

## INFORMATION TO USERS

This was produced from a copy of a document sent to us for microfilming. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help you understand markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure you of complete continuity.
2. When an image on the film is obliterated with a round black mark it is an indication that the film inspector noticed either blurred copy because of movement during exposure, or duplicate copy. Unless we meant to delete copyrighted materials that should not have been filmed, you will find a good image of the page in the adjacent frame. If copyrighted materials were deleted you will find a target note listing the pages in the adjacent frame.
3. When a map, drawing or chart, etc., is part of the material being photographed the photographer has followed a definite method in "sectioning" the material. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.
4. For any illustrations that cannot be reproduced satisfactorily by xerography, photographic prints can be purchased at additional cost and tipped into your xerographic copy. Requests can be made to our Dissertations Customer Services Department.
5. Some pages in any document may have indistinct print. In all cases we have filmed the best available copy.

University  
Microfilms  
International

300 N. ZEEB RD., ANN ARBOR, MI 48106

8209433

**Mertens, Ruth Ann**

**EFFECTS OF DIFFERING TYPES OF INFORMATION UPON CLINICAL  
JUDGMENTS**

*The University of Oklahoma*

**PH.D. 1981**

**University  
Microfilms  
International** 300 N. Zeeb Road, Ann Arbor, MI 48106

PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark .

1. Glossy photographs or pages \_\_\_\_\_
2. Colored illustrations, paper or print \_\_\_\_\_
3. Photographs with dark background \_\_\_\_\_
4. Illustrations are poor copy \_\_\_\_\_
5. Pages with black marks, not original copy \_\_\_\_\_
6. Print shows through as there is text on both sides of page \_\_\_\_\_
7. Indistinct, broken or small print on several pages
8. Print exceeds margin requirements \_\_\_\_\_
9. Tightly bound copy with print lost in spine \_\_\_\_\_
10. Computer printout pages with indistinct print \_\_\_\_\_
11. Page(s) \_\_\_\_\_ lacking when material received, and not available from school or author.
12. Page(s) \_\_\_\_\_ seem to be missing in numbering only as text follows.
13. Two pages numbered \_\_\_\_\_. Text follows.
14. Curling and wrinkled pages \_\_\_\_\_
15. Other \_\_\_\_\_

University  
Microfilms  
International



THE UNIVERSITY OF OKLAHOMA  
GRADUATE COLLEGE

EFFECTS OF DIFFERING TYPES OF INFORMATION  
UPON CLINICAL JUDGMENTS

A DISSERTATION  
SUBMITTED TO THE GRADUATE FACULTY  
in partial fulfillment of the requirements for the  
degree of  
DOCTOR OF PHILOSOPHY

By  
RUTH ANN MERTENS  
Norman, Oklahoma  
1981

EFFECTS OF DIFFERING TYPES OF INFORMATION  
UPON CLINICAL JUDGMENTS

APPROVED BY

B. E. Reahnd  
Francis Rouse  
William H. Gram  
Alan D. Saunders  
Dorothy A. Foster

DISSERTATION COMMITTEE

ACKNOWLEDGMENTS

My sincere appreciation is extended to the following members of my dissertation committee for their valuable contributions to this study: Dr. Robert Ragland, Dr. Avi Scherman, Dr. Dorothy Foster, Dr. Bill Graves, Dr. Alan Nicewander, and Dr. Albert Smouse. I would like to express a special note of thanks to my chairman, Dr. Robert Ragland, for his continuing support and encouragement throughout my graduate studies. I am also especially grateful to Dr. Avi Scherman for his help and support as committee chairman during Dr. Ragland's sabbatical leave.

In addition to the committee members, I wish to thank a number of others for their generous and valuable contributions to the conduct of this research project. I particularly wish to thank Dr. Oscar Parsons of the University of Oklahoma Health Sciences Center for his interest, advice, and his assistance in recruitment of subjects. Appreciation is also extended to all of those professional psychologists at the OUHSC, within the Guidance Center system of the state of Oklahoma, and in other agencies and private practice who gave generously of their limited time and considerable expertise to serve as subjects. Special thanks is given to Ray Hand for his assistance and support and to David Mitchell for serving as counselor for this project. I am

also particularly grateful to RC for allowing videotaping of her private counseling session for research purposes.

Finally, I wish to thank my husband Henry, for his unfailing encouragement and support throughout my graduate studies and my children, Kim and Kristen, for their patience and love at times when they wished their mother didn't have to work on a dissertation!



ABSTRACT

Three groups of ten professional psychologists were provided with differing types of information about an actual counseling client: one group received minimal, or stereotype, information about the client (Stereotype Group); a second group received minimal information and viewed videotaped excerpts from a counseling session with the client (Stereotype-Video Group); and a third received both of these types of information plus test protocols from a standard test battery administered to the client (Stereotype-Video-Diagnostics Group). All subjects then made a number of clinical judgments about the client. These included: 1) predicting how the client described herself on a personality checklist (Predictive Task), 2) evaluating the client on a number of clinical dimensions (Evaluative Task), and 3) diagnosis of the client. Judges were also asked to rate their confidence in the accuracy of each judgment made.

No significant effect of type of information upon accuracy in performance of either the Predictive or Evaluative Tasks was obtained. Type of information did appear to have an effect upon diagnosis of the client and upon confidence in judgments. Type of information was found to have a significant effect upon the extent to which judges relied upon assumed similarity in predicting how the client saw herself.

## Effects of Information

It was concluded that type of information may have complex differential effects upon different types of clinical judgments.

## TABLE OF CONTENTS

	Page
ACKNOWLEDGMENTS . . . . .	iii
ABSTRACT. . . . .	v
LIST OF TABLES. . . . .	ix
LIST OF FIGURES . . . . .	x
INTRODUCTION. . . . .	1
METHODS . . . . .	9
Subjects . . . . .	9
Experimental Conditions and Procedures . . . . .	9
Judgmental Tasks . . . . .	11
RESULTS . . . . .	13
Predictive and Evaluative Judgment Tasks . . . . .	13
Confidence in Predictive and Evaluative Judgments. . . . .	14
Assumed and Real Similarity. . . . .	15
Effects of Information Type Upon Accuracy, Confidence, Real and Assumed Similarity. . . . .	17
Appropriateness of Confidence. . . . .	18
Effects of Information Type on Diagnosis . . . . .	20
Effects of Information Type Upon Other Tasks. . . . .	21
DISCUSSION. . . . .	23
Effects of Information Type Upon Accuracy of Predictive and Evaluative Judgments . . . . .	23
Confidence and Appropriateness of Confidence in Predictive and Evaluative Judgment. . . . .	26
Effects of Information Type Upon Assumed and Real Similarity. . . . .	29
Effects of Information Type Upon Diagnosis . . . . .	29
SUMMARY . . . . .	32
REFERENCES. . . . .	47

	Page
APPENDICES. . . . .	53
A. Prospectus . . . . .	53
B. Information for Prospective Parti- cipants. . . . .	114
C. Legally Effective Informed Consent of Client . . . . .	117
D. Consent for Videotaping of Client. . . . .	124
E. Institutional Review Board Executive Summary. . . . .	127
F. Instructions to Experimental Subjects Stereotype Group . . . . .	134
G. Instructions to Experimental Subjects Stereotype-Video Group . . . . .	137
H. Instructions to Experimental Subjects Stereotype-Video-Diagnostics Group . . . . .	141
I. Evaluative Dimensions Instrument . . . . .	145
J. Demographic Information Questionnaire. . . . .	152
K. Trait Rating Instrument (Judges to Fill Out for Client). . . . .	155
L. Trait Rating Instrument (Judges to Fill Out for Self). . . . .	162
M. Client Information (Stereotype Information) . . . . .	168
N. Client Rating of Likability and Similarity and Client Diagnosis . . . . .	170
O. Data . . . . .	172

LIST OF TABLES

Table Number		Page
1	Demographic Data for Groups of Subjects. . . . .	34
2	Means and Standard Deviations of Accuracy Scores Presented by Type of Task and Type of Information. . . . .	35
3	Confidence in Predictive and Evaluative Judgments for the Three Information Conditions. . . . .	36
4	Assumed Similarity. . . . .	37
5	Mean Assumed Similarity, Real Similarity, and Predictive Accuracy as a Function of Information Type . . . . .	38
6	Correlation of Accuracy with Confidence as a Function of Type of Information . . . . .	39
7	Primary Diagnosis of the Client by Information Condition . . . . .	40
8	Primary and Other Diagnoses of the Client by Information Condition. . . . .	41
9	Mean Rated Liking of the Client on a Five-Point Scale as a Function of Information Type. . . . .	42
10	Mean Perceived Similarity to the Client on a Five-Point Scale as a Function of Information Type . . . . .	43

LIST OF FIGURES

Figure		Page
1	Mean Accuracy of Predictions and Evaluations as a Function of Information Type . . . . .	44
2	Mean Confidence in Predictions and Evaluations as a Function of Information Type . . . . .	45
3	Mean Assumed Similarity, Real Similarity, and Predictive Accuracy as a Function of Information Type . . . . .	46

Effects of Differing Types of Information  
Upon Clinical Judgments

The manner in which different kinds of information about a client affect the clinician's judgment about that client is an issue of importance in the field of psychodiagnostic testing and psychotherapy. The rationale for administration of psychological tests would appear to be that testing provides additional information with which to increase understanding of the client in one or both of two important ways: first, from an "external" perspective in terms of the actual personality dynamics of the client as it has import for making clinical judgments (e.g., therapy approach, hospitalization, court-related issues, and others), and, secondly, from an "internal" reference point in terms of understanding the client's self-perceptions. This latter type of understanding may relate to the process of "empathy" in therapy which Rogers (1957) and Carkhuff (1967), among others, have envisioned as important to positive process change.

On the other hand, some therapists question the value of psychological testing, objecting that, far from aiding in understanding the client, diagnostic material impedes understanding through the creation of response sets, resulting in stereotyping or pigeon-holing of clients.

Previous research examining the effect of different types or amounts of information, such as diagnostic test material, upon the ability of judges to make accurate judgments about a subject has yielded equivocal results (Hamlin, 1954; Borke and Fiske, 1957; Giedt, 1955; Soskin, 1954a, 1954b, 1959; Kostlan, 1954; Horowitz, 1962, Stelmachers and McHugh, 1964; Hjelle, 1968; Taft, 1966; Rodin, 1975; Cline, 1955; Cline and Richards, 1958; and Cline and Richards, 1960). Studies have sometimes found predictions based on minimal or "stereotype" information to be as accurate as, or more accurate than, judgments based on more complex data, such as test material. The information provided to judges in many of these studies has, however, been of questionable value to the judgments required. The judgment tasks themselves have typically not been relevant to the judgments actually required of the practicing clinician. In most cases, judges have been asked to make predictions based upon descriptive or diagnostic material without ever having seen the individual whom they were judging. Shrauger and Altrocchi (1964) pointed out that selection effects may occur in studies in which the person judged is not actually observed by the judges.

Rodin (1975) also speculated that failure of a number of earlier studies to find increased predictive accuracy as



a function of increased information may have been due to the lack of relevance of information given judges to the judgments required of them. She found that when judges were allowed to self-select information they deemed useful in making given judgments, they were able to predict with greater accuracy than a "stereotype" group. However, judgment tasks which she employed seem of questionable relevance to those required in clinical settings.

Few generalizations relating to the effects of differing types of information upon predictive accuracy can be made from the findings reviewed above. The differences among studies in judgments required and other experimental procedures employed were rarely systematic and make direct comparison of results extremely difficult. However, one relatively consistent finding seems to be that certain types of information, such as diagnostic labels or projective test data, can result in response biases in perceptions of the client. Another finding which seems to recur is that predictive accuracy can be about as good, or sometimes better, when responses are based upon minimal or "stereotype" information as when based upon more complex information, such as test data.

The current study examines again the effect of differing kinds of information upon accuracy of clinical judgments. An

attempt has been made to employ experimental procedures which more closely approximate the conditions routinely existing in clinical practice both through inclusion of diagnostic test material and observation of an actual counseling session with a client.

As noted by Reed and Jackson (1975), since the mid nineteen-sixties, researchers have focused upon the development of descriptive models of information processing, rather than upon the accuracy of judgments. Many studies have been concerned with the application of probability notions, or Bayesian theory, to judgment processes. Reviews of this area were made by Peterson and Beach (1967) and Slovic and Lichtenstein (1971). Fishbein and Ajzen (1975) cite these reviews as support for their view that "Bayes' theorem is a reasonably good descriptive model of human information processing" (p. 184).

On the other hand, Kahneman and Tversky (1973), found that judges in their studies did not act in accord with normative, or Bayesian theory, as they relied too much on individuating information and paid little attention to "base rates" or prior probability. The findings of Kahneman and Tversky (1973) were supported by Lyon and Slovic (1976); Nisbett, Borgida, Crandall and Reed (1976); and Slovic, Fischhoff, and Lichtenstein (1977).

Rosenhan's 1973 study of psychiatric diagnostic judgment called that process into serious question and precipitated much discussion and analysis (Crown, 1975; Farber, 1975; Millon, 1975; Rosenhan, 1975; Spitzer, 1975; Weiner, 1975) and Davis (1976; 1979) discussed the role of base rate application in psychodiagnosis and concluded that the misdiagnoses made by clinicians in the Rosenhan (1973) study were attributable to their failure to consider appropriate base rate data. In a review of the literature on stereotyping, McCauley, Stitt, and Segal (1980) conclude, ". . . people sometimes underuse data and rely too much on base rate, sometimes underuse base rate and rely too much on individuating information, and sometimes attend to both according to the Bayesian prescription" (p. 200).

In a series of studies (Kahneman and Tversky, 1972; 1973; and Tversky and Kahneman, 1974) Kahneman and Tversky found that their judges did not perform in a Bayesian manner when making predictive versus evaluative judgments. Although tasks of a predictive nature require greater extrapolation from the data than do tasks of an evaluative nature, judgments were found to be equally extreme when subjects were asked to predict a remote objective criterion on the basis of sketchy information as when they were asked merely to evaluate the information itself. Kahneman and Tversky

explain their findings on the basis of the "representativeness heuristic." That is, judges will make that prediction which seems to be "representative" of the data no matter how incomplete the information or how irrelevant the information may be as a basis for prediction.

In practice, clinicians are frequently called upon not only to evaluate a client (or make a comparison on some dimensions with a reference group), but also to make some prediction about performance. Such judgments may be risky not only in that they involve very large inferences on the basis of insufficient or perhaps irrelevant data, but such predictions may also involve great risk in terms of practical consequences to the individual involved (e.g., how likely is a client to commit a violent crime, etc.).

The current study will examine the manner in which differing types of information affect judgments of a predictive and of an evaluative nature made in a clinical context. These tasks will be described in detail in the discussion of experimental procedure.

Finally, the relationship between confidence in a judgment and the kind of information upon which the clinician bases that judgment will be examined. Little (1961) found evidence for the existence of a stable "level of general confidence" which is characteristic of individuals.

Oskamp (1965) investigated the relationship between amount of information and confidence in judgments, and found that as amount of information increased, confidence increased dramatically while accuracy of judgments did not increase significantly. Judgments were, however, made only on the basis of case-summary material, without the judges ever having viewed the subject of their judgments. The present study will allow for viewing, in a clinical context, the client to be judged.

Several hypotheses related to effect of differing types of information upon predictive and evaluative judgments and confidence in those judgments were formulated. First, it was hypothesized that allowing for viewing of the client in an actual counseling session would make available highly relevant client-related cues. This might result in improvement in predictive accuracy for the Stereotype-Video and Stereotype-Video-Diagnostics Information Groups as compared to the Stereotype Information Group, who did not view the client.

It was further hypothesized that diagnostic material might result in a set to "objectively evaluate" a client rather than an attempt to see the client as she saw herself. "Objective" evaluation of the client was required in the evaluative task while understanding of the clients' self-

perception was necessary in the predictive task. For the evaluative task, then, the Stereotype-Video-Diagnostics Information Group was expected to achieve greatest accuracy, followed by the Stereotype-Video Information Group, with the Stereotype Information Group least accurate.

Conversely, it was hypothesized that for the predictive task, the Stereotype-Video Information Group would achieve the greatest predictive accuracy, followed by the Stereotype-Video-Diagnostics Information Group, with the Stereotype Information Group again being least accurate.

Finally, as confidence has previously been shown to increase with increased information, it was hypothesized that the Stereotype-Video-Diagnostics Group would be most confident in their judgments, as they possessed the most information about the client, followed by the Stereotype-Video and Stereotype Information Groups.

## METHOD

Subjects

The subjects were 30 professional counselor/therapists with experience in individual therapy and diagnostic testing. Included were 18 males and 12 females. Demographic information concerning subjects is summarized in Table 1. Subjects were randomly assigned in equal numbers to three experimental conditions, as follows.

Experimental Conditions and Procedure

Stereotype condition. Those subjects in the Stereotype Condition were given minimal information about an actual client including sex (female), age (21), college major (computer science), and the fact that she was a counseling/psychotherapy client at a university counseling center.

Stereotype - Video condition. Subjects in this condition received the minimal information about the same client and also viewed a 15-minute videotape of segments taken from a counseling session with the client. The tape was composed of three five-minute segments taken at random from the beginning, middle, and end of the client's first session with her counselor after interview with an intake worker.

Stereotype-Video-Diagnostics condition. Subjects in

this group received the minimal information about the client, viewed the videotape, and additionally were provided with diagnostic material about the client. Viewing of the videotape and study of the diagnostic material were counterbalanced for order of presentation. Diagnostic data was presented in protocol form in order to approximate actual clinical conditions. For the Rorschach Inkblot Test, a summary table was also provided. The battery was composed of: 1) Wechsler Adult Intelligence Scale (WAIS), 1) Rorschach Inkblot Test, 3) Thematic Apperception Test, 4) Minnesota Multiphasic Personality Inventory (MMPI), and 5) Human Figure Drawings. This battery was chosen to be representative of a typical test battery employed by many of the subjects. Because of limited time availability of the professionals involved, a time limit of 1½ hours was placed on subjects in this group for completion of their tasks. All subjects in the group appeared to be able to comfortably complete the procedure within this time limit.

Client. The female client about whom information was provided was an actual counseling client who voluntarily agreed to participate. The client evidenced concerns over dependency and achievement, and appeared anxious and depressed. Extensive precautions to safeguard the identity of the client were taken. All subjects were required to



make judgments about the same client.

Trait rating instrument. The client was also asked to rate herself on five-point scales on 15 personality dimensions chosen from among those commonly appearing as assessment dimensions on widely-used personality tests, such as "anxious," "caring," "independent," etc. An equal number of positive and negative terms were included. Rating was done at the conclusion of the counseling session which was taped for viewing as a part of the experimental procedure.

#### Judgmental Tasks

After having acquired the information appropriate to their respective groups, judges were asked to complete a number of judgmental tasks concerning the client. Previously, three trained psychologists not otherwise participating in the experiment had sorted descriptions of the tasks as to whether they were predictive or evaluative in nature as compared to a criterion example of the type reported by Kahneman and Tversky (1973). Only those tasks sorted in the same way by all three judges were retained in the experimental procedure. A description of the different judgmental tasks follows.

Predictive tasks. All subjects were asked to predict the client's responses on the Trait Rating Instrument.

Evaluative tasks. S's were asked to rate the client on 15 five-point scales representing dimensions upon which counselors are frequently asked to characterize clients, such as anxiety, depression, trust, etc. Judges' responses were compared to those of the client's own therapist after twelve therapy sessions.

Confidence tasks. Counselors were asked to indicate their confidence in each judgment in terms of their estimate of the probability that they had made an accurate judgment. Thus, confidence estimates were made on a scale ranging from 20 through 100 with 20 indicating a random chance of being accurate and 100 indicating absolute certainty of accuracy.

Other tasks. Upon completion of the predictive, evaluative, and confidence tasks, judges performed a number of other tasks: 1) completion of the demographic questionnaire, 2) rating how well they liked the client as a person, 3) rating how similar the client was to themselves, 4) diagnosis of the client, and 5) filling out the Trait Rating Instrument to describe themselves.

## RESULTS

Predictive and Evaluative Judgment Tasks

An absolute difference score was calculated for each judge between the judge's predicted responses for the client on the Trait Rating Instrument and the client's actual responses (Cronbach, 1955; Cronbach and Gleser, 1953; and Gage and Cronbach, 1955). This score was arrived at by summing the difference without regard to sign between the client's response and the judge's prediction of her response on each item of the Trait Rating Instrument. These difference scores can be thought of as representing the judge's predictive accuracy with regard to the client on the Trait Rating Instrument. These scores also indicate how well the judges were able to see the client as she saw herself. Range of possible scores is zero (highest accuracy) to 60 (lowest accuracy). A score of 30 would be expected by chance.

Similarly, an absolute difference score was calculated for each judge by summing the difference between the responses of the client's therapist and the judge for each of the evaluative dimensions (Evaluative Dimensions Instrument). These scores can be conceptualized as representing

the ability of the judge to accurately evaluate the client from an external perspective. Again, possible scores range from zero (highest accuracy) to 60 (lowest accuracy) with a score of 30 representing chance accuracy.

Mean accuracy scores and standard deviations on the two judgment tasks for the three information conditions are shown in Table 2 and Figure 1. Accuracy of all judgments was greater than that expected as a result of chance. Although there was little effect of type of information upon accuracy for either the predictive or evaluative tasks, judges in all information conditions seem to have been more accurate in performance of the evaluative task than the predictive one.

.....  
-----  
Table 2 about here  
-----  
.....

-----  
Figure 1 about here  
-----

#### Confidence in Predictive and Evaluative Judgments

Confidence estimates for the predictive and evaluative tasks for the three groups receiving differing types of information are shown in Table 3 and in Figure 2. It appears that confidence in judgments increased with

increasing information, especially for the evaluative task. Judges differed widely in the "stable level of general confidence" (Little, 1961) which was characteristic of them, with confidence estimates ranging from 21 to 97.

.....  
-----  
Table 3 about here  
-----  
.....

-----  
Figure 2 about here  
-----

#### Assumed and Real Similarity

Cronbach (1955) noted the importance of taking into account the possible effects of assumed and real similarity between judge and subject when interpreting differences in predictive accuracy. He found that differences between judges in predictive accuracy may be a function of differences in assumed or real similarity between judge and subject rather than due to differences in information, experience, or other variables.

Table 4 depicts and compares the concepts of assumed and real similarity and predictive accuracy.

.....  
-----  
Table 4 about here  
-----

An index of assumed similarity was computed for each judge between the judge's own self-description and the judge's prediction of the client's self-description. This score was arrived at by summing the difference without regard to sign between the judge's response and the judge's prediction of the client's response on each item of the Trait Rating Instrument. These difference scores represent the extent to which a judge assumed that the client described herself in the same way the judge described himself or herself, or the extent to which the judge based predictions on "assumed similarity" (Cronbach, 1955).

An index of real similarity was also calculated for each judge. This was accomplished by adding the absolute difference between the judge's self-description and the client's self-description on each item of the Trait Rating Instrument. These scores indicate the extent to which the client and judge actually viewed themselves similarly.

Data on assumed and real similarity as compared to predictive accuracy are shown in Table 5 and Figure 3. Judges in the three information conditions did not differ from one another in similarity of self-descriptions to the client's self-description (real similarity). Neither, as shown previously, did they differ in accuracy of their predictions.

It is apparent from Figure 3, however, that type of information did have an effect on assumed similarity, as the group receiving only stereotype information appears to have made much greater use of "assumed similarity" (Cronbach, 1955) than those judges receiving more individualized information about the client. This was confirmed by the statistical analysis reported below.

.....  
-----  
Table 5 about here  
-----  
.....

-----  
Figure 3 about here  
-----

Effects of Information Type Upon Accuracy, Confidence, Real and Assumed Similarity

Accuracy scores for the predictive and evaluative tasks, their respective confidence estimates, and indices of real and assumed similarity were submitted to a one-way multivariate analysis of variance in order to test for overall effect of type of information. A Wilk's lambda of 0.326 was obtained, yielding  $F(12, 44) = 2.76$  with  $p = .007$ . Thus, the multivariate analysis indicates that type of information did have an overall significant effect. Inspection of the six univariate analyses of variance reveals

that a significant F was obtained in only one such analysis. The analysis of effect of information type upon assumed similarity yielded  $F(2, 27) = 9.97$  with  $p = .0006$ . Individual comparisons showed the Stereotype Group to differ significantly from both the Stereotype-Video Group ( $F_{1,27} = 19.44$ ;  $p = .0001$ ) and the Stereotype-Video-Diagnostics Group ( $F_{1,27} = 7.90$ ;  $p = .009$ ). Thus, judges receiving minimal information alone did rely upon "assumed similarity" to a significantly greater extent than did judges provided with more information about the client.

The univariate analyses of variance of accuracy scores for the predictive and evaluative tasks did not yield significant results ( $F_{2, 27} = 0.56$ ;  $p = 0.58$  and  $F_{2, 27} = 1.20$ ;  $p = 0.32$ , respectively). Therefore, judges provided with minimal information alone were as accurate in their judgments as judges who viewed the client and who additionally studied her test protocols.

Analyses of variance of confidence estimates for the predictive and evaluative tasks also did not obtain significant results. The analysis of effect of information type upon confidence in evaluations did approach significance, however, with a resultant  $F(2, 27) = 2.81$  and  $p = .077$ .

#### Appropriateness of Confidence

Correlations between accuracy scores for the predictive



and evaluative tasks and their respective confidence estimates are shown in Table 6. These correlations represent the extent to which confidence in judgments was warranted. Here, negative correlations represent greater appropriateness of confidence, in which a small difference score is associated with high confidence and vice versa. Although none of the obtained correlations was statistically significant from zero, nor did any differ significantly from the other, it is of interest to note that the only negative correlation was found in the Stereotype-Video Information condition for the predictive task. This indicates that judges who viewed the client but did not receive diagnostic material seemed less confident in their ability to predict how the client saw herself when, in fact, their predictions were less accurate. Thus, their confidence in their predictions was more appropriate.

Confidence in accuracy of their predictions appeared less warranted for those judges receiving test data than for those who only viewed the client. However, both of the groups who viewed the client evidenced greater appropriateness of confidence than did the group who did not view her (Stereotype Group).

The groups did not appear to differ in appropriateness of confidence in their ability to evaluate the client as

they believed her to be "in reality." All groups obtained low positive correlations indicating largely unwarranted confidence in the accuracy of their judgments on this task.

---

Table 6 about here

---

#### Effects of Information Type on Diagnosis

Table 7 presents frequency of various primary diagnoses as a function of type of information upon which the diagnosis was based. Table 8 reports the same information for both primary and secondary diagnoses summed. These data did not lend themselves to statistical examination, as they violated the requirements of the chi-square test with regard to minimum cell frequencies (Siegel, 1956, p. 178).

Some interesting differences in diagnoses made by judges receiving differing types of information should be noted, however. Only one primary diagnosis of depression or affective disorder was made by judges viewing the client but not receiving test data, whereas six judges who both saw the client and received test results diagnosed her in this way. In this respect, judges who received test data appeared to be more similar in their perception of the client to those who had never viewed her (Stereotype Group), five of whom gave a primary diagnosis of depression.

For summed diagnoses, four judges receiving only stereotype information diagnosed the client as obsessive-compulsive, whereas only one judge who actually viewed the client saw her in this way. Also, only one person who did not view the client diagnosed her as either hysterical or dependent personality, while nine judges who had seen her gave one of these diagnoses, four of which were primary diagnoses. Thus, type of information appears to have had some effect upon judge's perception of the client.

.....  
-----  
Table 7 about here  
-----  
.....

-----  
Table 8 about here  
-----

#### Effects of Information Type Upon Other Tasks

Means and standard deviations of judges' perceived liking of the client and perceived similarity to the client are seen in Table 9 and Table 10. Ratings were made on five-point scales.

Type of information had no effect upon either acknowledged liking of or perceived similarity to the client. Judges in the Stereotype-Video Group saw themselves as being slightly less similar to the client than did judges in the

other groups. Real similarity indices reflect that this was, in fact, the case.

.....  
-----  
Table 9 about here  
.....  
-----

.....  
-----  
Table 10 about here  
.....  
-----

## DISCUSSION

Effects of Information Type Upon Accuracy of Predictive and Evaluative Judgments

It had been hypothesized that, under the conditions in the current study, providing individualized clinical information to judges in the Stereotype-Video and Stereotype-Video-Diagnostics groups would result in their being more accurate in their judgments about the client than judges who received stereotype information alone. This hypothesis was not supported by results of this study. Judges were able to perform both the predictive and evaluative tasks as accurately based upon minimal information as they were with the additional information provided by viewing the counseling session and studying the test material. These findings are in agreement with those of previous studies which have found judgments based upon minimal information to be as accurate as those based on more individualized data, in most cases (Soskin, 1959; Stelmachers and McHugh, 1964). In the present study, this was true even though judges were able to see a counseling session with the client and with tasks which were more clinically relevant. That Rodin (1975) found increased accuracy with increased information about the client may be attributable to the fact that judges in

her study were allowed to self-select information upon which to base their judgments. This was not the case in the current study, although the test information provided was representative of the type of data routinely used in practice by many of the judges. Complex differences among studies in types of judgment tasks employed, type of information provided, client-related variables, and conditions under which judgments were made make interpretation of differences in results extremely difficult.

It could be argued that the analog nature of the present study contributed to the findings of no information effect upon accuracy of judgment by removing judges from the actual clinical setting. This would not appear to be the case, however, since viewing of the client on videotape does appear to have had an effect upon diagnostic perceptions of her and upon assumed similarity to her.

It should be noted that the current study employed only measures of average relative accuracy across a number of predictive and evaluative judgments. Thus, ways in which differing types of information might have differentially affected accuracy in specific predictions and specific evaluations were not examined here. It may well be that differing information types might have affected accuracy in different specific judgments in different ways.

Although this question was beyond the scope of the current study, it is one which should certainly be given consideration in further research.

Soskin (1954, 1959) found projective material to result in increased perception of maladjustment in the client. Thus, it has been further hypothesized that diagnostic material in general might result in a set to "objectively evaluate" a client rather than an attempt to understand the client as she saw herself, which might be facilitated by merely viewing her. This would have resulted in greater accuracy of the Stereotype-Video-Diagnostics Group on the evaluative task and greater accuracy of the Stereotype-Video Group on the predictive task. No significant differences in accuracy between the groups on either of the two tasks was obtained in the current study. It is of interest to note, however, that, while they were not more accurate in their judgments on the predictive task, judges in the Stereotype-Video Group did seem to evidence greater appropriateness of confidence in their ability to accurately understand how the client saw herself. That is, judges in this group were more confident that their predictive judgments were accurate when they were, in fact, accurate and less confident when their judgments were, in fact, less accurate.

Inspection of Figure 1 indicates that judges in all groups appear to be more accurate in their evaluations than in their predictions. It is not possible to ascertain on the basis of the present data, however, whether differences in accuracy on the two tasks are attributable to: 1) the differences in degree of extrapolation required from the available data (prediction or evaluation), 2) the different perspective on the client which each task involves (external or internal), or 3) merely the specific differences in the two tasks employed in the study. In regard to this latter possibility, the greater accuracy on the evaluative task may simply reflect that psychologists are better able to predict the responses of another psychologist (client's therapist), expressed in clinical terms, than they are to predict the responses of a client in more general descriptive terms.

#### Confidence and Appropriateness of Confidence in Predictive and Evaluative Judgments

It had been hypothesized that increased information would result in increased confidence, in line with the findings of Oskamp (1965). Although differences in confidence between the information conditions were not found to be statistically significant, they did closely approach



significance with the evaluative task. The trend of increased confidence in evaluations with increasing information is clearly seen in Figure 2. A similar trend, though less pronounced, is also seen for the predictive task. These findings are generally in agreement with those of Oskamp (1965), who found a large increase in confidence with increased amounts of case-summary material. In the present study, this effect was obtained with experimental procedures designed to more closely approximate clinical practice, including observation of the client's counseling session. The great variability in individual "stable level of confidence" (Little, 1961) noted in the confidence estimates undoubtedly contributed to lack of statistical significance in the current study.

Further research in the area of confidence in clinical judgments should attempt to control for the effect of differing "stable levels of confidence" among individual judges, perhaps through the use of an analysis of covariance technique.

It appears from Figure 2 that addition of test data made a large contribution to the apparent increase in confidence. Addition of test data did not result in overall increased accuracy of judgments, however (Figure 1). It

would seem, then, that test material gave judges a false sense of confidence in their ability to make accurate judgments about the client. This is reflected in the data on appropriateness of confidence seen in Table 6, and is also in agreement with similar findings by Oskamp (1965).

It had been hypothesized that judges in the Stereotype-Video Group would be more appropriately confident in their judgments on the predictive task, while those in the Stereotype-Video-Diagnostics Group would evidence greater appropriateness of confidence on the evaluative task. Although differences in appropriateness of confidence were not statistically significant, they were in the hypothesized direction for the predictive task. As noted earlier, judges who only viewed the client, but did not receive diagnostic material, seemed to be more appropriately confident in their ability to see the client as the client saw herself than judges who also received test data. Thus, judges who did not see test data seemed to have a better understanding of when they were right and when they were wrong in their predictions.

That information type appears to have had some effect on appropriateness of confidence for predictions and not for evaluations is also of interest, though a clear explanation of this difference is not possible within the

framework of this study. Again, this may be a function of the content- and criterion-related differences between the predictive and evaluative tasks employed here.

#### Effects of Information Type Upon Assumed and Real Similarity

Judges in the three information conditions did not differ significantly in similarity of their self-descriptions to the client's self-description (real similarity). However, type of information did have a significant effect upon assumed similarity. (See Table 5 and Figure 3.)

Judges receiving only minimal information (Stereotype Group) assumed similarity between their self-descriptions and the self-description of the client to a much greater extent than did judges receiving more individualized information about the client. It appears that, lacking any extensive personal information about the client, judges in this group were forced to base their ideas about how she saw herself to a large extent upon their own self-descriptions, assuming their own responses to be typical of the responses of human beings in general.

#### Effects of Information Type Upon Diagnosis

Effects of information type upon diagnosis of the client may be seen in Tables 7 and 8. As previously noted, Shrauger and Altrocchi (1964) discussed the fact that selection effects may obtain when the person judged is not actually

seen, but judged on the basis of descriptive material alone. Such effects related to viewing or non-viewing of the client are readily apparent here in terms of diagnosis. Four judges who had received minimal information about the client, but who had not viewed her, diagnosed her as obsessive-compulsive, whereas only one judge who had seen her diagnosed her in this way. This appears to be directly related to the stereotype associated with computer science majors as rigid, compulsive individuals. Though judges in all three information conditions were given the information that the client was a computer science major, those judges who were additionally able to view the client seem to have been able to free themselves from this stereotype.

Conversely, only one person who did not view the client gave a diagnosis of either hysterical or dependent personality, while nine judges who had seen her diagnosed her in one of these ways. Again, it appears that the Stereotype Group was operating from the notion of a computer science major as an individual who is independent and unemotional, whereas judges who actually saw the client perceived her in a much different way. These findings appear to be consistent with those of Kahneman and Tversky (1972, 1973) and Tversky and Kahneman (1974) regarding the effects of reliance upon the representativeness heuristic when making

judgments. Those subjects in the Stereotype Group appear to have made that diagnosis (obsessive-compulsive) which was "representative" of the data (computer science major). In this case, however, the client was not typical, in this respect, of the stereotype, and the resulting judgment was in error.

It is also noteworthy that providing judges with test data appears to have influenced their diagnostic perception of her. Only one judge who viewed the client but did not receive test data about her diagnosed her primarily as depressed, while six judges who both viewed her and received her test results gave the primary diagnosis of depression. Thus, addition of test data seems to have contributed to the perception of the client as depressed. Five judges who had never seen the client also gave a primary diagnosis of depression. Thus, judges who had test data were similar in their perception of the client to judges who had never seen her.

## SUMMARY

In the current study, providing clinicians with an opportunity to view the client in a counseling session and with test data regarding the client did not result in improved accuracy in predictive and evaluative judgments about her. There did appear to be a trend, however, for greater amounts of information about the client to result in inappropriate increases in clinicians' confidence in their judgments about the client.

Type of information about the client did seem to have an effect upon diagnosis of the client, with those clinicians not having viewed her describing her as more obsessive-compulsive, and those having viewed her diagnosing her more frequently as dependent or hysterical personality. Those clinicians receiving test data about the client more often viewed her as depressed than did those clinicians who did not receive test data.

Finally, it was found that clinicians who received only stereotype information about the client relied to a much greater extent upon "assumed similarity" in making judgments about her than did those provided with more individuating information.

Thus, type of information about a client seems to affect perception of the client in complex ways which require further clarification.

TABLE 1  
Demographic Data for Groups of Subjects

	Stereotype Group (N=10)	Stereotype- Video Group (N=10)	Stereotype- Video- Diagnostics Group (N=10)
Mean Age	37.2 (SD=11.1)	33.8 (SD=4.7)	34.5 (SD=8.5)
Sex	7 Male 3 Female	5 Male 5 Female	6 Male 4 Female
Degree	5 Ph.D. 5 Masters	7 Ph.D. 3 Masters	7 Ph.D. 3 Masters
Mean Years Experience	9.2 (SD=4.8)	7.5 (SD=3.7)	6.9 (SD=2.6)
Mean Hours Course- Work in Testing	13.6 (SD=5.6)	14.1 (SD=8.2)	14.1 (SD=5.5)
Mean Number Ror- schachs Given	75.5 (SD=53.7)	44.3 (SD=36.9)	35.2 (SD=26.9)
Mean Number MMPI Given	97.9 (SD=152.1)	239.6 (SD=295.4)	71.4 (SD=87.2)
Mean Number Wechslers Given	125.5 (SD=155.0)	149.0 (SD=147.6)	45.3 (SD=45.2)



TABLE 2  
Means and Standard Deviations  
of Accuracy Scores  
Presented by Type of Task  
and  
Type of Information

	Stereotype Group (N=10)	Stereotype-Video Group (N=10)	Stereotype- Video- Diagnostics Group (N=10)
Predictive Task			
Mean	18.7	19.7	18.5
Standard Deviation	1.95	3.13	2.91
Evaluative Task			
Mean	14.9	13.1	15.0
Standard Deviation	1.79	2.69	4.27

NOTE: Range of possible scores is zero to 60. A score of 30 represents chance accuracy. Lower scores indicate greater accuracy.

TABLE 3

Confidence in Predictive and Evaluative Judgments  
for the Three Information Conditions

	Stereotype Group (N=10)	Stereotype-Video Group (N=10)	Stereotype- Video- Diagnostics Group (N=10)
<b>Predictions</b>			
Mean	59.9	58.9	65.5
Standard Deviation	16.3	17.5	10.7
<b>Evaluations</b>			
Mean	61.6	66.6	80.8
Standard Deviation	21.6	22.1	10.3

NOTE: Range of possible confidence scores is 20 to 100.  
A score of 20 represents minimal confidence, while  
a score of 100 represents certainty.

TABLE 4

Assumed Similarity

Judge rates self	Judge predicts client's self-rating
(J)	(JC)

$J - JC = \text{assumed similarity}$

Real Similarity

Judge rates self	Client rates self
(J)	(C)

$J - C = \text{real similarity}$

Predictive Accuracy

Judge predicts client's self-rating	Client rates self
(JC)	(C)

$JC - C = \text{predictive accuracy}$

TABLE 5  
 Mean Assumed Similarity, Real Similarity,  
 and Predictive Accuracy  
 as a Function of Information Type

	Stereotype Group (N=10)	Stereotype- Video Group (N=10)	Stereotype- Video- Diagnostics Group (N=10)
Assumed Similarity			
Mean	15.5	24.6	21.3
Standard Deviation	4.35	5.79	3.88
Real Similarity			
Mean	20.1	22.4	20.6
Standard Deviation	2.02	3.97	1.83
Predictive Accuracy			
Mean	18.7	19.7	18.5
Standard Deviation	1.95	3.13	2.91

NOTE: Range of possible scores is from zero to 60. A score of 30 represents chance similarity or accuracy. Lower scores indicate greater similarity or accuracy.

TABLE 6  
 Correlation of Accuracy with Confidence  
 as a Function of  
 Type of Information

	Stereotype Group (N=10)	Stereotype- Video Group (N=10)	Stereotype- Video- Diagnostics Group (N=10)
Predictive Task	.497	-.357	.162
Evaluative Task	.200	.290	.307

NOTE: Negative scores represent greater appropriateness of confidence.

TABLE 7  
 Primary Diagnosis of the Client by  
 Information Condition

Diagnosis	Stereotype Group (N=10)	Stereotype Video Group (N=10)	Stereotype- Video- Diagnostics Group (N=10)
Depression/ Affective Disorder	5	1	6
Adjustment Reaction to Adult Life	2	2	1
Anxiety Neurosis	1	1	2
Obsessive Compulsive Personality (or traits)	1	1	0
Dependent Personality	0	2	0
Hysterical Personality	0	1	1
Neurosis (Unspecified)	0	1	0
No Diagnosis	1	1	0

TABLE 8  
 Primary and Other Diagnoses of the Client by  
 Information Condition

Diagnosis	Stereotype Group	Stereotype- Video Group	Stereotype- Video- Diagnostics Group
Depression/ Affective Disorder	8	4	9
Adjustment Reaction to Adult Life	2	3	1
Anxiety Neurosis	1	3	3
Obsessive Compulsive Personality (or traits)	4	1	0
Dependent Personality	1	3	2
Hysterical Personality	0	2	2
Neurosis (unspecified)	0	1	0
Developmental Conflicts	0	0	1
Borderline Schizophrenia	0	0	1
No Diagnosis	1	1	0

TABLE 9  
Mean Rated Liking of the Client  
on a Five-Point Scale  
as a Function of Information Type

	Stereotype Group (N=10)	Stereotype- Video Group (N=10)	Stereotype- Video- Diagnostics- Group (N=10)
Mean	3.8	3.7	4.0
Standard Deviation	.63	.48	.47



TABLE 10  
Mean Perceived Similarity to the Client  
on a Five-Point Scale  
as a Function of Information Type

	Stereotype Group (N=10)	Stereotype- Video Group (N=10)	Stereotype- Video- Diagnostics Group (N=10)
Mean	2.7	2.3	2.7
Standard Deviation	.48	.48	.94

Figure 1

Mean Accuracy of Predictions and Evaluations  
as a Function of Information Type

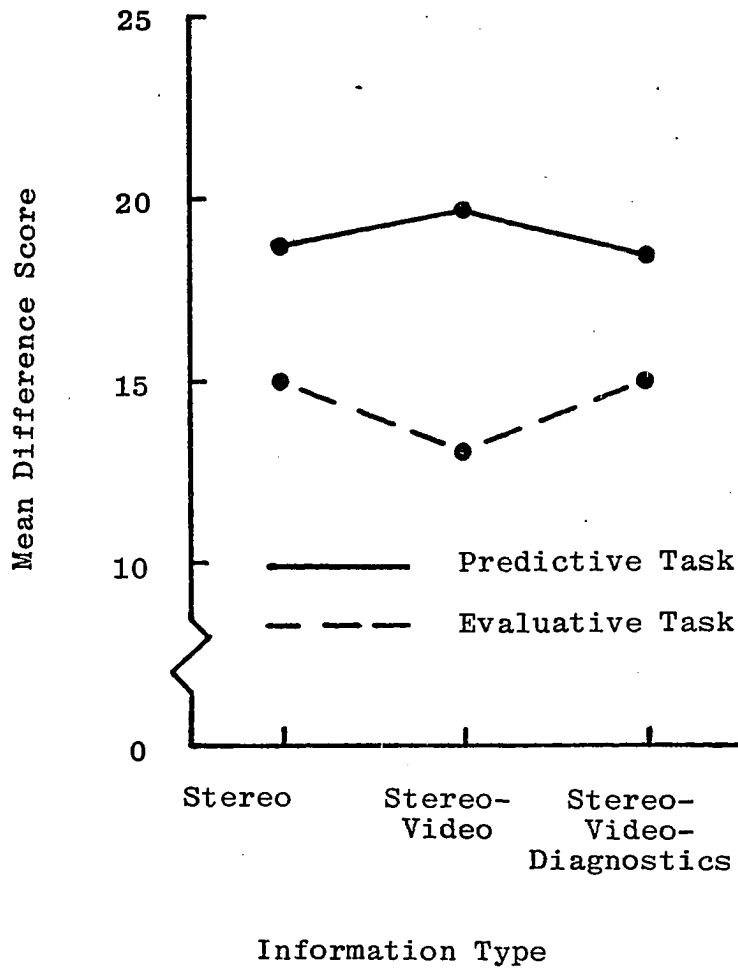


Figure 2

Mean Confidence in Predictions and Evaluations  
as a Function of Information Type

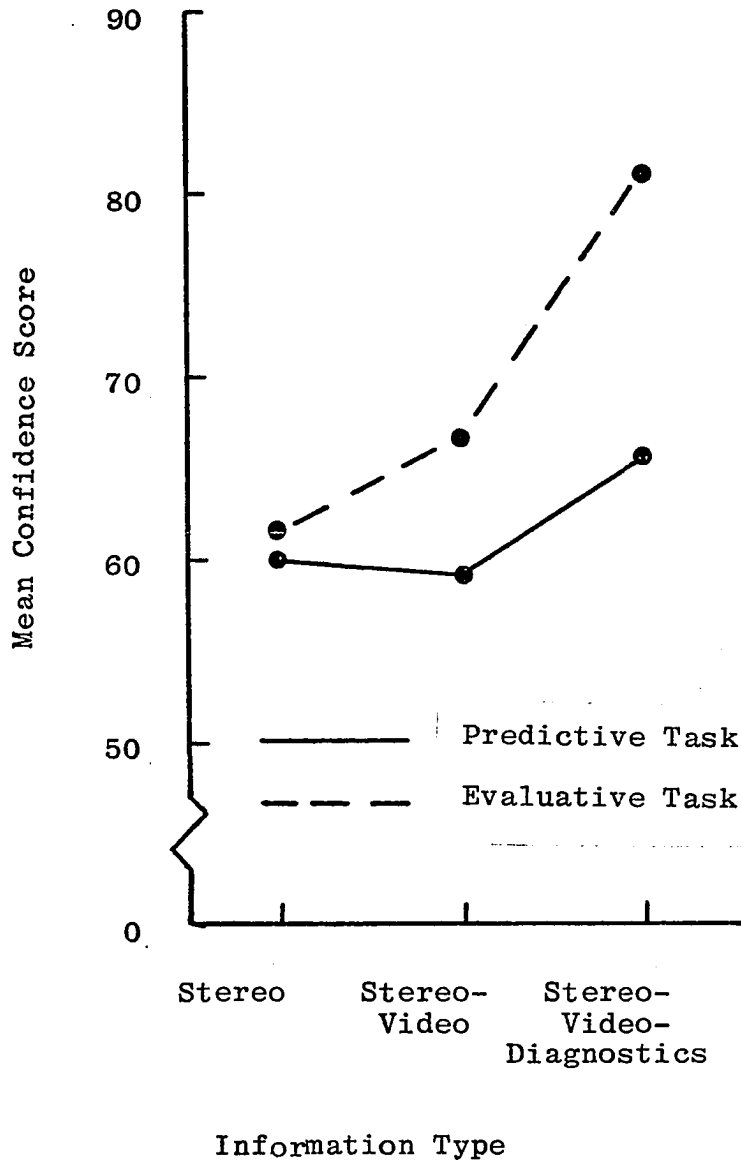
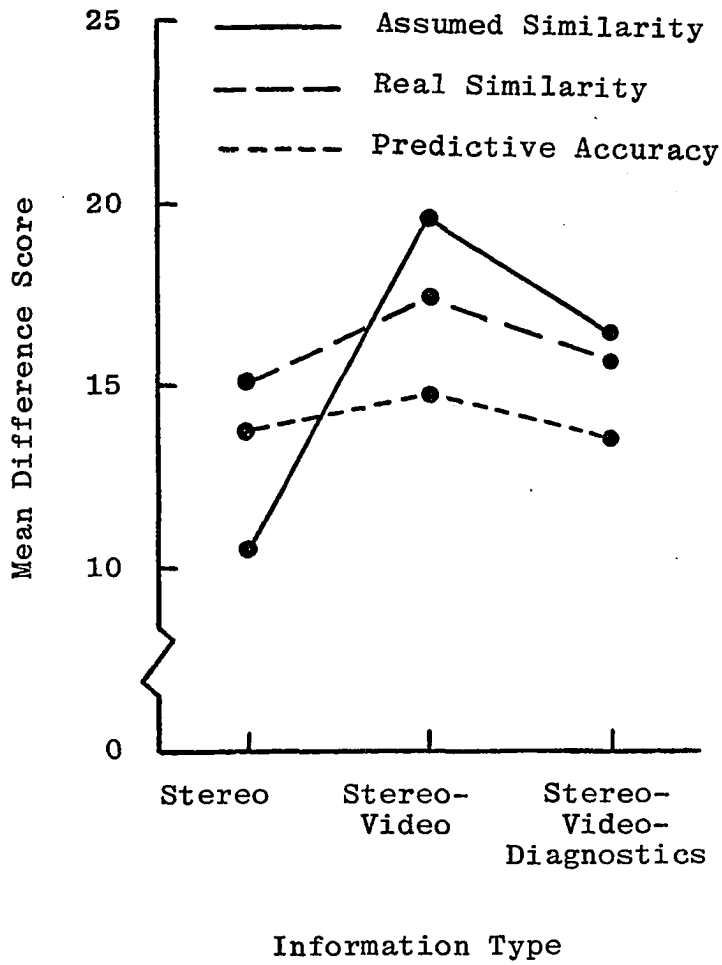


Figure 3

Mean Assumed Similarity, Real Similarity, and Predictive Accuracy as a Function of Information Type



## REFERENCES

- Borke and Fiske, D. W. Factors influencing the prediction of behavior from a diagnostic interview. Journal of Consulting Psychology, 1957, 21, 78-80.
- Cline, V. B. Ability to judge personality assessed with a stress interview and sound-film technique. Journal of Abnormal and Social Psychology, 1955, 50, 183-187.
- Cline, V. B., and Richards, J. M., Jr. Variables related to accuracy in interpersonal perception. USN Annual Report, 1958, ONR 171-146, Department of Psychology, University of Utah.
- Cline, V. B., and Richards, J. M., Jr. Accuracy of interpersonal perception: A general trait? Journal of Abnormal and Social Psychology, 1960, 60, 1-7.
- Cronbach, L. J. Processes affecting scores on "understanding of others" and "assumed similarity." Psychological Bulletin, 1955, 52, 177-193.
- Cronbach, L. J., and Gleser, G. C. Assessing similarity between profiles. Psychological Bulletin, 1953, 50, 456-473.
- Crown, S. "On being sane in insane places": A comment from England. Journal of Abnormal Psychology, 1975, 84, 453-455.

- Davis, D. A. On being detectably sane in insane places: Base rates and psychodiagnosis. Journal of Abnormal Psychology, 1975, 84, 442-452.
- Davis, D. A. What's in a name?: A Bayesian rethinking of attributional biases in clinical judgment. Journal of Consulting and Clinical Psychology, 1979, 47, 1109-1114.
- Farber, I. E. Sane and insane: Constructions and misconstructions. Journal of Abnormal Psychology, 1975, 84, 589-620.
- Fishbein, M., and Ajzen, I. Belief, attitude, intention, and behavior: An introduction to theory and research. Reading, Mass.: Addison-Wesley, 1975.
- Gage, N. L., and Cronbach, L. J. Conceptual and methodological problems in interpersonal perception. Psychological Review, 1955, 62, 411-422.
- Giedt, F. H. Comparison of visual, content, and auditory cues in interviewing. Journal of Consulting Psychology, 1955, 19, 407-416.
- Hamlin, R. M. The clinician as judge: Implications of a series of studies. Journal of Consulting Psychology, 1954, 18, 233-238.
- Hjelle, L. A. Accuracy of personality and social judgments as functions of familiarity. Psychological Records, 1968, 22, 311-319.

- Horowitz, M. J. A study of clinicians' judgments from projective test protocols. Journal of Consulting Psychology, 1962, 3, 251-256.
- Kahneman, D., and Tversky, A. Subjective probability: A judgment of representativeness. Cognitive Psychology, 1972, 3, 430-454.
- Kahneman, D., and Tversky, A. On the psychology of prediction. Psychological Review, 1973, 84, 237-251.
- Kostlan, A. A. A method for the empirical study of psychodiagnosis. Journal of Consulting Psychology, 1954, 18, 83-88.
- Little, K. B. Confidence and Reliability. Educational and Psychological Measurement, 1961, 21, 95-101.
- Lyon, D., and Slovic, P. Dominance of accuracy information and neglect of base rates in probability estimation. Acta Psychologica, 1976, 40, 286-298.
- McCauley, E., Stitt, C. L., and Segal, M. Stereotyping: From prejudice to prediction. Psychological Bulletin, 1980, 87, 195-208.
- Millan, T. Reflections in Rosenhan's "On being sane in insane places." Journal of Abnormal Psychology, 1975, 84, 456-461.

- Nisbett, R. E., Borgida, E., Crandall, R., and Reed, H.  
Popular induction: Information is not necessarily informative. In J. Carroll and J. Payne (Eds.), Cognition and social behavior. Hillsdale, N.J.: Erlbaum, 1976.
- Oskamp, S. Overconfidence in case-study judgments. Journal of Consulting Psychology, 1965, 29, 261-265.
- Peterson, C. R., and Beach, L. R. Man as an intuitive statistician. Psychological Bulletin, 1967, 68, 29-46.
- Reed, P. J., and Jackson, D. N. Clinical judgment of psychopathology: A model for inferential accuracy. Journal of Abnormal Psychology, 1975, 54, 475-482.
- Rodin, M. J. The effect of behavioral context on information selection and differential accuracy in a person perception task. Journal of Social Psychology, 1975, 97, 83-94.
- Rogers, C. R. The necessary and sufficient conditions of therapeutic personality change. Journal of Consulting Psychology, 1957, 21, 95-103.
- Rosenhan, D. L. The contextual nature of psychodiagnosis. Journal of Abnormal Psychology, 1975, 84, 462-474.
- Shrauger, S., and Altrocchi, J. The personality of the perceiver as a factor in person perception. Psychological Bulletin, 1964, 62, 289-308.



- Siegel, S. Nonparametric statistics for the behavioral sciences. New York: McGraw-Hill, 1956.
- Slovic, P., and Lichtenstein, S. Comparison of Bayesian and regression approaches to the study of information processing. Organizational Behavior and Human Performance, 1971, 6, 649-744.
- Slovic, P., Fischhoff, B., and Lichtenstein, S. Behavioral decision theory. Annual Review of Psychology, 1977, 28, 1-39.
- Soskin, W. F. Bias in postdiction from projective tests. Journal of Abnormal and Social Psychology, 1954, 49, 69-74.
- Soskin, W. F. Frames of reference in personality assessment. Journal of Clinical Psychology, 1954, 10, 107-114.
- Soskin, W. F. Influence of four types of data on diagnostic conceptualization in psychological testing. Journal of Abnormal and Social Psychology, 1959, 58, 69-78.
- Spitzer, R. L. On pseudoscience in science, logic in remission, and psychiatric diagnosis: A critique of Rosenhan's "On being sane in insane places." Journal of Abnormal Psychology, 1975, 84, 442-452.

- Stelmachers, Z. T., and McHugh, R. B. Contribution of stereotyped and individualized information to predictive accuracy. Journal of Consulting Psychology, 1964, 28, 234-242.
- Taft, R. Accuracy of empathic judgments of acquaintances and strangers. Journal of Personality and Social Psychology, 1966, 3, 600-604.
- Truax, C., and Carkhuff, R. R. Toward effective counseling and psychotherapy. Chicago: Aldine, 1967.
- Tversky, A., and Kahneman, D. Judgment under uncertainty: Heuristics and biases. Science, 1974, 185, 1124-1131.
- Weiner, B. "On being sane in insane places": A process (attributional) analysis and critique. Journal of Abnormal Psychology, 1975, 84, 433-441.

APPENDIX A  
PROSPECTUS

## PROSPECTUS

## CHAPTER I

## INTRODUCTION AND STATEMENT OF THE PROBLEM

The present study is designed to investigate several meaningful aspects of clinical judgment. An examination will be made of the manner in which different types of information available to the clinician affect the accuracy of his predictive and evaluative judgments and his confidence in these judgments. The following section of the paper will relate these variables (type of information, type of judgment, and confidence) to previous relevant research, and a rationale for the current investigation will be developed.

Differing Types of Information

First, attention is directed to the effects of differing types, or amounts, of information upon the counselor's perceptions of the client. The rationale for administration of psychological tests to clients would appear to be that testing provides additional information with which to better understand the client in one or both of two important ways: (a) from an "external" perspective in terms of the "actual" personality dynamics of the client as it has import for making clinical judgments (e.g., therapy approach, hospitalization, court-related decisions, etc.), and (b) from an

"internal" reference point in terms of understanding how the client sees himself or herself. This latter type of understanding has been conceived by some to relate to the process of "empathy" in therapy which Rogers (1957) and Carkhuff (1967) have envisioned as of such great importance to positive process change. Lesser (1961) found "predictive empathy" (the ability of the counselor to predict accurately certain responses of the client, such as his self-rating on personality traits) to be significantly related to client-perceived empathy. Carr (1974) concluded from the results of his study that where client and therapist were similar in precision of conceptual differentiation among people (differential predictive accuracy) and where client and therapist commonly endorsed a large number of self-descriptive concepts, client/therapist communication was more successful and resulted in the client's perception of positive treatment effects. If testing contributes to our understanding of the client's self-perception, then, it indeed is of positive value.

On the other hand, many therapists question the value of psychological testing, objecting that far from aiding in understanding of the client, diagnostic material actually impedes understanding of him through the creation of response sets, resulting in "stereotyping" and

"pigeon-holing" of clients. This notion would appear to receive some support from Soskin's (1954) finding that data from projective tests predisposed judges to overestimate maladjustment of a client. If test material does result in certain biases in clinical judgment, the clinician should become aware of these biases, gain an understanding of how they operate, and correct for their influence upon his judgmental processes.

While a relatively large number of studies in the area of interpersonal or social perception (primarily done in the 1950's to early 1960's) have dealt with the effect of differing types of information upon the judge/counselor's ability to make accurate predictions about the subject/client, results of these studies have been equivocal.

Hamlin (1954) reviewed a series of ten studies of judgment utilizing projective material (Albee and Hamlin, 1949; Albee and Hamlin, 1950; Bialick, 1950; Bialick and Hamlin, 1954; Cummings, Hamlin, Albee and Leland, 1950; Hamlin, Berger, and Cummings, 1952; Hamlin and Newton, 1952; Newton, 1954; Newton and Goodman, 1952; and Whitmyre, 1953). He concluded that the accuracy of such judgments is related to the simplicity or complexity of the material judged and to the conditions provided for making the judgments. He found that small, mechanically scored or judged units of

information tend to result in poor predictive accuracy, as do overly large complex units, unless conditions for judgment approach actual clinical procedures.

In a 1957 study, Borke and Fiske found no significant differences in the accuracy of judges' predictions about a client whether they interviewed the client directly, observed the client from behind a one-way mirror, listened to an audio recording of the interview, or read a verbatim transcript. Borke and Fiske interpret their findings to mean that clinicians in their study relied primarily on content cues when making their predictions. Segel (1952) and Giedt (1955) reported similar findings. It should be noted, however, as Borke and Fiske point out, that the heavy reliance on content cues found in these studies is interpretable in terms of demand characteristics of the experiment, as judges were attempting to predict a subject's performance on a verbal task, such as a Q-sort. They suggest that if judges had been asked to make different kinds of inferences, such as those related to affect of the client, they might have utilized other (non-content) cues to a far greater extent. This question of the effect of different kinds of information upon different kinds of judgments will be addressed in the current study.

The effect of differing types of information upon

judges' perceptions was assessed in a series of studies by Soskin (1954a; 1954b; 1959). In the 1954a study, Soskin found that trained clinicians initially had a more unfavorable stereotype of an experimental subject than did graduate students provided with minimal information about the subject. However, it was also found that providing the graduate students with projective test information (TAT) resulted in an increase in "unfavorableness" of their judgments. In addition, training in the use of the TAT did not result in increased accuracy of judgments by the students. Soskin concludes that "projective tests predispose (clinicians) toward an overestimation of maladjustive trends in postdiction situations." Soskin himself calls attention to the fact that clinicians are infrequently called upon to postdict specific situational behaviors of the types employed in this study, and points out that the information available from the projective devices may not have been relevant to the postdictions required of the judges.

In a second 1954 study, Soskin investigated frames of reference in personality assessment. He studied a number of clinical judgments made in the assessment of first-year graduate students who had been accepted for admission to APA-approved clinical training programs. He concluded that judges markedly differed in their descriptions of the



experimental population as a function of the type of information with which the judges were provided, the relationship between judge and subject, and the purpose for which judgments were being made. The extreme complexity of the experimental design and lack of adequate experimental control in this study, however, make reliable interpretation of the findings extremely difficult.

The 1959 Soskin study also is subject to a number of methodological criticisms, predominantly related to questionable relevance of available cue configurations to the judgments required. Judges were provided with either 1) biographical data; 2) a Rorschach protocol; 3) data from a battery of objective and projective tests; or 4) the opportunity to view the subject in a number of role-play situations. Results indicated that none of the three types of critical data improved accuracy scores beyond the level achieved by study of biographical facts alone. After studying the Rorschach protocol, in fact, Rorschach judges showed a significant decrease in accuracy on postdiction items compared to their "stereotype" predictions (those based on minimal information). As in previous studies, use of Rorschach information also resulted in predictions more characteristic of "maladjustment."

Of interest is the fact that significant group

differences in perception of the subject by the role-play and the Rorschach judges existed at the initial stage (stereotype stage) of the experiment when they were supplied only with biographical data and before exposure to the respective experimental conditions (i.e., viewing role-plays or studying Rorschach data). Soskin speculates that this difference may be due to different "response sets" operating in the two groups based upon knowledge of the experimental procedures to follow for their particular groups. That is, "mere knowledge that their next task would demand further blind appraisal based on the Rorschach alone (might have) heightened the sensitivity or the responsiveness of Rorschach judges to adverse signs in the biographical data." This difficulty can be circumvented by utilizing a separate "Stereotype Group" for obtaining stereotype base lines rather than asking each judge in the experimental groups to serve as his own control. This procedure will be followed in the present investigation.

Kostlan (1954) provided judges with different types of information and observed the effect upon predictive accuracy. Experimental conditions were: 1) Rorschach missing, 2) MMPI missing, 3) Sentence Completion Test missing, 4) Social Case History missing, or 5) minimal data only. Kostlan found that predictions based upon minimal data

were better than chance and that Social Case Histories were required to increase accuracy above that achieved by minimal data alone. Batteries which included both the MMPI and the Social Case History provided the greatest accuracy of prediction. Patterson (1955) criticized the Kostlan article on a number of statistical points, but Kostlan (1955) defended the conclusions of his earlier study.

In a study designed in part to improve upon the methodology in the Kostlan (1954) study, Horowitz (1962) also examined clinical judgment based upon projective test protocols. Subjects were required to predict responses to a 64-item Q-sort chosen to represent personality areas commonly covered in psychological test reports. It should be noted that the criterion to be predicted here was not the response of the client himself or herself on the measure as in the typical study, but the Q-sort description of the client made by a therapist who had seen the patient for a minimum of twenty hours in psychotherapy. The three prediction sources evaluated were: 1) brief biographical information concerning the client, 2) biographical information plus a test battery, 3) base rate descriptions obtained by having a group of 15 therapists perform the Q-sort on their respective 15 clients and construction of an "average client" (or base rate) Q-sort from this. Predictions on

the basis of biographical information alone were made by both a group of trained clinicians and a group of naive judges. All other judgments were made only by trained clinicians.

Horowitz found that base rate predictions were as accurate as those based upon projective data. (This is, of course, interpreted as support for the efficacy of "statistical prediction.") However, access to test information did increase the accuracy of clinicians' predictions beyond that achieved through biographical information alone. Indeed, without the test information clinicians' predictions were significantly less accurate than the base rate predictions. Horowitz concludes that projective data were useful only insofar as they enabled the clinicians to sub-categorize patients. She points to the need for further investigations into the effect of information upon clinical judgment, and suggests that: "It would be valuable to learn exactly what information the clinician does use, as well as what sort of inferences he makes from it and how accurate these inferences are."

Luft (1950) made a similar suggestion in calling for "post mortems" in prediction studies in an attempt to verify implicit hypotheses and "thus (improve) the basis for clinical theory and judgment." In line with this idea,

Kleinmuntz (1963a, 1963b, 1963c) had a clinician "think aloud" into a tape recorder as he made judgments about the adjustment of college students on the basis of their MMPI profiles, and then used these reports to construct a computer program modelling the clinicians' judgmental processes. Other researchers (Oskamp, 1962) have requested judges, upon completion of their predictions, to write out a brief description of the manner in which they made their inferences.

A study by Stelmachers and McHugh (1964) investigated the effect of seven types of information upon predictive accuracy. Four of the information types were of a stereotype nature: 1) age and sex, 2) education and occupation, 3) whether the subject was a "well-adjusted normal," a "person with long-standing chronic illnesses," a "psychiatric patient," or a "delinquent," and 4) a combined input consisting of the sum of the above three stereotype inputs. The remaining three types of information were comprised of more individualized data: 1) free description, in which intimate acquaintances of the subject wrote short personal descriptions about him, 2) trait cluster data, in which the intimate acquaintances were asked to select from a set of adjectives those five most descriptive of the subject, and 3) total input, consisting of a Biographical Questionnaire, Interest and Activities Questionnaire, Sentence Completion

Test, and Case History Material. Upon the basis of selected inputs, clinicians and nursing students made predictions of the responses of four subjects on certain MMPI items and bipolar traits. Predictive accuracy was found to be greatest based upon the total input information described above, although it was significantly superior only to the "Age and Sex Stereotype," and the "Free Description" inputs. In fact, the "Combined Stereotype" input resulted in next greatest accuracy and was not statistically significantly inferior to the "Total Input." Stelmachers and McHugh conclude from their results that predictive accuracy depended little upon the type of information upon which the judgments were based. They interpreted achieved predictive accuracy in terms of reliance upon response sets of "Social Desirability," "Normality," and "Assumed Similarity." They also determined base-rate predictors for relevant populations for the subjects, and found this base rate data to result in predictive accuracy equal to that based upon the best information for all but the most "deviant" subject (that subject most atypical of the base rate populations), who was predicted more accurately by clinicians with access to test information. It should be noted, however, that in neither case (base rate data or clinical predictions for the atypical subject) was accuracy significantly above chance level.

Stelmachers and McHugh view their results as lending support to the relative efficacy of "statistical prediction" methods, but state that it seems to be "worthwhile to utilize extensive individualized information in most cases if this information is fairly objective, and largely unmediated, systematically collected, and partially produced by the subject himself..."

In two related studies, Hjelle (1968) and Taft (1966) investigated the accuracy of judgments about subjects as a function of the judges' familiarity with the subjects. It may be assumed that increased familiarity with an individual affords one quantitatively more information and a greater variety of types of information about the individual. Results of both studies indicated that predictive accuracy increased with increasing familiarity, although judgments of all subjects were significantly more accurate than would be expected by chance. That one can more accurately judge someone with whom he is more familiar hardly seems a startling finding, but Taft (1966) points to two factors which might mitigate against increased accuracy of prediction with increased familiarity. First, in his 1959 study, Taft found that increasing amount of information beyond some optimal point may actually interfere with efficacy of use

of existing information and lead to decrease in predictive accuracy. This finding seems to parallel that of Hamlin (1954), discussed above, concerning optimum complexity of projective materials to be judged. Taft also points to the fact that increased familiarity with an individual might result in certain emotional biases which could interact with familiarity and possibly "reduce the contribution which familiarity might make to the accuracy of judgments." This was not found to be the case, however. Judges appeared to be able to use the greater information associated with greater familiarity with the subjects to make more accurate judgments about them.

In a more recent study, Rodin (1975) studied the predictive accuracy of judges who were allowed to self-select their information. In her study, Rodin also calculated "Information Coefficients" for the available pieces of information which indicated "objective information value" of an item. Rodin speculated that failure of a number of earlier studies to find evidence for increased predictive accuracy as a function of increased information may have been due to the lack of relevance of information given the judges to the judgments required of them. She hypothesized that allowing judges to choose their own information according to individual information preferences would result in



their being able to make more accurate judgments. Results of the two experiments she reported support this hypothesis. She found that when allowed to self-select information deemed helpful to them, judges were able to predict with greater accuracy than a stereotype group. She also found that predictive accuracy was not correlated with the "objective information value" of the information items chosen. Judges choosing that information objectively rated "best" were less accurate than judges with objectively "inferior" information. She concludes that "the value of information to individual judges is not well-described by objective information value measures." While Rodin's methodology results in improved cue-criterion relevance, the judgment tasks involved are of questionable relevance to those required in clinical settings.

Few generalizations relating to the effects of differing types of information upon predictive accuracy can be made from the findings reviewed above. The differences among studies in terms of judgments required, criteria to be predicted, and other experimental procedures employed were rarely systematic and make direct comparison of results extremely difficult. However, one relatively consistent finding seems to be that certain types of information can result in response biases in the perception of the client.

Another finding which seems to recur is that predictive accuracy is about as good (or sometimes better) when responses are based upon minimal or "stereotype" information as when based upon more complex information, such as test data. However, the fact that Rodin (1975) did obtain increased predictive accuracy with improved cue-criterion relationships supports her contention that much of the past failure to find increased accuracy with increased information may be a function of the irrelevance of the information to the judgments required. It is the feeling of the present author that the one thing which is strikingly apparent from a review of the foregoing studies is that most of them employ questionable cue-criterion relationships and lack relevance to the judgments actually required of the practicing clinicians.

Indeed, one is almost tempted to ask the question, "Where is the 'Person' in person-perception?" Most studies requesting judges to make judgments concerning a subject have not even allowed the judges to view the individual about whom they are to make decisions, but have asked for judgments based solely on diagnostic or descriptive information. Shrauger and Altrocchi (1964) comment upon this difficulty:

"...an important source of individual variability in the judgment process may well be differences in cues chosen as bases for judgments. Selection can

occur in studies in which the person judged is not actually observed by the judges, but the data about him are presented in brief vignettes or descriptions of his personal attributes."

As mentioned earlier, those studies which have provided for viewing of the subject/client by the judge/counselor have suffered from what would appear on the whole to be highly tenuous cue-criterion relationships. That is, the relevance of the information given to the judgments required has been questionable. For example, judges have been asked to view films of interviews with subjects (not therapy clients) made at shopping centers (Cline, 1955; Cline and Richards, 1958; Cline and Richards, 1960), and then asked to fill out standard psychological test protocols as they believed the targets had done. Rodin (1975) justifiably criticized the cue-criterion relationship in the Cline and Richards study (1960), but herself employs judgment tasks of dubious clinical relevance and does not allow for viewing of the individual to be judged. One may question why, given the opportunity to watch films of individuals at a supermarket talking about various topics, judges should be expected to accurately predict their responses to psychological tests. Another example of poor cue-criterion selection is the Soskin (1959) study in which judges are asked to postdict

specific behaviors of an individual based on either differing types of descriptive or diagnostic material or (but not in addition to) on observation of the subject (not a client) role-playing a number of vignettes in which he played such roles as a "new young superintendent of a boy's correctional school discussing school policy with veteran staff members, etc.". The subject, however, was not seen by judges in the "role" of himself. Not only do such contrived role-plays appear clinically irrelevant, but the items which judges were asked to postdict seem to be equally far-removed from considerations actually encountered in a clinical setting. Consider, for example, a typical postdiction item from the Soskin (1959) study involving a subject "David":

On one occasion two foreign students who had completed their studies and were leaving the school arranged a small party for their friends. Some mild spirits were available to those interested, although as nearly as can be ascertained, nothing stronger than wine was served. In this school, opinion is divided on the matter of social drinking. Some students object strongly, others are more tolerant of moderate social drinking. Members of both groups were present at the party.

- a) At first, David was a little uncomfortable, but he joined in, swapped stories with the rest and drank.
- b) After a couple of drinks, he began a rather unconvincing act of being slightly intoxicated, began to slur his speech, walk uncertainly, etc.
- c) He conspicuously refused a proffered drink, and shortly after reprimanded the host for serving liquor to such a gathering.
- d) He remained throughout the party, but made careful note of the persons who accepted drinks and those who told questionable stories and subsequently dropped them from his list of social acquaintances.

It would appear that such predictions are rarely asked of practicing clinicians.

In order to assess experimentally the effects of differing types of information upon the counselor's judgments about a client, the present investigation will include the following information conditions:

- 1) A Stereotype Condition, in which counselors are provided with minimal information about an actual client, such as age, sex, educational or occupational status, and the fact that he/she is a

is a counseling client.

- 2) A Stereotype-Video Condition, in which counselors view an actual video-taped counseling session involving the client in addition to receiving the minimal information outlined above, and
- 3) A Stereotype-Video-Diagnostics Condition, in which counselors receive the minimal information, view the videotape, and additionally are provided with diagnostic material obtained from the client.

Specific hypotheses relating to the information conditions will be treated at another point in this paper.

In an effort to make research in the area more clinically relevant, as has been strongly urged by a number of authors (Horowitz, 1962; Kostlan, 1964; Oskamp, 1962; Rodin, 1975; Shrauger and Altrocchi, 1964), and to improve experimental cue-criterion relationships, an attempt has been made in the present study to ask judges to make clinically-relevant judgments on the basis of clinically-relevant material, both through use of diagnostic material and observation of an actual counseling session with an actual client. The clinically-related judgment tasks to be employed will be discussed in detail in the appropriate section of the paper, but it should be mentioned at this point that these

tasks were constructed to more closely approximate the kinds of judgments or conceptualizations which clinicians routinely make in clinical practice.

#### Confidence

A second major area of consideration in the current investigation involves the clinician's confidence in judgments. A number of studies have been concerned with this variable.

Forer and Tolman (1952) asked clinical psychologists to rate the items of a sentence completion test (blank form) as to how potentially clinically useful they felt each to be. They also asked the judges to rate the confidence they felt in each of their judgments. They found that individual clinicians varied from one another in general level of confidence in their judgments in a consistent, systematic way. Judges who were generally highly confident in their judgments were found to be most confident when in disagreement rather than agreement with the group. Highly confident judges also tended to be more extreme in their judgments.

Block and Petersen (1955) measured the confidence of subjects in psychophysical judgments and concluded that confidence in "such judgments is correlated with personality characteristics of the judges." They found overly confident subjects to be more rigid and dogmatic, while overly cautious

subjects appeared more introspective and inclined toward self-abasement. Individuals with more realistic confidence in their judgments, on the other hand, seemed to be more self-reliant and socially perceptive.

Little (1961) also found evidence for the existence of a stable "level of general confidence" characteristic of individuals. He asked 48 experienced clinical psychologists to make certain judgments about a subject based on a test protocol (either the Rorschach, TAT, Make-a-Picture-Story Test, or MMPI). Twelve test subjects from four diagnostic categories were employed. Re-test on the task was performed in 4-6 weeks after the original test. Little obtained:

- 1) estimates of generalized confidence levels of each judge,
- 2) measures of reliability of decision (proportion of items left unchanged from test to retest, and 3) correlations between confidence in a judgment and whether that item changed from test to re-test.

Little reported the following findings with regard to confidence and reliability of judgments: first, the relative ranking of judges in generalized confidence was quite stable across judgments. Secondly, there was no significant relationship between average confidence of judges and reliability of their judgments. Last, decisions which remain unchanged from test to re-test were those in which the judges had greater original confidence.



Analysis revealed no significant mean differences in generalized confidence levels among judges working with different types of protocols or with patients of different diagnoses.

On the basis of his findings, Little calls for a distinction between "confidence" and "certainty." "Confidence" would thus be defined as a judge's "generalized confidence in his decisions," whereas "certainty" would refer to the "momentary certainty or uncertainty (of the judge) produced by the difficulty of the particular task for him." He concludes that:

. . . there is substantial evidence that confidence is unrelated to either the accuracy or reliability of judgments, but is quite a stable personality characteristic, whereas certainty is quite positively related to reliability of judgments and is much more determined by the immediate difficulty of judgment.

Findings supporting the view that confidence is a general personality trait substantially uncorrelated with validity or reliability of decisions are also reported by Trow (1923), Seward (1928), D. M. Johnson (1939, 1955) and L. C. Johnson (1957).

Wolff (1955) studied the generality of two types of "certainty" ("subjective" and "behavioral") and their relationship to manifest anxiety. Subjective certainty was

defined as "the degree of conviction experienced consciously by any given subject." Behavioral certainty referred to the "amount of information requested by subjects before making choices." Sixty college women performed a series of tasks scored for either behavioral or subjective certainty: 1) a paired-associates learning task, 2) a story completion task, 3) an angle estimation task, 4) a question-answering task, 5) an adjective check-list task, and 6) completion of the booklet form of the MMPI.

Findings indicated little generality of uncertainty across the highly varied and unusual tasks employed. Little support was found for a relationship between manifest anxiety and uncertainty as measured.

Goldberg (1959) examined the effect of clinical experience upon accuracy of diagnostic judgment and judges' confidence in such judgments. He asked staff psychologists, psychology trainees, and nonprofessional judges (secretaries) to attempt to diagnose organic brain damage on the basis of Bender Gestalt test protocols. The criterion of accuracy was independent diagnosis by a competent neurological team. He found that the three groups of judges did not differ in their ability to make such diagnoses accurately. (an "expert" in use of the Bender was able to make more accurate

diagnoses than the clinicians.) However, differences in confidence in judgments were found as a function of level of experience of the judges. Goldberg found that his most experienced judges were least confident in their judgments, whereas untrained judges were found to be the highest in confidence. He concluded that ". . . in such ambiguous decision situations increased experience may have the effect of increasing cautious skepticism."

Oskamp (1962) also investigated accuracy of and confidence in predictions as a function of clinical experience and training. Judges were 44 clinical psychologists (about half staff members and half trainees) and 28 undergraduate psychology majors. Judges were asked to decide on the basis of MMPI profiles whether a patient was psychiatrically or medically hospitalized. Oskamp found that although clinicians were more accurate in their judgments than inexperienced judges, the magnitude of the difference in accuracy of the two groups was not great enough to be significant. As did Goldberg (1959), Oskamp found that confidence was significantly decreased by general experience and that additionally "appropriateness of confidence" (correlation between confidence in a judgment and accuracy of that judgment) increased significantly with increasing experience. This prompted Oskamp to venture the following definition of

clinical expertise: "In short, the expert must know when he is apt to be right and when he is more apt to be wrong."

In a study which is especially relevant to the present experiment in terms of the variables examined, Oskamp (1965) studied the effects of differing amounts of information upon judges' confidence and accuracy. Oskamp gave his judges four sets of cumulatively increasing amounts of information about a case subject. Stage 1 information included only minimal biographical data. Stage 2 added material about the subject's childhood through twelve years of age. Stage 3 information covered the subject's high school and college years, while Stage 4 material consisted of history of his adult life. Upon the basis of the information currently available to them, judges were asked to attempt to answer correctly multiple choice items concerning the case subjects and to indicate their confidence in each prediction. Oskamp found that as amount of information increased, confidence soared while accuracy did not increase significantly. It should be noted once again, however, that judgments were made only on the basis of case-summary material, without the judges ever having viewed the subject of their judgments.

The current study allows for more clinically-meaningful examination of the question of the effect of differing amounts or types of information upon the judge's confidence through

allowing the judge to observe in an actual clinical session the individual to be clinically judged. Confidence is also related to type of judgment, as will be seen in the following section of the paper.

#### Prediction vs. Evaluation

In the present study, the question of whether clinicians are equally confident in their predictions and their evaluations is also considered. The distinction between "prediction" and "evaluation" is discussed by Kahneman and Tversky (1972, 1973) and Tversky and Kahneman (1971, 1974) as it relates to judgmental heuristics. These authors have found that when people are asked to make predictions or judgments under conditions of uncertainty, they do not appear to "follow the calculus of chance or the statistical theory of prediction." Instead they seem to rely on a few heuristics (i.e., representativeness, availability, adjustment from an anchor) which "sometimes yields reasonable judgments and sometimes lead to severe and systematic errors." Kahneman and Tversky (1973) reported a series of experiments the results of which illustrate the reliance of judges upon the "representativeness heuristic" and have clear implications concerning clinical judgments.

In one study a group of college students (base-rate group) serving as judges were asked to guess the percentage

of U.S. graduate students enrolled in each of nine areas of specialization. A second group of subjects (similarity group) were given a brief personality sketch of an individual and then asked to indicate (by ranking) how similar they believed the individual described to be to the average graduate student in the nine fields of specialization. A third group of subjects (prediction group) were given the personality sketch along with additional information designed to call into question the validity and reliability of the information contained in the sketch (i.e., the fact that the sketch was written several years previously and that it was based solely on projective tests). Judges in the prediction group were then also asked to rank the nine fields of specialization in order of the likelihood that the individual described in the sketch was a graduate student in each of the fields.

They found that the judgments of likelihood were virtually identical to the judgments of similarity, but differed significantly from the base-rate estimates. Following the prediction task, judges were asked to estimate the percentage of correct choices which could be achieved with several types of information (a confidence measure). The judges' estimate was only 23 percent for predictions based on projective tests as compared to 53 percent, for example, for predictions based on "high school seniors' reports of their interests and

plans." It would seem that the subjects (graduate students in psychology) had little confidence in judgments based on projective tests. As Kahneman and Tversky point out, however: "Nevertheless the graduate students relied on a description derived from such tests and ignored the base rates." Since their expected accuracy was low, judges should have weighted base-rate data heavily according to the rules of statistical prediction, but this was not done. To quote Kahneman and Tversky:

In general, three types of information are relevant to statistical prediction: (a) prior or background information (e.g., base rates of fields of specialty); (b) specific evidence concerning the individual case (e.g., the personality sketch of the individual); (c) the expected accuracy of prediction (e.g., the estimated probability of hits). A fundamental rule of statistical prediction is that expected accuracy controls the relative weights assigned to specific evidence and to prior information. When expected accuracy decreases, predictions should become more regressive, that is, closer to the expectations based on prior information. In the case of Tom W., expected accuracy was low, and prior probabilities

should have been weighted heavily. Instead, our subjects predicted by representativeness, that is, they ordered outcomes by their similarity to the specific evidence, with no regard for prior probabilities.

In another study, Kahneman and Tversky further explored the effect of prior versus individuating evidence upon predictions. They found that when judges were given no information about an individual, they based their predictions about him on base-rate data. However, when given the least bit of worthless individuating evidence concerning the individual, judges ignored base-rate data altogether and made extreme predictions.

Kahneman and Tversky also found that subjects' judgments were equally extreme when they were asked to predict a remote objective criterion on the basis of sketchy information as when they were asked merely to evaluate the information itself. Subjects were given a description of a college freshman in which he was described by a counselor in a very positive way. An "evaluation" group was asked to evaluate the descriptions by estimating ". . . the percentage of students in the entire class whose descriptions indicate a higher intellectual ability." "Prediction" groups were given the same descriptions and asked ". . . to predict the



GPA achieved by the student at the end of his freshman year and his class standing in percentiles."

That they were asked to make predictions on the basis of highly incomplete data of questionable relevance to the prediction did not cause judges to normalize their predictions or use base rate data, as would be prudent. Kahneman and Tversky explain their findings on the basis of the "representativeness heuristic." That is, judges will make that prediction which seems to be "representative" of the data no matter how incomplete the information or how irrelevant the information may be as a basis for prediction. Holzworth and Doherty (1974) also report findings that show subjects' judgments to be based on the "representative heuristic."

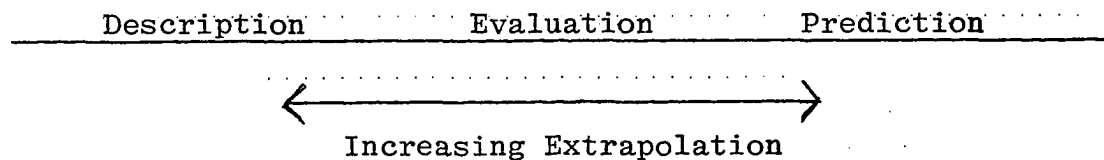
It is felt that the distinction between "evaluating" an input and "predicting" an outcome is highly relevant to the clinician, who is frequently asked to evaluate a client and in addition make some prediction about the client's future performance.

A hypothetical, simplified example of the distinction between "evaluative" and "predictive" judgments, applied to a clinical situation, might be:

A clinician has available the information that a particular client obtained an F+% of 25 on the Rorschach. The clinician is asked:

- 1) To evaluate how that  $F+\%$  compares to that of the average member of some stated population (e.g., normals, psychotics, etc.).
- 2) To predict how effective a parent the client will be to a three-year-old child or how likely the client is to commit a violent crime, etc.

The current author feels that the distinction between evaluations and predictions made by Kahneman and Tversky might be conceived of in terms of a judgment continuum ranging from "descriptive" on the one end, through some intermediate point "evaluative" to "predictive" at the opposite extreme. As one moves from left to right across the continuum, the degree of extrapolation required from the data at hand increases. That is, the judgment required becomes less and less closely tied to the actual cue configuration and greater "leaps of inference" are required.



In their 1955 paper on methodological problems in interpersonal perception, Gage and Cronbach make a similar analysis:

"Understanding another person may be regarded as

having two stages, which suggest two continua for classifying investigations. First, the judge must take in information about the Other, perhaps by observing him, perhaps by dealing with him over a period of time; the first continuum therefore deals with the degree of acquaintance of the Judge with the Other. Second, the Judge must interpret the information in order to arrive at a predictive statement; the second continuum therefore deals with the degree of extrapolation or inference required between Input and Outtake. An experiment may be designed to make great demands on the intake process (little acquaintance) or the interpretive process (much extrapolation) or both, or neither."

As also pointed out by Gage and Cronbach, Meehl (1954) makes a similar consideration in discussing two possible meanings of the term "clinical intuition." First, it can refer to the situation in which the clinician cannot be articulate about the evidence for his diagnosis. Secondly, it may allude to the instance in which he cannot "show in what manner a particular hypothesis was arrived at from the stated evidence."

The distinction made by Meehl (1960) of three classes

of functions performed by the psychodiagnostician appears to deal in part with the degree of extrapolation from available information necessitated by each. He lists the following diagnostic functions:

. . . formal diagnosis (the attachment of a nosological label); prognosis (including "spontaneous" recoverability, therapy-stayability, recidivism, response to therapy, indications for differential treatment), and personality assessment other than diagnosis or prognosis. This last may be divided somewhat arbitrarily into phenotypic and genotypic: the former being the descriptive or surface features of the patient's behavior including his social impact; the latter covering personality structure and dynamics, and basic parameters of a constitutional sort.

Meehl's concept of "prognosis" would appear to correspond to judgments characterized as "predictive" on the above-described continuum, while "phenotypic" and "genotypic" resemble "descriptive" and "evaluative," respectively.

According to this conceptualization, one would conceive of the previous example in a slightly different way:

A clinician has available the information that a particular client obtained an  $F+\%$  of 25 on the

Rorschach. The clinician is asked:

1) Descriptive

To compare that F+% to the average F+% of some stated population (e.g., normals, psychotics, etc.)

2) Evaluative

To characterize the individual on a continuum of emotional disturbance or ability to test reality appropriately (e.g., "The individual is severely emotionally disturbed." Or, "The individual evidences extremely poor reality testing.")

3) Predictive

To predict how effective a parent the client will be to a three-year-old child or how likely the client is to commit a violent crime, etc.

On this continuum, what Kahneman and Tversky term "evaluation" is conceived to be more "descriptive" in nature. The predictive class of judgment is obviously the most risky based upon such limited information and the prediction should be normalized, or not made at all, because of this. It should be noted that such a prediction is risky not only in that it involves very large inferences on the basis of

questionable cue-criterion relationships, but it is also risky in terms of practical consequences to the individual involved. However, in line with the representativeness heuristic, one would expect the clinician to predict in agreement with what seems "representative" of the data (e.g.,  $F+%=25=$ pathological). Thus, he would be anticipated to make predictions consistent with pathology, merely because they "fit" that data at hand no matter how insufficient the data may be to justify such predictions.

Cronbach (1955) referred to exactly this same phenomenon when he stated: "Evidently the fault of the clinician is in too little central tendency of judgment. . ." and made this cautionary injunction:

"If the judge is forced to base his judgment on inadequate cues or if the available personality theory and situational knowledge do not permit trustworthy inference, then he should treat people as if they were very nearly alike."

Cronbach also stresses the need for investigation in this area:

"Systematic errors such as overoptimism and over-differentiation may be corrected fairly easily. It is important for studies of clinical judgment to measure these errors as separate components,

and for clinicians to train themselves to avoid these errors."

In an extremely interesting and stimulating paper, Einhorn and Hogarth (1978) discuss confidence in judgments which are subject to heuristic biases and relate such apparently unwarranted confidence to three main factors:

- 1) "lack of search for and use of disconfirming evidence,"
- 2) "lack of awareness of environmental effects on outcomes,"
- and 3) "the use of unaided memory for coding, storing, and retrieving outcome information."

They make the following suggestions for alleviation of these problems: first, "formal instruction in experimental design, and teaching the logic of control groups and baseline predictions," second, the use of a model of the environment as per Hammond (1971) to help gain awareness of environmental effects on outcomes; and third, keeping a "box score" tabulation of judgments and outcomes. This latter suggestion was made previously by Goldberg (1968) in his discussion of the learning of clinical inference. To quote Goldberg:

"For learning to occur, some systematic feedback regarding the accuracy of the judgmental response must be linked to the particular cue configuration which led the clinician to make that judgment. But in clinical practice feedback is virtually

non-existent, and in the relatively rare cases where feedback does occur the long interval of time between the prediction and the feedback serves to ensure that the initial cue configuration leading to the prediction has disappeared from the clinician's memory."

Goldberg postulated that three conditions must exist if complex clinical inferences are to be learned: 1) some kind of feedback is a "necessary, though not necessarily a sufficient condition for learning to occur;" 2) the order of cases must be rearranged so that clinical hypotheses can be "immediately verified or discounted;" and 3) it may be necessary "to tally the accuracy of one's hypotheses" by pencil-and-paper means. These suggestions would certainly seem to be of interest in the training of future clinicians.



## CHAPTER II

## DESIGN OF STUDY

Subjects. Subjects will consist of 30 professional counselor/therapists with experience in individual therapy and in diagnostic testing. Each clinician will be required to provide demographic data relating to professional experience, as in Oskamp (1962), and theoretical orientation. Subjects will be randomly assigned in equal numbers to the following conditions: 1) Stereotype Condition, 2) Stereotype-Video Condition, and 3) Stereotype-Video-Diagnostics Condition.

Experimental Conditions and Procedure. Those subjects in the Stereotype Condition will be given minimal information about an actual counseling client, such as age, sex, educational or occupational status, and the fact that he/she is a counseling client. In addition to receiving the minimal information about the client, subjects in the Video Only Condition will view an actual video-taped counseling session involving the client. Subjects in a Video-Diagnostics Condition will receive the minimal information, view the videotape, and additionally be provided with diagnostic material obtained from the client. Viewing of the videotape and study of the diagnostic material will be counterbalanced for order of presentation. Diagnostic data will be presented in protocol form in order to approximate actual clinical conditions.

Tests to be included in the battery are: 1) Wechsler Adult Intelligence Scale, 2) Rorschach Inkblot Test, 3) Thematic Apperception Test, 4) MMPI, and 5) Human Figure Drawings. This battery is chosen to be representative of a typical test battery commonly employed by many of the subjects. Judges will be allowed to follow their customary procedures for interpretation of the diagnostic material, including having sufficient time for study of the data, as suggested by Hamlin (1954).

The client will also be asked to rate himself/herself on five-point scales on a number of personality dimensions (Trait Rating Instrument). These dimensions were chosen from among those commonly used by psychologists to describe personality characteristics and frequently employed as assessment dimensions on widely-used personality tests, and include such terms as "Orderly," "Anxious," and "Independent." Terms included will be counterbalanced so that an equal number of favorable and unfavorable terms are rated. In a small pilot study, the Trait Rating Instrument detected differences in predictive accuracy for judgments concerning more-well-known and less-well-known Others, and between a trained Judge and untrained Judge.

Client. The client viewed will be an actual counseling client in a counseling session. The client will be chosen

to present a problem of mild to moderate proportions, such as a situational disturbance, mild to moderate depression and/or anxiety, etc. At any rate, severely disturbed clients will not be considered for presentation. All possible steps will be taken to safeguard the confidentiality of the client. The counselor who takes part in the videotaped counseling session will not otherwise participate in the experiment and will have no knowledge concerning the design or purpose of the experiment other than being told that the general purpose of the study is to gain knowledge in the area of clinical judgment. In addition, the counselor will be told, prior to taking part in the taping, that the conduct of the experiment will at no time require judgments of any kind to be made of his/her counseling style or ability. He/she will be informed that the videotape in which he/she takes part will be used merely as an "analog client" for experimental purposes.

Although the analog nature of the study removes the judges one step from the actual stimulus-object (client) which they are required to judge, it is felt that the additional control of cue configurations gained through the analog method is important enough in the current study to outweigh this disadvantage. In fact, Gage and Cronbach (1955) recommend just such a procedure in the investigation of interpersonal perceptual accuracy: "Another device is

to use a 'standard Other,' requiring everyone whose accuracy is tested to make predictions for the same individual or group."

Judgmental tasks. After having acquired the information appropriate to their respective groups (1) minimal information, 2) minimal information + viewing videotaped session, or 3) minimal information, videotape, and diagnostics), judges will be required to complete a number of judgmental tasks concerning the client about whom they received information. Some tasks will be of a "high extrapolation" ("predictive") and some of a "low extrapolation" ("evaluative") nature of each task to be performed, three trained psychologists not otherwise taking part in the experiment will sort descriptions of the tasks as to whether they should properly be considered predictive (or requiring less extrapolation from the data) or evaluative (or requiring less extrapolation from the data), as compared to a criterion example of the type reported by Kahneman and Tversky. Only those tasks sorted in the same way by all three psychologists will be retained in the experimental procedure.

1) High Extrapolation (Predictive) Tasks. Experimental subjects will be asked to rate the client on a number of personality trait dimensions as they believe the client rated himself/herself. That is, they will be asked to predict the

client's responses on the Trait Rating Instrument. Counselors will also be asked to rate themselves on the Trait Rating Instrument in order to explore the possibility that "Assumed Similarity" or "Real Similarity" between counselor and client is correlated with the counselor's predictive accuracy (Cronbach, 1955; Cronbach and Gleser, 1953; Gage and Cronbach, 1955). For each counselor, a difference score will be calculated between his predicted responses for the client and the client's actual responses, as per Cronbach (1955), Cronbach and Gleser (1953), and Gage and Cronbach (1955). This score can be thought of as representing the counselor's predictive accuracy with regard to the client on the Trait Rating Instrument. These difference scores will be submitted to a one-way analysis of variance in order to test the following null hypothesis:

$$\begin{array}{rcccl} \text{Stereotype} & = & \text{Stereotype-Video} & = & \text{Stereotype-} \\ \text{Predictive} & & \text{Predictive} & & \text{Video-} \\ \text{Accuracy} & & \text{Accuracy} & & \text{Diagnostics} \\ & & & & \text{Predictive} \\ & & & & \text{Accuracy} \end{array}$$

This hypothesis tests the effect of differing types of information upon the counselor's ability to see the client as the client sees himself. It permits one to ascertain whether different types of information result in any improvement or deficit in predictive accuracy beyond that which can be achieved by knowledge of minimal or "stereotype"

information alone.

Some previous studies have found "stereotype" information to result in more accurate predictions than more complex information regarding the client. However, these studies have not allowed for viewing of the client in an actual counseling session, as does the current study. It is predicted that this availability of highly relevant client-related cues will result in improvement in predictive accuracy of the Stereotype-Video Group as compared to the Stereotype Group (that group with only minimal information). The Stereotype-Video-Diagnostics Group will also be exposed to the client cue configuration made available to the Stereotype-Video Group, of course, but expectations about this group's performance require scrutiny.

The familiarity studies reviewed earlier (Hjelle, 1968; Taft, 1966) indicated that accuracy of prediction increased with increasing familiarity (presumably involving the availability of more information). In the current study, the Stereotype-Video-Diagnostics Group possesses the greatest amount and variety of information about the client. Therefore, we might expect that this group would be most "familiar" with the client, and hence most able to make accurate predictions concerning the client. This hypothesis might receive some support from results of the Borke and Fiske

(1957) study which indicated that clinicians relied heavily on content cues for verbal prediction tasks. Since the Stereotype-Video-Diagnostics Group of the current study is provided with the most content cues (content of session + test data), it might be speculated that they could predict most accurately the client's responses on the Trait Rating Instrument, a verbal prediction task.

On the other hand, Taft (1959) found that "beyond a certain point, more information is a handicap and may even interfere with the correct use of existing information." He speculates that in such cases, judges may give "too much weight to some data, and far too little to other, more relevant data." Perhaps this would account for the findings of Soskin (1954a, 1954b, 1959) and of Weinberg (1957) that projective data resulted in biases in perception of the individual to be judged. It may be that judges tended to give too much weight to the projective data, which certainly does appear to have been of questionable relevance to the judgments required in these studies. This is further supported by the findings of Kahneman and Tversky (1973, 1974) and others, that judges frequently tend to "overpredict" on the basis of insufficient or irrelevant data. With improved cue-criterion relevance, Rodin (1975) did find increased information about the subject to result in increased

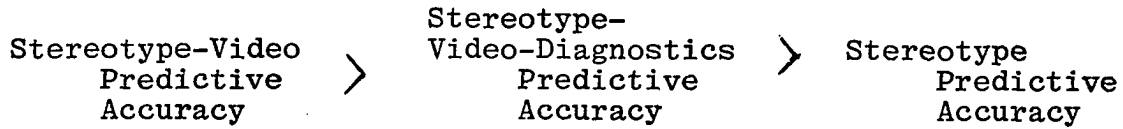
accuracy of predictions. Horowitz (1962) also found that access to test data improved predictive accuracy above stereotype accuracy; however, the criterion for her predictions was not the actual response of the client, as in the "prediction" required in the present study, but the response of the client's therapist, as in the current "evaluation" task.

It would seem that determination of the relevance of the particular information possessed by each group of judges to the specific judgment required of them would provide for formulation of the most probable hypothesis.

For the Stereotype-Video-Diagnostics Group, then, the question could become, "How relevant is the diagnostic data to prediction of the client's self-perception?" While information of the type contained in the diagnostic material may be relevant to understanding him/her from an "external" reference point (i.e., what are the "actual" personality dynamics of the client), this may have little to do with understanding his view of himself, as asked in the current task. In fact, test material may result in a "set" to "objectively evaluate" the client, rather than encouraging the judge to consider how the client sees himself, thus resulting in inferior performance of the required task.

Therefore, the following experimental hypothesis is tentatively made:





Results will be discussed in relation to this hypothesis and the literature reviewed earlier.

2) Low Extrapolation (Evaluative) Tasks. "Evaluative" tasks are those which received consensual validation as requiring less extrapolation as compared to "predictive" tasks by a panel of three psychologists, as described earlier. In these tasks, counselors will be asked to evaluate the client, based on what they know about the client, on five-point scales corresponding to various "evaluative dimensions."

For example:

"From what you know about the client, how anxious a person do you believe he/she is?"

.....

---

Extremely	Anxious	Slightly	Not Very	Not at All
Anxious		Anxious	Anxious	Anxious

"From what you know about the client, how open and expressive of feelings do you think he/she is, in general?"

.....

---

Extremely	Open	Somewhat	Not Very	Not at All
Open		Open	Open	Open

These "evaluative dimensions" are chosen to represent variables according to which counselors are frequently asked to characterize clients in a broad way, i.e., is the client an "anxious" person?"; "is the client an open, expressive person?," etc.

The criterion for determining accuracy of a judge's evaluation of the client is defined as the corresponding evaluation of the client made by the client's own therapist after 12 sessions of therapy. Although this is, of course, in actuality a measure of inter-judge reliability, it is felt that it can provide for a meaningful comparison when viewed in terms of "accuracy." It would be assumed that the greater opportunity of the client's therapist to observe him personally upon a greater number of occasions in more varied circumstances would allow for greater accuracy of evaluation. This receives support from Luft's (1950) finding that a patient's therapist was more accurate in making predictions about him than were 34 of 35 clinical judges serving as subjects. The findings of Hjelle (1968) and Taft (1966) that greater familiarity with the individual to be judged resulted in greater predictive accuracy also support the use of the therapist's rating of his client as a criterion for predictive accuracy. Finally, Horowitz (1962) employed the therapist's rating of his client as the criterion in

her study of accuracy of clinical judgment.

A difference score will be calculated for each counselor between his responses concerning the client and the responses of the client's therapist about him. This score is taken to represent the accuracy of the judge's evaluation of the client on the Evaluative Dimensions Instrument described above. These difference scores will be submitted to a one-way analysis of variance in order to test the following null hypothesis:

$$\begin{array}{ccccc} \text{Stereotype} & = & \text{Stereotype-Video} & = & \text{Stereotype-} \\ \text{Evaluative} & & \text{Evaluative} & & \text{Video-Diagnostics} \\ \text{Accuracy} & & \text{Accuracy} & & \text{Evaluative} \\ & & & & \text{Accuracy} \end{array}$$

This hypothesis tests the effect of differing types of information upon the counselor's ability to accurately evaluate the client on a number of dimensions commonly employed by psychologists in evaluating clients. It permits one to ascertain whether different types of information result in any improvement or deficit in evaluative "accuracy" beyond that which can be achieved by knowledge of minimal or "stereotype" information alone.

If, as postulated above, diagnostic material results in a "set" to "objectively evaluate" a client rather than attempt to see the client as the client sees himself, the Stereotype-Video-Diagnostics Group would have the advantage

in the current task which requires just such judgments.

Therefore, the following experimental hypothesis is made:

Stereotype- Video-Diagnostics	>	Stereotype-Video	>	Stereotype
Evaluative		Evaluative		Evaluative
Accuracy		Accuracy		Accuracy

Results will be discussed in relation to the above hypothesis and previous relevant findings.

3) Confidence Tasks. Counselors will be required to indicate their confidence in each judgment they make regarding the client according to the method of Adams (1957) and Oskamp (1962, 1965). They defined the confidence scale in terms of expected percentages of successful judgments.

As confidence has been shown to increase dramatically with increased information, the following hypotheses are advanced relative to the effect of differing types of information upon judges' confidence in predictions and evaluations:

Stereotype- Video-Diagnostics	>	Stereotype-Video	>	Stereotype
Predictive Confidence		Predictive Confidence		Predictive Confidence

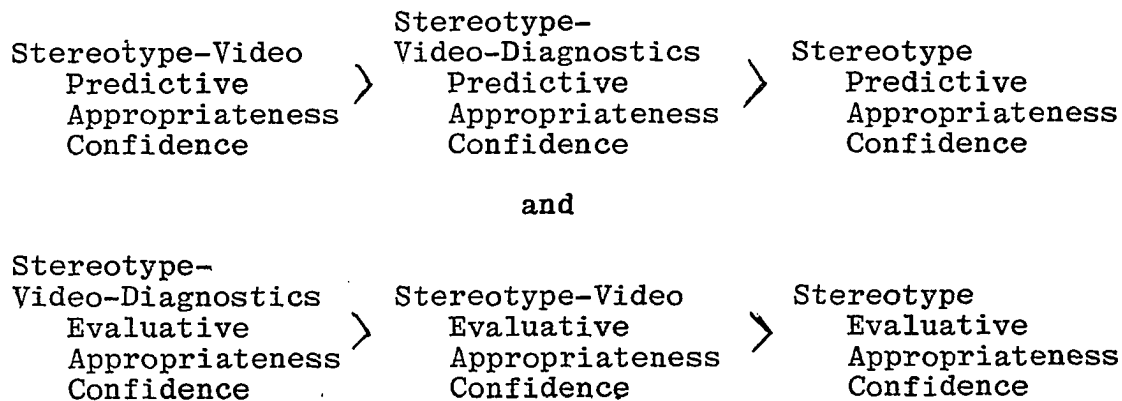
and

Stereotype- Video-Diagnostics	>	Stereotype-Video	>	Stereotype
Evaluative Confidence		Evaluative Confidence		Evaluative Confidence

Test of these hypotheses will be made by submitting

confidence scores to appropriate statistical analyses.

Where both confidence scores and accuracy scores are available for a particular type of judgment, examination will be made of the "appropriateness" of judges' confidence (i.e., is greater confidence in a judgment associated with greater accuracy of that judgment?). Both confidence and accuracy scores will be available for responses on the Trait Rating Instrument (predictive task) and for responses on the Evaluative Dimensions Instrument (evaluative task). In line with the previous hypotheses concerning the effect of differing information upon accuracy of predictions and evaluations:



That is, differing types of information are expected to affect appropriateness of confidence differentially depending upon what type of judgment task is involved (i.e., one greater or lesser extrapolation from the data).

Judges will finally be requested to answer the following questions:

(1) How well did you like the client as a person?  
.....

---

Very	Fairly	Slightly	Not Very	Not at
Much	Well		Much	All

(2) How much like you was the client?  
.....

---

Very	Quite a Bit	Somewhat	Not Much	Not at All
Much	Like Me	Like Me	Like Me	Like Me

Like me

(3) What is your diagnosis of the client?

## References

- Adams, J. K. A confidence scale defined in terms of expected percentages. American Journal of Psychology, 1957, 70, 432-436.
- Albee, G. W., and Hamlin, R. M. An investigation of the reliability and validity of judgments of adjustments inferred from drawings. Journal of Clinical Psychology, 1949, 5, 396-398.
- Bialick, I. A method of applying clinical judgment to problems of Rorschach validation: The quality of Rorschach whole responses and intelligence. Unpublished master's thesis, University of Pittsburgh, 1950.
- Bialick, I., and Hamlin, R. M. The clinician as judge: Details of procedure in judging projective materials. Journal of Consulting Psychology, 1954, 18, 239-242.
- Block, J., and Petersen, P. Some personality correlates of confidence, caution, and speed in a decision situation. Journal of Abnormal and Social Psychology, 1955, 51, 34-41.
- Borke, H., and Fiske, D. W. Factors influencing the prediction of behavior from a diagnostic interview. Journal of Consulting Psychology, 1957, 21, 78-80.

- Carr, J. E. Perceived therapy outcome as a function of differentiation between and within conceptual dimensions. Journal of Clinical Psychology, 1974, 30, 282-285.
- Cline, V. B. Ability to judge personality assessed with a stress interview and sound-film technique. Journal of Abnormal and Social Psychology. 1955, 50, 183-187.
- Cline, V. B., and Richards, J. M., Jr. Variables related to accuracy in interpersonal perception. USN Annual Report, 1958, ONR 171-146, Department of Psychology, University of Utah.
- Cline, V. B., and Richards, J. M., Jr. Accuracy of interpersonal perception: A general trait? Journal of Abnormal and Social Psychology, 1960, 60, 1-7.
- Cronbach, L. J., and Gleser, G. C. Assessing similarity between profiles. Psychological Bulletin, 1953, 50, 456-473.
- Cronbach, L. J. Processes affecting scores on "understanding of others" and "assumed similarity." Psychological Bulletin, 1955, 52, 177-193.
- Cummings, S. T. The clinician as judge: Judgments of adjustment inferred from Rorschach single-card performance. Journal of Consulting Psychology, 1954, 18, 243-247.



- Einhorn, H. J., and Hogarth, R. M. Confidence in judgment: Persistence of the illusion of validity. Psychological Review, 1978, 85, 395-416.
- Forer, B. R., and Tolman, R. S. Some characteristics of clinical judgment. Journal of Consulting Psychology, 1952, 16, 347-352.
- Gage, N. L., and Cronbach, L. J. Conceptual and methodological problems in interpersonal perception. Psychological Review, 1955, 62, 411-422.
- Giedt, F. H. Comparison of visual, content, and auditory cues in interviewing. Journal of Consulting Psychology, 1955, 19, 407-416.
- Goldberg, L. R. The effectiveness of clinicians' judgments: The diagnosis of organic brain damage from the Bender-Gestalt Test. Journal of Consulting Psychology, 1959, 23, 25-33.
- Goldberg, L. R. Simple models or simple processes? Some research on clinical judgments. American Psychologist, 1968, 23, 483-496.
- Hamlin, R. M. The clinician as judge: Implications of a series of studies. Journal of Consulting Psychology, 1954, 18, 233-238.

- Hamlin, R. M., and Newton, R. L. Comparison of a schizophrenic and a normal subject, both rated by clinicians as well-adjusted on the basis of "blind" Rorschachs. Paper read at Eastern Psychological Association, Atlantic City, N.J., March 1952.
- Hamlin, R. M., Albee, G. W., and Leland, E. M. Objective Rorschach "signs" for groups of normal, maladjusted and neuropsychiatric subjects. Journal of Consulting Psychology, 1950, 14, 276-282.
- Hamlin, R. M., Berger, B., and Cummings, S. T. Changes in adjustment following psychotherapy as reflected in Rorschach signs. In W. Wolff, and J. A. Precker (Eds.), Success in psychotherapy. New York: Grune and Stratton, 1952. Pp. 94-111.
- Hammond, K. R. Computer graphics as an aid to learning. Science, 1971, 172, 903-908.
- Hjelle, L. A. Accuracy of personality and social judgments as functions of familiarity. Psychological Records, 1968, 22, 311-319.
- Holzworth, R. J., and Doherty, M. E. Inferences and predictions: Normative vs. representative responding. Bulletin of the Psychonomic Society, 1974, 3, 300-302.

- Horowitz, M. J. A study of clinicians' judgments from projective test protocols. Journal of Consulting Psychology, 1962, 3, 251-256.
- Johnson, D. M. Confidence and speed in the two category judgment. Archives of Psychology, 1939, 34 (No. 241), 1-52.
- Johnson, D. M. The psychology of thought and judgment. New York: Harper, 1955.
- Kahneman, D., and Tversky, A. Subjective probability: A judgment of representativeness. Cognitive Psychology, 1972, 3, 430-454.
- Kahneman, D., and Tversky, A. On the psychology of prediction. Psychological Review, 1973, 4, 237-251.
- Johnson, L. C. Generality of speed and confidence in judgment. Journal of Abnormal and Social Psychology, 1957, 54, 264-266.
- Kleinmuntz, B. MMPI decision rules for the identification of college maladjustments: A digital computer approach. Psychological Monographs, 1963, 77 (Whole No. 577). (a)
- Kleinmuntz, B. Personality test interpretation by digital computer. Science, 139, 416-418. (b)
- Kleinmuntz, B. Profile analysis revisited: A heuristic approach. Journal of Counseling Psychology, 1963, 10, 315-324. (c)

- Kostlan, A. A. A method for the empirical study of psycho-diagnosis. Journal of Consulting Psychology, 1954, 18, 83-88.
- Kostlan, A. A. A reply to Patterson. Journal of Consulting Psychology, 1955, 19, 486.
- Lesser, W. M. The relationship between counseling progress and empathic understanding. Journal of Counseling Psychology, 1961, 8, 330-336.
- Little, K. B. Confidence and reliability. Educational and Psychological Measurement, 1961, 21, 95-101.
- Luft, J. Implicit hypotheses and clinical predictions. Journal of Abnormal and Social Psychology, 1950, 45, 756-760.
- Meehl, P. E. Clinical versus statistical prediction: A theoretical analysis and a review of the evidence. Minneapolis: University of Minnesota Press, 1954.
- Meehl, P. E. The cognitive activity of the clinician. American Psychologist, 1960, 15, 19-27.
- Newton, R. L. The clinician as judge: Total Rorschachs and clinical case material. Journal of Consulting Psychology, 1954, 18, 248-250.

- Newton, R. L., and Goodman, H. W. A comparison of Basic Rorschach Scores with judgments of adjustment based on clinical case material. Paper read at Eastern Psychological Association, Atlantic City, N.J., March, 1952.
- Oskamp, S. The relationship of clinical experience and training methods to several criteria of clinical prediction. Psychological Monographs, 1962, 76 (No. 28, Whole No. 547).
- Oskamp, S. Overconfidence in case-study judgments. Journal of Consulting Psychology, 1965, 29, 261-265.
- Patterson, C. H. Diagnostic accuracy or diagnostic stereotype? Journal of Consulting Psychology, 1955, 19, 483-485.
- Rodin, M. J. The effect of behavioral context on information selection and differential accuracy in a person perception task. Journal of Social Psychology, 1975, 97, 83-94.
- Rogers, C. K. The necessary and sufficient conditions of therapeutic personality change. Journal of Consulting Psychology, 1957, 21, 95-103.
- Seward, G. H. Recognition time as a measure of confidence. Archives of Psychology, 1928, 16 (No. 99), 1-54.
- Shrauger, S., and Altrocchi, J. The personality of the perceiver as a factor in person perception. Psychological Bulletin, 1964, 62, 289-308.

- Soskin, W. F. Bias in postdiction from projective tests. Journal of Abnormal and Social Psychology, 1954, 49, 69-74.
- Soskin, W. F. Frames of reference in personality assessment. Journal of Clinical Psychology, 1954, 10, 107-114.
- Soskin, W. F. Influence of four types of data on diagnostic conceptualization in psychological testing. Journal of Abnormal and Social Psychology, 1959, 58, 69-78.
- Stelmachers, Z. T., and McHugh, R. B. Contribution of stereotyped and individualized information on predictive accuracy. Journal of Consulting Psychology, 1964, 28, 234-242.
- Taft, R. Multiple methods of personality assessment. Psychological Bulletin, 1959, 56, 333-352.
- Taft, R. Accuracy of empathic judgments of acquaintances and strangers. Journal of Personality and Social Psychology, 1966, 3, 600-604.
- Trow, W. C. The psychology of confidence. Archives of Psychology, 1923, 10 (No. 67), 1-47.
- Truax, C., and Carkhuff, R. R. Toward effective counseling and psychotherapy. Chicago: Aldine, 1967.
- Tversky, A., and Kahneman, D. Belief in the new law of small numbers. Psychological Bulletin, 1971, 76, 105-110.

- Tversky, A., and Kahneman, D. Judgment under uncertainty: Heuristics and biases. Science, 1974, 185, 1124-1131.
- Wolff, W. M. Certainty: Generality and relation to manifest anxiety. Journal of Abnormal and Social Psychology, 1955, 50, 59-62.
- Whitmyre, J. W. The significance of artistic excellence in the judgment of adjustment inferred from human figure drawings. Journal of Consulting Psychology, 1953, 17, 421-424.

APPENDIX B

Information for Prospective  
Participants



Information for Prospective Participants

I am Ruth Ann Mertens, a doctoral student in Counseling Psychology at the University of Oklahoma.

I am conducting a research study on the ability of psychologists to understand the people they are trying to help. To do this, it is necessary to make videotapes of real clients working with actual counselors. As you are a client here at the Counseling Center, I am asking you to consider participating in this project. In addition to allowing videotaping of your counseling session, you would be required to take a battery of standard psychological tests. The videotape would then be shown to other professional counselors and psychologists to measure how well they can understand a client. Your responses on the psychological tests would also be made available to the professionals to see if this information helps in understanding a client.

Participation in this project will in no way affect the type or quality of counseling which you will receive here at the Counseling Center. The counseling will not be an experiment, but will be traditional, using standard, accepted techniques. Results of the psychological tests will be discussed with you for purposes of increased self-understanding, and you will receive a fee of \$25 as remuneration for the time

spent in taking the tests if you are selected to participate in the study.

You will not be eligible to participate in the project if you commute to the University from the Oklahoma County area. Also, I am interested in videotaping only clients with certain types of concerns. Because of this, if you believe you are interested in participating, I will need to discuss with you briefly why you think you would like to see a professional counselor.

If you think you are interested, or if you have any questions, please indicate so at this time, and I will explain more fully.

APPENDIX C

Legally Effective Informed  
Consent of Client

Legally Effective Informed Consent for Client

It is important that you understand fully the purpose of this project and your participation in it. Please feel free to ask any questions or make any comments now or at any time during the project. As I read this statement, I will periodically stop so that you may ask questions. Please feel free to interrupt me at any time. This form explains fully:

- 1) The nature and purpose of this study
- 2) How the tapes of counseling and the test responses will be used
- 3) Any possible risks to you
- 4) The obligations of the investigator
- 5) Your responsibilities and your rights.

1. Nature and Purpose of This Study

If you agree to participate in this project, you will receive counseling from a qualified counselor here at the University of Oklahoma. You are in no way obligated to participate in this project in order to receive counseling at the Counseling Center. You will receive exactly the same counseling whether or not you participate in the project. I would like to stress that this is not an experiment on a type of counseling. The help you receive will in no way be experimental.

The counseling session will be conducted in a standard office here in the counseling center with only you and your counselor present. The videotape equipment is a standard part of the office equipment and will be operated by your counselor. At the end of your counseling session, you will be asked to fill out a brief questionnaire about how you see yourself. This will take about five minutes.

You will also be asked to take a battery of well-known, standard psychological tests, including both intellectual and personality tests. This battery will be administered by me and should take approximately six hours to complete. Testing will be done in two sessions of approximately three hours each, which can be scheduled at mutually agreeable times. Upon completion of the testing, you will be paid a fee of \$25 as compensation for your time. At a later date, I will schedule an appointment with you for discussion of the test results. You may find these helpful in increasing your self-understanding. Your counselor will also be present at our meeting.

Before any taped material or test responses are shown to the professional counselors and psychologists who will participate in the study, you will be allowed to review all of the taped material and will have an opportunity to discuss the test results with me. At this time, you may ask that any

part of the tapes be edited out and/or that any testing material not be shown to the professionals. I will review the tapes with you and also may suggest that some of the tapes be edited. You are not required to review or edit the tapes, but you have the right to do so.

## 2. How the Tapes Will Be Used

If after viewing the tapes and discussing the test results you choose to allow this material to be used in the project, it will be shown only to qualified, trained professional counselors and psychologists who hold the M.D., Ph.D., M.A. or M.S. degrees.

Before any subject will be allowed to view the tape or the testing material, they will be asked to read and sign this form (give copy of Appendix to client). After these people view the tape and/or see the test responses, they will fill out the same brief questionnaire you completed after the counseling session seen in the tape as they believe you filled it out.

The purpose of this study is to measure how different types of information about a client help a counselor to understand the client. If counselors with access to your test responses are closer to your self-description than those without your test data, it may indicate that such test data contribute to a counselor's understanding more fully how a

client sees himself. This information is important for counselors if they are to be best able to understand their clients. This is the entire purpose of the project: to attempt to measure how well people can understand another individual based on the kind and amount of information they have about him/her.

### 3. Possible Risks to You

It is unlikely that anyone viewing the tape or the test data will know you or ever come into contact with you. The tapes will not be shown in Norman, Oklahoma or in any other place you do not wish them to be shown. They will be shown only to professional counselors and psychologists who have a thorough understanding of your right to confidentiality. You will not be identified by your name in any material used. However, there does exist a possibility that you may be recognized by someone viewing the tapes. Therefore, you will want to ask that any information contained in the tape or test material which you believe could be personally identifying or troublesome be edited out. Furthermore, after reviewing the tape and discussing the test results, you may withdraw from the project completely and the tapes will be destroyed. I do not wish for there to be any harm to you in any way. Therefore, if I believe there to be any reasonable chance of negative effects, I will not use the tape or test material.

If this is the case, I will inform you that the material will not be used.

4. Obligations of the Investigator

It is my obligation:

- A. To pay you \$25 upon your completion of the psychological test battery.
- B. To discuss with you the results of the test battery for purposes of your self-understanding.
- C. To completely inform you of all your rights in this project.
- D. To allow you to withdraw from this project at any time prior to and including your review of the tapes and discussion of the test material.
- E. To protect your identity as much as is possible as outlined in this statement and the Instructions of the Experimental Subjects (Appendix).

5. Client Responsibilities and Rights

- A. You have the right to discontinue this project any time up to and including your review of the tapes and discussion of the test data.
- B. It is your responsibility to complete the questionnaire and psychological test battery as honestly and fully as you are capable of doing.



- C. It is your responsibility to decide whether or not you wish for the counselors and psychologists to view the tape and see the test material.
- E. It is your responsibility to decide whether you consider there to be any undue risk of harm to yourself by having the professionals view the tape and the test material. If, at any time up to and including your review of the material you believe there to be risk of harm, it is your responsibility to inform me and withdraw from the project.

By my signature, I state that I have read and fully understand the purpose and nature of this, the Clinical Judgment Study I.

.....  
\_\_\_\_\_  
Signature

\_\_\_\_\_  
Date

APPENDIX D

Consent for Videotaping  
of Client

Client Name \_\_\_\_\_ Date \_\_\_\_\_  
Address \_\_\_\_\_ Soc. Sec. # \_\_\_\_\_  
Date of Birth \_\_\_\_\_

I, the undersigned client, do hereby authorize Ruth Ann Mertens, Doctoral student in the Counseling Psychology program, College of Education, the University of Oklahoma, to videotape the counseling session between myself and my counselor at the University Counseling Center on this date for purposes of research. I understand that portions of the tape of my counseling session will be shown to professionals in mental health for purposes of research in the area of counseling skills.

I further understand that I am under no obligation to participate in this project. Furthermore, I understand that I will be allowed to view these tapes, and may at that time withhold my permission for the use of these tapes in research.

By my first signature, I do authorize the videotaping of my counseling session at the University of Oklahoma Counseling Center. I do also state that I have read and understand the Legally Effective Informed Consent form labeled Clinical Judgment Study I.

1. \_\_\_\_\_  
Witness \_\_\_\_\_ Date \_\_\_\_\_ Client \_\_\_\_\_ Date \_\_\_\_\_  
Address \_\_\_\_\_  
\_\_\_\_\_

2. \_\_\_\_\_  
Witness \_\_\_\_\_ Date \_\_\_\_\_ Client \_\_\_\_\_ Date \_\_\_\_\_  
Address \_\_\_\_\_  
\_\_\_\_\_

By my second signature, I do hereby release the video-tape and test data to be used for research purposes.

1. \_\_\_\_\_  
Witness \_\_\_\_\_ Date \_\_\_\_\_ Client \_\_\_\_\_ Date \_\_\_\_\_  
Address \_\_\_\_\_  
\_\_\_\_\_

2. \_\_\_\_\_  
Witness \_\_\_\_\_ Date \_\_\_\_\_ Client \_\_\_\_\_ Date \_\_\_\_\_  
Address \_\_\_\_\_  
\_\_\_\_\_

APPENDIX E

Institutional Review Board

Executive Summary

To: The University of Oklahoma Institutional Review Board--  
Norman Campus  
c/o Mark Elder

Office of Research Administration  
Buchanan Hall, Room #314

From: Ruth Ann Mertens  
Counseling Psychology Program  
College of Education  
842-3319 (home telephone)

- I. A. Project Title: Clinical Judgment Study I
  - B. Principal Investigator: Ruth Ann Mertens, Counseling Psychology Program, 842-3319.
  - C. Sponsor's Name: Robert Ragland, Ph.D., 325-5975
  - D. Proposed Starting Date: September 15, 1980
- II. Executive Summary: Clinical Judgment I

The purpose of this study is to examine the manner in which different types of information available to the counselor affect the accuracy of his clinical judgments and his confidence in these judgments. To do this, it is necessary to make videotapes of a real client working with an actual counselor. I propose to tape such a counseling session at the Counseling Center, University of Oklahoma. The client will also be required to complete a questionnaire describing himself on dimensions commonly employed by psychologists to

characterize clients. An example of such a questionnaire is included in this proposal and is labeled Appendix F. The client will further be asked to complete a battery of well-known and widely-used psychological tests including: 1) the Wechsler Adult Intelligence Scale (WAIS), 2) the Rorschach Inkblot Test, 3) the Thematic Apperception Test (TAT), 4) the Minnesota Multiphasic Personality Inventory (MMPI), and 5) Human Figure Drawings. I will personally administer these tests, and will discuss the results of the battery with the client accompanied by his counselor. These tests may be valuable to the client in helping him/her achieve better self-understanding. The client will also be paid a \$25 fee to compensate him/her for the time spent in taking the test battery. Information about the client will then be made available to professional counselors and psychologists holding the M.D., Ph.D., M.A., or M.S. degrees and experienced in both counseling/psychotherapy and in psychological testing. These professionals will receive different types of information about the client. One group will be told only the age and sex of the client and that he/she is a student and a counseling client. A second group of professionals will receive this "minimal information," but will additionally view a segment of the videotaped counseling session involving the client. A third group will have access to the

information given the two other groups, but will also be given the test protocols of the client. The professional counselors will then be asked to perform a number of clinical tasks based upon the information they have received about the client. They will be asked: 1) to fill out the self-description inventory (Appendix F) as they believe the client filled it out for him/herself; 2) to complete a questionnaire evaluating the client as they believe the client's own therapist evaluated him/her after a number of counseling sessions (Appendix G); and 3) to complete a form making certain clinical predictions about the client (Appendix H). The professional counselors will also be asked to indicate how confident they are in each of the judgments they make. Differences in accuracy of clinical judgments made by the three groups of professionals will be assumed to arise from the different types of information upon which the judgments were based. This will allow us to consider which types of information lead to the greatest accuracy in which types of clinical judgments.

Because the client is required to describe himself genuinely, and respond genuinely to psychological tests, an actual client, rather than an actor playing the client's part, is required. Also, the experimental subject must have the opportunity to view the client in order to have access



to critical non-verbal communication of the client. If subjects are to be able to make accurate clinically-relevant judgments about a client, they must have access to actual clinically-relevant cues.

To protect the confidentiality of the client, his/her name will be deleted from the taped material and test data. Also, the client will be allowed to review and edit the tapes and discuss his test results with the investigator before releasing any material for research use. The client will be advised that he may withdraw from the project at any time prior to his/her release of the material.

A further protection of the client is that the experimental subjects viewing the tapes will all be trained, qualified professionals with a thorough understanding of the client's right to confidentiality. (See Appendix E.) Furthermore, the tapes will not be shown in Norman, Oklahoma or in any other place in which the client requests they not be shown, so that there is little likelihood of recognition of him/her by a subject.

Also, I will personally screen clients for this study with respect to the type and severity of the problem for which they are seeking counseling. Only clients with less serious problems will be used. Examples of such problems might be a situational disturbance, mild to moderate

depression or anxiety, etc. In no case will a seriously-disturbed individual be considered for participation.

Finally, my doctoral advisor will screen any material (tapes and test data) before its use in the experiment. A more detailed discussion of client safeguards will be found in the Informed Consent accompanying this proposal.

Any risk to the client is considered to be minimal. There are no physical risks involved. Also, if the client provides legally effective informed consent, it indicates his/her understanding that professional counselors will have access to the counseling information and that he/she perceives no harm in this process. Since the client will be allowed to review the tapes and edit any part and will discuss the test results with the investigator, and since the client will have the opportunity to remove him/herself entirely from the project, there exists little possibility of psychological or social risk of harm.

The client may benefit by his/her participation in the study in the following ways: 1) the administration to the client at no charge of intellectual and personality tests and the discussion of their results with the client for purposes of his self-understanding and growth; 2) receipt of the \$25 fee as compensation for time spent in taking the test battery; 3) the opportunity to participate in an

interesting project which may prove valuable in the training of professional counselors; and 4) possible credit toward research participation requirements in introductory psychology classes.

Experimental subjects should run no risk, as all data pertaining to them is anonymous. The primary benefit to them is that stated in 3) above. The social risk to the counselor in the videotape is very slight, as only the client will be seen in the tape. However, the remote possibility does exist of the counselor being recognized by his voice. Therefore, I will also obtain Legally Effective Informed Consent from the counselor.

APPENDIX F

Instructions to Experimental Subjects

Stereotype Group

Instructions to Experimental Subjects

Introduction to the Task

The purpose of this study is to measure the ability of counselors/psychotherapists to understand a client. You will be given some information about a client. After you have read the information, you will be asked to make some clinical judgments about the client. Therefore, it is important that you try to understand the client as well as possible based upon the information you have about her.

Subject Protection

If you have any questions or comments concerning this project, please contact me, Ruth Ann Mertens, about this matter. My name and telephone number are listed at the bottom of these instructions. You are to tear them off and retain them for future reference.

Your responses in this study will be kept completely confidential and will not be associated with your name.

If you have read this statement and fully understand and agree to the instructions, you are asked to date and sign this statement. If you do not completely understand

or completely agree to abide by these instructions, you are asked to notify me at this time.

Name: \_\_\_\_\_  
Address: \_\_\_\_\_  
Signature \_\_\_\_\_  
Date \_\_\_\_\_

-----  
Tear off here and retain.

Clinical Judgment Study I

Experimenter: Ruth Ann Mertens, Counseling Psychology  
Program, College of Education, University of  
Oklahoma.  
Home Address: 2420 N.W. 45th, Oklahoma City,  
Oklahoma 73112  
Home Phone: 842-3119

Additional Sources of Information:

Mark Elder	Robert Ragland, Ph.D.
Office of Research Adminis- tration	College of Education Room 308
Buchanan Hall, Room 314	University of Oklahoma
University of Oklahoma	Norman, Oklahoma 73019
Norman, Oklahoma 73019	(405) 325-5975
(405) 325-4757	

APPENDIX G

Instructions to Experimental Subjects

Stereotype-Video Group

## Instructions to Experimental Subjects

### Introduction to the Task

The purpose of this study is to measure the ability of counselors/psychotherapists to understand a client. You will be shown a videotape of an actual counseling session and given some information about the client. After you view the videotape and read the information, you will be asked to make some clinical judgments about the client. Therefore, it is important that you watch and listen carefully so that you will understand the client as well as possible.

### Confidentiality of Information

The material you are about to see is considered very confidential. The client has been very gracious in allowing her private counseling session to be used for research purposes. Any identifying information about the client has been changed or edited out of the material in order to protect the client's identity. It is very unlikely that you will ever come into contact with this person. However, under no circumstances should you ever discuss the person or material with anyone.

### Subject Protection

If you have any questions or comments concerning this project, please contact me, Ruth Ann Mertens, about this



matter. My name and telephone number are listed at the bottom of these instructions. You are to tear them off and retain them for future reference.

Your responses in this study will be kept completely confidential and will not be associated with your name.

If, when the tape is played, you recognize the client, please notify me immediately and the tape will be stopped.

If you have read this statement and fully understand and agree to the instructions, you are asked to date and sign this statement. If you do not understand or completely agree to abide by these instructions, you are asked to notify me at this time.

Name: \_\_\_\_\_

Signature

Address: \_\_\_\_\_

Date

-----  
Tear off here and retain.

Clinical Judgment Study I

Experimenter: Ruth Ann Mertens, Counseling Psychology Program

College of Education, University of Oklahoma

Home Address: 2420 N.W. 45th, Oklahoma City,

Oklahoma 73112

Home Phone: 842-3119

Additional sources of information:

Mark Elder

Office of Research Administration

Buchanan Hall, Room 314

University of Oklahoma

Norman, Oklahoma 73019

(405) 325-4757

Robert Ragland, Ph.D.

College of Education

Room 308

University of Oklahoma

Norman, Oklahoma 73019

(405) 325-5975

APPENDIX H

Instructions to Experimental Subjects

Stereotype-Video-Diagnostics Group

Instructions to Experimental SubjectsIntroduction to the Task

The purpose of this study is to measure the ability of counselors/psychotherapists to understand a client. You will be shown a videotape of an actual counseling session with a client and given test protocols of the client you view in the tape. You will also be given some information about the client. After you view the videotape, read the information, and study the test material, you will be asked to make some clinical judgments about the client. Therefore, it is important that you watch and listen carefully so that you will understand the client as well as possible.

Confidentiality of Information

The material you are about to see is considered very confidential. The client has been very gracious in allowing her private counseling session to be used for research purposes. Any identifying information about the client has been changed or edited out of the material in order to protect the client's identity. It is very unlikely that you will ever come into contact with this person. However, under no circumstances should you ever discuss the person or material with anyone.

Subject Protection

If you have any questions or comments concerning this project, please contact me, Ruth Ann Mertens, about this matter. My name and telephone number are listed at the bottom of these instructions. You are to tear them off and retain them for future reference.

Your responses in this study will be kept completely confidential and will not be associated with your name.

If, when the tape is played, you recognize the client, please notify me immediately and the tape will be stopped.

If you have read this statement and fully understand and agree to the instructions, you are asked to date and sign this statement. If you do not understand or completely agree to abide by these instructions, you are asked to notify me at this time.

Name: \_\_\_\_\_  
Address: \_\_\_\_\_  
Signature \_\_\_\_\_  
Date \_\_\_\_\_

-----  
Tear off here and retain.

Clinical Judgment Study I

Experimenter: Ruth Ann Mertens, Counseling Psychology Program  
College of Education, University of Oklahoma

Home Address: 2420 N.W. 45th, Oklahoma City,  
Oklahoma 73112

Home Phone: 842-3119

Additional sources of information:

Mark Elder

Office of Research Administration

Buchanan Hall, Room 314

University of Oklahoma

Norman, Oklahoma 73019

(405) 325-4757

Robert Ragland, Ph.D.

College of Education

Room 308

University of Oklahoma

Norman, Oklahoma 73019

(405) 325-5975

APPENDIX I

Evaluative Dimensions Instrument

Evaluative Dimensions Instrument

At this time, I would like for you to describe the client about whom you were given information in the way you think is most typical of her on each of the characteristics below.

For each characteristic listed below, place an X by the statement which you feel generally describes the client best on that particular characteristic. Do not try to mark the items in the same way you believe the client marked them about herself. Do not use the checklist to describe yourself. Try only to describe the client most accurately as you believe she really is.

1. From what you know about the client, how depressed a person do you think she is?

- Very depressed                    \_\_\_\_\_
- Depressed                            \_\_\_\_\_
- Slightly depressed                \_\_\_\_\_
- Not very depressed                \_\_\_\_\_
- Not at all depressed              \_\_\_\_\_

2. From what you know about the client, how withdrawn a person do you think she is?

- Very withdrawn                    \_\_\_\_\_
- Withdrawn                            \_\_\_\_\_
- Slightly withdrawn                \_\_\_\_\_



- Not very withdrawn .....  
\_\_\_\_\_
- Not at all withdrawn .....  
\_\_\_\_\_
3. From what you know about the client, how expressive of feelings and emotions do you think she is?
- Very expressive .....  
\_\_\_\_\_
- Expressive .....  
\_\_\_\_\_
- Slightly expressive .....  
\_\_\_\_\_
- Not very expressive .....  
\_\_\_\_\_
- Not at all expressive .....  
\_\_\_\_\_
4. From what you know about the client, how anxious a person do you think she is?
- Very anxious .....  
\_\_\_\_\_
- Anxious .....  
\_\_\_\_\_
- Slightly anxious .....  
\_\_\_\_\_
- Not very anxious .....  
\_\_\_\_\_
- Not at all anxious .....  
\_\_\_\_\_
5. From what you know about the client, how controlling of others do you think she is?
- Very controlling .....  
\_\_\_\_\_
- Controlling .....  
\_\_\_\_\_
- Slightly controlling .....  
\_\_\_\_\_
- Not very controlling .....  
\_\_\_\_\_
- Not at all controlling .....  
\_\_\_\_\_

6. From what you know about the client, how much do you think she needs the approval of others?
- |                              |       |       |
|------------------------------|-------|-------|
| Needs approval very much     | ..... | ..... |
|                              | _____ | _____ |
| Needs approval               | ..... |       |
|                              | _____ |       |
| Needs approval slightly      | ..... |       |
|                              | _____ |       |
| Doesn't need approval much   | ..... |       |
|                              | _____ |       |
| Doesn't need approval at all | ..... |       |
|                              | _____ |       |
7. From what you know about the client, how angry a person do you think she is?
- |                  |       |       |
|------------------|-------|-------|
| Very angry       | ..... | ..... |
|                  | _____ | _____ |
| Angry            | ..... |       |
|                  | _____ |       |
| Slightly angry   | ..... |       |
|                  | _____ |       |
| Not very angry   | ..... |       |
|                  | _____ |       |
| Not at all angry | ..... |       |
|                  | _____ |       |
8. From what you know about the client, how much do you think she tends to take responsibility for others?
- |                                    |       |       |
|------------------------------------|-------|-------|
| Takes responsibility a great deal  | ..... | ..... |
|                                    | _____ | _____ |
| Takes responsibility               | ..... |       |
|                                    | _____ |       |
| Takes responsibility somewhat      | ..... |       |
|                                    | _____ |       |
| Doesn't take much responsibility   | ..... |       |
|                                    | _____ |       |
| Doesn't take responsibility at all | ..... |       |
|                                    | _____ |       |
9. From what you know about the client, how generally well-adjusted do you think she is?

Very well-adjusted .....  
Well-adjusted .....  
Slightly well-adjusted .....  
Not well-adjusted .....  
Not at all well-adjusted .....

10. From what you know about the client, to what extent does she exhibit disturbed thought processes?

Thought processes very disturbed .....  
Thought processes disturbed .....  
Thought processes slightly disturbed .....  
Thought processes not very disturbed .....  
Thought processes not at all disturbed .....

11. From what you know about the client, how trusting of others do you think she is?

Very trusting .....  
Trusting .....  
Slightly trusting .....  
Not very trusting .....  
Not at all trusting .....

12. From what you know about the client, to what extent is she characterized by dependency needs?

Great dependency needs .....  
Dependency needs .....  
Slight dependency needs .....

- Not many dependency needs .....  
\_\_\_\_\_
- No dependency needs .....  
\_\_\_\_\_
13. From what you know about the client, how stable a person  
do you think she is?
- Very stable .....  
\_\_\_\_\_
- Stable .....  
\_\_\_\_\_
- Slightly stable .....  
\_\_\_\_\_
- Not very stable .....  
\_\_\_\_\_
- Not at all stable .....  
\_\_\_\_\_
14. From what you know about the client, how rigid a person  
do you think she is?
- Very rigid .....  
\_\_\_\_\_
- Rigid .....  
\_\_\_\_\_
- Slightly rigid .....  
\_\_\_\_\_
- Not very rigid .....  
\_\_\_\_\_
- Not at all rigid .....  
\_\_\_\_\_
15. From what you know about the client, how intelligent a  
person do you think she is?
- Very intelligent .....  
\_\_\_\_\_
- Intelligent .....  
\_\_\_\_\_
- Slightly intelligent .....  
\_\_\_\_\_
- Not very intelligent .....  
\_\_\_\_\_
- Not at all intelligent .....  
\_\_\_\_\_

Now that you have completed the checklist describing the client as you believe she really is, please go back and indicate for each characteristic on the checklist how confident you are that you marked it to describe the client accurately. You will state this confidence estimate in terms of the probability that your marking of the item accurately describes the client as she really is.

This means that your confidence estimates will be made on a scale ranging from 20 through 100, but using only the even tens (i.e., 20, 30, 40, 50, 60, 70, 80, 90, 100). Thus, if you think the chances of your having marked the item to accurately describe the client are 70 out of 100, write 70 in the blank to the far right of each item. If you think that your chances are 8 out of 10, or 80%, write 80 in the blank. If you are certain that your marking of an item accurately describes the way the client really is, you should write 100. If you feel that your marking of an item has no more than a random chance of being accurate, you should write 20 in the blank.

APPENDIX J

Demographic Information

Questionnaire

Demographic Information

Please furnish the following information for demographic/statistical purposes. Note that this information is anonymous.

1. Age \_\_\_\_\_
2. Sex: M \_\_\_\_\_ F \_\_\_\_\_
3. Specialty area (clinical psychology, social work, etc.)  
.....  
\_\_\_\_\_
4. Degree held (M.A., M.S., Ph.D., etc.) \_\_\_\_\_
5. Number of years counseling-psychotherapy experience \_\_\_\_\_
6. Approximate number of hours of coursework taken in  
psychological testing \_\_\_\_\_
7. Approximate number of hours of coursework taken in  
projective techniques \_\_\_\_\_
8. Number of years experience in psychological testing/  
evaluation \_\_\_\_\_
9. Number of years experience with projective techniques  
.....  
\_\_\_\_\_
10. Number of years experience with MMPI \_\_\_\_\_
11. Number of years experience with WAIS \_\_\_\_\_
12. Approximate number of Rorschachs given \_\_\_\_\_
13. Approximate number of MMPI's given \_\_\_\_\_
14. Approximate number of WAIS given \_\_\_\_\_

15. Place an X by those tests which you routinely use in psychological evaluations: Rorschach\_\_\_, WAIS\_\_\_, MMPI\_\_\_, TAT\_\_\_, Human Figure Drawings\_\_\_.
16. Do you now engage, or have you engaged in the last five years, in psychotherapy as a routine part of your work?  
Yes \_\_\_ No \_\_\_
17. Do you now, or have you within the last five years, done psychological testing/evaluation as a routine part of your work? Yes \_\_\_ No \_\_\_
18. What is your basic theoretical orientation (psychoanalytic, behavioral, Gestalt, etc.)? \_\_\_\_\_



APPENDIX K

Trait Rating Instrument  
(Judges to Fill Out for Client)

Trait Rating Instrument

The client about whom you have been given information was asked to describe herself using the following check list.

Here are the instructions which she was given:

"I would like for you to describe yourself as you are right now. Do not describe yourself as you would like to be or as you think other people would describe you. Do not describe yourself as you think you used to be or as you might be in the future. Please describe yourself only as you think you are at this time.

For each characteristic listed below, place an X by the statement which you feel describes you best on that particular characteristic.

Do not discuss with me how you are marking the characteristics as you mark them. When you have finished, turn your paper over and let me know you have finished marking it."

At this time, please mark the check list as you think the client about whom you were given information marked it. For each characteristic, place an X by the statement which you believe the client marked in describing herself. That is, using the checklist, try to describe the client as you

believe she described herself. Do not describe yourself.  
Try only to describe the client as you believe she described  
herself.

1. Ambitious

Very much like me	.....	.....
Like me	.....	.....
Slightly like me	.....	.....
Not much like me	.....	.....
Not at all like me	.....	.....

2. Independent; self-sufficient

Very much like me	.....	.....
Like me	.....	.....
Slightly like me	.....	.....
Not much like me	.....	.....
Not at all like me	.....	.....

3. Energetic

Very much like me	.....	.....
Like me	.....	.....
Slightly like me	.....	.....
Not much like me	.....	.....
Not at all like me	.....	.....

4. Orderly

Very much like me	.....	.....
Like me	.....	.....



- |     |                         |       |       |
|-----|-------------------------|-------|-------|
|     | Very much like me       | ..... | ..... |
|     | Like me                 | ..... |       |
|     | Slightly like me        | ..... |       |
|     | Not much like me        | ..... |       |
|     | Not at all like me      | ..... |       |
| 9.  | Cheerful; even-tempered |       |       |
|     | Very much like me       | ..... | ..... |
|     | Like me                 | ..... |       |
|     | Slightly like me        | ..... |       |
|     | Not much like me        | ..... |       |
|     | Not at all like me      | ..... |       |
| 10. | Anxious; fearful        |       |       |
|     | Very much like me       | ..... | ..... |
|     | Like me                 | ..... |       |
|     | Slightly like me        | ..... |       |
|     | Not much like me        | ..... |       |
|     | Not at all like me      | ..... |       |
| 11. | Non-conforming          |       |       |
|     | Very much like me       | ..... | ..... |
|     | Like me                 | ..... |       |
|     | Slightly like me        | ..... |       |
|     | Not much like me        | ..... |       |
|     | Not at all like me      | ..... |       |

12. Unhealthy; frequently ill

Very much like me .....  
\_\_\_\_\_

Like me .....  
\_\_\_\_\_

Slightly like me .....  
\_\_\_\_\_

Not much like me .....  
\_\_\_\_\_

Not at all like me .....  
\_\_\_\_\_

13. Caring; nurturing

Very much like me .....  
\_\_\_\_\_

Like me .....  
\_\_\_\_\_

Slightly like me .....  
\_\_\_\_\_

Not much like me .....  
\_\_\_\_\_

Not at all like me .....  
\_\_\_\_\_

14. Physically attractive

Very much like me .....  
\_\_\_\_\_

Like me .....  
\_\_\_\_\_

Slightly like me .....  
\_\_\_\_\_

Not much like me .....  
\_\_\_\_\_

Not at all like me .....  
\_\_\_\_\_

15. Often am the one "in charge"

Very much like me .....  
\_\_\_\_\_

Like me .....  
\_\_\_\_\_

Slightly like me .....  
\_\_\_\_\_

Not much like me .....  
\_\_\_\_\_

Not at all like me .....  
\_\_\_\_\_

Now that you have completed the checklist as you believe the client completed it, please go back and indicate for each characteristic on the checklist how confident you are that you marked it in the same way that the client did. You will state this confidence estimate in terms of the probability that your marking of the item is the same as that of the client.

This means that your confidence estimates will be made on a scale ranging from 20 through 100, but using only the even tens (i.e., 20, 30, 40, 50, 60, 70, 80, 90, 100). Thus, if you think the chances of your having market the item the same as the client marked it are 70 out of 100, write 70 in the blank to the far right of each item. If you think that your chances are 8 out of 10, or 80%, write 80 in the blank. If you are certain that your marking of an item is the same as the way the client marked it, you should write 100. If you feel that your marking of an item has no more than a random chance of being the same as the client marked it, you should write 20 in the blank.

APPENDIX L

Trait Rating Instrument  
(Judges to Fill Out for Self)



Trait Rating Instrument

I would like for you to describe yourself as you are right now. Do not describe yourself as you would like to be or as you think other people would describe you. Do not describe yourself as you think you used to be or as you might be in the future. Please describe yourself only as you think you are at this time.

For each characteristic listed below, place an X by the statement which you feel describes you best on that particular characteristic.

Do not discuss with me how you are marking the characteristics as you mark them. Let me know when you have finished.

## 1. Ambitions

Very much like me      .....  
\_\_\_\_\_

Like me      .....  
\_\_\_\_\_

Slightly like me      .....  
\_\_\_\_\_

Not much like me      .....  
\_\_\_\_\_

Not at all like me      .....  
\_\_\_\_\_

## 2. Independent; self-sufficient

Very much like me      .....  
\_\_\_\_\_

Like me      .....  
\_\_\_\_\_

Slightly like me      .....  
\_\_\_\_\_

- Not much like me .....  
\_\_\_\_\_
- Not at all like me .....  
\_\_\_\_\_
- 3. Energetic
- Very much like me .....  
\_\_\_\_\_
- Like me .....  
\_\_\_\_\_
- Slightly like me .....  
\_\_\_\_\_
- Not much like me .....  
\_\_\_\_\_
- Not at all like me .....  
\_\_\_\_\_
- 4. Orderly
- Very much like me .....  
\_\_\_\_\_
- Like me .....  
\_\_\_\_\_
- Slightly like me .....  
\_\_\_\_\_
- Not much like me .....  
\_\_\_\_\_
- Not at all like me .....  
\_\_\_\_\_
- 5. Socially withdrawing; don't like to be with others
- Very much like me .....  
\_\_\_\_\_
- Like me .....  
\_\_\_\_\_
- Slightly like me .....  
\_\_\_\_\_
- Not much like me .....  
\_\_\_\_\_
- Not at all like me .....  
\_\_\_\_\_
- 6. Not given to display of emotions; usually don't show feelings
- Very much like me .....  
\_\_\_\_\_
- Like me .....  
\_\_\_\_\_

- Slightly like me .....  
\_\_\_\_\_
- Not much like me .....  
\_\_\_\_\_
- Not at all like me .....  
\_\_\_\_\_
- 7. Suspicious; do not trust others
- Very much like me .....  
\_\_\_\_\_
- Like me .....  
\_\_\_\_\_
- Slightly like me .....  
\_\_\_\_\_
- Not much like me .....  
\_\_\_\_\_
- Not at all like me .....  
\_\_\_\_\_
- 8. Intelligent
- Very much like me .....  
\_\_\_\_\_
- Like me .....  
\_\_\_\_\_
- Slightly like me .....  
\_\_\_\_\_
- Not much like me .....  
\_\_\_\_\_
- Not at all like me .....  
\_\_\_\_\_
- 9. Cheerful; even-tempered
- Very much like me .....  
\_\_\_\_\_
- Like me .....  
\_\_\_\_\_
- Slightly like me .....  
\_\_\_\_\_
- Not much like me .....  
\_\_\_\_\_
- Not at all like me .....  
\_\_\_\_\_
- 10. Anxious; fearful
- Very much like me .....  
\_\_\_\_\_
- Like me .....  
\_\_\_\_\_

- Slightly like me .....  
\_\_\_\_\_
- Not much like me .....  
\_\_\_\_\_
- Not at all like me .....  
\_\_\_\_\_
- 11. Non-conforming
- Very much like me .....  
\_\_\_\_\_
- Like me .....  
\_\_\_\_\_
- Slightly like me .....  
\_\_\_\_\_
- Not much like me .....  
\_\_\_\_\_
- Not at all like me .....  
\_\_\_\_\_
- 12. Unhealthy; frequently ill
- Very much like me .....  
\_\_\_\_\_
- Like me .....  
\_\_\_\_\_
- Slightly like me .....  
\_\_\_\_\_
- Not much like me .....  
\_\_\_\_\_
- Not at all like me .....  
\_\_\_\_\_
- 13. Caring; nurturing
- Very much like me .....  
\_\_\_\_\_
- Like me .....  
\_\_\_\_\_
- Slightly like me .....  
\_\_\_\_\_
- Not much like me .....  
\_\_\_\_\_
- Not at all like me .....  
\_\_\_\_\_
- 14. Physically attractive
- Very much like me .....  
\_\_\_\_\_
- Like me .....  
\_\_\_\_\_



APPENDIX M

Client Information  
(Stereotype Information)

Client Information

You will be asked to make some clinical judgments about a client. This client is a twenty-one (21)-year-old female. She is a student at a state university in the field of computer science. She is being seen as a client at the counseling center of her university.

APPENDIX N

Client Rating of Likability and Similarity

and

Client Diagnosis



Please answer the following questions about the client.  
 For the first two questions, place an X above one of the five  
 statements on the scale which best answers the question. For  
 the third question, write out your diagnosis of the client.

(1) How well did you like the client as a person?  
 .....

---

Very	Fairly	Slightly	Not Very	Not at
Much	Well		Much	All

(2) How much like you was the client"  
 .....

---

Very	Quite a Bit	Somewhat	Not Much	Not at All
Much	Like Me	Like Me	Like Me	Like Me
Like Me				

(3) What is your diagnosis of the client?

APPENDIX O

Data

Stereotype Group

Subjects:	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>	<u>7</u>	<u>8</u>	<u>9</u>	<u>10</u>
Accuracy Score - Predictive Task	19	20	19	20	18	18	18	20	14	21
Accuracy Score - Evaluative Task	14	17	12	14	17	15	13	17	14	16
Confidence Score - Predictive Task	84	67	47	56	71	55	71	71	27	50
Confidence Score - Evaluative Task	95	61	58	41	82	47	77	73	21	61
Assumed Similarity Score	11	19	20	12	23	17	14	9	14	16
Real Similarity Score	20	23	20	20	23	17	18	19	22	19

Stereotype-Video Group

Subjects:	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>	<u>7</u>	<u>8</u>	<u>9</u>	<u>10</u>
Accuracy Score - Predictive Task	21	21	17	20	19	24	20	17	24	14
Accuracy Score - Evaluative Task	15	12	14	14	9	8	13	16	15	5
Confidence Score - Predictive Task	63	65	62	59	73	33	30	90	60	54
Confidence Score - Evaluative Task	55	85	79	80	82	35	27	94	66	63
Assumed Similarity Score	20	24	26	25	22	19	35	25	32	18
Real Similarity Score	24	18	23	18	21	21	29	28	24	18

Stereotype-Video-Diagnostics Group

Subjects:	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>	<u>7</u>	<u>8</u>	<u>9</u>	<u>10</u>
Accuracy Score - Predictive Task	21	22	16	16	21	19	14	18	22	16
Accuracy Score - Evaluative Task	15	13	12	12	19	15	16	11	25	12
Confidence Score - Predictive Task	61	55	78	59	83	65	66	69	72	47
Confidence Score - Evaluative Task	82	72	93	68	97	80	65	83	87	81
Assumed Similarity Score	23	22	17	17	30	20	22	19	24	19
Real Similarity Score	20	22	19	17	21	19	22	21	22	23

Likability of ClientRating on Five-Point Scale

Subjects:	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>	<u>7</u>	<u>8</u>	<u>9</u>	<u>10</u>
Stereotype Group	5	4	3	3	4	4	4	3	4	4
Stereotype- Video Group	4	3	4	4	4	4	3	4	4	3
Stereotype- Video- Diagnostics Group	4	4	4	5	4	4	3	4	4	4

Similarity to ClientRating on Five-Point Scale

Subjects:	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>	<u>7</u>	<u>8</u>	<u>9</u>	<u>10</u>
Stereotype Group	3	2	3	3	2	3	3	3	3	2
Stereotype- Video Group	3	2	2	3	2	2	2	2	3	2
Stereotype- Video- Diagnostics Group	2	4	3	3	1	3	2	2	4	3