of the six lectures.

Following the lecture, interview five from the film was shown, and subjects completed the 30 questions relating to that interviewee. Subjects were given the correct answers for the items. The experimenter determined by a show of hands the two or three people who had attained the highest scores and asked the high scorers to indicate how they had applied the principle presented in the lecture. After the remarks of the high scorers and other remarks from the group, the experimenter briefly restated the principle of the particular training program.

Interview six from the film was shown and the same procedure as for interview five was followed. After the final statement of the principle, subjects returned their packets to the file folders and left the room for a short break. Total time for the training session ranged from 50 to 60 minutes.

During the final hour subjects again saw the first three filmed interviews and completed the tests in Packet III. The procedure followed in the final hour was the same as that for the first hour of the session. Before the first interview was shown, experimenter reminded the group,

Just before the last break you were introduced to a principle for improving judgments which you make about other people and you had an opportunity to practice that principle. Please remember and use that principle as you complete these tests. You are not expected to be perfect, but do the best that you can.

Results

Before the data were analyzed for the comparisons of the study, the criterion measures were examined for reliability. Test-retest coefficients and internal consistency coefficients were computed for the tests, Accuracy in Judging People and Ability to Judge Differences Between People.

The test-retest reliabilities were obtained from the Control Group, $n = 22$. The control group took both administrations of the two tests under the same conditions as the other groups of the study. The only difference for the control group was that they received no training between administrations of the test. Correlation coefficients were computed using the Pearson r . Values for r appear in Table 1. Internal consistency measures were based on all the subjects who completed the initial administration of the two criterion measures, $n = 137$, and were computed by Kuder-Richardson formula 20. The values obtained

Table 1: Reliabilities for the Two Criterion Measures

The overall null hypothesis of no change in performance resulting from the administration of the various training programs was not rejected. Table 3 gives the mean and standard deviation for each of the seven groups of the study for both the administration of the instrument preceding training and the administration following training. An examination of Table 3 suggests that further analysis of the data from this instrument is not likely to be productive.

Table 3: Means and Standard Deviations for the Test, Accuracy in Judging People, for All Groups of the Study

Group	N	Pretest Mean and Standard Deviation	Posttest Mean and Standard Deviation	Mean Difference
Control	22	33.68 5.44	34.86 5.42	+ 1.22
Observing-Self	21	32.33 6.30	33.38 5.19	+ 1.05
Observing-Other	18	33.11 6.25	31.67 4.30	- 1.44
Inferring- Individual Differences	17	31.47 5.11	31.53 2.92	+ .06
Inferring-Pooling	18	33.44 5.95	32.89 5.28	- .55
Recording-Rating	22	32.68 5.23	32.73 6.64	+ .05
Combination	19	31.47 4.48	32.68 6.14	+ 1.21

The results of the analysis of covariance for the test, Ability to Judge Differences Between People, are reported in Table 4.

Table 4: Covariance Analysis for the Test, Ability to Judge Differences Between People

Source	df	Pre-score		Post-scores		df	Adjusted MS Scores
Among	6	248.04	161.26	428.99	6	331.47	55.25
Within	130	4160.69	1968.94	3656.75	129	2725.00	21.12
Total	136	4408.73	2130.20	4035.74			
$F = 2.62, df = 6, 129, F_{.05} = 2.17$							

For the analysis presented in Table 4 the null hypothesis of no change in performance produced by training was rejected. However, in examining the data to determine if they met the assumptions for the analysis, it was discovered that the assumption of equal slope of the regression lines for the various groups could not be met. Thus, the use of the analysis of covariance is rendered suspect. Examination of Table 5, however, suggests that two of the programs, Inferring-Pooling and Combination, did lead to improvements in performance which appear to be substantial.

Since Table 5 did suggest that additional analysis would be profitable, two further assessments were made. The data were analyzed by separate analysis of variance for the prescores and the postscores. Simple analysis of variance for the prescores yielded a value, $F = 1.24$ with $df = 6, 130$, which is not statistically significant. The analysis of variance for the postscores yielded a value, $F = 2.67$ with $df = 6, 130$, which is statistically significant ($F_{.025} = 2.52$). The F_{\max} test applied to the prescores to test for homogeneity of variance yielded a value of 3.33 with $k = 7$ and $df = 20$ which is not statistically significant ($F_{\max, .05} = 3.93$). For the postscores the F_{\max} test yielded a value of 6.20 which is statistically significant and indicates heterogeneous variance.

Table 5: Means and Standard Deviations for the Test, Ability to Judge Differences Between People, for all Groups of the Study

Group	N	Pretest Mean and Standard Deviation	Posttest Mean and Standard Deviation	Mean Difference
Control	22	37.05 5.10	37.91 5.12	+ 0.86
Observing- Self	21	34.52 6.30	36.19 8.03	+ 1.67
Observing- Other	18	38.33 8.05	38.67 4.98	+ 0.34
Inferring- Individual Differences	17	38.00 5.22	39.53 3.91	+ 1.53
Inferring- Pooling	18	37.22 5.41	42.28 3.20	+ 5.06
Recording- Rating	22	38.95 4.41	38.95 4.85	0.00
Combination	19	37.21 4.53	40.47 4.05	+ 3.26

While the postscores fail to meet the assumption of homogeneous variance, the F value obtained was sufficiently large to reject the overall null hypothesis (See Lindquist, p. 76ff). Following the rejection of the overall null hypothesis, comparisons were made between the means of the groups. The study was not concerned with comparing the various training programs with each other, but was concerned with comparing each of the six different training programs to the control condition. Therefore, of the 21 possible t-tests to compare means only six were computed.

The test selected for the mean comparisons was one suggested by Behrens and discussed in Fisher and Yates (1953, p. 47) and in Lindquist (1953, p. 97). The test is appropriate when the numbers in

the groups are small and heterogeneity of variance is present. Table 6 presents the results of the comparisons of the means following the F-test for the posttraining administration of the test, Ability to Judge Differences Between People. Table 6 shows that only Inferring-Pooling differs significantly from the control condition. The values for t in Table 6 were referred to tabled values in Fisher and Yates (1953, p. 52).

Table 6: Comparison of Each Training Condition with the Control Condition for the Second Administration of the Test, Ability to Judge Differences Between People

Condition	N	Mean	Variance	t-Value
Control	22	37.91	26.18	
Observing-Self	21	36.19	64.46	.83
Observing-Other	18	38.67	24.84	.48
Inferring-Individual Differences	17	39.53	15.26	1.12
Inferring-Pooling	18	42.28	10.21	3.29*
Recording-Rating	22	38.95	23.57	.69
Combination	19	40.47	16.37	1.79

* Significant $p < .01$

A second more powerful analysis which took cognizance of the actual differences between the groups on the premeasures and of the correlation between performance on the first and second administrations of the criterion instrument was applied to the data. The second analysis was a t-test for the difference between differences (Walker and Lev, 1953, p. 158ff). The mean difference in performance between the first

and second administrations of the test and the standard error of the mean difference (for correlated means) was computed for the control group and for each of the six training groups. The difference between pretraining and posttraining performance for each of the trained groups was compared to the difference between pretraining and post-training performance found for the control group. The comparisons were made using a t-test for presumed unequal variances. Appropriate degrees of freedom for interpretation of the t-values obtained were computed using Welch's formula (Walker and Lev, 1953, p. 158). The means, standard deviations, sample sizes and mean differences used to make these tests appear in Table 5. Table 7 presents the t-values which were obtained.

Table 7: t-Values for the Comparison of Differences Between the Differences

Control Condition Compared With:	t	df
Observing-Self	.48	
Observing-Other	.24	
Inferring-Individual Differences	.42	
Interring-Pooling	2.69*	28
Recording-Rating	.70	
Combination	2.50*	40

* Significant $p < .02$

Table 7 indicates that the Inferring-Pooling program and the Combination program were effective in improving the performance of the groups on the criterion measure.

Summary of Results

For this study there were twelve specific comparisons to be made. The comparisons were divided into two groups with six comparisons for each of the two criterion measures.

When the test, Accuracy in Judging People, was used none of the comparisons from before to after training was found to be significant. Training as conceived and presented in this study does not appear to be appropriate for modifying the kind of behavior which is required to make the judgments required by this test. Whether different training or the same training presented over a longer period of time would be effective remains an open question.

When the test, Ability to Judge Differences Between People, was used as the criterion measure, significant changes in performance were obtained for some training programs.

The results indicate that the Inferring-Pooling program was effective for improving accuracy in judging others. For the other training program which emphasized the inferring aspect of judging behavior, Inferring-Individual Differences, the results do not suggest that the program was effective for improving performance.

For the training programs which emphasized the focus for observing, Observing-Self and Observing-Other, the results do not suggest that the programs were effective.

For the recording aspect of judging behavior the results do not suggest that the Recording-Rating program was effective.

The results concerning the effectiveness of the Combination program are ambiguous. The Combination program group did not have a posttraining mean which was significantly different from the posttraining mean for the Control group, but it did have a mean difference between pretraining performance and posttraining performance which was significantly larger than the comparable difference for the Control group.

The Combination training program itself is an ambiguous one. It is structured in that it presents an analysis of the three steps performed by the subject in the judging situation. The program is unstructured, however, in that it presents several principles and relies on the subject's discretion as to when to use which principle.

Statements during the training period for this group and for other groups indicate that there is a favorite principle which is used in making judgments. Something akin to the Pooling principle appears to be the principle most commonly used. The effectiveness of the Combination program may be the result of judges ignoring other principles presented in the program and using the Pooling principle. This explanation is conjectural and not a finding of the study.

The major finding of this first study was that the Inferring-Pooling program is effective in increasing the accuracy with which people make judgments of others. A second study of smaller scope was planned and executed to ascertain the generalization of this findings. The following section is a report of the second study.

CHAPTER III

STUDY TWO

Study Two was designed as an extension of Study One. In the second study three groups were used. One group was treated as a control group and received no training. A second group received the Inferring-Pooling training program which was identical to the program in Study One. The third group received a combination of the Observing-Other and the Inferring-Pooling training programs from Study One, hereafter referred to as Other-Pooling group (See Appendix D for the complete account of this training program).

The Other-Pooling program was constructed as a brief presentation of a complete and rationally consistent program for the training of persons in judging others more accurately. This training program presented the three-part analysis of judging behavior, observing, inferring and recording, and emphasized the principle of the Observing-Other program for making observations and the principle of the Inferring-Pooling program for making inferences.

The purposes of the second study were to determine whether or not the effectiveness of the Inferring-Pooling program found in the first study would generalize to a different group of subjects and to test the effects of the Other-Pooling program.

Comparisons of the Study

A comparison was made between the pretraining and posttraining criterion scores of the Inferring-Pooling group. The first study suggested that this comparison would show that the Inferring-Pooling group improved significantly in ability to judge others.

A second comparison was made between the pretraining and post-training criterion scores of the Other-Pooling group.

The third comparison was made between the two trained groups, Inferring-Pooling and Other-Pooling, to determine which one had improved more.

Subjects

All subjects for this study were males who were taking evening courses through the Extension Service of this university. The subjects were college graduates working toward advanced degrees in business administration. The age range of subjects was from about 22 years to about 52 years. The subjects were engaged in engineering or personnel work, and most subjects were employed by one of two manufacturing concerns.

The Control group was a class of ten men taking an introductory course in personnel administration. The Inferring-Pooling group was a class of ten men taking a course in technical and business report writing. The Other-Pooling group was 11 men from a larger class of about 50 men who were taking a course in organization theory. The 11 men were selected to participate in the study since they all worked at one location and volunteered to meet with the experimenter at said location.

Criterion Instrument

The criterion instrument used in this study was the test, Ability to Judge Differences Between People, discussed in the previous chapter.

Procedures

As in the first study, the procedure for these groups was divided into three phases.

During the first phase the group saw the first filmed interview, completed the first twelve items of the test, Accuracy in Judging People; saw the second filmed interview, completed items numbered 31 through 42; and saw the third filmed interview, and completed items 61 through 72. Subjects next completed the test, Ability to Judge Differences Between People. The test, Accuracy in Judging People, was not intended as a criterion measure. The test was included in the procedure so that there would be a natural break between the various interviews and subjects would not see the three interviews consecutively.

The second phase was the training phase. During the second phase the Control group viewed filmed interview five and completed items numbered 1 through 12 on the training instrument (see Appendix C) and viewed filmed interview number six and completed items 31 through 42 on the training instrument. The two experimental groups received training following the same procedures as were used in the first study.

The third phase was the posttesting of the groups and was identical in procedure to the first phase.

Unlike the first study, not all of the groups in the present study accomplished the complete procedure in a single session. The Control group completed the first phase during one hour and finished the other two phases one week later. The Inferring-Pooling group completed all procedures during one session. The Other-Pooling group completed the first phase during one hour and finished the other two phases two weeks later. The Control and Inferring-Pooling groups were tested in a single location in Benton Harbor, Michigan. The Other-Pooling group was pre-tested at a location in Niles, Michigan, and completed the last two phases of the procedure in Mishawaka, Indiana.

Results

Two aspects of the data were examined for interest before the analysis concerning changes in the groups was begun. A coefficient of the test-retest reliability of the instrument was computed and the pre-training means of the subjects in the first and second studies were compared.

The test-retest coefficient, with one week delay, was computed by product moment correlation. The subjects in the Control group were used in this correlation, $n = 10$. The coefficient obtained was .91, which is consistent with the Pearson r of .79 found in the first study (see Table 1).

A test of the difference between the means of the groups used in the first study and the groups used in the second study was computed and found to be not statistically significant (see Table 8).

Table 8: Comparison of the Pretraining Means on the Test, Ability to Judge Differences Between People, for All Subjects, Study One and Study Two

	Study One	Study Two
N	137	31
Mean	37.31	35.48
s^2	33.74	45.91
t-Value	1.50	

The first two comparisons suggested for the study were made using the t-test for correlated means. Each posttraining mean was compared to its own pretraining mean to determine if a significant change had occurred. Reference to Table 9 shows that the three groups were so nearly equal on pretraining means that analysis of the data by more complex statistics was not necessary.

Table 9: Matched t-Tests for the Three Groups of the Study

Group	N	Pre-Mean and s^2	Post-Mean and s^2	M_d s_d	t-Value
Control	10	35.40 60.49	33.90 78.10	1.50 1.19	- 1.26
Inferring- Pooling	10	35.40 74.71	38.80 73.07	3.40 1.02	3.33*
Other-Pooling	11	35.64 16.25	39.27 15.62	3.63 1.26	2.89*

* Significant $p < .02$

The t-values for both the Inferring-Pooling and the Other-Pooling groups are significant. The third comparison of the study compared the mean difference for the Inferring-Pooling group with the mean difference for the Other-Pooling group and found no significant difference.

Summary of the Results

The Inferring-Pooling program in the second study was effective in increasing the accuracy with which individuals judge other people.

The Other-Pooling program was effective in increasing the accuracy with which individuals judge other people.

The comparison concerned with the relative improvement obtained by the Inferring-Pooling and Other-Pooling programs found no significant difference in the effectiveness of the two programs.

The results of Study One and Study Two are discussed in the next section of the report.

CHAPTER IV

DISCUSSION OF THE STUDIES

Each of the two studies reported above made comparisons within the particular study. The purpose of the present chapter is to discuss the substantive findings of the studies and to relate these findings to some broader considerations.

One general area of consideration is the relationship of the findings to the analysis of judging behavior which was made in Chapter I. The analysis was that when an individual is making judgments of another person, he observes the other, makes inferences about the other's behavior, and records some portion of his inference in answering specific questions posed by the experimenter. While the training programs were formulated following this analysis, the effects of training in the two studies do not constitute a test of the validity of the analysis. The analysis is descriptive and provides a framework for specification of training in the area of interpersonal judging accuracy.

Two additional areas of consideration for this discussion are the weaknesses of the present studies and the implications for future research of the present studies.

Training to Emphasize Observation

Neither of the two training programs concerned with providing a focus for making observations of the other person, Observing-Self and Observing-Other, appears to be effective in improving accuracy in making judgments about others. The hypothesis of no change in performance could not be rejected for either the "Self" program or the "Other" program for either of the criterion measures.

The ineffectiveness of the training programs concerned with observing can be stated only as tentative results for three reasons. First, to accept the ineffectiveness as fact would be to accept the null hypothesis. Second, this study tested the effects of the training programs by using indirect criterion measures, i.e., measures of judging accuracy. Direct measures of the effects of training programs would be those which determined whether or not the trainees had made improvements as observers of behavior. The analysis of judging behavior indicates that it is necessary to gather information about others in order to make judgments of them. However, it cannot be stated that improvement in observing behavior is sufficient to produce improvement in judging others accurately. Third, the results of the Lundy (1956) study reported in Chapter I suggest that the Observing-Other program should be effective in improving accuracy in judging others. The present study cannot be considered a replication of Lundy's work, but the two studies do appear to be closely related and to have produced conflicting results.

Training to Emphasize Inference

Training with the Inferring-Individual Differences program does not appear to be effective in improving judging accuracy. The hypothesis of no change could not be rejected for this program for either of the two criterion measures. The ineffectiveness of this program is also accepted tentatively but is supported by Crow's (1957) findings that increased responsiveness to individual differences is likely accompanied by decreased accuracy in judging others.

This author's conjecture about the finding is that the training is not effective. The principle presented in the program may be sound, but the time spent in developing the principle as a basis for inference was quite short. Determining what is unique about an individual is not simply a matter of identifying what features or personality traits set the

individual apart from others but is also a matter of being aware of unusual organizations of perhaps ordinary features and traits. It appears that the brief training period, one hour, was not sufficient to modify the trainees' usual inferential processes. The possibility does exist that more extensive training based on this principle may be effective.

The Inferring-Pooling training program was shown to be effective. Results of the first study showed that Inferring-Pooling was the only single-principle program which did lead to improvement in accurately judging others. In the second study the Inferring-Pooling program proved to be effective with a somewhat different sample of trainees. The two studies together give strong support to the position that Inferring-Pooling is a generally effective program for training to increase the accuracy with which people judge other people.

The major reservation concerning the generality of the Inferring-Pooling program is that the improvement was measured by the test, Ability to Judge Differences Between People, but was not apparent when the test, Accuracy in Judging People, was the criterion measure. Thus, the effectiveness of the program appears to have some generality across samples of people but not across kinds of judgments. This observation will be pursued later as a possible explanation for the effectiveness of the "Pooling" program.

With the demonstration that "Pooling" as a method of making inferences is effective in improving interpersonal judging accuracy, the question is Why? Why should the Inferring-Pooling program be effective?

One possible answer to the question is based on observations made by the investigator during discussions with subjects. Pooling appears to be the principle which is commonly used by people in making decisions about other people. People evidently learn to compare a relatively unknown person to a reference group of known persons as a

usual practice for making judgments demanded by social development. The Inferring-Pooling program quite possibly owes its effectiveness to condoning a normal practice and to emphasizing that the practice can be employed systematically.

Another aspect of the Inferring-Pooling program which may help to explain its effectiveness is the criterion instrument which demonstrated the effectiveness. The Pooling principle requires that the subject make comparisons of the stimulus person with many other persons. If the subject is forming a pool of known people who are similar to the unknown person, then he is making many comparisons of the stimulus person to others. The criterion instrument, Ability to Judge Differences Between People, which indicated that "Pooling" is an effective training principle is an instrument which requires the comparison of a specific set of people and a ranking of the people within the set on each of a series of specific statements. Thus the "Pooling" principle appears to be quite similar to the judging process necessary to complete the criterion measure.

The second criterion instrument, Accuracy in Judging People, does not require direct person-to-person comparisons, but emphasizes the relationship of an individual to an absolute scale.

The other training programs do not necessarily advocate the comparing of person with person. The two training programs concerned with gathering information emphasize that one person should be the focus. These two programs avoid indicating how information should be combined; rather, they emphasize the focus of observation and the retention of information gathered during the observations.

The Inferring-Individual Differences program emphasizes more the relationships of characteristics within the individual than the relationships between various specific individuals. This is not to say that other people are ignored when one attempts to determine what it is about one

individual that makes him different from others. What is said is that the comparisons are of a much more general sort than the comparisons made if the "Pooling" principle is employed.

The Inferring-Pooling program advocates stereotyping. "Pooling" requires the judge to simplify the stimulus person who is being judged, and in effect advocates that the judge ignore the individuality of the stimulus person. The advocacy of stereotyping may appear to be ill-advised because of the negative connotations which have become associated with the word, stereotype. The training program, however, is explicit in its attempt to overcome the aspect of rigidity which for many is the offensive aspect of stereotypes. Instruction is given that reference groups for judging should be formed of persons well-known to the judge, and that the reference groups should be formed on the basis of several variables rather than a single variable such as skin color, religious preference or socio-economic status.

As mentioned in the first chapter of this report, several investigators have indicated that accuracy in judging others is a complex ability with ability to predict for specific others and ability to predict for generalized others as the major components. These factors have been conceptualized as a result of the analysis of criterion measures with little if any attention given to the individual behaviors which are related to the abilities. Crow (1960) in a paper in defense of stereotyping has indicated that often the judging accuracy which accrues from the use of stereotyping, predicting for generalized others, is considered artifactual. He indicates that most investigators would prefer that prediction be for the specific other, that it be dependent on the immediate sense impressions and that it exclude information which the judge brings to the judging situation. Crow disagrees with those who consider stereotype accuracy artifactual.

This investigator is in substantial agreement with Crow's argument that information which the judge brings to the situation should be used. An extension offered from this study is that training to improve accuracy in judging others should take as principles of training those principles which permit the judge to make systematic use of his information.

Making accurate judgments is not a matter of having all of the information about an individual or of using all of the information available about an individual. In making judgments about people, it is required to simplify the stimulus. Stereotyping as presented in the Inferring-Pooling program is a process which makes use of the judge's information in simplifying the stimulus. Further, these studies indicate that Inferring-Pooling leads to improved accuracy in judging others.

Training to Emphasize Recording

Only the test, Accuracy in Judging People, which was in rating scale form could be used to assess the Recording-Rating program since the test, Ability to Judge Differences Between People, was in multiple-choice form. Apparently training on the proper use of rating scales does not significantly improve judging accuracy. Because this finding depends on the acceptance of the null hypothesis and because of the brevity of the training program, the answer is accepted tentatively.

Cronbach (1955) in his analysis of the components of accuracy scores has shown that systematic differences among people concerning the use of rating scales do influence the accuracy scores obtained. The result of the present study does not seriously challenge Cronbach's analysis. One fact which was not sufficiently considered before the study was performed and which may account for the ineffectiveness of the training program is the systematic error of the judged. Many of the items on the test, Accuracy in Judging People, are ones for which

the right answer was determined by the interviewees' self descriptions. The judge with knowledge of the proper use of rating scales may not be able to use his knowledge to advantage unless he is able to predict correctly the way in which the interviewees differed in their use of the rating scales.

Training to Emphasize the Analysis of Judging Behavior

Two separate training programs emphasizing the analysis of judging behavior were used in the two studies reported above. Both of these programs were effective in improving the accuracy scores of trainees on the test, Ability to Judge Differences Between People.

As indicated in Chapter II of this report the Combination program is difficult to assess very precisely since subjects were instructed to use their own discretion in employing the various principles presented in the program. The Combination program was effective, and to some extent this result indicates that a complete description of judging behavior does have merit as a training procedure. However, without knowledge of how trainees used their discretion any discussion of why the program was effective would be too highly conjectural to be of much importance.

The training program to emphasize the analysis of judging behavior which was presented in the second study was formulated to be less ambiguous than the Combination program of the first study. The program for the second study, identified as Other-Pooling, presented the analysis of the judging situation and presented the Observing-Other principle for making observations and the Inferring-Pooling principle for making inferences. The interrelationships between these principles were discussed and their relationships to the analysis were stressed. It might be argued that since previous results showed the "Pooling" principle to be effective and since the Other-Pooling program presented "Pooling"

along with one additional principle, the effectiveness of the Other-Pooling program indicates that an additional principle in the program does not diminish the effectiveness of the "Pooling" principle. This argument cannot be directly challenged by the data of this study. However, it appears to this author that presenting an integrated set of principles within the analysis of judging behavior rather than a single relatively isolated principle provides more complete training and is quite likely regarded as a more rational approach by trainees.

Further Research

Three major weaknesses were apparent in the studies reported and discussed. Correction of these three weaknesses in further research should lead to more definite information about the practical question of the effectiveness of the particular training programs of the studies and the general question of the adequacy of the conceptualization of judging behavior.

One weakness of the studies is the short duration of the training programs. The training was limited to a simple presentation by lecture and two short practice periods. More extensive development of the principles of the training programs and more extensive time spent by judges in practicing the principles and obtaining knowledge of their performance appears to be desirable.

A second weakness is that the sample sizes in the second study were small, $N = 10$ or 11 , and that in the second study only one of the original training programs was retested. Further research with larger groups would be desirable for confirming the effectiveness or ineffectiveness of the programs presented in the first study.

A third weakness is that training programs were evaluated with regard to their effectiveness for increasing accuracy in judging others and not with regard to their effectiveness in improving the specific

ability stressed in the training program. For example, it was not determined whether trainees in the Observing-Other training program actually did make more accurate observations of the person who was to be judged. Direct assessment of the training programs, the development of criterion instruments to permit direct assessment, and the determination of the relationship between the specific abilities emphasized in the training programs and accuracy in judging others are all possibilities for further research in the area of training to improve accuracy in interpersonal judgment.

There are some additional problems for further research which do not arise from weaknesses of the present studies but from impressions made during the course of the research.

Research is needed to determine how relatively naive individuals go about making judgments of other people. What kinds and amounts of information do people want or feel they need when they are asked to make judgments of others? To what extent do the criterion instruments usually employed in this area of research parallel the kinds of judging situations and judgments which normally arise in interpersonal situations? What processes do people use for combining information about others in order to arrive at judgments?

CHAPTER V

SUMMARY AND CONCLUSIONS

Previous research in the area of training to improve accuracy in judging others has been remiss in specifying training and in relating training to a conceptualization of the problem of improving accuracy in judging others. In the present study an analysis of the activities performed by subjects in the usual situation in which accuracy in judging others is measured was made to provide a framework for the specification of training to improve interpersonal judging accuracy.

The analysis of the individual's behavior in the research situation indicates that there are three aspects to the judging task. The judge gathers information about the stimulus person within the limits of the situation; he makes inferences about the stimulus person; and he records his judgments by answering specific questions posed by the experimenter. From this analysis six training programs were developed and assessed for effectiveness in improving accuracy in judging others.

Two of the programs were concerned with training to improve judges' observations of others. One program, Observing-Self, instructed trainees to focus on their own reactions to the stimulus person when gathering information, and the second program, Observing-Other, instructed trainees to focus on the other person when gathering information.

Two of the programs were concerned with training to improve judges' inferences about others. One program, Inferring-Individual Differences, instructed trainees to concentrate on the individual's unique characteristics when making inferences, and the second program, Inferring-Pooling, instructed trainees to concentrate on the characteristics

of the individual which made him similar to people who were well-known to the judge when making inferences.

One program, Recording-Rating, instructed trainees on avoiding three common rater errors in the use of rating scales.

The sixth program, Combination, instructed trainees on the three aspects of judging behavior, i.e., observing, inferring and recording, presented all of the principles of the other five programs and instructed trainees to use their discretion in choosing which principles to use.

All training programs were approximately one hour in duration and consisted of a brief lecture, two practice periods with knowledge of results given, and a discussion session.

The programs were assessed by using seven groups, one for each training program and one control group, of introductory psychology students ranging in number from 17 to 22. Each group completed two criterion measures of accuracy in judging others, participated in one of the training programs or the control condition, and again completed the criterion measures. Each group was in session for three consecutive hours with one hour devoted to each of the three phases of the session. Sound-color films were used to present the stimulus persons to the judges.

Separate analyses were made for each of the criterion instruments. Analysis of covariance for the instrument, Accuracy in Judging People, suggested no change in performance had occurred after training. Analysis of covariance was inappropriate for the instrument, Ability to Judge Differences Between People. Analyses of variance and t-tests of the differences between differences were used to compare each trained group with the control group. Results of the analyses indicated that the programs, Combination and Inferring-Pooling, were effective in improving accuracy in judging others. None of the other programs appeared to be effective.

A second study using three groups, N = 10 or 11, of engineers and personnel workers from an industrial population was completed. The procedures of the experimental session were the same as for the first study with the exception that only the test, Ability to Judge Differences Between People, was used as a criterion instrument.

Study Two confirmed the effectiveness of the Inferring-Pooling program and suggested that a modification of the Combination program was effective.

The results of the two studies were discussed in regard to the three general aspects of the analysis of judging behavior.

The results indicate that:

1. Training to improve observation does not appear to improve accuracy in judging others;
2. Training to emphasize unique features of an individual when making inferences does not appear to improve accuracy in judging others;
3. Training to emphasize the use of individually determined reference groups, a kind of stereotyping, leads to improvement in judging accuracy;
4. Training in the proper use of rating scales does not appear to improve judging accuracy; and
5. Training with some combinations of principles does appear to improve accuracy in judging others.

Additional discussion was addressed to the advantages of the specification of training programs and the feasibility of training to improve judging accuracy. The suggestion was made that training is most effective when instruction emphasizes processes which are similar to the processes required to complete criterion instruments.

REFERENCES

- Allport, F. H. and Musgrave, R. S. Teleonomic description in the study of behavior. Charact. and Pers., 1941, 9, 326-343.
- Bronfenbrenner, U., Harding, J. and Gallway, Mary. The measurement of skill in social perception. In: D. C. McClelland (Ed), Talent and Society. New York: Van Nostrand, 1958. Pp. 29-111.
- Bruner, J. S. and Tagiuri, R. The perception of people. In: G. Lindzey (Ed), Handbook of Social Psychology, Vol. 2. Cambridge, Mass.: Addison-Wesley, 1954, Pp. 634-654.
- Cline, V. B. Ability to judge personality assessed with a stress interview and sound film technique. J. abnorm. soc. Psychol. 1955, 50, 183-187.
- Cline, V. B. and Richards, J. M. Variables related to accuracy in interpersonal perception. First Annual Report. Office of Naval Research, Contract NR 171-146. University of Utah. 1958.
- Cline, V. B. and Richards, J. M. Accuracy of interpersonal perception--a general trait? J. abnorm. soc. Psychol. 1960, 60, 1-7.
- Cottrell, L. S. Some neglected problems in social psychology, Amer. socio. Rev., 1950, 15, 705-713.
- Cottrell, L. S. and Dymond, Rosalind F. The empathic responses: a neglected field for research. Psychiatry, 1949, 12, 355-359.
- Cronbach, L. J. Processes affecting scores on "understanding of others" and "assumed similarity." Psychol. Bull. 1955, 52, 177-193.
- Crow, W. J. The effect of training upon accuracy and variability in interpersonal perception. J. abnorm. soc. Psychol. 1957, 57, 355-359.
- Crow, W. J. In defense of the use of stereotypes. Paper read at the American Psychological Association, Chicago, Sept., 1960.

- Crow, W. J. and Farson, R. E. The effect of training upon accuracy and variability in judging others. Paper read at the Western Psychological Association, San Jose, California, April, 1960.
- Dymond, Rosalind F. A scale for measurement of empathic ability. J. consult. Psychol. 1949, 13, 127-133.
- Fisher, R. A. and Yates, F. Statistical Tables, 4th edition. New York: Hafner Publishing Co., 1953.
- Gage, N. L. and Cronbach, L. J. Conceptual and methodological problems in interpersonal perception. Psychol. Rev. 1955, 62, 411-422.
- Ghiselli, E. E. and Brown, C. W. Personnel and Industrial Psychology. New York: McGraw-Hill, 1955.
- Guilford, J. P. An experiment in learning to read facial expression. J. abnorm. soc. Psychol. 1929, 24, 191-202.
- Guilford, J. P. Psychometric Methods. New York: McGraw-Hill, 1954.
- Jenness, A. The recognition of facial expression. Psychol. Bull. 1932, 29, 324-350.
- Johnson, D. M. The Psychology of Thought and Judgment. New York: Harper, 1955.
- Lindquist, E. F. Design and Analysis of Experiments in Psychology and Education. Boston: Houghton Mifflin, Co., 1953.
- Luft, J. Implicit hypotheses and clinical predictions. J. abnorm. soc. Psychol. 1950, 45, 756-759.
- Lundy, R. M. Assimilative projection in interpersonal perceptions. J. abnorm. soc. Psychol. 1956, 52, 33-38.
- Martin, H. W. Effects of practice on judging various traits of individuals. Psychol. Bull. 1938, 35, 690. (Abstract)
- Peiper, W. J. Study on and revision of a test on accuracy in judging people. Unpublished research project. Michigan State University, 1960.
- Rabin, A. I. Szondi's pictures: identification of a diagnosis. J. abnorm. soc. Psychol. 1950, 45, 392-395.

Stone, G. C., Gage, N. L., and Leavitt, G. S. Two kinds of accuracy in predicting another's responses. J. soc. Psychol. 1957, 45, 245-254.

Taft, R. The ability to judge people. Psychol. Bull. 1955, 52, 1-23.

Walker, Helen M. and Lev, J. Statistical Inferences. New York: Henry Holt and Co., 1953.

APPENDICES

APPENDIX A

Training Lectures for Study One

1. Observing-Self
2. Observing-Other
3. Inferring-Individual Differences
4. Inferring-Pooling
5. Recording-Rating
6. Combination

Training Lecture 1: Observing-Self

Purpose:

The purpose of my few remarks this evening is to help you to be a better judge of people. When we are done you will not be an expert at judging people and you will not have any magical ability. However, to the extent that you apply what you learn you should be able to do a better job of judging people.

In order for what we are going to discuss to be of any benefit to you, you must learn it individually and apply it individually. We cannot exhaust this topic in the brief time that we will be together, but we can explore at least one basic approach to improving the accuracy with which you make judgments about people.

Principle:

When you are asked to make judgments about another person, you find it very difficult to do if you do not know the other person. In order to make judgments you must have some information. If you are a psychologist and are asked to make judgments, one thing that you might do is give the individual some tests, combine the results of the tests in certain ways and arrive at your judgment. In the usual circumstances in which you might be asked to make judgments about people, however, there are several reasons why you cannot or would not want to give someone a battery of tests. Another thing that you might do is find out ahead of time what kind of judgments you are going to be asked to make and on which people. Then you can set about gathering the specific information that you think will be most useful. Such favorable circumstances, however, do not come along very often. You seldom know the judgments that you are going to be asked to make. The problem is what general

method is best for gathering information. What is the most useful way of gathering information when you are not sure what kinds of judgments you will be asked to make?

Let's make this analysis. Here is a person. What do you know that is more like a person than anything else? Another person. What person do you know better than any other person? Yourself. Therefore, when you are gathering information about another person in order to make a judgment it seems sensible to use yourself as a measuring instrument.

This might make good sense and might sound like a good idea, but the difficulty is how to do it. First of all let us state specifically what the principle is and then try to see how it is applied.

The principle is: When gathering information about another person which is to be used to make judgments you should pay attention to yourself.

Now what does this mean and how do you do it?

When a person does something he does it because he is a certain way, he feels a certain way, and he has certain reasons. If, when you are gathering information about the other person, you pay attention to why you would do such things and how you feel about those things, then you are using yourself as an instrument to measure the other. If you hear a person talk or see him doing something, you should answer the question, If I were in his shoes why would I do that or what kind of person would I be if I could do or did do that?

Another way in which you can use yourself as an instrument is simply to note how the other person makes you react. You have been alive a long time and have met and known many people, and you react to them. You have a certain sensitivity to people. Some people make you feel good, some don't; some people make you nervous, some don't; some people you respect, some you don't. When you are using yourself as an instrument you should pay attention to how the other person affects you.

The principle again is: pay attention to yourself.

Training Lecture 2: Observing-Other

Purpose:

The purpose of my few remarks this evening is to help you to be a better judge of people. When we are done you will not be an expert at judging people and you will not have any magical ability. However, to the extent that you apply what you learn you should be able to do a better job of judging people.

In order for what we are going to discuss to be of any benefit to you, you must learn it individually and apply it individually. We cannot exhaust this topic in the brief time that we will be together, but we can explore at least one basic approach to improving the accuracy with which you make judgments about people.

Principle:

When you are asked to make judgments about another person, you find it very difficult to do if you do not know the other person. In order to make judgments you must have some information. If you are a psychologist and are asked to make judgments, one thing that you might do is give the individual some tests, combine the results of the tests in certain ways and arrive at your judgment. In the usual circumstances in which you might be asked to make judgments about people, however, there are several reasons why you cannot or would not want to give someone a battery of tests. Another thing that you might do is find out ahead of time what kind of judgments you are going to be asked to make and on which people. Then you can set about gathering the specific information that you think will be most useful. Such favorable circumstances, however, do not come along very often. You seldom know the judgments that you are going to be asked to make. The problem is what general method is

best for gathering information. What is the most useful way of gathering information when you are not sure what kinds of judgments you will be asked to make?

A good principle to keep in mind when gathering information is: Pay attention to the other person. Paying good attention doesn't mean just staying awake. It means that when you are with the other person or watching the other person you should be actively engaged in gathering information. You should watch the other person carefully and you should listen to the other person carefully.

This advice to pay attention to the other person is just advice; to make it useful as a way of gathering information you must know how to apply it as a principle.

When a person does something or says something, he does it or says it in a certain way. Usually the way that he does it is something that he has learned is a good way in the kind of life he leads, the way he lives. By listening and watching you can find out clues from the behavior which may help you later in making judgment. Those of you who have read Sherlock Holmes will recognize this approach.

A second way in which the principle becomes useful in application is when there is a difference between the way the person is talking and the way that the person is behaving. If a person is talking very calmly but you notice that his hands are shaking and that his knees are knocking together, you have gained some valuable information.

To repeat, the principle is: pay attention to the other person; what does he do and how does he do it; what does he say and how does he say it.

Training Lecture 3: Inferring-Individual Differences

Purpose:

The purpose of my few remarks this evening is to help you to be a better judge of people. When we are done you will not be an expert at judging people and you will not have any magical ability. However, to the extent that you apply what you learn you should be able to do a better job of judging people.

In order for what we are going to discuss to be of any benefit to you, you must learn it individually and apply it individually. We cannot exhaust this topic in the brief time that we will be together; but we can explore at least one basic approach to improving the accuracy with which you make judgments about people.

Principle:

When you are asked to make a judgment about another person you generally go through a process of making an inference. That is, in order to make the judgment you must recall what you know about the person, put the information together in some way and then arrive at a conclusion about how the other person is or what his behavior is likely to be.

When a psychologist is asked to make judgments about another person he goes through the same process. He, however, has had many years of training and experience in putting information together. The psychologist usually uses personality theory to put together his information and make inferences; or he may have some statistical ways of putting together the information and making inferences. Most of you are not now and will never be trained psychologists. All of you do make judgments about people and it may be that already you have been or at some time will be in a position where you must make judgments about other

people. What you would like to know is some principle that you can use for combining information; some principle for making inferences.

A good principle to use is the principle of individual differences. What is the principle and how do you use it? In everyday language the principle simply points out that everyone is different from everyone else. When you are making inferences about another you should keep in mind and consider carefully what it is about the other that makes him different from other people. The unique characteristic may be something as obvious as a physical disability which would certainly be important in making judgments about an individual's athletic ability. The unique quality may be a personality characteristic or a deeply held conviction which colors the individual's whole personality. Quite frequently, the unique quality of the individual is his combination of traits, convictions and appearance.

Whatever the unique quality is, you should keep it firmly in mind, and when you are asked to make judgments about the person, your inference should be guided by the question: Since this person is unique in this way, what is the most reasonable thing to expect of him?

When you are making inferences, concentrate on putting together what you know about the person so that his difference will be clear. How is he unique?

Training Lecture 4: Inferring-Pooling

Purpose:

The purpose of my few remarks this evening is to help you to be a better judge of people. When we are done you will not be an expert at judging people and you will not have any magical ability. However, to the extent that you apply what you learn you should be able to do a better job of judging people.

In order for what we are going to discuss to be of any benefit to you, you must learn it individually and apply it individually. We cannot exhaust this topic in the brief time that we will be together, but we can explore at least one basic approach to improving the accuracy with which you make judgments about people.

Principle:

When you are asked to make a judgment about another person you generally go through a process of making an inference. That is, in order to make the judgment you must recall what you know about the other person, put the information together in some way and then arrive at a conclusion about how the other person is or what his behavior is likely to be. When a psychologist is asked to make judgments about another person he goes through the same process. He, however, has had many years of training and experience in putting information together. The psychologist usually uses some personality theory to put together his information and make inferences; or he may have some statistical ways of putting together the information and making inferences. Most of you, however, are not now and will never be trained psychologists. All of you do make judgments about people and it may be that if you are in a position of leadership you must make judgments about people.

What you would like to know is some principle that you can use for combining information; some principle for making inferences.

A good general principle is the pooling principle. What is this principle and how do you use it?

In the course of your living you have obtained a great deal of information about many people. The pooling principle simply suggests that you use this information when making inferences about a person with whom you have had little contact. When you are attempting to make inferences about a person whom you do not know well, one of the things which you can do is to form a pool of people whom you do know well who are like the unknown person. You take what you do know about the person, form a pool of people you know well, and then make your predictions or judgments based on the pool. The important things to remember in making these pools are to use people you know well and to use all of the information you have about the person you are trying to judge. You may form a pool that leads to wrong predictions if you use just one piece of information about the person, such as, his skin color, his religious preference or any other single piece of information. You may also form some pools that lead to wrong predictions if you use people whom you do not know well.

The principle for making inferences is, then, to use the information which you have about people whom you know well when you are making judgments about people whom you know less well.

Training Lecture 5: Recording-Rating

Purpose:

The purpose of my few remarks this evening is to help you to be a better judge of people. When we are done you will not be an expert at judging people and you will not have any magical ability. However, to the extent that you apply what you learn you should be able to do a better job of judging people.

In order for what we are going to discuss to be of any benefit to you, you must learn it individually and apply it individually. We cannot exhaust this topic in the brief time that we will be together, but we can explore at least one basic approach to improving the accuracy with which you make judgments about people.

Principle:

Many times you are asked to record your judgments about people, e.g. letters of recommendation and ratings. Psychologists have found as the result of many years of study by many different investigators that when we are asked to record our judgments we quite frequently make certain kinds of errors. We are going to consider these errors. The errors are so frequent that they have special names. They are called (1) The error of Central Tendency, (2) The error of Leniency, and (3) The Halo Error.

Let us start with the error of Central Tendency and see what it is and how to avoid making it. The error of Central Tendency is made when an individual avoids rating individuals either very high or very low and thus clusters his ratings around the central or average point on the scale. (Draw examples; seven people on a trait). In some cases this may not be an error, but to expect it to be true in all cases is an error.

You should be aware that this kind of error can happen and make your rating so that everyone doesn't turn out to be average.

The second kind of error is called the error of Leniency. This error represents the tendency on the part of the judge to rate individuals above the average on certain traits, usually desirable traits. (Draw example; again seven people on a trait.) Again it is possible that putting everyone at the top of the scale is not an error. It may be that you are rating the three prettiest girls you know and in that case you would naturally expect that all of them would be at the top of the scale. However, to expect that everyone will always be at the top of the scale which you are using in making ratings is an error. In making ratings you must be aware that such an error is possible and guard against it.

The Halo Error is the third error which judges commonly make. This error occurs when we judge our fellows in terms of a general mental attitude toward them; and there is, dominating this mental attitude toward the personality as a whole, a like mental attitude toward particular qualities. That is, we make our ratings on the basis of a general impression. For example, say that we know that an individual is very bright, has a high IQ; this one fact may dominate our attitude about that person and lead us to rate that person high on many other traits where he is not actually high. The one fact that he is very bright has put a "Halo" around everything else about him. (Example: one person rated on 5-6 traits.) Again we must be aware that this kind of error can occur. To guard against it we must, when making ratings, consider the particular individual trait-by-trait and not let one good trait be too influential in our thinking about the person. Again, there is the possibility that it is not an error.

(Summarize by naming and defining the three kinds of errors: Central Tendency, Leniency, and Halo.)

Training Lecture 6: Combination

Purpose:

The purpose of my few remarks this evening is to help you to be a better judge of people. When we are done you will not be an expert at judging people and you will not have any magical ability. However, to the extent that you apply what you learn you should be able to do a better job of judging people.

In order for what we are going to discuss to be of any benefit to you, you must learn it individually and apply it individually. We cannot exhaust this topic in the brief time that we will be together, but we can explore at least one basic approach to improving the accuracy with which you make judgments about people.

Principle:

In the judging tasks which you have just completed, you were doing three things. You were gathering information about the person in the film, you were using that information to make inferences or guesses about the person and finally, you were recording some of your impressions by answering the questions which were on the forms. What I want to do now is give you some principles which should help you do the three things which I mentioned. It is not possible to use all of these principles at the same time; but by knowing these principles, you should be better able to pick the principles to use in any certain situation.

When you are gathering information about the other person there are essentially two things which you can do. You can pay attention to yourself and you can pay attention to the other person. If you watch yourself closely you are concerned with how the other person makes you feel and what you would do if you were in the other person's place.

If you are paying attention to the other person, you are trying to be a recording instrument; you are trying to observe everything that is done and said and you are trying to remember what is done and said.

When you are trying to put together what you know about the other person in order to make judgments there are again two general things that you can do. You can concentrate on what it is about the person which is different, how he is unique with regard to all other people; or you can concentrate on what there is about the person that makes him like other people you know. If you concentrate on what is different you are in effect saying this is what makes that person different and this difference would make him behave like this or would make his personality like this. If you concentrate on how this person is like most other people or like this particular group of people whom you know well, you are in effect getting a broader base to your judgments but you may be losing some of the specific details.

When you write down your judgments there are three kinds of errors which you are quite likely to make. These errors are so common with people who are writing down their judgments that they have special names. There is the error of Central Tendency, the error of Leniency and the Halo error. The error of Central Tendency occurs when you rate everybody or most people as average on whatever your judging task happens to be. The error of Leniency occurs when you refuse to make an unfavorable judgment about anybody. The Halo error occurs when you make all of your judgments about another person on the basis of one or two very good things that you know about the person.

Very briefly then, these are the principles. What you must do is figure out in each situation which principles will permit you to make the most accurate judgments.

APPENDIX B

Criterion Instruments

1. Accuracy in Judging People
2. Ability to Judge Differences Between People

Smith - Wakeley
April, 1961
From: V. B. Cline

ACCURACY IN JUDGING PEOPLE

DIRECTIONS: You are about to see a brief filmed interview with each of several persons. At the end of each interview, the projector will be stopped. You will then attempt to make the 30 predictions about the person listed below by recording your answers on the separate answer sheet. The correct answers are known from what the person actually said or the way acquaintances actually rated the person. Be sure to answer all questions even if you feel you are guessing.

ONE: Mrs. P.

Mrs. P. responded to each of the statements below by checking one of the following alternatives:

- (1) "Strongly disagree"
- (2) "Disagree"
- (3) "Neither agree nor disagree"
- (4) "Agree"
- (5) "Strongly agree"

For each statement mark on the separate answer sheet, the number of the response you think she made.

1. God will punish those who disobey his commandments and reward those who obey Him (either in this life or in a future life).
2. There exists an evil intelligence, personage, or spirit in the universe often referred to as Satan or the Devil.
3. No one who has experienced God like I have could doubt His existence.
4. If there is a "God", it is only an impersonal creative force in the universe.
5. I believe that after death we will ultimately regain our bodies and in a real sense be resurrected.
6. People don't necessarily need to believe in God in order to live good lives and have a high system of ethics and morals.
7. When in doubt, I have usually found it best to stop and ask God for guidance.
8. While God may exist, it is quite difficult for me to accept such a fact without some definite proof.
9. I am unable to accept the idea of "Life after death," at least not until we have definite evidence there is such a thing.
10. God does marvelous things which are called miracles by some.
11. I have sometimes been very conscious of the presence of God.
12. While I am responsible for my own actions, I feel that God has some definite purpose or role for me to fulfill in life.

Mrs. P. was rated on each of the traits below by persons who knew her well. They used the following rating system:

- (1) "Rather UNlike" or "Very UNlike her"
- (2) "A little UNlike her"
- (3) "A little like her"
- (4) "Rather like her"
- (5) "Very like her"

Rate Mrs. P. as you think she was rated by others.

- 13. COOPERATIVE
- 14. STUBBORN
- 15. CONFIDENT
- 16. EGOTISTICAL
- 17. FRIENDLY
- 18. SHY
- 19. AFFECTIONATE
- 20. REBELLIOUS
- 21. CAREFUL
- 22. IMPRACTICAL
- 23. AMBITIOUS
- 24. UNREALISTIC

People who know Mrs. P. also made judgments about other aspects of her behavior. Rate her on each of the following statements as you think she was rated by these people.

25. People who know her say that:

- (1) she is a somewhat lazy person.
- (2) she is about like the average housewife.
- (3) she is an exceptionally hard working and energetic person.

26. With regard to paying bills and debts:

- (1) she is usually late or delinquent.
- (2) she is only average in this regard.
- (3) she is quite punctual and conscientious.

27. With her children:

- (1) she maintains quite firm and strict discipline.
- (2) she is about average in this regard.
- (3) they usually get their way.

28. At a party she:

- (1) can "unbend" and have a lot of fun.
- (2) feels a little out of place and doesn't quite know what to do and say.
- (3) tends to be a "social climber."

29. People who know her say that she is:

- (1) a very generous and warm hearted person.
- (2) friendly to your face but critical behind your back.
- (3) rather selfish and self-centered.

30. With regard to her emotional and mental health:
- (1) she is an exceptionally sound and stable person.
 - (2) she is average in this regard.
 - (3) is very excitable and easily upset by things that happen around her.

TWO: Mr. W.

Mr. W. as well as the other interviewees answered the same values-belief questionnaire that Mrs. P. did. For each statement below, mark on the separate answer sheet, the number of the response you think he made:

- (1) "Strongly disagree"
 - (2) "Disagree"
 - (3) "Neither agree nor disagree"
 - (4) "Agree"
 - (5) "Strongly agree"
31. God will punish those who disobey his commandments and reward those who obey Him (either in this life or in a future life).
32. There exists an evil intelligence, personage, or spirit in the universe often referred to as Satan or the Devil.
33. No one who has experienced God like I have could doubt His existence.
34. If there is a "God", it is only an impersonal creative force in the universe.
35. I believe that after death we will ultimately regain our bodies and in a real sense be resurrected.
36. People don't necessarily need to believe in God in order to live good lives and have a high system of ethics and morals.
37. When in doubt, I have usually found it best to stop and ask God for guidance.
38. While God may exist, it is quite difficult for me to accept such a fact without some definite proof.
39. I am unable to accept the idea of "Life after death," at least not until we have definite evidence there is such a thing.
40. God does marvelous things which are called miracles by some.
41. I have sometimes been very conscious of the presence of God.
42. While I am responsible for my own actions, I feel that God has some definite purpose or role for me to fulfill in life.

Mr. W. as well as the other interviewees was also rated on the same personality traits as Mrs. P. For each trait, mark the number of the rating that you think was given to him:

- (1) "Rather UNlike" or "Very UNlike him"
- (2) "A little UNlike him"
- (3) "A little like him"
- (4) "Rather like him"
- (5) "Very like him"

- 43. COOPERATIVE
- 44. STUBBORN
- 45. CONFIDENT
- 46. EGOTISTICAL
- 47. FRIENDLY
- 48. SHY
- 49. AFFECTIONATE
- 50. REBELLIOUS
- 51. CAREFUL
- 52. IMPRACTICAL
- 53. AMBITIOUS
- 54. UNREALISTIC

As in the case of Mrs. P. and the other interviewees, Mr. W. was judged on other aspects of his behavior. Rate him on the following statements as you think he was rated by other people.

- 55. With regard to his personality he is considered a somewhat:
 - (1) secure, warm, and easy-going person.
 - (2) insecure and highstrung person.
 - (3) deceitful person who takes advantage of people.
- 56. With regard to music, art, etc., he has:
 - (1) rather little interest for these sorts of things.
 - (2) he has no strong tastes and enjoys almost all good art and music.
 - (3) he has a number of intense likes and dislikes.
- 57. When it comes to food:
 - (1) he will eat about anything on the table.
 - (2) he may not like something but he will eat it and say nothing so he will not hurt the feelings of others.
 - (3) he is rather fussy about what he eats and how it is prepared.
- 58. When it comes to having "emotional scenes" with people:
 - (1) he avoids them because they make him feel most uncomfortable.
 - (2) he is about average in this regard.
 - (3) he enjoys getting people all worked up (angry, tearful, etc.).
- 59. When it comes to making major decisions in his life
 - (1) he has few or no self-doubts, knows exactly where he is going.
 - (2) he tends to "stew" about things, change his mind back and forth before making final decisions.

- (3) he finds it almost impossible to decide things and tends to rely mainly on his wife or close friends.
60. When it comes to basic likes and dislikes in fields of work he would most enjoy:
- (1) manual or mechanical work.
 - (2) giving men "hell" as a boss in industry.
 - (3) creating floral displays, artistic work.

THREE: Mrs. N.

Mrs. N. responded to each of the statements below by checking one of the following alternatives:

- (1) "Strongly disagree"
- (2) "Disagree"
- (3) "Neither agree nor disagree"
- (4) "Agree"
- (5) "Strongly agree"

For each statement mark on the answer sheet, the number of the response you think she made.

- 61. God will punish those who disobey his commandments and reward those who obey Him (either in this life or in a future life).
- 62. There exists an evil intelligence, personage, or spirit in the universe often referred to as Satan or the Devil.
- 63. No one who has experienced God like I have could doubt His existence.
- 64. If there is a "God," it is only an impersonal creative force in the universe.
- 65. I believe that after death, we will ultimately regain our bodies and in a real sense be resurrected.
- 66. People don't necessarily need to believe in God in order to live good lives and have a high system of ethics and morals.
- 67. When in doubt, I have usually found it best to stop and ask God for guidance.
- 68. While God may exist, it is quite difficult for me to accept such a fact without some definite proof.
- 69. I am unable to accept the idea of "Life after death," at least not until we have definite evidence there is such a thing.
- 70. God does marvelous things which are called miracles by some.
- 71. I have sometimes been very conscious of the presence of God.
- 72. While I am responsible for my own actions, I feel that God has some definite purpose or role for me to fulfill in life.

Mrs. N. was rated on each of the traits below by persons who knew her well. They used the following rating system:

- (1) "Rather UNlike" or "very UNlike her"
- (2) "A little UNlike her"
- (3) "A little like her"
- (4) "Rather like her"
- (5) "Very like her"

Rate Mrs. N. as you think she was rated by others.

- 73. COOPERATIVE
- 74. STUBBORN
- 75. CONFIDENT
- 76. EGOTISTICAL
- 77. FRIENDLY
- 78. SHY
- 79. AFFECTIONATE
- 80. REBELLIOUS
- 81. CAREFUL
- 82. IMPRACTICAL
- 83. AMBITIOUS
- 84. UNREALISTIC

People who know Mrs. N. also made judgments about other aspects of her behavior. Rate her on each of the following statements as you think she was rated by these people.

- 85. People who know her consider her as being:
 - (1) very conscientious and responsible.
 - (2) about average in this regard.
 - (3) at times a little irresponsible.
- 86. She is the sort of person who:
 - (1) is always complaining about her aches and pains (trying to get sympathy from people).
 - (2) is about average in this regard.
 - (3) keeps these things to herself and almost never complains.
- 87. With regard to money, she:
 - (1) tends not to budget very wisely.
 - (2) is about average in this regard.
 - (3) handles and budgets it very well.
- 88. In attending meetings and social gatherings, she is:
 - (1) usually a few minutes late.
 - (2) is always on time.
 - (3) is about average (sometimes on time, sometimes a little late).
- 89. People who know her say that she is:
 - (1) rather self-centered.
 - (2) quite unselfish and interested in pleasing others.

- (3) the sort who tries to please others but with a view in mind of getting something out of them.

90. Emotionally she is:

- (1) a somewhat unstable woman.
- (2) about average in this regard.
- (3) a very stable, well-balanced woman.

Smith, H. C.
Winter, 1961

Test of Ability to Judge Differences Between People

Part 1

DIRECTIONS

You are going to see five minute film interviews with Mrs. P., Mr. W., and Mrs. N. When the film is over you will be asked questions about their behavior and attitudes. When you answer the questions on the IBM sheets use only spaces 1, 2, and 3. The numbers correspond to the order in which the interviews appeared. That is Mrs. P. (1) Mr. W. (2), and Mrs. N. (3). In other words if the answer to a particular statement is Mr. W. or the person in the second interview mark space 2 on the answer sheet. Do all the items and try not to leave any blank.

(1) Mrs. P., (2) Mr. W., (3) Mrs. N., filled out a rating scale showing their agreement or disagreement with a series of statements about religious questions. Which person fits the following statement?

1. Most strongly agreed that "God will punish those who disobey his commandments and reward those who obey Him (either in this life or in a future life)."
2. Most strongly agreed with the statement "no one who has experienced God like I have could doubt his existence."
3. Believed "That after death we will ultimately regain our bodies and in a real sense be resurrected."
4. Strongly agreed that "When in doubt, I have usually found it best to stop and ask God for guidance."
5. Most strongly disagreed that "I am unable to accept the idea of 'life after death' at least not until we have definite evidence there is such a thing."
6. Would also disagree with the above statement.
7. Strongly agreed that "I have sometimes been very conscious of the presence of God."
8. Would also agree with the above statement.
9. Agree that "While I am responsible for my own actions, I feel that God has some definite purpose or role for me to fulfill in life."

(1) Mrs. P., (2) Mr. W., (3) Mrs. N. were each given pairs of words and asked to select the one which they thought was a better description of themselves. Who made the following choices?

10. Efficient in the pair efficient - precise.
11. Loyal in the pair loyal - sophisticated.
12. Arrogant in the pair arrogant - apathetic.
13. Practical in the pair practical - wholesome.
14. Reliable in the pair reliable - feminine.
15. Determined in the pair determined - relaxed.
16. Egotistical in the pair egotistical - stingy.
17. Nervous in the pair nervous - selfish.
18. Rational in the pair rational - painstaking.
19. Courageous in the pair courageous - rational.
20. Assertive in the pair assertive - reckless.
21. Tactful in the pair tactful - enthusiastic.
22. Conservative in the pair conservative - excitable.
23. Conscientious in the pair conscientious - excitable.
24. Dependable in the pair dependable - excitable.
25. Unselfish in the pair unselfish - cool.
26. Mannerly in the pair mannerly - humorous.
27. Complicated in the pair complicated - hasty.
28. Moody in the pair moody - complaining.
29. Warm in the pair warm - reflective.

- 30. Contented in the pair contented - progressive.
- 31. Understanding in the pair understanding - timid.
- 32. Restless in the pair restless - unemotional.
- 33. Snobbish in the pair snobbish - hostile.
- 34. Appreciative in the pair appreciative - sharp-witted.
- 35. Severe in the pair severe - hard headed.
- 36. Cool in the pair cool - timid.
- 37. Capable in the pair capable - obliging.
- 38. Poised in the pair poised - moderate.
- 39. Active in the pair active - artistic.

(1) Mrs. P., (2) Mr. W., (3) Mrs. N. were given a series of incomplete sentences to complete. Which of the three completed each of the following sentences in the way underlined?

- 40. "Sex is too often considered filthy."
- 41. "When I'm criticized I appreciate it."
- 42. "I felt most dissatisfied when I didn't do the right thing."

(1) Mrs. P., (2) Mr. W., (3) Mrs. N. were given a series of true-false items. Who answered true to these items?

- 43. My daily life is full of things that keep me interested.
- 44. At parties I am more likely to sit by myself or with just one other person than to join in with the crowd.
- 45. I believe a person should never taste an alcoholic drink.
- 46. I am happy most of the time.
- 47. I seem to be about as capable and smart as most others around.
- 48. Policemen are usually honest.

- 49. Sometimes I enjoy hurting people I love.
- 50. I like to visit places where I have never been before.
- 51. I find it hard to set aside a task that I have undertaken, even for a short time.
- 52. I like to read adventure stories.
- 53. I am nearly always on time for appointments.
- 54. I enjoy bull sessions where everyone talks about sex.
- 55. I fall in love rather easily.

(1) Mrs. P., (2) Mr. W., and (3) Mrs. N. were rated by their friends on a series of personality traits. Which was rated as follows?

- 56. least cooperative
- 57. most unfriendly
- 58. least affectionate
- 59. least rebellious
- 60. most practical
- 61. least ambitious
- 62. least careful

Friends of (1) Mrs. P., (2) Mr. W., and (3) Mrs. N. also gave thumbnail sketches of them. Which was described by friends as follows?

- 63. "Somewhat insecure and highstrung"
- 64. "Exceptionally hard working and energetic"
- 65. "In a state of rebellion against all religions"
- 66. "Very conscientious and responsible"
- 67. "Handles money and budgets it extremely well"

- 68. "Maintains quite firm and strict discipline with children"
- 69. "Rather fussy about what he (or she) eats and how it is prepared"
- 70. "Avoids emotional scenes because they make this person feel most uncomfortable"
- 71. "Always on time"
- 72. "Unselfish and interested in pleasing others"
- 73. "Most enjoys creating floral displays and artistic work"
- 74. "Generous and warm hearted"

APPENDIX C

Instrument Used in Training

FIVE: Mr. G.

Mr. G. responded to each of the statements below by checking one of the following alternatives:

- (1) "Strongly disagree"
- (2) "Disagree"
- (3) "Neither agree nor disagree"
- (4) "Agree"
- (5) "Strongly agree"

For each statement mark on the separate answer sheet, the number of the response you think he made.

- 1. God will punish those who disobey his commandments and reward those who obey Him (either in this life or in a future life).
- 2. There exists an evil intelligence, personage, or spirit in the universe often referred to as Satan or the Devil.
- 3. No one who has experienced God like I have could doubt His existence.
- 4. If there is a "God", it is only an impersonal creative force in the universe.
- 5. I believe that after death we will ultimately regain our bodies and in a real sense be resurrected.
- 6. People don't necessarily need to believe in God in order to live good lives and have a high system of ethics and morals.
- 7. When in doubt, I have usually found it best to stop and ask God for guidance.
- 8. While God may exist, it is quite difficult for me to accept such a fact without some definite proof.
- 9. I am unable to accept the idea of "Life after death," at least not until we have definite evidence there is such a thing.
- 10. God does marvelous things which are called miracles by some.
- 11. I have sometimes been very conscious of the presence of God.
- 12. While I am responsible for my own actions, I feel that God has some definite purpose or role for me to fulfill in life.

Mr. G. was rated on each of the traits below by persons who knew him well. They used the following rating system:

- (1) "Rather UNlike" or "Very UNlike him"
- (2) "A little UNlike him"
- (3) "A little like him"
- (4) "Rather like him"
- (5) "Very like him"

Rate Mr. G. as you think he was rated by others.

13. COOPERATIVE
14. STUBBORN
15. CONFIDENT
16. EGOTISTICAL
17. FRIENDLY
18. SHY
19. AFFECTIONATE
20. REBELLIOUS
21. CAREFUL
22. IMPRACTICAL
23. AMBITIOUS
24. UNREALISTIC

People who knew Mr. G. also made judgments about other aspects of his behavior. Rate him on each of the following statements as you think he was rated by these people.

25. In his social life he:
 - (1) enjoys "dress-up" parties.
 - (2) entertains informally frequently at home.
 - (3) is rather inactive and only rarely goes to social activities.
26. People who know him describe him as being:
 - (1) easy to get along with.
 - (2) somewhat irresponsible.
 - (3) somewhat dominating.
27. He tends to be:
 - (1) quite ambitious.
 - (2) about average in this regard.
 - (3) somewhat indifferent and lazy.
28. When he gets in an argument with his wife he:
 - (1) shouts loudly and obscenely at her.
 - (2) may occasionally strike her.
 - (3) raises his voice a little but maintains good control.
29. In handling his 12-year-old son he is:
 - (1) rather strict.
 - (2) fairly easy going.
 - (3) not too interested.
30. When he goes to a party he:
 - (1) tends to flirt with women other than his wife.
 - (2) enjoys himself but is not much noticed.
 - (3) sometimes gets drunk and then becomes the life of the party.

SIX: Mrs. L.

Mrs. L. responded to each of the statements below by checking one of the following alternatives:

- (1) "Strongly disagree"
- (2) "Disagree"
- (3) "Neither agree nor disagree"
- (4) "Agree"
- (5) "Strongly agree"

For each statement mark on the separate answer sheet, the number of the response you think she made.

- 31. God will punish those who disobey his commandments and reward those who obey Him (either in this life or in a future life).
- 32. There exists an evil intelligence, personage, or spirit in the universe often referred to as Satan or the Devil.
- 33. No one who has experienced God like I have could doubt His existence.
- 34. If there is a "God", it is only an impersonal creative force in the universe.
- 35. I believe that after death we will ultimately regain our bodies and in a real sense be resurrected.
- 36. People don't necessarily need to believe in God in order to live good lives and have a high system of ethics and morals.
- 37. When in doubt, I have usually found it best to stop and ask God for guidance.
- 38. While God may exist, it is quite difficult for me to accept such a fact without definite proof.
- 39. I am unable to accept the idea of "Life after death," at least not until we have definite evidence there is such a thing.
- 40. God does marvelous things which are called miracles by some.
- 41. I have sometimes been very conscious of the presence of God.
- 42. While I am responsible for my own actions, I feel that God has some definite purpose or role for me to fulfill in life.

Mrs. L. was rated on each of the traits below by persons who knew her well. They used the following rating system:

- (1) "Rather UNlike" or "Very UNlike her"
- (2) "A little UNlike her"
- (3) "A little like her"
- (4) "Rather like her"
- (5) "Very like her"

Rate Mrs. L. as you think she was rated by others.

- 43. COOPERATIVE
- 44. STUBBORN
- 45. CONFIDENT
- 46. EGOTISTICAL
- 47. FRIENDLY
- 48. SHY
- 49. AFFECTIONATE
- 50. REBELLIOUS
- 51. CAREFUL
- 52. IMPRACTICAL
- 53. AMBITIOUS
- 54. UNREALISTIC

People who knew Mrs. L. also made judgments about other aspects of her behavior. Rate her on each of the following statements as you think she was rated by these people.

- 55. In handling money she:
 - (1) budgets and handles it very well.
 - (2) tends to buy impulsively.
 - (3) tends to be a little tight.
- 56. With regard to her intelligence (compared with people in general) she is:
 - (1) above average.
 - (2) just average.
 - (3) a little below average.
- 57. When it comes to recreation or hobbies she particularly enjoys:
 - (1) record collecting.
 - (2) gardening.
 - (3) swimming.
- 58. When her children won't mind her she usually:
 - (1) punishes them severely.
 - (2) lets them get away with it.
 - (3) goes to her room and cries.
- 59. If somebody kept pestering her she would:
 - (1) get mad and probably "tell them off."
 - (2) joke and kid them about it.
 - (3) Shrug it off.
- 60. If her parents were to tell her to do something that she didn't want to do she would:
 - (1) just not do it.
 - (2) go ahead and do it rather than make a big fuss.
 - (3) make a big fuss about it but finally do it.

APPENDIX D

Training Lecture for Study Two (Other-Pooling)

Purpose:

The purpose of my few remarks this evening is to help you to be a better judge of people. When we are done you will not be an expert at judging people and you will not have any magical ability. However, to the extent that you apply what you learn you should be able to do a better job of judging people. What I am going to say applies particularly to the kind of task which you completed last week with the questions based on the movies of interviews; however, with very little change you will find that what we discuss tonight can be applied in your day-to-day relations. If you want what we discuss this evening to be of use to you, you must practice it. Like virtually every other ability, the ability to make accurate judgments can be improved through practice.

We cannot exhaust this topic of judging others in the brief time which we have this evening, but we can explore some basic principles and attempt to practice these principles briefly.

Principles:

The last time I was with you, you were presented with a very difficult task. You were asked to make some judgments about another person. This is always a difficult task, but when you see the person for only about five minutes in a moving picture the job is even tougher. In that task you were doing essentially three things. You were gathering information about the person in the film, you were using the information to make inferences or guesses about the person and finally, you were recording some of your impressions by answering the questions which were on the forms. What I want to do now is to give you some principles which should help you in this process. There are two principles to be discussed and the way that these two principles go together will also be of interest to us.

First, there is a definite principle which you can follow when gathering information about the other person. The principle is: pay attention to the other person. This sounds simple, but in our usual interactions with other people it is frequently difficult to apply. The principle means more than just staying awake when you are gathering information. It means being actively engaged in gathering as much information as you can. You are attempting to record everything that is done and said so that you can recall and use the information. You must watch and listen to know what is done and what is said, and equally important, to determine how it is done and how it is said.

When a person does something or says something, he does it or says it in a certain way with certain gestures and certain words. Usually the way that he does it is a way that he has learned is a good way in the kind of life he leads. By listening and watching actively you can find clues for the other's behavior which may help you later in making judgments.

When you have gathered as much information as you can about the other person, the task of making judgments has just begun. You must now put the information together in some way and arrive at some judgments. The process of putting together the information and making judgments can be called making inferences. You never have all of the information about another person. You are always working with samples of the other person's behavior. One principle which you can use to put together the information which you have gathered is the pooling principle.

In the course of your living you have obtained a great deal of information about many people. The pooling principle suggests that you use this information when making inferences about a person with whom you have had little contact.

When you are attempting to make inferences about a person whom you do not know well, one of the things which you can do is to form a pool of people whom you do know well who are like the unknown person.

You take what you do know about the person, form a pool of people you know well and then make your predictions or judgments based on the pool. What are some of the variables which you can use to form the pool? There are several: age, sex, education, background, vocabulary, method of expression. The important thing is to use as many variables as possible. You may get into some inaccurate stereotyping if you use just one variable, and you may get into some inaccurate stereotyping if you use ready-made stereotypes such as those based on skin color, nationality, religion, etc. The important things are to use as many variables as possible and to form your pool from people whom you know well.

The one thing which you cannot do in a research situation such as you have participated in but which you can do in ordinary situations is to test your judgments. That is, once you have gathered information and made your judgments, you can continue to gather information and determine whether your judgments are right or wrong.

It is probably obvious to you the way in which the two principles we have discussed tend to work together. One principle says to pay attention to the other person when you are gathering information to be used in making judgments. The second principle says that you can make use of pooling when attempting to make inferences about another person. While you are observing you are identifying variables which you can use when you form the pool of people to use in making your prediction. The two principles do not necessarily need to be separated. That is, you can make inferences as you go along rather than gathering all of the information and then attempting to make all of your inferences. If you do make inferences as you go along, however, you must realize that they are tentative and subject to be changed as new information becomes available.

These then are a couple of principles which you can use in making inferences about others. Now we are going to take a little time and practice the principles. The idea is to see how you as an individual can make use of these principles.

APPENDIX E

Table 10: Comparison of the Tests, Accuracy in Judging People
and Ability to Judge Differences Between People

Table 10: Comparison of the Tests, Accuracy in Judging People and Ability to Judge Differences Between People

	Accuracy Test	Difference Test
N	137	137
Number of items	72 five-choice 18 3-choice 90 total	74 3-choice
Chance Mean	20.4	24.7
Obtained Mean	32.6	37.3
Difference	12.2	12.6
Obtained Standard Deviation	5.45	5.81
Pearson r	.141	

ROOM USE ONLY

~~JUL 2 1963~~

~~SEP 2 4 65~~

~~SEP 2 4 65~~

~~SEP 2 4 65~~ P13

~~MAR 28 1963~~ 117

~~SEP 2 4 65~~ FIVE
P24

~~APR 15 1968~~ 114

~~JUL 1 1961~~ 113

~~JUL 26 1971~~ 164

~~JUL 26 1971~~ 244

MICHIGAN STATE UNIV. LIBRARIES



31293103433102