THE DEVELOPMENT AND VALIDATION OF A SCALE TO MEASURE EMPATHY

Thesis for the Degree of Ph. D.
MICHIGAN STATE UNIVERSITY
James Lee Chapman
1966

This is to certify that the

thesis entitled

THE DEVELOPMENT AND VALIDATION OF

A SCALE TO MEASURE EMPATHY

presented by

James Lee Chapman

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Guidance and Counseling

Norman Kagan

Major professor

Date May 20, 1966

JAN 13 1968 22 pd F5
FEB 13 1968 157
JUL 10 1968 5 Aug 7/15/68
JUL 17 1969 24
JUL 24 1969 130
OCT 15 1969
OCT 28 1968 132
NOV 27 1968 130
JAN 8 1969 192
MAR 5 1970 pd 32
APR 20 1970 1140
APR 2 1971 58
20 1971 142
NOV 9 1971 206
94
JUN 8 1973 158
JUN 20 1973 161
DEC 5 1974 212
346
338
A 3 5 0
206

ABSTRACT

THE DEVELOPMENT AND VALIDATION OF A SCALE TO MEASURE EMPATHY

by James Lee Chapman

The purpose of this study was to develop and validate an instrument that tested a subject's ability to identify the emotion or emotions expressed by another in a series of videotaped excerpts from actual counseling interviews.

The instrument was composed of selected scenes of videotaped recordings of counseling interviews. Each scene was followed by four to seven descriptive adjectives. Subjects were to respond to each adjective on a continuum indicating how strongly they possessed the feeling described by the adjective at the end of each scene. Two criterion groups, judged high and low in empathic regard, responded to the instrument and generated items from a list of 57 adjectives expressing various feelings. The final list of four to seven descriptive adjectives was produced from the list of 57 adjectives.

The instrument was administered to 53 subjects in two universities for the purpose of validation. In order to cross-validate the results of the item analysis of the instrument, it was administered to another 88 persons enrolled in three National Defense Education Act Summer Institutes (1964).

The instrument was scored for "right" and "wrong" answers on 172 scorable items from the original 280. Correct and incorrect answers were determined on the basis of the responses given by the criterion

groups of high and low empathizers. The total sample was increased from 88 to 148 by adding the data from two additional National Defense Education Act Summer Institutes (1964). The purpose of scoring the instrument and increasing the sample was to determine the predictive validity and reliability of the instrument. In addition, a t-test for significance among the sub-groups of the sample and a factor analysis to determine commonality of meaning in adjectives used were made possible by scoring the instrument.

The desired data from the subject's responses were placed on IBM cards for statistical analysis. Chi-square was used in the statistical treatment of the data for validation and cross-validation purposes with the test of significance being established at .20 or higher. The two-by-two chi-square contingency was employed on the response scale. The response scale was split at the point at which the median occurred in the response pattern of each item. Analysis of variance to ascertain predictive validity; t-test to determine significant differences among means of the sample sub-groups; Kuder-Richardson formula 20 to establish reliability and factor analysis for commonality of meaning were also used in working with the data.

A total of 65 items of the original 280 included met the test of significance for the purpose of validation, and nine items were found to be significant in cross-validating the instrument.

These nine items which held up on cross-validation included the adjectives "bitter" (three of the items), "disgusted," "scared," "helpless," "confused," "mixed-up," "lonely," and "satisfied." The large number of negative adjectives results from the fact that the adjectives

in general were descriptive of negative affect. There was no major concentration on any one client or episode although nearly all of the items were toward the end of the scale. Because the nine items represent about a fifth of the 65, and one would expect about a fifth of the items to show significance by chance at the 20% level, it would appear that there is nothing more than chance relationship between the item responses and the criterion.

The predictive validity of the instrument was nonexistent. The analysis of variance values for three effects (sex grouping, high-low grouping and interaction) were low enough to indicate that should the sample be increased to infinity, significance would not be produced. The correlation of reliability was .441. The t-test did not find a significant difference between the highest and lowest means of the two sub-groups involved. There were five factors operative in a ten-factor matrix. A total of 29 scenes and 62 items were found in the five factors.

THE DEVELOPMENT AND VALIDATION OF A SCALE TO MEASURE EMPATHY

By

James Lee Chapman

A THESIS

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

DOCTOR OF PHILOSOPHY

College of Education
Department of Guidance, Personnel Services
and Educational Psychology

1966

ACKNOWLEDGMENTS

I wish to express sincere appreciation to my adviser, Dr. Norman Kagan, for his kind assistance and helpful suggestions in the preparation and presentation of this thesis. I am also grateful to Drs. Marion Kinget, James Costar and Floyd Parker for their assistance as members of my guidance committee. In addition, I express gratitude to Drs. Walter Johnson and David Krathwohl who were members of the guidance committee earlier in my program.

I wish to thank my wife, Mary, and son, Ronald, who were willing to be without a husband and father on many occasions throughout my entire doctoral program.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER I

INTRODUCTION

> What would be desirable preparation for the person who is to receive training as a therapist? . . . If the student is to become a therapist, the more he has been able to achieve of empathic experiencing with other individuals, the better will his preparation be. (Rogers, 1951, p. 437)

Theorists of both the psychoanalytic and client-centered schools of counseling have emphasized that one important therapist variable, independent of technique, is the therapist's ability to understand the client empathically (Cartwright and Lerner, 1963). Although, therorists and practitioners of different positions agree upon the importance of the therapist's empathic ability, studies investigating this variable are few when compared to those relating to intelligence, interests and personality development. The confusion which surrounds the definition of empathy could be one explanation for the dearth of studies in this area. The inability to replicate human behavior prior to the advent of motion picture and television facilities in university and research settings may be another.

Since one important therapist's variable in counseling is the therapist's ability to understand the client empathically, the extent to which a counselor possesses empathic regard could be one major factor leading to success in counseling. The instrument developed in the present study is an attempt to provide an aid to assist counselor educators in predicting counseling success by potential counselors.

Background of the Study

Studies relating to the concept of empathy, prior to 1956, have been confusing since theoretical and operation definitions do not agree. Cronback (1955) indicates that studies of "social perception" have been chiefly concerned with differences among perceivers either in terms of their accuracy or in terms of their tendency to view others as similar to themselves. These studies have usually been built around a particular operation in which a judge "predicts" how another person will respond in a given situation or to a test. The judge's response is then compared with characteristics of the judge himself. Often, for example, both persons describe themselves on a personality inventory, and the judge is then asked to fill out the inventory as he thinks the other person did. The extent to which the prediction agrees with the other person's actual response is taken as a measure of the judge's accuracy of social perception (or "empathy," "social sensitivity," "diagnostic competence"). Scores obtained in this manner are difficult to interpret, and several investigators have reported low consistency for them (Crow, 1954; Gage, 1953; Stone and Leavit, 1954).

Recently attempts have been made by Astin (1957), O'Hern (1964), and Stefflre (1962) to study empathy by confronting the subject with a simulated counseling situation which has been recorded on audio tape. The recording is made from a "staged" counseling interview and is designed to be stopped at the appropriate time. The subject is required to choose a proper response from a number of alternatives given. The situation is structured in such a way that the behavioral variable or emotion in question can be revealed and responded to by the subject. Kelly and Friske

(1951) have employed this approach in assessing behavioral variables of clinical psychology trainees in Veteran Administration hospitals.

Buchheimer (1964) is currently conducting research on empathy by using an actual counseling situation. The situation is recorded on film and the subject is requested to respond verbally at selected points in the interview as though he were the counselor. The subject's response is recorded on an audio tape recorder. The results of most studies of empathy continue to indicate that the definition of this variable, as espoused by the researcher, is not operationally consistent with the design. The confusion between theoretical definitions of empathy and operational definitions have contributed to a lack of insight into this important variable. It becomes necessary that the nature of this variable be investigated due to the important part it plays in counseling and other interpersonal relationships.

The possibility of uniting theoretical and operational definitions of empathy has been enhanced with the availability and use of closed circuit television. Researchers are able to record human behavior as it occurs in the counseling interview. A new research tool has been developed in the College of Education at Michigan State University through use of closed circuit television which is designed to assist in the study of empathy (Kagan, Krathwohl and Miller, 1963). The technique which is known as Interpersonal Process Recall (IPR) can be classified in the general area of stimulated recall methodology and permits identification of the nature of emotions expressed by clients in recorded behavioral episodes.

Need for the Study

In an extensive review of studies relating to empathy, Taft (1955) describes the different methods of measuring empathy, classifications of the tests used, the reliability of tests and concludes the article by listing the characteristics (i.e., age, sex, intelligence, training in psychology, esthetic ability, sensitivity and emotional stability) of those who should be good judges of others. Although these characteristics are given, one is still faced with the question as to whether empathic ability is being tested or simply the ability to predict basic personality traits of individuals or groups from a reservoir of stereotypic concepts and attitudes. The subject seems to be required in most of these tests of "empathy" to calculate and analyze the facts presented and come to a decision that requires a great amount of cognitive ability but little, if any, affective sensitivity ability.

The point is made by Allport (1937) that, "the theory of empathy is a peculiar blend, and must be regarded both as a theory of inference and a theory of intuition depending upon the coloring given it by different authors."

If one looks at empathy as having at least two dimensions, that of inference and intuition, the need continues for an instrument to be developed which measures a subject's affective sensitivity ability by ascertaining the momentary psychological state of another person. In order to ascertain the affective state of another, one must be capable of determining the feelings at the moment they are expressed. At the present time there is not a dependable test for empathy for as Buchheimer (1963) indicates, the empathic process has been studied from only a limited point of view.

The Problem

The purpose of this study is to develop and validate an instrument that tests a subject's ability to identify the emotion or emotions that are being expressed by another in a series of videotaped scenes from actual counseling interviews. The specific problem is to determine whether or not such an affective sensitivity instrument will significantly correlate with counseling effectiveness.

Definition of Terms Used

Instrument. In this study, the term "instrument" shall refer to the device used to obtain an objective score of affective sensitivity. The instrument itself consists of selected scenes of videotaped recordings of actual counseling interviews. Each scene is followed by four to seven descriptive adjectives, from which subjects are to respond to each word on a continuum from feeling strongly to not possessing the feeling at all.

Affective sensitivity. For the purpose of this study, affective sensitivity shall refer to the subject's ability to identify the momentary feelings of a client as revealed through the videotaped recording of the client.

Affective sensitivity score. The raw score a subject earns on the instrument shall be referred to as the affective sensitivity score. The higher scores shall indicate a greater degree of agreement with criterion groups of known empathizers.

Sample. Students enrolled in five summer institutes established by the National Defense Education Act in different areas of the United States comprised the sample.

Known empathizers. Those persons who by reputation were judged by their instructors or colleagues to be high in empathic regard are referred to as known empathizers.

Interpersonal Process Recall. The IPR technique operates in the following manner: A counselor and client conduct a counseling interview within the studios of the closed circuit television installation. The studio is curtained, the cameras are pre-set and unmanned so that a minimum of distraction exists. The interview is enacted and the two participants are videotape recorded on a split screen with head and torso head-on views enlarged as much as the screen permits. Immediately after the interview is concluded, the client witnesses a playback of the interview in the presence of another trained counselor. The trained counselor encourages the client to describe his feelings, interpret statements, and translate body movements at various times during the replayed interview. The trained counselor or subject may stop the playback and discuss recalled feelings and elaborate on meanings.

Limitations

The study is limited by the following:

1) There was no follow-up or replication of the study.
2) The criterion groups of high and low empathizers were selected by "reputation" among their peers or instructors rather than by an objective measure.
3) No systematic or equated stratification of samples was attempted.

Assumptions

The following assumptions are implicit in the study:

1) That client affective behavior in a counseling interview is measurable.
2) That client behavior observed and recorded at Michigan State University is not markedly different from that observed in other areas of the country.
3) That counselor educators can judge between effective and less effective counselors based upon their conduct during the counselor's professional education.
4) That colleagues can judge between counselors who are high and low in affective sensitivity.

Hypotheses

Counselor education has as one major goal the education of effective counselors. There seems to be general agreement in the field that many personal traits including one's attitudes toward others are included in counselor effectiveness (Weitz 1957 and O'Hern 1964). The present study investigates only one of these traits, that of affective sensitivity, under the assumption that this trait can be measured. The hypotheses to be tested in this investigation are:

1) A scale based upon actual counseling interviews will differentiate between persons high and low in empathic regard.
2) A scale based upon actual counseling interviews will produce a significant correlation between scores received on the affective sensitivity scale and counselor effectiveness.

Organization of the Thesis

Studies related to the concept of empathy and their findings will be summarized in Chapter II. The basic design and statistical procedures that were used in the construction of the instrument will be presented in Chapter III. A description of the treatment of the data and results of the analysis will be reported in Chapter IV. Chapter V will contain the summary, conclusions and recommendations.

CHAPTER II

PREVIOUS APPROACHES TO THE DEFINITION AND MEASUREMENT OF EMPATHY

The literature on empathy is found in various areas of psychology, from industrial to social and clinical psychology. Empathy is a respected concept in the German school of Verstehende Psychologie, which holds that "psychic processes are not to be described but to be intuitively understood" (English and English, 1958). Empathy or Einfühlung, in this context, is regarded as a faculty of immediate knowledge. This connotation has little acceptance in American psychology, in which empathy is partly synonymous with clinical inference. Sarbin (1960) pointed out the apparent distinction between clinical inference, which follows a general inference model, and intuition or immediate knowledge as a nonmediated process. He found that careful consideration of the elements of the intuitive process proved that it conformed to the logical model of inference, the underlying postulates being immediately available from past experience. The absence of a sound theory which would explain the nature of empathic interpersonal communication either as an intuitive, cognitive model or as an inferential, learning model has permitted the operational meanings or usages of the term "empathy" to become increasingly vague and contradictory.

Definitions: Empathy and Affective Sensitivity

The problem of defining empathy is an epistemological one--how do we "know" others in a counseling relationship? Attempts to define empathy have been as diverse and as inconclusive as the efforts of philosophers to describe the nature of knowledge. To establish an operational basis for the development of the present scale, various definitions and synonyms for empathy were compared to generate the testable concept of affective sensitivity.

The term, Einfuhlung, later translated by Titchener as empathy, was introduced in 1897 by Theodore Lipps to describe the process of aesthetic appreciation as a form of inner imitation. This imitative characteristic is distinctive of those usages of the term which postulate an identity between the empathizer and his object. Sensitivity to the affect of another, according to such a definition, depends on analogy, identification, or real similarity. This connotation is illustrated in the common sense idea of "putting yourself in the other fellow's place" which Allport (1954) found as a point of agreement in many definitions of empathy.

An important theoretical disparity exists between this group of definitions and those which stress the detachment and objectivity of the empathizer. Typical of the second class of definitions for empathy is the one given by English and English (1958):

> Apprehension of the state of mind of another person without feeling (as in sympathy) what the other feels. While the empathic process is primarily intellectual, emotion is not precluded, but it is not the same emotion as that of the person with whom one empathizes. The parent may empathize with the child's puny rage, feeling pity or amusement, whereas in sympathy he would feel rage along with the child. The attitude in empathy is one of acceptance and understanding; of an implicit "I see how you feel."

Carl Rogers (1959) emphasized the importance in counseling of the empathic understanding of the other person's "internal frame of reference." He described the state of being empathic as one of perceiving:

> . . . the internal frame of reference of another with accuracy, and with the emotional components and meanings which pertain thereto, as if one were the other person, but without ever losing the "as if" condition. Thus it means to sense the hurt or pleasure of another as he senses it, and to perceive the causes thereof as he perceives them, but without ever losing the recognition that it is as if I were hurt or pleased, etc. If this "as if" quality is lost, then the state is one of identification (1959, pp. 210-211).

The element of detachment, or as Rogers calls it, the "as if" condition, is present in the psychoanalytic definition of empathy, as given by Fenichel (1945, p. 511):

> . . . empathy consists of two acts: (a) an identification with the other person, and (b) an awareness of one's own feelings after the identification, and in this way an awareness of the object's feelings.

Katz (1963) advocated a four phase model for empathy, based on the formula given by Reik in Listening with the Third Ear. This finer division of the continuum from identification to objectivity suggests that the more closely one inspects the elements of an empathic act, the further one can amplify the definition. The semantic controversy as to whether empathy is immediate, unmediated knowledge gained through a process of identification, or a logical process of inference culminating in objective awareness, may depend on the duration of the act or sequence of acts denoted as being empathic. An immediate focusing of attention on the affective state of another simultaneous with his behavioral expression of that state may be only one brief facet of interpersonal sensitivity. The nature of this particular operation is essentially passive and receptive, and if it is of extended duration and constancy it becomes sympathetic rather than

empathic behavior. Buchheimer (1963) points out that a sympathetic person does not need to interact with the other person. In sympathy, he says, there is a paralleling of thought and feeling between the two or more individuals, whereas empathy connotes a converging of behavior. Thus an empathic process may begin with a receptive state of sensitivity to the affect being communicated by another, but the state rapidly shifts into covert inferential activity and into overt operations intended to elicit meaningful or needed communication from the other; and this would be merely one phase of a cyclical process of interaction leading to increasing congruence.

The plethora of different usages for the term empathy in the literature points to the need for a clear analysis of the nature and duration of the empathic act. Some authors, to side-step the semantic confusion, have contrived fresh paraphrasings of a more generic type, such as "interpersonal perception" (Gage and Cronbach, 1955) or "ability to judge people" (Taft, 1955). With a similar intent, though hopefully taking finer aim in conceptualizing, this study has attempted to steer clear of the theoretical shoals inherent in the discussion of empathy by purporting to measure only the more circumscribed trait of affective sensitivity. This trait is defined as the ability to detect and describe the immediate affective state of another; or in terms of communication theory, the ability to receive and decode affective communication.

Affective sensitivity would be a component of both sympathy and empathy, depending on the nature and duration of the interaction. The demonstration of affective sensitivity would depend upon an opportunity to observe--though not necessarily to interact with--the other. A high

degree of affective sensitivity would not suffice to ensure effectiveness as a counselor; a low score, however, on a measure of affective sensitivity would suggest unsuitability for counselor training, indicating that the person perceives the feelings of others inaccurately, perhaps on the basis of stereotyping, assumed similarity, projection, or other distorting premises.

The Measurement of Empathy

Investigators attempting to measure empathy have used approaches as varied as the efforts to define the term empathy. The earliest ostensible study of empathy was conducted by Gordon in 1934. He instructed his subjects to tell whether the right or the left arm was raised in four photographs of a Mexican male, and four reversals of these images. This procedure was an operationalizing of the original Lippsian concept of aesthetic empathy, in its postulation of an identity of feeling in the musculature of the empathic perceiver. Subjects were rated as empathic if they overtly or covertly attempted to emulate the attitude of the photographed image.

The current wave of interest in devising instruments for estimating and correlating empathic ability mounted in the 1950's. As yet, however, only one normative test has been reported in the Fifth Mental Measurements Yearbook (Buros, 1959), a group paper and pencil device entitled The Empathy Test (Kerr and Speroff, 1954). Designed for use in industrial selection, the test was intended to measure "the ability to put yourself in the other person's position, establish rapport, and anticipate his feelings, reactions and behaviors" (1954, p. 269). The researchers then translated this definition into operational terms the ability to rank as

the average person would rank them: (a) the popularity of musical forms, (b) reading and buying preferences for magazines, and (c) the annoyance magnitude of such experiences as hearing a person chew gum. Validation criteria were obtained from national survey results and related research findings. Robert L. Thorndike commented on The Empathy Test as follows:

> There appears to be no inherent validity in the operations called for in this test, and so its validity must be established empirically through its ability to predict socially important criteria, or its relationships to other variables that would make it a meaningful construct. The manual reports several studies presenting evidence on the validity of the test, and certain of these appear quite impressive. However, the relatively few studies by persons not associated with the author have tended to yield predominately negative results. Unless the positive results reported in the manual are verified in the findings of other workers, this test cannot be recommended as either a useful practical device or a contribution to the description and understanding of an individual (Buros, 1959, p. 120).

In a similar endeavor to measure "massempathy," Norman and Ainsworth (1954) administered two forms of the GAMIN Personality Inventory at an interval of two weeks apart. On the first administration, the subjects answered for themselves; on the second, they responded as they thought most other people of their own age and sex would respond. The criterion for scoring the responses on the second form as being "massempathic" was based on the scores of the total group on the first form. While the authors were careful to define their terms appropriately, the demonstration of ability to conform to the average response pattern on a personality scale cannot be used to predict sensitivity in individual interaction. As in the case of Kerr and Speroff's test the subject is required to take the role of the generalized other, rather than that of a specific other.

Gage and Cronbach (1955) delineated four components of experimental design in studies of interpersonal perception:

a. The Judge whom the experimenter is attempting to measure.
b. The Other(s) whom the Judge is asked to interpret.
c. The Input, or information concerning the Other which is available to the Judge.
d. The Outtake, i.e., the statement or predictions about the Other obtained from the Judge. (p. 412)

The two studies discussed above attempted to assign a score to the judges on the basis of their sensitivity to the generalized other. Let us now turn to those research studies in which empathy for a specific other has been estimated and attempt to distinguish among these on the components of input and outtake, which between them determine the nature and availability of validating criteria.

Predictive studies of empathy. One of the most popular approaches to the measurement of empathy was devised by Dymond (1949), who defined empathy as "the imaginative transposing of oneself into the thinking, feeling and acting of another and so structuring the world as he does." As a measure of this ability she found the correspondence between subjects' ratings of others known to them through group interaction and the actual self-ratings made by the others. This approach thus emphasized the outtake component by requiring judges to make global predictions about others. A refinement of the original test appeared in 1950; it required rankings on a 5-point scale on each of six traits: superior-inferior, friendly-unfriendly, leader-follower, shy-self-assured, sympathetic-unsympathetic and secure-insecure.

Other studies employing a predictive criterion followed in the wake of Dymond's articles. Bender and Hastorf (1950) asked subjects to predict the responses of acquaintances on three personality scales and found low positive correlations between obtained and predicted scores. Cowden (1955) had 29 married couples take the California Test of

Personality and then predict how his or her partner would answer the questions; he considered the accuracy of prediction to be a measure of empathy.

Predictive studies have an inherent weakness in the tendency of subjects to avoid negative or unfavorable ratings of themselves and others. The results usually produce a high clustering of scores. The authors cited above have attempted to apply corrections for the biasing factors of projection and assumed similarity. Hastorf, Bender and Weintraub (1955) found that subjects were cautious in predicting, and used the middle "often" and "seldom" alternatives. They remark:

> In the light of this finding, we learn why S's who themselves had definite preferences and close associates who reply with midscale responses tended to obtain high refined empathy scores. In view of these adventitious findings it would then become possible to predict an S's "empathic ability" without even scoring his predictions. We must therefore impugn the refined empathy score as a true measure of empathic ability. (p. 343)

A similar conclusion was reached by Lindgren and Robinson (1953) who found in evaluating Dymond's empathy rating scale that predictions tended to conform to a cultural stereotype or norm. They found the scale too cumbersome for use as a normative test and, finally, questioned the validity of the Dymond scale in that a high score for predicting test responses might not denote actual empathy. In a similar indictment of tests stressing the predictive component, Cronbach (1955, p. 191) comments:

> Social perception research has been dominated by simple, operationally defined measures. Our analysis has shown that any such measure may combine and thereby conceal important variables, or may depend heavily on unwanted components. Only by careful subdivision of global measures can an investigator hope to know what he is dealing with. Our analysis makes especially clear that the investigator of social perception must develop more explicit theory regarding the constructs he intends to study, so that he can reduce his measures to the genuinely relevant components.

Thus, the ability to predict global attributes of another is a complex and fortuitous variable to investigate, and studies of this type have not contributed valid or reliable methods for assessing interpersonal sensitivity.

The authors of two further studies attempted to avoid the problems of using a predictive criterion for empathy. Lifton (1958) asked the members of a class in music education to rank their fellow members as "someone--sensitive to the feelings others may have about the problems that face them." Rosalind Dymond Cartwright appended to the article the comment that Lifton was measuring only the reputation for empathy. Lesser (1958) devised an Empathic Understanding Scale of 16 items such as the following: "I understand the client's feelings well"; or "I seem to see the client as he sees himself." A corresponding version of the scale allowed the client to rate the counselor. The scales were completed by counselor-client pairs, and client progress in therapy was estimated by the amount of change of the self-concept toward the ideal self, measured by Q-technique. The correlations between client progress and Empathic Understanding as measured by the counselor self-rating scale or by the client rating scale were not significant and were, in fact, negative. However, as in the case of the previously described rating scales, the mean ratings by both clients and counselors were significantly beyond the midpoint of the scale. Thus, the validity of the Empathic Understanding Scale must be held in question. Neither Lifton's nor Lesser's approach to the rating of empathic ability can be adapted to yield normative data.

Situational tests of empathy. The most felicitous approaches for the generation of measures of empathy or sensitivity are those which stress the input component in their experimental design. These studies provide a situation or situations involving combinations of visual auditory or kinesthetic stimuli so as to provide a standard experience to elicit empathic responses from the subject.

Reid and Snyder (1947) played phonographic recordings of passages from actual interviews, asking the subjects to name the feelings expressed in each of 40 client statements within a 15-second time limit. They stated that approximately 80 percent of the 15 subjects agreed 50 percent of the time on the designation of feelings. The authors did not attempt, however, to define the "correct" answers or to develop an evaluative instrument.

Astin (1947) devised a situational test of empathy by having ten client statements recorded on audiotape by a professional actor and requiring subjects to respond to the statement as if they were the counselor. The performance on this task of eight subjects who were counselors was significantly superior to that of the eight non-counselor subjects.

Weinstein and McCandless (1959) used a more restricted visual situation to derive a measure of empathy in medical students: they showed a slide of a doctor talking to a mother in a sickroom. Subjects were required to choose from multiple-choice items those which best described what was happening in the picture. The empathy score was determined by the correspondence between the student's designation of the mother's feeling and the doctor's concern. The validity of this procedure is suspect, as it seems to postulate that an empathic person will project his empathic needs onto others.

O'Hern (1964) constructed an instrument to measure the sensitivity of trainee counselors. The instrument consisted of 30 different client problems portrayed by actors and recorded on audiotape. Thirteen individuals with varied backgrounds responded to the stimulus tape. This group included full-time counselors, graduate students in counseling, and undergraduates. The refined instrument included those responses that met modal jury agreement with variance no greater than one in either direction. The four answers for each response were weighted according to degrees of rightness and degrees of wrongness as determined by a panel of seven judges who had had counseling experience, adhered to various counseling philosophies, and were involved in counselor education. The final instrument contained 29 client problems and 114 responses. A total of 212 students, enrolled in seven different Guidance and Counseling Institutes, participated in the study. O'Hern found no significant relationship between degree of sensitivity possessed by potential counselors and grades received in counseling practicum courses. However, the potential counselors judged most effective in counseling practice scored significantly higher on the O'Hern scale than those judged least effective. An instrument such as this would lend itself to the development of standardized norms if further refinement and validation studies were carried out.

It is questionable whether tests based solely on the verbal text of an interview or on the audio recording alone are sufficiently complete to give all the cues and clues needed for empathic understanding. Sullivan pointed out:

> A verbatim record of an interview, until it has been heavily annotated, is almost invariably remarkably misleading. I have had some recordings of interviews which I have regarded as astonishingly good teaching material, but when I have sprung these

> on intelligent colleagues, I have often found them barking up trees I hadn't seen--if, indeed, such trees were ever there, and I came to realize that they weren't. In other words the complete meaning of a conversation is not to be found in the verbatim verbal context of the communication, but is reflected in all sorts of subtle interplay. For example, very slight changes of tone suggesting the faintest hint of irritation on the part of the psychiatrist often switch the patient from an attempt at concealment to a very reasonable compromise between what he thinks it is safe to tell and what the facts may have been. Such things do not appear in the most perfect verbal record. Thus, to give a third person a notion of all that happened in an interview, one would have to annotate the written record by adding the impressions that went with different statements, explaining why things were put as they were, and so on; only in this way could the richness of the interchange . . . begin to be apparent. (Sullivan, 1954, p. 19)

As it becomes easier to make visual as well as auditory records, the possibility of using the total stimuli in a situation becomes more possible, and Sullivan's interpreter becomes unnecessary and undesirable. Two studies by Buchheimer (1964) and Rank (1964), as well as the present one, have moved in this direction. Buchheimer used a filmed counseling interview in which the subject free responded verbally at selected points in the film as if he were the counselor, as one phase of his total study of empathy.

Buchheimer's results are not yet available. His free response situation presents obvious scoring problems, and the test requires a language laboratory or similar situation to group administer. The approach of having the subject play the role of the counselor has obvious and desirable face validity. It should, perhaps, be considered more a "work sample" type test in the prediction of counselor success than a measurement of empathy <u>per se</u>. This has its advantage in that work sample tests have proven to have greater predictive validity in a number of comparisons than aptitude variables. Thus measures of predictor variables

such as intelligence tests or special ability and aptitude tests have been repeatedly outshone by work sample tests or measures of previous related work in fields other than counseling. On the other hand, however, a predictor variable may add cases which might be missed by a work sample test. Subjects that have not yet matured or who have had sample test but may do quite well in a predictor test of related variables. A test of empathy would appear to fall in this category and thus be of value.

Still closer to the approach proposed in this study, and indicative of the fact that the approach may prove fruitful, is the work of Rank (1964). Rank explored the relationship between trainee perceptions of film-presented counseling and post-training counseling competence. Test subjects responded from "strongly agree" to "strongly disagree" concerning statements about the client and the counselor presented on movie film. His "Film Test of Counselor Perceptions" produced cross-validation predictive correlations of .41 on the criterion of trainee competence after a practicum. The test also measured significant increase in scores during the practicum. The test was scored with a complex weighting system. Rank plans to revalidate and improve this scale. At present it is limited to one counselor and client in a variety of scenes.

Summary of Literature

From this survey of the status of research in empathy it appears that there is a dearth of valid tests for the assessment of sensitivity to affective communication. Of all the possibilities, Rank's comes closest to being practical. Still, there is a need for a test which will be convenient to administer, provide normative scores, and present a sufficient number and variety of situations for the elicitation of

responses which demonstrate sensitivity to the affect being expressed. The use of videotaped sequences from actual counseling interviews appears to offer a highly realistic yet standardized mode of presentation of affective stimuli in items which would discriminate between those persons sensitive and those not sensitive to the expression of feeling in others. The variety of clients and counselors recorded in the IPR process would appear to be a rich source of material from which to create such test items. The availability of the IPR interpretation by the client seems to add a dimension beyond that found in previous work to date.

CHAPTER III

DESIGN OF THE INSTRUMENT AND METHODOLOGY

The counseling interview protocols used in constructing the instrument were taken from a larger research project conducted in the College of Education at Michigan State University (N.D.E.A. Grant # 7-32-0410-216). The affective sensitivity scale was developed from the counseling interviews obtained by employing the IPR technique.

Construction of the Instrument

Selection of videotaped scenes. The scale was constructed by reviewing videotapes of counselor trainees involved in the counselor training part of the IPR study. A team of two of the principal investigators and four of the graduate assistants reviewed the tapes and selected scenes which, in view of the interrogation and the content itself, indicated that some describable client emotion was present. Typically these were changes in mood depicting the onset of client anger, fear, embarrassment, or avoidance behaviors. In most scenes selected there was a movement on the part of the client from one feeling state to another. The criteria were (a) that some emotion was displayed in the content and (b) that the nature of the emotion had been revealed during interrogation. In some instances the display of emotion was very subtle, and in other instances quite blatant. There was an attempt to choose scenes which ranged along the continuum of obviousness. The

purpose in selecting scenes in this manner was to differentiate between high and low empathizers. The final form contained 41 such scenes involving 11 different clients and counselors. The number of scenes for each client varied from two to six so that there was a variable exposure to the different clients and counselors. Both male and female clients were included. In general they were clients experiencing normal problems of interpersonal conflicts, social maturity and educational planning. The scenes varied in the emotional depth of the content and the kinds of problems covered. The counselors were both male and female and varied considerably in their skill. The use of male and female counselors was done so that differences of affective sensitivity between both groups could be determined. Although most were beginning counselors, some experienced counselors were included. Several scenes of two married women as clients were included, but all other scenes were of high-school age students.

Selection of criterion groups for generating items. The two criterion groups were composed of subjects selected from several different populations. The first group of subjects consisted of seven professional counselors at Michigan State University, all of whom held a doctorate in the area of guidance and counseling, had professional counseling experience and were judged to possess a high degree of empathy by other faculty members of the Guidance and Counseling Department of Michigan State University. The second group of subjects consisted of five doctoral candidates judged to possess a high degree of empathy by members of the Department of Guidance and Counseling of Michigan State University. The third group of subjects consisted of nine

beginning counselors judged to be high in empathy by supervisors who assisted in training each during a term of counseling practicum. These three groups of subjects (N=21) composed the criterion group of high empathizers.

The criterion group of low empathizers (N=16), judged so by present or former instructors who were asked to make such judgment due to reputation in evaluating people, was composed of six beginning counselors, four superintendents of schools, four high school counselors, and two undergraduates majoring in education. The composition of the two criterion groups is consolidated in the following table:

Table 1 Composition of the criterion groups

High Empathizers		Low Empathizers	
Professional Counselors	7	Superintendent of Schools	4
Doctoral Students	5	Beginning Counselors	6
Beginning Counselors	9	Secondary School Counselors	4
		Undergraduates	2
Total	21	Total	16

Generating items. The two criterion groups were shown the selected excerpts in a studio of the closed circuit television facility of Michigan State University. The videotape recording was stopped after each scene. The subjects were instructed to try to feel as the client last felt, and responses were obtained from a list of 57 adjectives expressing various feelings (Appendix A). The list of adjectives was chosen from those compiled by Osgood (1953) as representing the major affective dimensions expressed in the English language. The subjects were instructed to go down the list of adjectives checking all the words which

they thought applied to the client's feelings whether expressed or not. Further, they were instructed to circle one check mark which corresponded to the word which they thought described most accurately the real feeling of the client. The same list of 57 adjectives was used for each scene.

Those words which received a minimum of three checks and one circle were selected from the responses of each criterion group. Those words which were chosen most frequently by the high criterion group and not by the low criterion group and vice versa were used to create items for the affective sensitivity scale. The method of selecting the words employed was by inspection. The words chosen by the criterion groups were used to construct an affective sensitivity scale of 280 items.

Selection of items for the Affective Sensitivity Scale. The affective sensitivity scale (Appendix B) was administered at two universities, one located in the midwest and one in the south, to a combined total of 53 persons who were either counselor educators or master's candidates. Instructions were given the subjects to try to feel as the client last felt in each scene. The videotape recording was stopped after each scene, and the subjects were asked to respond to each adjective from a list provided for every scene on the following scale:

1. I have this kind of feeling strongly.
2. I have this kind of feeling somewhat.
3. I have this kind of feeling only a little.
4. I have this kind of feeling not at all.

There were four counselor educators of the total 53 subjects who took the affective sensitivity scale. They were judged high in empathy by reputation. The 49 master's candidates were rated on a normal

distribution curve from high to low in empathy. The 14 master's candidates, in addition to the four counselor educators, that were rated the highest and the 18 that were rated the lowest comprised the high and low groups. These two groups were the validation groups. Table 2 gives a categorization of the 53 subjects from which the validation groups were taken:

Table 2 Categorization of subjects from which validation groups were obtained

University	High Group	Middle Group	Low* Group
A	9	9	9
B	9	8	9
Totals	18	17	18

*High and low groups refer to the top and bottom 33% of the sample based on ratings obtained from instructors. Only the high and low groups were used for validation purposes.

Items that significantly differentiated between the high and low groups were being sought in the investigation. The response pattern for each item was obtained by simple addition and the median determined in the following manner:

Figure 1 Illustrations on determining the median

An Example:	Response Categories			
Item #15	1	2	3	4
High Group (N=18)	13	1	1	3
Low Group (N=18)	6	8	1	3
Totals	19	9	2	6

(In this example the median occurs between responses 1 and 2.)

An Example:	Response Categories			
Item #30	1	2	3	4
High Group (N=18)	6	7	2	3
Low Group (N=18)	2	3	8	5
Totals	8	10	10	8

(In this example the median occurs between responses 2 and 3.)

An Example:	Response Categories			
Item #45	1	2	3	4
High Group (N=18)	3	8	1	6
Low Group (N=18)	2	1	1	14
Totals	5	9	2	20

(In this example the median occurs between responses 3 and 4.)

The study included all items which might discriminate so long as they were consistent in keying. Thus a dual criterion was used in selecting items as validated: (a) they had a chi-square significance at the .20 level or better and (b) the direction of difference between

high and low empathizers was the same as in the data for the group used in selecting adjectives for the instrument. Analysis was determined by a two-by-two chi-square contingency after the median had been established. The formula used was (Siegel, 1956):

$$X^2 = \frac{N\left(\left|AD-BC\right| - \frac{N}{2}\right)^2}{(A+B)\quad(C+D)\quad(A+C)\quad(B+D)}$$

A total of 65 items of the original 280 included in the scale met the test of significance at .20 or higher. This low level of significance was employed because of the small number of subjects and the exploratory nature of the study. Those items found to be significant are listed in Table 4 in Chapter IV (p. 35) and those which are not significant at .20 are given in Appendix C (p. 81). The frequency and significance of each item are also listed.

Methodology for Cross-Validation

Selection of the sample. Students enrolled in three National Defense Education Act Summer Institutes (1964) comprised the sample of 88 persons who responded to the instrument for the purpose of cross-validation. The 88 Institute enrollees were rated on a normal distribution curve from high to low in empathy by their instructors and practicum supervisors. The 28 enrollees rated the highest and the 28 enrollees rated the lowest comprised the high and low groups. These two groups were the cross-validation groups. Table 3 gives a categorization of the 88 enrollees from which the cross-validation groups were taken:

Table 3 Categorization of enrollees from which cross-validation groups were obtained

University	High Group	Middle Group	Low* Group
A	10	10	10
B	10	10	10
C	8	12	8
Totals	28	32	28

*High and low groups refer to the top and bottom 33% of the sample based on ratings obtained from instructors. Only the high and low groups were used for cross-validation purposes.

Procedures for obtaining the data. A kinescope film of the original videotape was made and the film was shown to the Institute enrollees. An exception was made in the case of one group which viewed the videotape recording. A member of the IPR research project staff went to each Institute and gave instructions for taking the instrument and answered questions relating to the entire procedure.

The director of each Institute was requested to have his staff rate enrollees for counseling effectiveness (Appendix D) and empathy (Appendix E). Data obtained from the 88 Institute enrollees and the ratings from the directors were analyzed in the following manner:

1. To cross-validate the results obtained on the 65 significant items, listed in Table 4, data received from the top and bottom 33% (N=56) of the sample (N=88) in the three Institutes were selected for item analysis purposes. The results are contained in Chapter IV in Table 5 (p. 41).

2. To determine the correlation between scores received on the affective sensitivity scale and counseling effectiveness, it was planned to obtain a Pearson Product-Moment correlation from data received from all enrollees of the three Institutes. However, this was inappropriate since too few items were significant under the procedures of step 1.

Validity

An item analysis was conducted to determine those items that significantly differentiated between those persons judged high and low in empathy. The item analysis was conducted on those items that significantly differentiated in order to determine the validity of the instrument. However, an item analysis that employs a two-by-two chi-square contingency considers each item independently of all other items. There is a possibility that items of a test considered separately will not produce enough significant items to prove the test to be valid, but a combination of a part or all of the items will. A total score was obtained for the instrument as each item was scored for "right" and "wrong" answers. Thus the total scale could be assigned correct and incorrect answers on the basis of the responses given by the criterion groups of high and low empathizers. A score of 1 was given for the correct response and 0 for an incorrect response.

The sample was increased from 88 to 148 by adding data obtained from two additional National Defense Education Act Summer Institutes (1964). The increased sample and the scoring of the 172 items permitted an analysis of variance to determine the predictive validity of the instrument.

Reliability

The data from the scored instrument was used to determine a correlation of reliability by employing the Kuder-Richardson formula 20. Further, a t-test was utilized to test for significance among the sub-groups of the five Institutes.

Factor Analysis

A ten-factor matrix was employed to ascertain commonality of meaning in adjectives used in the 172 items.

Summary

The videotape scenes used in the instrument were selected by a panel of six counselors. Two criterion groups, judged high and low in empathic regard, responded to the instrument from a list of 57 adjectives representing major affective dimensions. Their choices of adjectives were used to comprise the affective sensitivity scale of 280 items.

The instrument was administered to 53 subjects in two universities. A total of 65 items of the original 280 included, met the test of significance at .20 or higher.

In order to cross-validate the results of the item analysis of the instrument, another 88 persons responded to the instrument. These subjects were enrolled in three National Defense Education Act Summer Institutes (1964).

The instrument was scored for "right" and "wrong" answers. A value of 1 was given for the correct response and 0 for an incorrect response on 172 scoreable items. The total number of subjects taking the test

was increased from 88 to 148 by administering the test to two additional Institutes. The purpose of scoring the instrument and increasing the sample was to determine the predictive validity and reliability of the instrument. In addition, a t-test for significance among the sub-groups of the sample and a factor analysis to determine commonality of meaning in adjectives used in the 172 items were made possible by scoring the instrument. The desired data from the subject's responses were placed on IBM cards for statistical analysis.

Analysis of variance, t-test, chi-square, Kuder-Richardson formula 20 and factor analysis were used in the statistical treatment of the data.

CHAPTER IV

ANALYSIS OF DATA

This chapter contains the analysis of the results obtained in the steps to validate, cross-validate, factor analyze and determine the predictive validity and reliability of the affective sensitivity scale.

Results of Validating the Instrument

The median response for each item on the four-point scale response was determined, and the scale dichotomized at the combined median for both high and low empathy groups so that a two-by-two chi-square analysis could be performed. The median was at the middle of the scale for 39 items, between the first and second responses for 15 items and between the third and fourth responses for 11 items.

The following table contains the results of the 65 items which were significant at .20 or higher by employing a two-by-two chi-square contingency after determining the point at which the median occurred in the response pattern (chi-squares of 215 items which were not significant are given in Appendix C (p. 81):

Table 4 Frequency distribution and significance of items on the affective sensitivity scale which failed to meet the .20 level of significance in validating the instrument[1]

Part A
(Median occurs between responses 2 and 3.)

Client and Scene	Item and Adjective	Group[2] Rating	Selected Response 1 or 2 %	Selected Response 3 or 4 %	Chi-square
I	3	High	33.33[3]	16.67	4.000
1	Lonely	Low	16.67	33.33	
I	18	High	27.78	22.22	1.800
2	Ambivalent	Low	16.67	33.33	
I	19	High	36.11	13.89	4.050
3	Hopeful	Low	19.44	30.56	
I	21	High	25.00	25.00	1.870
3	Refreshed	Low	13.89	36.11	
I	33	High	11.11	38.89	2.000
4	Ambivalent	Low	22.22	27.78	
I	35	High	22.22	27.78	2.000
4	Embarrassed	Low	11.11	38.89	
I	38	High	38.89	11.11	4.028
5	Resistant	Low	22.22	27.78	
II	52	High	38.89	11.11	2.000
2	Surprised	Low	27.78	22.22	
II	54	High	30.56	19.44	2.786
2	Anxious	Low	16.67	33.33	

[1]A chi-square value of 1.642 or higher was needed for the item to be significant at .20.

[2]Group rating refers to the top or bottom 33% of the sample based on ratings obtained from instructors and supervisors.

[3]The figure given in each cell is the percent of total response (N=36).

Table 4 (Continued)

Part A

Client and Scene	Item and Adjective	Group Rating	Selected Response 1 or 2 %	Selected Response 3 or 4 %	Chi-square
II	59	High	30.56	19.44	1.778
3	Anxious	Low	19.44	30.56	
II	60	High	33.33	16.67	1.800
3	Guilty	Low	22.22	27.78	
III	69	High	47.22	2.78	2.090
1	Controlled	Low	38.89	11.11	
III	77	High	38.89	11.11	4.208
2	Bored	Low	22.22	27.78	
IV	90	High	16.67	33.33	1.800
2	Capable	Low	27.78	22.22	
IV	94	High	38.89	11.11	4.208
2	Annoyed-Angry	Low	22.22	27.78	
IV	95	High	38.89	11.11	5.600
2	Abused	Low	19.44	30.56	
IV	100	High	27.78	22.22	4.985
3	Talkative	Low	44.44	5.56	
V	111	High	20.59	32.35	4.480
1	Disgusted-Fed-up	Low	35.29	11.76	
V	113	High	20.59	32.35	4.480
1	Upset-Disturbed	Low	35.29	11.76	
V	116	High	23.53	29.41	3.265
1	Bitter	Low	35.29	11.76	
V	117	High	47.06	5.88	3.278
1	Cautious	Low	29.41	17.65	
V	135	High	5.88	47.06	2.101
4	Erratic-Disorganized	Low	14.71	32.35	

Table 4 (Continued)

Part A

Client and Scene	Item and Adjective	Group Rating	Selected Response 1 or 2 %	Selected Response 3 or 4 %	Chi-square
VI	167	High	48.48	6.06	2.418
4	Happy	Low	30.30	15.15	
VI	177	High	36.36	18.18	5.241
5	Disgusted-Fed-up	Low	12.12	33.33	
VII	179	High	27.27	27.27	3.182
1	Comfortable-At ease	Low	36.36	9.09	
VII	180	High	6.06	48.48	5.215
1	Regretful	Low	21.21	24.24	
VII	185	High	27.27	27.27	1.866
2	Abused	Low	12.12	33.33	
VII	191	High	15.15	39.39	2.238
2	Scared	Low	24.24	21.21	
VIII	194	High	48.48	6.06	2.418
1	Embarrassed	Low	30.30	15.15	
VIII	199	High	27.27	27.27	1.866
2	Frustrated	Low	33.33	12.12	
VIII	203	High	42.42	12.12	2.200
2	Foolish-Riciculous	Low	24.24	21.21	
IX	223	High	39.39	15.15	4.911
4	Defeated	Low	45.45	.00	
X	243	High	42.42	12.12	4.891
3	Worried	Low	18.18	27.27	
X	247	High	27.27	27.27	1.866
3	Bitter	Low	12.12	33.33	
XI	265	High	18.18	36.36	1.782
2	Helpless	Low	6.06	39.39	

Table 4 (Continued)

Part A

Client and Scene	Item and Adjective	Group Rating	Selected Response 1 or 2 %	Selected Response 3 or 4 %	Chi-square
XI 3	271 Courageous-Daring	High Low	45.45 24.24	9.09 21.21	3.487
XI 3	272 Stubborn-Resistant	High Low	18.18 27.27	36.36 18.18	2.347
XI 4	278 Sensible-Reasonable	High Low	39.39 42.42	15.15 3.03	2.451

Part B
(Median occurs between responses 1 and 2.)

Client and Scene	Item and Adjective	Group Rating	Selected Response 1 %	Selected Response 2 or 3 or 4 %	Chi-square
I 3	26 Controlled	High Low	47.22 36.11	2.78 13.89	3.200
I 4	30 Controlled	High Low	38.89 47.22	11.11 2.78	2.090
I 5	40 Reluctant	High Low	47.22 36.11	2.78 13.89	3.200
II 1	46 Anxious	High Low	47.22 38.89	2.78 11.11	2.090
III 1	71 Confused-Mixed-up	High Low	50.00 41.67	0.00 8.33	3.273
IV 1	82 Cautious	High Low	38.89 27.78	11.11 22.22	2.000
IV 3	99 Discouraged	High Low	44.44 50.00	5.56 0.00	2.118
IV 4	108 Controlled	High Low	50.00 44.44	0.00 5.56	2.118

Table 4 (Continued)

Part B

Client and Scene	Item and Adjective	Group Rating	Selected Response 1 %	Selected Response 2 or 3 or 4 %	Chi-square
V	123	High	47.06	5.88	1.889
2	Anxious	Low	47.06	0.00	
V	133	High	29.41	23.53	2.555
4	Proud	Low	38.24	8.82	
VI	173	High	39.39	15.15	2.451
5	Disappointed	Low	42.42	3.03	
VIII	192	High	54.55	0.00	2.555
1	Confused-Mixed-up	Low	39.39	6.06	
IX	214	High	54.55	0.00	2.555
2	Optimistic	Low	39.39	6.06	
IX	221	High	42.42	12.12	3.793
3	Bitter	Low	45.45	0.00	
X	232	High	42.42	12.12	3.793
1	Discouraged	Low	45.45	0.00	

Part C
(Median occurs between responses 3 and 4.)

Client and Scene	Item and Adjective	Group Rating	Selected Response 1 or 2 or 3 %	Selected Response 4 %	Chi-square
I	23	High	0.00	50.00	2.118
3	Trapped	Low	5.56	44.44	
I	24	High	5.56	44.44	2.118
3	Annoyed-Angry	Low	0.00	50.00	
II	49	High	2.78	47.22	2.090
2	Bitter	Low	11.11	38.89	
III	76	High	2.78	47.22	2.090
2	Confused-Mixed-up	Low	11.11	38.89	

Table 4 (Continued)

Part C

Client and Scene	Item and Adjective	Group Rating	Selected Response 1 or 2 or 3 %	Selected Response 4 %	Chi-square
IV	86	High	8.33	41.67	3.273
1	Forced	Low	0.00	50.00	
IV	88	High	5.56	44.44	2.118
1	Embarrassed	Low	0.00	50.00	
IV	106	High	19.44	30.56	2.215
4	Careful	Low	8.33	41.67	
VII	183	High	6.06	48.48	1.774
1	Lonely	Low	0.00	45.45	
X	251	High	0.00	54.55	2.555
4	Bitter	Low	6.06	39.39	
XI	268	High	6.06	48.48	1.774
2	Satisfied	Low	0.00	45.45	
XI	276	High	15.15	39.39	2.451
4	Ambivalent	Low	3.03	42.42	

Results of Cross-Validating the Instrument

The median was at the middle of the scale for five items, between the first and second responses for one item and between the third and fourth responses for three items. The following table contains the results of the cross-validation analysis of the 65 items. Of the 65, nine items were found to be significant at .20 or higher by employing a two-by-two chi-square contingency (chi-squares of 56 items which were not significant are given in Appendix F (p. 101):

Table 5 Frequency distribution and significance of items of the affective sensitivity scale which were significant at .20 or higher obtained in cross-validating the instrument[1]

Part A
(Median occurs between responses 2 and 3.)

Client and Scene	Item and Adjective	Group Rating[2]	Selected Response 1 or 2 %	Selected Response 3 or 4 %	Chi-square
V	116	High	17.86[3]	32.14	5.793
1	Bitter	Low	33.93	16.07	
VI	177	High	41.07	8.93	4.139
5	Disgusted-Fed-up	Low	28.57	21.43	
VII	191	High	5.36	44.64	10.220
2	Scared	Low	25.00	25.00	
X	247	High	32.14	17.86	2.585
3	Bitter	Low	21.43	28.57	
XI	265	High	32.14	17.86	12.469
2	Helpless	Low	8.93	41.07	

Part B
(Median occurs between responses 1 and 2.)

Client and Scene	Item and Adjective	Group Rating	Selected Response 1 %	Selected Response 2 or 3 or 4 %	Chi-square
III	71	High	10.71	39.29	2.947
1	Confused-Mixed-up	Low	21.43	28.57	

[1]A chi-square value of 1.642 or higher was needed for the item to be significant at .20.

[2]Group rating refers to the top or bottom 33% of the sample based on ratings obtained from instructors and supervisors.

[3]The figure given in each cell is the percent of total responses (N=56).

Table 5 (Continued)

Part C
(Median occurs between responses 3 and 4.)

Client and Scene	Item and Adjective	Group Rating	Selected Response 1 or 2 or 3 %	Selected Response 4 %	Chi-square
VII	183	High	35.71	14.29	1.697
1	Lonely	Low	42.86	7.14	
X	251	High	28.57	21.43	7.376
	Bitter	Low	44.64	5.36	
XI	268	High	46.43	3.57	4.383
	Satisfied	Low	35.71	14.29	

Predictive Validity Results

An examination of Tables 4 and 5 indicates the instrument does not discriminate consistently across samples of high and low empathic counselors when each item is considered independently of all other items. In order to consider the data from a different perspective a total score was obtained for the instrument. Those items that discriminated between high and low counselor criterion groups were scored on a scale 0-1. A score of 1 was given for the correct response and 0 for an incorrect response. A total of 172 items from the original 280 were scoreable, in that the high counselor criterion group selected the item and the low criterion group did not or vice versa. The items that were not scored were equally attractive to both criterion groups and would not discriminate between the two. The scored instrument is given in Appendix G (p. 109).

Once a total score was obtained, an analysis of variance was employed to determine the predictive validity of the instrument. If the instrument discriminated between persons judged high and low in empathy, then one could state that it possessed predictive validity.

The summary data on the 172 scored items not only considered high and low groupings but male and female as well. This was done to determine whether or not there were significant differences between sexes. These data are contained in Table 6:

Table 6 Summary data of scored instrument

Type or Group	Male (3)		Female (4)		Totals	
	K	$\bar{X}$	K	$\bar{X}$	K	$\bar{X}$
High (1)	35	92.571	9	96.778	44	93.432
Low (2)	33	90.575	11	91.818	44	90.886
Totals	68	91.603	20	94.050	88	92.159

The differences in means between high-low and male-female groupings were:

$$D_{1,2} = \bar{X}_H - \bar{X}_L$$

$$D_{1,2} = \underline{2.555}$$

$$D_{3,4} = \bar{X}_M - \bar{X}_F$$

$$D_{3,4} = \underline{2.457}$$

where:

D = difference

1 = high; 2 = low; 3 = male; 4 = female

The analysis of variance operates on the assumption of equal N's among the different sub-groups. This assumption is not met in the data presented in Table 6. Therefore, an adjustment term (Wert, Neidt, and Ahmann, 1954) was employed to correct for disproportionality among the

sub-groups. The formula for the adjustment term is:

$$\frac{(ad-bc)^2}{(K_1K_2K_3K_4)} \left[(K_1)\ (K_2)\ (D_{1,2})^2 + (K_3)\ (K_4)\ (D_{3,4})^2 \right] - 2(D_{1,2})\ (D_{3,4})\ (ad-bc)$$

$$N \left[1 - \frac{(ad-bc)^2}{(K_1K_2K_3K_4)} \right]$$

The alphabetical symbols of Figure 2 correspond to the numerical values of Table 6:

Figure 2 Table of symbols

Type or Group	Male (3) K	Male (3) $\bar{X}$	Female (4) K	Female (4) $\bar{X}$	Totals K	Totals $\bar{X}$
High (1)	a		b		K_1	
Low (2)	c		d		K_2	
Totals	K_3		K_4			

The following sum of squares and their formulae (Wert, Neidt and Ahmann, 1954) were employed in computing the analysis of variance:

1) Sum of squares for total

$$SS_T = \Sigma X^2 - \frac{(\Sigma X)^2}{N}$$

2) Sum of squares for sex

$$SS_{Sex} = \frac{(\Sigma X_m)^2}{K_m} + \frac{(\Sigma X_f)^2}{K_f} - \frac{(\Sigma X)^2}{N}$$

3) Sum of squares for high-low

$$SS_{H-L} = \frac{(\Sigma X_H)^2}{K_H} + \frac{(\Sigma X_L)^2}{K_L} - \frac{(\Sigma X)^2}{N}$$

4) Sum of squares for interaction

$$SS_I = \frac{(\Sigma X_{HM})^2}{K_{HM}} + \frac{(\Sigma X_{HF})^2}{K_{HF}} + \frac{(\Sigma X_{LM})^2}{K_{LM}} + \frac{(\Sigma X_{LF})^2}{K_{LF}} - \frac{(\Sigma X)^2}{N}$$

5) Sum of squares for within

$$SS_W = SS_T - (SS_{Sex} + SS_{HL} + SS_I)$$

The material in Table 7 contains the data for the analysis of variance with means adjusted for disproportionality by sex, high-low groupings and interaction of all possible combinations:

Table 7 Analysis of variance of instrument scores with means adjusted for disproportionality

Source of Variation	Degrees of Freedom (N-1)	Sum of Squares Unadjusted	Sum of Squares Adjusted	Mean Square
Sex	1	92.543	80.640	80.640
High-low	1	142.533	130.730	130.730
Interaction	1	171,475.435	171,487.238	171,487.238
Within	87	-164,910.738		- 1,895.526
Totals	90	6,799.773		

Adjustment Term = 11.803*

For Sex: $F_{1, 87} = \frac{80.640}{-1,895.526} = -.0043$**

For High-Low: $F_{1,87} = \frac{130.730}{-1,895.526} = -.0690$**

For Interaction: $F_{1,87} = \frac{171,487.238}{-1,895.526} = 90.469$**

*The adjustment term is added algebraically to main effects (sex and high-low) and subtracted from minor effect (interaction).

**An analysis of variance value of F(.05) = 253 was needed for the instrument to discriminate on a source of variation.

Reliability Results

The scoring of the instrument permitted a correlation of reliability to be determined. The Kuder-Richardson formula 20 for reliability was .441. The correlation of reliability obtained (.441) is high enough to indicate the instrument possesses average consistency of measurement when compared with other tests of personality variables. (This would not be the case where tests of aptitude and achievement are involved since correlations of reliability are usually higher in these two areas.) Although the reliability of the instrument is acceptable, the instrument does not possess the ability to differentiate between persons judged high and low in empathy other than that produced by chance. A score obtained in using the instrument could not be relied upon since it is not known whether or not affective sensitivity is measured.

A t-test was employed to determine any significant difference between the highest mean (104.7) and the lowest mean (82.) of all the NDEA Institutes sub-groups. These data were the same used for purposes of determining the validity and reliability of the instrument. This was done in order to answer the question, "are we dealing with any combination of effects that is significant." The data in Table 8 contains the means of each NDEA Institute according to high-low and male-female groupings.

There was no significant difference between the highest mean of 104.7 (Institute D, high-female) and the lowest of 82 (Institute B, low-female).[1] Since the difference between the highest and lowest means did not produce significance, it could be concluded that all other possible combinations of means would also fail to produce significance.

---

[1] t (α = .05) = 2.447 had to be reached to obtain significance. The difference produced a t score of only .013.

Table 8 NDEA Institute means

Institute	High				Low			
	Male		Female		Male		Female	
	K	X	K	X	K	X	K	X
A	9	87.6	1	99	8	85.3	2	90.5
B	10	92.9	0	--	9	92	1	82
C	6	98	2	86.5	5	89	3	90
D	5	96.8	3	104.7	5	93.8	3	96
E	5	90.2	3	95	6	94.2	2	94.5
Totals	35	92.6	9	96.3	33	90.6	11	91.8

Factor Analysis Results

The data for the 172 items, which were scoreable from the original 280, were factor analyzed. The purpose of employing factor analysis to the data was to determine commonality of meaning in adjectives used in the items for a maximum of ten factors. A factor had to possess a value of .450 or higher to meet a level of significance of .05. The selecting of factors for each item was as follows:

Figure 3 Illustration on determining factors

An Example:

Item Number	1	2	3	4	5	6	7	8	9	10
4	.649	.115	.106	.006	.029	.065	.065	.004	.049	.005
9	.033	.071	.023	.144	.187	.083	.344	.312	.209	.174

(In the example given, factor 1 of item 4 was selected as the factor for that item. There was no factor that was significant in item 9.)

A particular factor was considered operative when at least eight items had values of .450 or higher from the 172 items. The cut-off point of eight items was determined by N-2, where N equals total number of factors or a total of ten in this factor analysis. There were five factors operative in the data analyzed, and these are given in Table 9. As a final step in factor analysis, all scenes that contained an item or items that were operative in a particular factor were considered for commonality of meaning. This was accomplished by viewing those scenes containing operative items and determining the common thread of feeling or thought contained in the combined group of scores. The descriptive statements of feeling or thought are found in the next section under "Analysis" (p. 51).

Analysis

Validating and cross-validating the instrument. The two-by-two chi-square analysis was performed on each item independently of all other items. This was done to determine those items that were empirically significant. Then, all significant items would have been used to produce a valid instrument. However, an analysis of Tables 4 and 5 reveals the instrument cannot be relied upon as being valid since the number of significant items obtained in cross-validation was no better than chance.

Predictive validity. The analysis of variance values obtained for sex grouping, high-low grouping and interaction were not significant. These three values were obtained to discover whether one or more of these effects may be significant. If only one had proven to be significant,

Table 9 Adjectives in operative factors

FACTOR 1		FACTOR 2		FACTOR 4		FACTOR 6		FACTOR 7	
Client and Scene	Item and Adjective	Client and Scene	Item and Adjective	Client and Scene	Item and Adjective	Client and Scene	Item and Adjective	Client and Scene	Item and Adjective
I	4	II	46	I	19	I	23	II	43
1	Depressed	1	*Anxious	3	Hopeful	3	*Trapped	1	*Baffled
I	6	II	54	I	20	I	32	II	56
1	Discouraged	2	*Anxious	3	Optimistic	4	*Baffled	2	*Frustrated
III	67	IV	82	I	22	I	33	III	72
1	Anxious	1	*Cautious	3	*Relieved	4	*Ambivalent	1	*Listless-Indifferent
III	68	IV	84	II	58	VI	169		
1	Worried	1	*Reluctant	3	*Daydreaming	4	Baffled	III	74
III	71	IV	85	V	121	VI	170	2	*Listless-Indifferent
1	*Confused-Mixed-up	1	Anxious	2	Annoyed-Angry	4	*Confused-Mixed-up	III	77
IV	99	V	139	VI	157	VIII	194	2	*Bored
3	*Discouraged	5	*Bitter	2	*Satisfied	1	*Embarrassed	III	79
IV	100	IX	226	VI	158	VIII	201	2	Reluctant
3	Talkative	4	*Annoyed-Angry	2	Appreciated	2	*Embarrassed	VII	180
IV	105	X	242	VI	160	XI	258	1	*Anxious
4	Defeated	3	Depressed	2	*Amused	1	*Embarrassed	IX	219
V	113			VI	171			3	Amused
1	Upset-Disturbed			4	*Helpful			IX	221
V	124			VIII	197			3	Bitter
2	Lonely			1	*Discouraged			IX	224
V	126			X	242			4	Bitter
3	Depressed			3	*Depressed			IX	225
V	127							4	Hopeless
3	*Confused-Mixed-up								
V	135								
4	Erratic-Disorganized								
V	136								
4	Guilty								
V	140								
5	*Discouraged								
V	142								
5	Depressed								
V	150								
6	Disgusted								
VIII	200								
2	Impatient								
IX	207								
1	Cautious								
IX	227								
4	Baffled								
X	233								
1	Baffled								
X	241								
2	Confused-Mixed-up								
X	246								
3	Confused-Mixed-up								
XI	259								
1	Forced								

*The opposite in meaning should be given since the value was negative.

this would have been evidence enough to continue the investigation in the direction of the significant effect. The values for the three effects were far enough below significance to indicate that had the sample been increased to infinity the values needed for significance could not have been reached.

Reliability. The correlation of reliability obtained (.441) has little significance since the instrument does not possess predictive validity. There would always be the question of exactly what variable was being measured by the instrument. The t-test score obtained, when comparison was made in the difference between the group possessing the highest mean and the group with the lowest, indicated the groups were homogeneously grouped in ability on the variable measured. The lowest mean was five points below chance and the highest mean 18 points above chance. A chance score was 86, which implies that the subjects responding to the instrument were average in their ability on the variable measured by the scale. The logical interpretation to come from the t-test score is that it is difficult to know where the items chosen lie on an absolute scale of empathy.

Factor analysis. The five factors that were operative went from a high of 24 words under factor 1 to 11 words in factors 4 and 7 to eight words in factors 2 and 6. There were 29 different scenes in the 62 items found under the five factors. Each operative factor was considered for commonality of meaning by considering all scenes found in a particular factor. The following statements give the common idea for all scenes under each factor as observed by the researcher:

Factor 1 - There was a generalized distrust of the future.
(The clients were anxious regarding their plans for the future.)

Factor 2 - There was a need for relatedness to others and self-acceptance.
(The clients felt a need to force themselves upon other people.)

Factor 4 - There was a positive personal regard to the extent of being in agreement with one's self and self-potential.
(The clients felt they were accepted and needed by others which enhanced positive feelings toward self.)

Factor 6 - There was a concept of self-certainty.
(The clients felt sure of themselves to the point of being decisive.)

Factor 7 - There was a free-floating reaction that bordered on indifference.
(The clients simply wanted "to live and let live.")

The statements describing each factor are independent of each other and in some cases opposite in meaning. This is to be expected since the descriptive adjectives used in the instrument were representative of the major affective dimensions expressed in the English language.

Summary

The results of the data obtained in validating and cross-validating failed to demonstrate that the affective sensitivity scale could reliably discriminate between persons judged high and low in empathy. The predictive validity of the instrument thus could not be demonstrated and the reliability was average. There was no commonality of meaning in the five factors considered operative in the factor analysis. However, each factor did possess a common idea when considered independently of each other.

CHAPTER V

SUMMARY, CONCLUSIONS AND DISCUSSION

Summary

The purpose of this study was to develop and validate an instrument that tested a subject's ability to identify the emotion or emotions expressed by another in a series of videotaped excerpts from actual counseling interviews.

The instrument was composed of selected scenes of videotaped recordings of counseling interviews. Each scene was followed by four to seven descriptive adjectives. Subjects were to respond to each adjective on a continuum indicating how strongly they possessed the feeling described by the adjective at the end of each scene. Two criterion groups, judged high and low in empathic regard, responded to the instrument and generated items from a list of 57 adjectives expressing various feelings. The final list of four to seven descriptive adjectives was produced from the list of 57 adjectives.

The instrument was administered to 53 subjects in two universities for the purpose of validation. In order to cross-validate the results of the item analysis of the instrument, it was administered to another 88 persons enrolled in three National Defense Education Act Summer Institutes (1964).

The instrument was scored for "right" and "wrong" answers on 172 scoreable items from the original 280. Correct and incorrect answers

were determined on the basis of the responses given by the criterion groups of high and low empathizers. The total sample was increased from 88 to 148 by adding the data from two additional National Defense Education Act Summer Institutes (1964). The purpose of scoring the instrument and increasing the sample was to determine the predictive validity and reliability of the instrument. In addition, a t-test for significance among the sub-groups of the sample and a factor analysis to determine commonality of meaning in adjectives used were made possible by scoring the instrument.

The desired data from the subject's responses were placed on IBM cards for statistical analysis. Chi-square was used in the statistical treatment of the data for validation and cross-validation purposes with the test of significance being established at .20 or higher. The two-by-two chi-square contingency was employed on the response scale. The response scale was split at the point at which the median occurred in the response pattern of each item. Analysis of variance to ascertain predictive validity; t-test to determine significant differences among means of the sample sub-groups; Kuder-Richardson formula 20 to establish reliability and factor analysis for commonality of meaning were also used in working with the data.

Findings. A total of 65 items of the original 280 included met the test of significance for the purpose of validation, and nine items were found to be significant in cross-validating the instrument.

These nine items which held up on cross-validation included the adjectives "bitter" (three of the items), "disgusted," "scared," "helpless," "confused," "mixed-up," "lonely," and "satisfied." The large number of

negative adjectives results from the fact that the adjectives in general were descriptive of negative affect. There was no major concentration on any one client or episode although nearly all of the items were toward the end of the scale. Because the nine items represent about a fifth of the 65, and one would expect about a fifth of the items to show significance by chance at the 20% level, it would appear that there is nothing more than chance relationship between the item responses and the criterion.

The predictive validity of the instrument was nonexistent. The analysis of variance values for three effects (sex grouping, high-low grouping and interaction) were low enough to indicate that should the sample be increased to infinity significance would not be produced. The correlation of reliability was .441. The t-test did not find a significant difference between the highest and lowest means of the two sub-groups involved. There were five factors operative in a ten-factor matrix. A total of 29 scenes and 62 items were found in the five factors.

Conclusions

The hypotheses tested in this investigation were:

1. A scale based upon actual counseling interviews will differentiate between persons judged high and low in empathic regard.
2. A scale based upon actual counseling interviews will produce a significant correlation between scores received on the affective sensitivity scale and counselor effectiveness.

The affective sensitivity scale did not differentiate between persons judged high and low in empathic regard as stated in the first hypothesis; hence the first hypothesis is rejected. Since the affective sensitivity scale proved to be invalid, it was inappropriate to test the second hypothesis.

Discussion

The affective sensitivity scale, as developed in this study, did not differentiate between persons judged high and low in empathic regard. This resulted in rejecting one major hypothesis which indicated that the scale would differentiate and as a result a second one could not be tested. It is true that variations in methodology (see section on Implications below) might yield different data; however, the methodology used appeared appropriate. Further, unlike so many past "empathy" studies, the use of videotape and insights provided through recall sessions, made possible the creation of an instrument and instructions for its use which seemed to come closer to a congruence of theoretical and operational definition than ever before. Operationally, the subject was asked only to "feel as the other." He was not asked to predict future client behavior or to determine the counselor's next best statement--both typical of other studies and both subject to the influence of knowledge and training in the subject of counseling or personality theory. These characteristics of the scale place it in the category of situational tests of empathy which is described in Chapter II. These studies provide situations involving combinations of visual, auditory or kinesthetic stimuli so as to provide a standard experience to elicit empathic responses from the subject (Reid and Snyder, 1947; Astin, 1957; O'Hern 1964; and Rank, 1964).

The contradictory meanings and vague usages of the term "empathy" were described in Chapter II by presenting definitions of this variable as found in the professional literature. The definition of affective sensitivity, as used in this study, has more compatibility with the

German term, Einfuhlung, and the definition of empathy stated by Katz (1963) than with those espoused by English and English (1948) and Fenichel (1945). The former definitions indicate that detecting the momentary affective state of another is essential rather than the detachment and objectivity of the empathizer, as held by the latter.

The question must be asked: what contributed to negative results? One possibility is that adjectives, without qualifying phrases, have different meanings for different people. Thus individuals have similar feelings in empathizing but may be responding to the adjectives in different ways. It is possible that the use of phrases might be more specific and therefore might give more consistency of interpretation than the adjectives alone. This point is supported by Katz (1963, p. 181):

> The empathizer is not concerned merely with feelings labeled "sibling rivalry," "loneliness," "fear of authority," or "ambivalence." He enters into a specific situation and imaginatively places himself in the moment, in time and space when his client responded spontaneously to an event and an object. He empathizes with the person in the situation and with the meanings, values and feelings spontaneously projected in that specific context.

A second possibility is that the level of interpretation varied with the experience of the counselor. Thus, what was most indicative for counselors at one level might not have been at another. For instance, a subject seeing a client overtly express fright might initially feel fright when he tried to feel as the person did. On closer examination, he might realize that actually the fright is fear of being angry and that underlying this is anger at the counselor which is being suppressed. In a real sense then, at the deepest level of interpretation, the most appropriate response might be anger. Thus the counselor who reads much deeper meanings into the scene may check different responses than the less

experienced counselors. If this is the case, one will have to define better the level of insight requested. This could enhance considerably the method of using videotaped recordings as compared to tests based solely on the verbal text of an interview or on audio recordings. The weakness of the latter methods is described by Sullivan (1954) in Chapter II.

It is possible that the experience of the counselor can be a hindrance to his ascertaining the feeling of the client. Some counselors indicate some of their best empathic endeavors occurred when they were just starting in the profession. The fixed routines and traditions of the profession may decrease the empathic ability of the counselor. The need to experiment or become involved with a person is not so great once a counselor is established. The counselor who possesses a high degree of empathy is willing to explore new and unprotected areas.

It is also possible that as the individual checked down the list, the original feeling that the individual had at the time of the end of the scene may have been lost. Other feelings suggested by the adjectives may have modified the original feeling. If this were true, items at the beginning of the list should cross-validate more readily. An examination of the nine items which did cross-validate, however, indicated that the next to the last or lower half of the list in seven out of the nine instances were the items which held up. It would appear, then, that this is not a valid hypothesis. On the other hand, because the cross-validation indicates only chance results, a combination of several factors may have masked such results.

The point is made by Katz (1963) that social psychologists view empathy as adaptive rather than reactive and as a skill acquired in the process of general social interaction. The process is a rehearsal for realistic and appropriate action where it is not the other person one wishes to emulate but the other person's ideas and image of one's self that is desired to be understood. When empathy is viewed in this manner, the results obtained in validating and cross-validating are not surprising for two reasons. First, the subjects were instructed to assume a psychological "set" of viewing the client and respond or react as the client responded in the interview. The process of interaction possibly could have been accomplished by requiring the subjects to reply verbally to the client and record the reply on an audio tape recorder. Second the subjects focused their attention upon the client. Perhaps, if Katz is right, they should have been required to ascertain the ideas and the image of the counselor.

There is a possibility that using most of the words many times in the instrument created an attitude of redundancy on the part of the subjects. This in turn could lead most of the subjects to become bored or flippant in the testing situation.

A form containing 280 items is an extremely long test. It often took nearly two hours to administer the test (particularly when mechanical adjustments were needed). Fatigue certainly could have been a factor.

The criterion problem is unavoidable. Although the counselor educators ranked their students on a normal distribution and each had instructions on rating procedures, it is not known whether the different universities were using similar criteria for ranking or how valid each

university's ranking was. Two of the professors pointed out that although they knew their students well, empathy was so elusive a quality for them they found themselves revising their rankings several times before submitting them. "The only valid criterion," one professor stated, "will be their--the counselors--long term influence of the clients they counsel." This was beyond the scope of this study.

The suggestion has been made that people do not differ significantly in their affective sensitivity, that this quality is distributed over a very narrow range. This idea stems from the apparent insight on recall of individuals who might otherwise have been thought to be socially insensitive. A number of such individuals indicate an awareness of what is going on, but seem to be incapable of action based upon this awareness. Frequently they are able to verbalize their frustration at this gap. Such persons might be able to perform well on a non-threatening test of such behavior even though they could not act on it.

Viewed in this fashion, effectiveness in counseling would result, not so much from a difference in perception, as in what the counselor does based on this perception. The effective counselor enters the phenomenological field of the client and then confronts the client or deals with the feelings in other ways appropriate to his theoretical base. The ineffective counselor, on the other hand, equally aware of the client's communication, avoids confrontation, changes the topic, or perhaps "uses" the client's feeling to satisfy his own needs. It is possible that the recognition of affective communication is basic to survival within a given society and is learned well quite early. The child suffers pain who is unable to recognize the meaning "behind"

father's words or unable to understand the meaning of the clenched teeth of his larger peers. Any teacher has experienced situations in which a classroom of 30 or more children will immediately come to order as he or she enters the room, thus making unnecessary the speech she was going to make about their poor conduct on previous days--the "speech" apparently having been already communicated by her posture, gait and expression on walking from the classroom door to her desk. Would the theater and motion picture industries have been so successful financially if audiences differed significantly in affective sensitivity?

But people do appear to differ in this quality. How could one explain the social boor's "clumsiness" and apparent lack of sensitivity? Where the behavior results in annoyance and irritation with the behavior, could one as readily attribute the cause to a need for this manner of attention rather than social insensitivity? It is not so much that the individual does not perceive accurately according to this point of view, as it is that he does not know what to do about it, or he does not behave in socially acceptable ways.

It is interesting to speculate that if research shows this view to be valid, then texts and training procedures in counselor as well as teacher education would require extensive revision. Students in these and other areas would need less to "learn" the meaning of the communication of others and more to overcome effectively their own need to avoid dealing with the communication.

Studies conducted in the area of social suggestion may give another reason why negative results were obtained. In a given group each individual receives and gives stimuli which condition the response pattern

for the entire group. Despite the instructions and frequent reminders, many individuals began rating before the end of the scene and so did not react to the last feeling in the scene. This at times meant they would laugh at their response or give a sign of boredom. This situation would have been corrected by having each subject respond in private to the instrument. However, the instrument was designed to be given in a group testing situation. To administer it individually was beyond the limits of this study.

The criteria under which the enrollees of each institute were selected were fairly homogeneous since minimum requirements for acceptance at an Institute were established by the U.S. Office of Education. Also, most counselor educators have some similar expectation for persons entering the field of counseling. This could have contributed to the means of the sub-groups being rather close to each other. The range of difference in empathic regard from institute to institute or within an institute may not be wide due to the factors of selectivity and expectations held in common by counselor educators.

The fact that "face validity" appeared to be operationally high and predictive validity empirically low or nonexistent is difficult to interpret. This may have been otherwise had a panel of judges been asked to view each scene and had the researcher used only those items where agreement was reached as to the correct response for the four-point scale. This would have meant: (a) each item used in the instrument would have an agreed upon correct response; (b) the score obtained by each subject would have determined high and low empathic groups and (c) the item could have been weighed on a four-point scale.

The scoring procedure that was employed had to be on a two-point scale since only direction of response was given by the criterion groups instead of both direction and degree of response.

Implications for further research. At this juncture the criterion problem, the similarity of the subject's reactions to the scale, and the necessity of getting subjects to react to the same portion of the scene are points of departure for additional study. The following recommendations are suggested:

1. The present scale or the proposed revisions of it should be administered to full-year rather than to summer Counseling Institutes so that counselor criterion ranking can be based on more extensive contact with the counselor educators making the ranking.

2. The IPR interrogations responses themselves, along with alternatives, could be used as items instead of simple adjectives. An accuracy score could thus be determined without the use of judges, and this score studied in relation to other criteria e.g. grades, ranking by practicum supervisors, client rating of counselor.

3. Replicate the study but utilize professional counselors as the only criterion group to rate the feeling(s) being exhibited by the client and to indicate the correct response.

4. Conduct another study where subjects respond to descriptive statements instead of a list of descriptive words or respond verbally to each episode.

5. Conduct another study where subjects respond in a private setting instead of a group setting.

APPENDIX A

LIST OF FIFTY-SEVEN ADJECTIVES

INSTRUCTIONS

In the next two hours you will be viewing some kinescopes of actual counseling sessions. We are trying to identify what feelings were being expressed by the clients in these interviews. We are concerned about what observers would judge to be the actual feeling.

On the next page you will find a list of 57 adjectives expressing various feelings. Parallel to these words are spaces which you can check for episodes shown on the TV monitor. For example, all the spaces under Arabic numeral (1) will refer to the first episode of Client I. After you have seen counseling episode (2) go down the adjective list checking all those words which you feel apply to the client's feelings whether expressed or not. Then, circle only one check mark which corresponds to the word which you think describes most accurately the feeling of the client. In other words we would like you to put yourself in the place of the client and as empathically as possible record what you think is the real feeling the client is experiencing.

Name _____

*Episode	1	2	3	4	5	6
1. Abused	___	___	___	___	___	___
2. Ambivalent	___	___	___	___	___	___
3. Amused	___	___	___	___	___	___
4. Annoyed - Angry	___	___	___	___	___	___
5. Anxious	___	___	___	___	___	___
6. Appreciated	___	___	___	___	___	___
7. Baffled	___	___	___	___	___	___
8. Bitter	___	___	___	___	___	___
9. Bored	___	___	___	___	___	___
10. Capable	___	___	___	___	___	___
11. Careful	___	___	___	___	___	___
12. Cautious	___	___	___	___	___	___
13. Comfortable - at ease	___	___	___	___	___	___
14. Confused - mixed up	___	___	___	___	___	___
15. Controlled	___	___	___	___	___	___
16. Courageous - Daring	___	___	___	___	___	___
17. Daydreaming	___	___	___	___	___	___
18. Defeated	___	___	___	___	___	___
19. Depressed	___	___	___	___	___	___
20. Determined	___	___	___	___	___	___
21. Disappointed	___	___	___	___	___	___

*This list of words was used for all episodes.

Episode	1	2	3	4	5	6
22. Discouraged	___	___	___	___	___	___
23. Disgusted - Fed-up	___	___	___	___	___	___
24. Embarrassed	___	___	___	___	___	___
25. Enthusiastic	___	___	___	___	___	___
26. Erratic - Disorganized	___	___	___	___	___	___
27. Exhausted	___	___	___	___	___	___
28. Flexible	___	___	___	___	___	___
29. Foolish - Ridiculous	___	___	___	___	___	___
30. Forced	___	___	___	___	___	___
31. Frustrated	___	___	___	___	___	___
32. Guilty	___	___	___	___	___	___
33. Happy	___	___	___	___	___	___
34. Helpful	___	___	___	___	___	___
35. Helpless	___	___	___	___	___	___
36. Hopeful	___	___	___	___	___	___
37. Hopeless	___	___	___	___	___	___
38. Impatient	___	___	___	___	___	___
39. Listless - Indifferent	___	___	___	___	___	___
40. Lonely	___	___	___	___	___	___
41. Optimistic	___	___	___	___	___	___
42. Patient	___	___	___	___	___	___
43. Protected	___	___	___	___	___	___
44. Proud	___	___	___	___	___	___

Episode	1	2	3	4	5	6
45. Refreshed	___	___	___	___	___	___
46. Regretful	___	___	___	___	___	___
47. Relieved	___	___	___	___	___	___
48. Reluctant	___	___	___	___	___	___
49. Satisfied	___	___	___	___	___	___
50. Scared	___	___	___	___	___	___
51. Sensible - Reasonable	___	___	___	___	___	___
52. Stubborn - Resistant	___	___	___	___	___	___
53. Surprised	___	___	___	___	___	___
54. Talkative	___	___	___	___	___	___
55. Trapped	___	___	___	___	___	___
56. Upset - Disturbed	___	___	___	___	___	___
57. Worried	___	___	___	___	___	___

APPENDIX B

AFFECTIVE SENSITIVITY SCALE

INSTRUCTIONS

You will be viewing some kinescopes of actual counseling sessions. We are trying to identify what feelings are being experienced by the clients in these interviews. We want to know what you as an observer judge to be the actual feeling(s).

On the following pages you will find lists of adjectives which describe various feelings. After viewing each episode, try to feel as the client last felt. Then, taking each adjective in turn rate it on the following scale:

1. I have this kind of feeling strongly.
2. I have this kind of feeling somewhat.
3. I have this kind of feeling only a little.
4. I have this kind of feeling not at all.

For example, if you view a specific episode in which the client states, "I am happy with my experiences at school," you are to try to feel as the client last felt. If you find yourself feeling very annoyed and defeated, you would rate the words annoyed and defeated #1 to indicate, "I have this kind of feeling strongly." Assuming annoyed has arabic numeral 7 preceding it and defeated has arabic numeral 8 preceding it, your answer sheet would appear as follows:

7. 1 ■ 2 _ 3 _ 4 _ 5 _ 8. 1 ■ 2 _ 3 _ 4 _ 5 _

If you feel as the client last felt, and you find yourself feeling somewhat happy, you would rate the word happy #2 to indicate, "I have this kind of feeling somewhat." Assuming happy has arabic numeral 9 preceding it, your answer sheet would appear as follows:

9. 1 _ 2 ■ 3 _ 4 _ 5 _

Please put yourself in the place of the client and, as honestly as possible, record the real feeling(s) the client is last experiencing. Be sure to rate each word.

CLIENT I Episode 1

1. Frustrated
2. Anxious
3. Lonely
4. Depressed
5. Ambivalent
6. Discouraged
7. Resistant
8. Confused - Mixed-up
9. Worried
10. Pushed

CLIENT I Episode 2

11. Satisifed
12. Abused
13. Anxious
14. Hopeful
15. Frustrated
16. Confused - Mixed-up
17. Baffled
18. Ambivalent

CLIENT I Episode 3

19. Hopeful
20. Optimistic
21. Refreshed
22. Relieved
23. Trapped
24. Annoyed - Angry
25. Anxious
26. Controlled
27. Depressed

CLIENT I Episode 4

28. Cautious
29. Appreciated
30. Controlled
31. Optimistic
32. Baffled
33. Ambivalent
34. Anxious
35. Embarrassed

CLIENT I Episode 5

36. Forced
37. Baffled
38. Resistant
39. Trapped
40. Reluctant
41. Upset - Disturbed
42. Annoyed - Angry

CLIENT II Episode 1

43. Baffled
44. Talkative
45. Surprised
46. Anxious
47. Annoyed - Angry
48. Embarrassed
49. Bitter

CLIENT II Episode 2

50. Forced
51. Embarrassed
52. Surprised
53. Trapped
54. Anxious
55. Scared
56. Frustrated

CLIENT II Episode 3

57. Frustrated
58. Daydreaming
59. Anxious
60. Guilty
61. Embarrassed
62. Confused - Mixed-up
63. Appreciated
64. Scared
65. Ambivalent

CLIENT III Episode 1

66. Determined
67. Anxious
68. Worried
69. Controlled
70. Defeated
71. Confused - Mixed-up
72. Listless - Indifferent
73. Annoyed

CLIENT IV Episode 1

81. Dependent
82. Cautious
83. Discouraged
84. Reluctant
85. Anxious
86. Forced
87. Worried
88. Embarrassed
89. Infantile

CLIENT III Episode 2

74. Listless - Indifferent
75. Controlled
76. Confused - Mixed-up
77. Bored
78. Helpful
79. Reluctant
80. Surprised

CLIENT IV Episode 2

90. Capable
91. Bitter
92. Disgusted - Fed-up
93. Determined
94. Annoyed - Angry
95. Abused

CLIENT IV Episode 3

95. Frustrated
97. Confused - Mixed-up
98. Baffled
99. Discouraged
100. Talkative
101. Annoyed - Angry
102. Abused

CLIENT V Episode 1

110. Anxious
111. Disgusted - Fed-up
112. Baffled
113. Upset - Disturbed
114. Confused - Mixed-up
115. Pushed
116. Bitter
117. Cautious

CLIENT IV Episode 4

103. Reluctant
104. Frustrated
105. Defeated
106. Careful
107. Depressed
108. Controlled
109. Hopeful

CLIENT V Episode 2

118. Abused
119. Disappointed
120. Frustrated
121. Annoyed - Angry
122. Comfortable - At ease
123. Anxious
124. Lonely

CLIENT V Episode 3

125. Inflexible
126. Depressed
127. Confused - Mixed-up
128. Stubborn - Resistant
129. Abused
130. Worried
131. Resentful

CLIENT V Episode 4

132. Lonely
133. Proud
134. Dependent
135. Erratic - Disorganized
136. Guilty
137. Anxious
138. Frustrated

CLIENT V Episode 5

139. Bitter
140. Discouraged
141. Resentful
142. Depressed
143. Annoyed - Angry
144. Abused
145. Ambivalent

CLIENT V Episode 6

146. Trapped
147. Bitter
148. Resentful
149. Annoyed - Angry
150. Disgusted
151. Abused

CLIENT VI Episode 1

152. Talkative

153. Anxious

154. Foolish - Ridiculous

155. Guilty

156. Embarrassed

CLIENT VI Episode 2

157. Satisfied

158. Appreciated

159. Guilty

160. Amused

CLIENT VI Episode 3

161. Impatient

162. Anxious

163. Stubborn - Resistant

164. Annoyed - Angry

165. Disappointed

166. Determined

CLIENT VI Episode 4

167. Happy

168. Relieved

169. Baffled

170. Confused - Mixed-up

171. Helpful

CLIENT VI Episode 5

172. Bitter
173. Disappointed
174. Annoyed - Angry
175. Exhausted
176. Defeated
177. Disgusted - Fed-up
178. Baffled

CLIENT VII Episode 2

185. Abused
186. Helpless
187. Ambivalent
188. Protected
189. Embarrassed
190. Anxious
191. Scared

CLIENT VII Episode 1

179. Comfortable - At ease
180. Anxious
181. Regretful
182. Ambivalent
183. Lonely
184. Disappointed

CLIENT VIII Episode 1

192. Confused - Mixed-up
193. Defeated
194. Embarrassed
195. Frustrated
196. Worried
197. Discouraged
198. Baffled

CLIENT VIII Episode 2

199. Frustrated

200. Impatient

1. Embarrassed[1]

2. Confused - Mixed-up

3. Foolish - Ridiculous

4. Bored

CLIENT IX Episode 2

11. Hopeful

12. Anxious

13. Refreshed

14. Optimistic

15. Courageous - Daring

16. Enthusiastic

CLIENT IX Episode 1

5. Helpful

6. Enthusiastic

7. Cautious

8. Surprised

9. Hopeful

10. Relieved

CLIENT IX Episode 3

17. Talkative

18. Disgusted - Fed-up

19. Amused

20. Annoyed - Angry

21. Bitter

22. Baffled

---

[1]This item began on a second answer sheet.

CLIENT IX Episode 4

23. Defeated
24. Bitter
25. Hopeless
26. Annoyed - Angry
27. Baffled
28. Discouraged
29. Disgusted - Fed-up

CLIENT X Episode 2

37. Daydreaming
38. Optimistic
39. Hopeful
40. Anxious
41. Confused - Mixed-up

CLIENT X Episode 1

30. Confused - Mixed-up
31. Upset - Disturbed
32. Discouraged
33. Baffled
34. Anxious
35. Scared
36. Ambivalent

CLIENT X Episode 3

42. Depressed
43. Worried
44. Helpless
45. Hopeless
46. Confused - Mixed-up
47. Bitter
48. Sensible - Reasonable

CLIENT X Episode 4

49. Defeated
50. Lonely
51. Bitter
52. Helpless
53. Upset - Disturbed
54. Anxious
55. Discouraged
56. Embarrassed

CLIENT XI Episode 1

57. Anxious
58. Embarrassed
59. Forced
60. Discouraged
61. Impatient
62. Scared

CLIENT XI Episode 2

63. Listless - Indifferent
64. Hopeful
65. Helpless
66. Impatient
67. Determined
68. Satisfied

CLIENT XI Episode 3

69. Anxious
70. Bored
71. Courageous - Daring
72. Stubborn - Resistant
73. Impatient
74. Ambivalent

CLIENT XI Episode 4

75. Protected

76. Ambivalent

77. Anxious

78. Sensible - Reasonable

79. Satisfied

80. Determined

APPENDIX C

ITEM ANALYSIS OF ITEMS FOUND NONSIGNIFICANT IN VALIDATING THE INSTRUMENT

FREQUENCY DISTRIBUTION AND SIGNIFICANCE OF ITEMS ON THE AFFECTIVE SENSITIVITY SCALE WHICH FAILED TO MEET THE .20 LEVEL OF SIGNIFICANCE IN VALIDATING THE INSTRUMENT[1]

Part A
(Median occurs between 2 and 3.)

Client and Scene	Item and Adjective	Group[2] Rating	Selected Response 1 or 2 %	Selected Response 3 or 4 %	Chi-square
I	5	High	30.56[3]	19.44	.450
1	Ambivalent	Low	25.00	25.00	
I	13	High	16.67	33.00	2.786[4]
2	Anxious	Low	30.56	19.44	
I	14	High	16.67	33.33	.554
2	Hopeful	Low	11.11	38.89	
I	22	High	27.78	22.22	1.003
3	Relieved	Low	19.44	30.56	
I	25	High	16.67	33.33	.468
3	Anxious	Low	22.22	27.78	
I	29	High	22.22	27.78	.468
4	Appreciated	Low	16.67	33.33	
I	32	High	16.67	33.33	.000
4	Baffled	Low	16.67	33.33	
I	39	High	25.00	25.00	.000
5	Trapped	Low	25.00	25.00	

[1]A chi-square value of 1.642 or higher was needed for the item to be significant at .20.

[2]Group rating refers to the top or bottom 33% of the sample based on ratings obtained from instructors and supervisors.

[3]The figure given in each cell is the percent of total responses (N=36).

[4]Eighteen items were listed as having failed the validation test even though the chi-square was significant because the direction of discriminating high from low empathizers reversed from the group used in constructing the items.

Part A (Continued)

Client and Scene	Item and Adjective	Group Rating	Selected Response 1 or 2 %	Selected Response 3 or 4 %	Chi-square
I	47	High	16.67	33.00	1.333
5	Annoyed-Angry	Low	8.33	41.67	
II	48	High	13.89	36.11	1.084
1	Embarrassed	Low	22.22	27.78	
II	50	High	36.11	13.89	2.857
2	Forced	Low	22.22	27.78	
II	53	High	19.44	30.56	1.003
2	Trapped	Low	27.78	22.22	
II	56	High	19.44	30.56	.120
2	Frustrated	Low	16.67	33.33	
II	57	High	30.56	19.44	.450
3	Frustrated	Low	25.00	25.00	
II	61	High	30.56	19.44	1.003
3	Embarrassed	Low	22.22	27.78	
II	62	High	33.33	16.67	.120
3	Confused-Mixed-up	Low	30.56	19.44	
II	63	High	13.89	36.11	.000
3	Appreciated	Low	13.89	36.11	
II	64	High	11.11	38.89	.800
3	Scared	Low	5.56	44.44	
III	67	High	19.44	30.56	.500
1	Anxious	Low	13.89	36.11	
III	68	High	38.89	11.11	.148
1	Worried	Low	36.11	13.89	
III	72	High	13.89	36.11	1.870
1	Listless-Indifferent	Low	25.00	25.00	

Part A (Continued)

Client and Scene	Item and Adjective	Group Rating	Selected Response 1 or 2 %	Selected Response 3 or 4 %	Chi-square
III	74	High	19.44	30.56	.120
2	Listless-Indifferent	Low	16.67	33.33	
III	77	High	38.89	11.11	4.208
2	Bored	Low	22.22	27.78	
III	78	High	13.89	36.11	.643
2	Helpful	Low	8.33	41.67	
III	79	High	25.00	25.00	.450
2	Reluctant	Low	19.44	30.56	
IV	85	High	25.00	25.00	4.500
1	Anxious	Low	8.33	41.67	
IV	87	High	19.44	30.56	2.215
1	Worried	Low	8.33	41.67	
IV	93	High	16.67	33.33	.000
2	Determined	Low	16.67	33.33	
IV	98	High	36.11	13.89	.500
3	Baffled	Low	30.56	19.44	
IV	101	High	27.78	22.22	.000
3	Annoyed-Angry	Low	27.78	22.22	
IV	103	High	22.22	27.78	.468
4	Reluctant	Low	16.67	33.33	
IV	104	High	30.56	19.44	1.003
4	Frustrated	Low	22.22	27.78	
IV	105	High	22.22	27.78	.114
4	Defeated	Low	19.44	30.56	
V	110	High	26.47	26.47	.537
1	Anxious	Low	29.41	17.65	
V	114	High	32.35	20.59	3.031
1	Confused-Mixed-up	Low	41.18	5.88	

Part A (Continued)

Client and Scene	Item and Adjective	Group Rating	Selected Response 1 or 2 %	Selected Response 3 or 4 %	Chi-square
V	126	High	29.41	23.53	.002
3	Depressed	Low	26.47	20.59	
V	127	High	38.24	14.71	.049
3	Confused-Mixed-up	Low	32.35	14.71	
V	128	High	32.35	20.59	.216
3	Stubborn-Resistant	Low	32.35	14.71	
V	129	High	26.47	26.47	.000
3	Abused	Low	23.53	23.53	
V	131	High	32.35	20.59	.216
3	Resentful	Low	32.35	14.71	
V	136	High	32.35	20.59	3.032
4	Guilty	Low	14.71	32.35	
V	138	High	35.29	17.65	.926
4	Frustrated	Low	38.24	8.82	
V	139	High	44.12	8.82	1.001
5	Bitter	Low	32.35	14.71	
V	140	High	35.29	17.65	.926
5	Discouraged	Low	38.24	8.82	
V	142	High	32.35	20.59	.747
5	Depressed	Low	35.29	11.76	
V	145	High	17.65	35.29	.971
5	Ambivalent	Low	23.53	23.53	
V	149	High	41.18	11.76	.551
6	Annoyed-Angry	Low	41.18	5.88	
VI	155	High	42.42	12.12	1.224
1	Guilty	Low	27.27	18.18	

Part A (Continued)

Client and Scene	Item and Adjective	Group Rating	Selected Response 1 or 2 %	Selected Response 3 or 4 %	Chi-square
VI 2	159 Guilty	High Low	15.15 24.24	39.39 21.21	2.238
VI 2	160 Amused	High Low	24.24 24.24	30.30 21.21	.259
VI 3	162 Anxious	High Low	30.30 21.21	24.24 24.24	.259
VI 3	163 Stubborn-Resistant	High Low	27.27 24.24	27.27 21.21	.036
VI 3	164 Annoyed-Angry	High Low	33.33 33.33	21.21 12.12	.550
VI 3	165 Disappointed	High Low	18.18 21.21	36.36 24.24	.609
VI 4	168 Relieved	High Low	36.36 27.27	18.18 18.18	.157
VI 5	172 Bitter	High Low	30.30 27.27	24.24 18.18	.066
VI 5	174 Annoyed-Angry	High Low	30.30 15.15	24.24 30.30	1.630
VI 5	178 Baffled	High Low	12.12 24.24	42.42 21.21	3.422
VII 1	181 Regretful	High Low	39.39 27.27	15.15 18.18	.550
VII 1	182 Ambivalent	High Low	24.24 27.27	30.30 18.18	.793
VII 1	184 Disappointed	High Low	30.30 21.21	24.24 24.24	.259

Part A (Continued)

Client and Scene	Item and Adjective	Group Rating	Selected Response 1 or 2 %	Selected Response 3 or 4 %	Chi-square
VII	186	High	39.39	15.15	.550
2	Helpless	Low	27.27	18.18	
VII	189	High	39.39	15.15	.550
2	Embarrassed	Low	27.27	18.18	
VII	190	High	36.36	18.18	.000
2	Anxious	Low	30.30	15.15	
VIII	193	High	39.39	15.15	3.478
1	Defeated	Low	18.18	27.27	
VIII	195	High	39.39	15.15	.269
1	Frustrated	Low	36.36	9.09	
VIII	196	High	36.36	18.18	.000
1	Worried	Low	30.30	15.15	
VIII	197	High	33.33	21.21	1.460
1	Discouraged	Low	18.18	27.27	
VIII	198	High	24.24	30.30	.259
1	Baffled	Low	24.24	21.21	
VIII	200	High	18.18	36.36	.000
2	Impatient	Low	15.15	30.30	
VIII	204	High	21.21	33.33	.203
2	Bored	Low	21.21	24.24	
IX	212	High	27.27	27.27	.330
2	Anxious	Low	18.18	27.27	
IX	215	High	36.36	18.18	.000
2	Courageous-Daring	Low	30.30	15.15	
IX	218	High	27.27	27.27	.036
3	Disgusted-Fed-up	Low	21.21	24.24	
IX	219	High	21.21	33.33	.550
3	Amused	Low	12.12	33.33	

Part A (Continued)

Client and Scene	Item and Adjective	Group Rating	Selected Response 1 or 2 %	Selected Response 3 or 4 %	Chi-square
IX 3	220 Annoyed-Angry	High Low	39.39 39.39	15.15 6.06	1.021
IX 4	224 Bitter	High Low	36.36 39.39	18.18 6.06	1.782
IX 4	226 Annoyed-Angry	High Low	36.36 36.36	18.18 9.09	.733
IX 4	229 Disgusted-Fed-up	High Low	33.33 27.27	21.21 18.18	.004
X 1	233 Baffled	High Low	36.36 36.36	18.18 9.09	.733
X 2	241 Confused-Mixed-up	High Low	18.18 15.15	36.36 30.30	.000
X 3	242 Depressed	High Low	33.33 24.24	21.21 21.21	.203
X 3	246 Confused-Mixed-up	High Low	18.18 9.09	36.36 36.36	.733
X 3	248 Sensible-Reasonable	High Low	39.39 33.33	15.15 12.12	.005
X 4	256 Embarrassed	High Low	18.18 12.12	36.36 33.33	.172
XI 1	257 Anxious	High Low	33.33 24.24	21.21 21.21	.203
XI 1	258 Embarrassed	High Low	36.36 33.33	18.18 12.12	.172

Part A (Continued)

Client and Scene	Item and Adjective	Group Rating	Selected Response 1 or 2 %	Selected Response 3 or 4 %	Chi-square
XI	259	High	30.30	24.24	.016
1	Forced	Low	24.24	21.21	
XI	260	High	15.15	39.39	.120
1	Discouraged	Low	15.15	30.30	
XI	261	High	24.24	30.30	.793
1	Impatient	Low	27.27	18.18	
XI	269	High	30.30	24.24	1.630
3	Anxious	Low	15.15	30.30	
XI	273	High	27.27	27.27	.930
3	Impatient	Low	30.30	15.15	
XI	277	High	21.21	33.33	.689
4	Anxious	Low	24.24	21.21	
XI	279	High	27.27	27.27	.036
4	Satisfied	Low	24.24	21.21	

Part B
(Median occurs between 1 and 2.)

Client and Scene	Item and Adjective	Group Rating	Selected Response 1 %	Selected Response 2 or 3 or 4 %	Chi-square
I	1	High	47.22	2.78	.364
1	Frustrated	Low	44.44	5.56	
I	2	High	47.22	2.78	.000
1	Anxious	Low	47.22	2.78	
I	4	High	47.22	2.78	1.125
1	Depressed	Low	41.67	8.33	
I	6	High	47.22	2.78	.364
1	Discouraged	Low	44.44	5.56	
I	8	High	47.22	2.78	.000
1	Confused-Mixed-up	Low	47.22	2.78	
I	9	High	44.44	5.56	.364
1	Worried	Low	47.22	2.78	
I	15	High	47.22	2.78	.000
2	Frustrated	Low	47.22	2.78	
I	16	High	47.22	2.78	.000
2	Confused-Mixed-up	Low	47.22	2.78	
I	17	High	44.44	5.56	.000
2	Baffled	Low	44.44	5.56	
I	20	High	44.44	5.56	.000
3	Optimistic	Low	44.44	5.56	
I	28	High	47.22	2.78	1.029
4	Cautious	Low	50.00	0.00	
I	36	High	38.89	11.11	.148
5	Forced	Low	36.11	13.89	
I	37	High	41.67	8.33	1.333
5	Baffled	Low	33.33	16.67	

Part B (Continued)

Client and Scene	Item and Adjective	Group Rating	Selected Response 1 %	Selected Response 2 or 3 or 4 %	Chi-square
II	43	High	38.89	11.11	1.178
1	Baffled	Low	30.56	19.44	
II	44	High	50.00	0.00	2.118
1	Talkative	Low	44.44	5.56	
II	45	High	47.22	2.78	.364
1	Surprised	Low	44.44	5.56	
II	51	High	47.22	2.78	1.029
2	Embarrassed	Low	50.00	0.00	
III	75	High	50.00	0.00	1.029
2	Controlled	Low	47.27	2.78	
IV	91	High	41.67	8.33	.232
2	Bitter	Low	44.44	5.56	
IV	92	High	44.44	5.56	.364
2	Disgusted-Fed-up	Low	47.22	2.78	
IV	96	High	47.22	2.78	1.029
3	Frustrated	Low	50.00	0.00	
IV	97	High	44.44	5.56	.000
3	Confused-Mixed-up	Low	44.44	5.56	
IV	102	High	47.22	2.78	.364
3	Abused	Low	44.44	5.56	
V	112	High	47.06	5.88	.016
1	Baffled	Low	41.18	5.88	
V	115	High	44.12	8.82	.117
1	Pushed	Low	41.18	5.88	
V	118	High	52.94	0.00	2.391
2	Abused	Low	41.18	5.88	
V	119	High	52.94	0.00	.000
2	Disappointed	Low	47.06	0.00	

Part B (Continued)

Client and Scene	Item and Adjective	Group Rating	Selected Response 1 %	Selected Response 2 or 3 or 4 %	Chi-square
V 2	120 Frustrated	High Low	52.94 47.06	0.00 0.00	.000
V 2	121 Annoyed-Angry	High Low	52.94 47.06	0.00 0.00	.000
V 2	124 Lonely	High Low	47.06 47.06	5.88 0.00	1.889
V 3	125 Inflexible	High Low	50.00 47.06	2.94 0.00	.916
V 3	130 Worried	High Low	52.94 47.06	0.00 0.00	.000
V 4	132 Lonely	High Low	52.94 47.06	0.00 0.00	.000
V 4	134 Dependent	High Low	50.00 44.12	2.94 2.94	.007
V 4	137 Anxious	High Low	52.94 47.06	0.00 0.00	.000
V 5	141 Resentful	High Low	52.94 44.12	0.00 2.94	1.159
V 5	143 Annoyed-Angry	High Low	50.00 44.12	2.94 2.94	.007
V 5	144 Abused	High Low	47.06 44.12	5.88 2.94	.294
V 6	146 Trapped	High Low	44.12 41.18	8.82 5.88	.117
V 6	147 Bitter	High Low	52.94 47.06	0.00 0.00	.000
V 6	148 Resentful	High Low	52.94 47.06	0.00 0.00	.000

Part B (Continued)

Client and Scene	Item and Adjective	Group Rating	Selected Response 1 %	Selected Response 2 or 3 or 4 %	Chi-square
V	150	High	50.00	2.94	.916
6	Disgusted	Low	47.06	0.00	
V	151	High	52.94	0.00	.000
6	Abused	Low	47.06	0.00	
VI	152	High	48.48	6.06	1.774
1	Talkative	Low	45.45	0.00	
VI	153	High	54.55	0.00	1.237
1	Anxious	Low	42.42	3.03	
VI	154	High	54.55	0.00	1.237
1	Foolish-Ridiculous	Low	42.42	3.03	
VI	156	High	54.55	0.00	.000
1	Embarrassed	Low	45.45	0.00	
VI	157	High	42.42	12.12	.024
2	Satisfied	Low	36.36	9.09	
VI	158	High	48.48	6.06	.196
2	Appreciated	Low	42.42	3.03	
VI	161	High	48.48	6.06	.038
3	Impatient	Low	39.39	6.06	
VI	166	High	48.48	6.06	.503
3	Determined	Low	36.36	9.09	
VI	171	High	45.45	9.09	.071
4	Helpful	Low	39.39	6.06	
VIII	201	High	48.48	6.06	1.774
2	Embarrassed	Low	45.45	6.80	
VIII	202	High	45.45	9.09	2.750
2	Confused-Mixed-up	Low	45.15	0.00	
IX	206	High	54.55	0.00	.000
1	Enthusiastic	Low	45.45	0.00	

Part B (Continued)

Client and Scene	Item and Adjective	Group Rating	Selected Response 1 %	Selected Response 2 or 3 or 4 %	Chi-square
IX	208	High	54.55	0.00	.000
1	Surprised	Low	45.45	0.00	
IX	209	High	54.55	0.00	1.237
1	Hopeful	Low	42.42	3.03	
IX	210	High	54.55	0.00	.000
1	Relieved	Low	45.45	0.00	
IX	211	High	51.52	3.03	.018
2	Hopeful	Low	42.42	3.03	
IX	213	High	48.48	6.06	.038
2	Refreshed	Low	39.39	6.06	
IX	216	High	54.55	0.00	.000
2	Enthusiastic	Low	45.45	0.00	
IX	217	High	42.42	12.12	1.540
3	Talkative	Low	42.42	3.03	
IX	225	High	51.52	3.03	.859
4	Hopeless	Low	45.45	0.00	
IX	228	High	51.52	3.03	.859
4	Discouraged	Low	45.45	0.00	
X	230	High	48.48	6.06	1.774
1	Confused-Mixed-up	Low	45.45	0.00	
X	231	High	48.48	6.06	1.774
1	Upset-Disturbed	Low	45.45	0.00	
X	234	High	54.55	0.00	.000
1	Anxious	Low	45.45	0.00	
X	235	High	48.48	6.06	.196
1	Scared	Low	42.42	3.03	
X	238	High	51.52	3.03	.018
2	Optimistic	Low	42.42	3.03	

Part B (Continued)

Client and Scene	Item and Adjective	Group Rating	Selected Response 1 %	Selected Response 2 or 3 or 4 %	Chi-square
X	239	High	54.55	0.00	.000
2	Hopeful	Low	45.45	0.00	
X	240	High	45.45	9.09	.768
2	Anxious	Low	42.42	3.03	
X	244	High	54.55	0.00	.000
3	Helpless	Low	45.45	0.00	
X	245	High	51.52	3.03	.599
3	Hopeless	Low	39.39	6.06	
X	249	High	54.55	0.00	.000
4	Defeated	Low	45.45	0.00	
X	250	High	54.55	0.00	.000
4	Lonely	Low	45.45	0.00	
X	252	High	54.55	0.00	.000
4	Helpless	Low	45.45	0.00	
X	253	High	45.45	9.09	.071
4	Upset-Disturbed	Low	39.39	6.06	
X	254	High	54.55	0.00	1.237
4	Anxious	Low	42.42	3.03	
X	255	High	54.55	0.00	.000
4	Discouraged	Low	45.45	0.00	
XI	263	High	48.48	6.06	.196
2	Listless-Indifferent	Low	42.42	3.03	
XI	275	High	51.52	3.03	.018
4	Protected	Low	42.42	3.03	

Part C
(Median occurs between 3 and 4.)

Client and Scene	Item and Adjective	Group Rating	Selected Response 1 or 2 or 3 %	Selected Response 4 %	Chi-square
I	7	High	2.78	47.22	1.029
1	Resistant	Low	0.00	50.00	
I	10	High	2.78	47.22	.000
1	Pushed	Low	2.78	47.22	
I	11	High	0.00	48.57	.972
2	Satisfied	Low	2.86	48.57	
I	12	High	2.78	47.22	.000
2	Anxious	Low	2.78	47.22	
I	27	High	0.00	50.00	1.029
3	Depressed	Low	2.78	47.22	
I	31	High	2.78	47.22	.000
4	Optimistic	Low	2.78	47.22	
I	34	High	0.00	50.00	1.029
4	Anxious	Low	2.78	47.22	
I	41	High	5.56	44.44	.364
5	Upset-Disturbed	Low	2.78	47.22	
I	42	High	8.33	41.67	1.125
5	Annoyed-Angry	Low	2.78	47.22	
II	55	High	2.78	47.22	.000
2	Scared	Low	2.78	47.22	
II	65	High	2.78	47.22	2.090
3	Ambivalent	Low	11.11	38.89	
III	66	High	0.00	50.00	.000
1	Determined	Low	0.00	50.00	
III	70	High	5.56	44.44	.232
1	Defeated	Low	8.33	41.67	

Part C (Continued)

Client and Scene	Item and Adjective	Group Rating	Selected Response 1 or 2 or 3 %	Selected Response 4 %	Chi-square
III	73	High	11.11	38.89	.177
1	Annoyed	Low	8.33	41.67	
III	80	High	0.00	50.00	.000
2	Surprised	Low	0.00	50.00	
IV	81	High	5.56	44.44	2.118
1	Dependent	Low	0.00	50.00	
IV	83	High	0.00	50.00	1.029
1	Discouraged	Low	2.78	47.22	
IV	84	High	8.33	41.67	.232
1	Reluctant	Low	5.56	44.44	
IV	89	High	2.78	47.22	1.125
1	Infantile	Low	8.33	41.67	
IV	107	High	0.00	50.00	1.029
4	Depressed	Low	2.78	47.22	
IV	109	High	2.78	47.22	1.029
4	Hopeful	Low	0.00	50.00	
V	122	High	2.94	50.00	.916
2	Comfortable-At Ease	Low	0.00	47.06	
VI	169	High	6.06	48.48	.038
4	Baffled	Low	6.06	39.39	
VI	170	High	6.06	48.48	.196
4	Confused-Mixed-up	Low	3.03	42.42	
VI	175	High	0.00	54.55	.000
5	Exhausted	Low	0.00	45.45	
VI	176	High	0.00	54.55	1.237
5	Defeated	Low	3.03	42.42	

Part C (Continued)

Client and Scene	Item and Adjective	Group Rating	Selected Response 1 or 2 or 3 %	Selected Response 4 %	Chi-square
VII	187	High	3.03	51.52	.018
2	Ambivalent	Low	3.03	42.42	
VII	188	High	0.00	54.55	.000
2	Protected	Low	0.00	45.45	
IX	205	High	0.00	54.55	.000
1	Helpful	Low	0.00	45.45	
IX	207	High	0.00	54.55	.000
1	Cautious	Low	0.00	45.45	
IX	222	High	0.00	54.55	.000
3	Baffled	Low	0.00	45.45	
IX	227	High	3.03	51.52	.859
4	Baffled	Low	0.00	45.45	
X	236	High	6.06	48.48	.038
1	Ambivalent	Low	6.06	39.39	
X	237	High	3.03	51.52	2.836
2	Daydreaming	Low	12.12	33.33	
XI	262	High	3.03	51.52	.599
1	Scared	Low	6.06	39.39	
XI	264	High	0.00	54.55	1.237
2	Hopeful	Low	3.03	42.42	
XI	266	High	3.03	51.52	.018
2	Impatient	Low	3.03	42.42	
XI	267	High	3.03	51.52	.018
2	Determined	Low	3.03	42.42	
XI	270	High	12.12	42.42	.024
3	Bored	Low	9.09	36.36	
XI	274	High	6.06	48.48	.196
3	Ambivalent	Low	3.03	42.42	

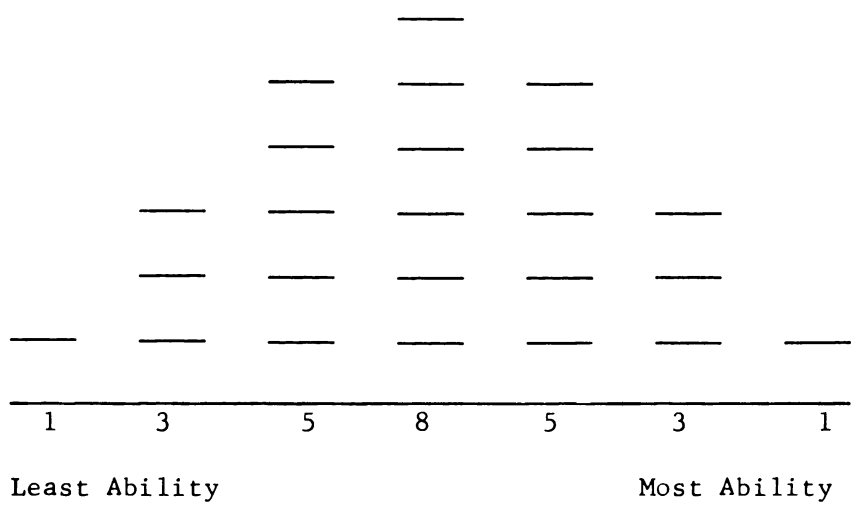
APPENDIX D

RATING SCALE FOR COUNSELING ABILITY

Instructions for Rating Counselor Trainees on Counseling Ability

Would you rate each person in the group in terms of his ability as a counselor. We are not concerned here with his ability as a general pupil personnel worker. We are interested only in counseling ability and skill.

As you rate, would you keep the following question in mind: "If I were to refer someone dear to me to a counselor, and needed to find a highly sensitive, skilled person how likely would I be to choose each of the members of this group?"



APPENDIX E

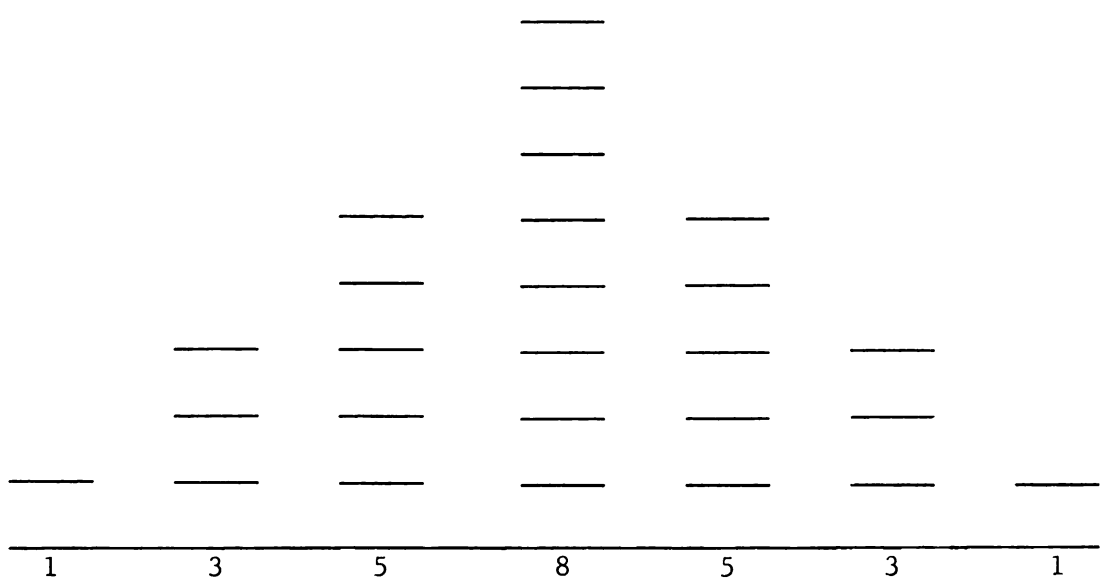
RATING SCALE FOR EMPATHY

Instructions for Rating Counselor Trainees on Empathy

We are asking that you rate a particular aspect of counseling ability, empathy. It is possible and likely that your ratings on this particular aspect will not correlate exactly with your overall ratings of the individuals as effective counselors. Because a number of definitions of empathy prevail, we are asking that you use this particular one:

Empathy is the ability or capacity to feel as another person does. In other words, if the client feels depressed, the counselor feels the client's depression. You can easily see from the example above that it is possible that some counselors may pick up the client's feelings but be unable to deal effectively with the feelings and hence be rated low in counselor effectiveness but high in empathy.

You have 28 clients to rate. We would like you to put their names into a normal distribution holding to the frequencies listed below.



1	3	5	8	5	3	1

Least Empathic Most Empathic

APPENDIX F

ITEM ANALYSIS OF ITEMS FOUND NONSIGNIFICANT IN CROSS-VALIDATING THE INSTRUMENT

FREQUENCY DISTRIBUTION AND SIGNIFICANCE OF ITEMS ON THE AFFECTIVE SENSITIVITY SCALE WHICH FAILED TO MEET THE .20 LEVEL OF SIGNIFICANCE IN CROSS-VALIDATING THE INSTRUMENT[1]

Part A
(Median occurs between 2 and 3.)

Client and Scene	Item and Adjective	Group Rating[2]	Selected Response 1 or 2 %	Selected Response 3 or 4 %	Chi-square
I	3	High	23.21[3]	26.79	.072[4]
1	Lonely	Low	25.00	25.00	
I	18	High	23.21	26.79	.000
2	Ambivalent	Low	23.21	26.79	
I	19	High	30.36	19.64	1.149
3	Hopeful	Low	23.21	26.79	
I	21	High	12.50	37.50	.000
3	Refreshed	Low	12.50	37.50	
I	33	High	16.07	33.93	1.198
4	Ambivalent	Low	23.21	26.79	
I	35	High	14.29	35.71	.717
4	Embarrassed	Low	19.64	30.36	
I	38	High	25.00	25.00	.072
5	Resistant	Low	26.79	23.21	
II	52	High	16.07	33.93	.080
2	Surprised	Low	17.86	32.14	
II	54	High	30.36	19.64	.000
2	Anxious	Low	30.36	19.64	

[1]A chi-square value of 1.642 or higher was needed for the item to be significant at .20.

[2]Group rating refers to the top or bottom 33% of the sample based on ratings obtained from instructors and supervisors.

[3]The figure given in each cell is the percent of total responses (N=56).

[4]Chi-square values were computed from frequency data, not from the percentage data reported here.

Part A (Continued)

Client and Scene	Item and Adjective	Group Rating	Selected Response 1 or 2 %	Selected Response 3 or 4 %	Chi-square
II	58	High	19.64	30.36	.000
3	Daydreaming	Low	19.64	30.36	
II	59	High	33.93	16.07	.350
3	Anxious	Low	37.50	12.50	
II	60	High	35.71	14.29	.381
3	Guilty	Low	39.29	10.71	
III	69	High	30.36	19.64	.074
1	Controlled	Low	28.57	21.43	
III	77	High	26.79	23.21	.286
2	Bored	Low	23.21	26.79	
IV	90	High	26.79	23.21	.072
2	Capable	Low	28.57	21.43	
IV	94	High	30.36	19.64	.000
2	Annoyed-Angry	Low	30.36	19.64	
IV	95	High	25.00	25.00	.000
2	Abused	Low	25.00	25.00	
IV	100	High	42.86	7.14	.487
3	Talkative	Low	39.29	10.71	
V	111	High	21.43	28.57	1.788[1]
1	Disgusted-Fed-up	Low	30.36	19.64	
V	113	High	33.93	16.07	1.524
1	Upset-Disturbed	Low	41.07	8.93	
V	117	High	25.00	25.00	.650
1	Cautious	Low	19.64	30.36	

[1]Some items will appear in this table where the chi-square value is higher than 1.642 because the response pattern of the cross-validation groups is opposite that of the criterion groups.

Part A (Continued)

Client and Scene	Item and Adjective	Group Rating	Selected Response 1 or 2 %	Selected Response 3 or 4 %	Chi-square
V 4	135 Erratic-Disorganized	High Low	16.07 23.21	33.93 26.79	1.198
VI 4	167 Happy	High Low	30.36 26.79	19.64 23.21	.292
VII 1	179 Comfortable-At ease	High Low	46.43 35.71	3.57 14.29	4.383
VII 1	180 Anxious	High Low	17.86 17.86	32.14 32.14	.000
VII 2	185 Abused	High Low	8.93 12.50	41.07 37.50	.424
VIII 1	194 Embarrassed	High Low	14.29 26.79	35.71 23.21	3.615
VIII 2	199 Frustrated	High Low	32.73 29.09	18.18 20.00	.147
VIII 2	203 Foolish-Ridiculous	High Low	10.71 21.43	39.29 28.57	2.947
IX 4	223 Defeated	High Low	71.43 71.43	28.57 28.57	.000
X 3	243 Worried	High Low	35.71 44.64	14.29 5.36	2.828
XI 3	271 Courageous-Daring	High Low	10.71 12.50	39.29 37.50	.100
XI 3	272 Stubborn-Resistant	High Low	48.21 12.50	1.79 37.50	29.944
XI 4	278 Sensible-Reasonable	High Low	23.21 17.86	26.79 32.14	.664

Part B
(Median occurs between 1 and 2.)

Client and Scene	Item and Adjective	Group Rating	Selected Response 1 %	Selected Response 2 or 3 or 4 %	Chi-square
I	26	High	3.57	46.43	1.469
3	Controlled	Low	8.93	41.07	
I	30	High	7.14	42.86	.487
4	Controlled	Low	10.71	39.29	
I	40	High	7.86	32.14	.076
5	Reluctant	Low	9.64	30.36	
II	46	High	8.93	41.07	.113
1	Anxious	Low	10.71	39.29	
IV	82	High	3.57	46.43	1.469
1	Cautious	Low	8.93	41.07	
IV	99	High	19.64	30.36	.074
3	Discouraged	Low	21.43	28.57	
IV	108	High	7.14	42.86	.163
4	Controlled	Low	5.36	44.64	
V	123	High	10.71	39.29	1.400
2	Anxious	Low	17.86	32.14	
V	133	High	10.71	39.29	.100
4	Proud	Low	12.50	37.50	
VI	173	High	16.07	33.93	.080
5	Disappointed	Low	17.86	32.14	
VIII	192	High	16.07	33.93	.000
1	Confused-Mixed-up	Low	16.07	33.93	
IX	214	High	19.64	30.36	.074
2	Optimistic	Low	21.43	28.57	
IX	221	High	10.71	39.29	1.191
3	Bitter	Low	5.36	44.64	
X	232	High	17.86	32.14	.760
1	Discouraged	Low	12.50	37.50	

Part C
(Median occurs between 3 and 4.)

Client and Scene	Item and Adjective	Group Rating	Selected Response 1 or 2 or 3 %	Selected Response 4 %	Chi-square
I	23	High	26.79	23.21	1.198
3	Trapped	Low	33.93	16.07	
I	24	High	32.14	17.86	.000
3	Annoyed-Angry	Low	32.14	17.86	
II	49	High	21.43	28.57	.000
1	Bitter	Low	21.43	28.57	
III	76	High	37.50	12.50	1.018
2	Confused-Mixed-up	Low	42.86	7.14	
IV	86	High	16.07	33.93	1.198
1	Forced	Low	23.21	26.79	
IV	88	High	23.21	26.79	.286
1	Embarrassed	Low	26.79	23.21	
IV	106	High	41.07	8.93	1.469
4	Careful	Low	46.43	3.57	
XI	276	High	25.00	25.00	8.187
4	Ambivalent	Low	42.86	7.14	

APPENDIX G

ITEMS SCORED AND CORRECT RESPONSE FOR OBTAINING TOTAL SCORE

ITEMS SCORED AND CORRECT RESPONSE
FOR OBTAINING TOTAL SCORE

Item Number	Correct Response*	Item Number	Correct Response
1	3 or 4	31	Not scored
2	Not scored**	32	3 or 4
3	Not scored	33	3 or 4
4	3 or 4	34	Not scored
5	1 or 2	35	Not scored
6	3 or 4	36	Not scored
7	Not scored	37	3 or 4
8	1 or 2	38	Not scored
9	3 or 4	39	1 or 2
10	Not scored	40	Not scored
11	Not scored	41	Not scored
12	Not scored	42	3 or 4
13	1 or 2	43	3 or 4
14	Not scored	44	3 or 4
15	3 or 4	45	Not scored
16	1 or 2	46	1 or 2
17	1 or 2	47	1 or 2
18	1 or 2	48	1 or 2
19	1 or 2	49	Not scored
20	1 or 2	50	3 or 4
21	Not scored	51	1 or 2
22	3 or 4	52	Not scored
23	3 or 4	53	Not scored
24	1 or 2	54	1 or 2
25	Not scored	55	Not scored
26	Not scored	56	3 or 4
27	Not scored	57	3 or 4
28	1 or 2	58	3 or 4
29	Not scored	59	Not scored
30	1 or 2	60	Not scored

*The data was dichotomized between responses 2 and 3 since direction of response was considered more important than degree of response. Thus, one of two responses was always correct.

**These items did not present a clear pattern of discrimination between high and low counselor criterion groups. They were used to constitute the 280 items as there was a slight pattern of discrimination.

Item Number	Correct Response	Item Number	Correct Response
61	Not scored	100	3 or 4
62	1 or 2	101	Not scored
63	1 or 2	102	1 or 2
64	Not scored	103	3 or 4
65	1 or 2	104	Not scored
66	Not scored	105	3 or 4
67	3 or 4	106	Not scored
68	3 or 4	107	3 or 4
69	Not scored	108	Not scored
70	3 or 4	109	Not scored
71	1 or 2	110	Not scored
72	1 or 2	111	1 or 2
73	Not scored	112	Not scored
74	3 or 4	113	3 or 4
75	1 or 2	114	Not scored
76	Not scored	115	Not scored
77	1 or 2	116	3 or 4
78	3 or 4	117	3 or 4
79	3 or 4	118	1 or 2
80	Not scored	119	Not scored
81	Not scored	120	1 or 2
82	1 or 2	121	1 or 2
83	3 or 4	122	3 or 4
84	1 or 2	123	3 or 4
85	3 or 4	124	3 or 4
86	Not scored	125	Not scored
87	3 or 4	126	3 or 4
88	Not scored	127	1 or 2
89	Not scored	128	1 or 2
90	3 or 4	129	3 or 4
91	1 or 2	130	Not scored
92	Not scored	131	Not scored
93	3 or 4	132	1 or 2
94	1 or 2	133	3 or 4
95	Not scored	134	Not scored
96	Not scored	135	3 or 4
97	Not scored	136	3 or 4
98	Not scored	137	Not scored
99	1 or 2	138	Not scored

Item Number	Correct Responses	Item Number	Correct Responses
139	1 or 2	178	Not scored
140	1 or 2	179	3 or 4
141	Not scored	180	3 or 4
142	3 or 4	181	Not scored
143	Not scored	182	Not scored
144	3 or 4	183	Not scored
145	1 or 2	184	3 or 4
146	3 or 4	185	Not scored
147	Not scored	186	3 or 4
148	Not scored	187	3 or 4
149	1 or 2	188	3 or 4
150	3 or 4	189	1 or 2
151	1 or 2	190	1 or 2
152	Not scored	191	3 or 4
153	Not scored	192	Not scored
154	3 or 4	193	3 or 4
155	1 or 2	194	1 or 2
156	1 or 2	195	3 or 4
157	3 or 4	196	Not scored
158	1 or 2	197	3 or 4
159	1 or 2	198	1 or 2
160	3 or 4	199	3 or 4
161	1 or 2	200	3 or 4
162	3 or 4	201	1 or 2
163	Not scored	202	Not scored
164	Not scored	203	Not scored
165	3 or 4	204	3 or 4
166	1 or 2	205	Not scored
167	1 or 2	206	3 or 4
168	3 or 4	207	3 or 4
169	1 or 2	208	Not scored
170	3 or 4	209	1 or 2
171	3 or 4	210	Not scored
172	Not scored	211	Not scored
173	3 or 4	212	3 or 4
174	1 or 2	213	3 or 4
175	3 or 4	214	Not scored
176	3 or 4	215	1 or 2
177	Not scored	216	1 or 2

Item Number	Correct Response	Item Number	Correct Response
217	Not scored	249	Not scored
218	Not scored	250	1 or 2
219	3 or 4	251	3 or 4
220	1 or 2	252	3 or 4
221	1 or 2	253	Not scored
222	3 or 4	254	1 or 2
223	3 or 4	255	1 or 2
224	1 or 2	256	1 or 2
225	3 or 4	257	Not scored
226	1 or 2	258	1 or 2
227	3 or 4	259	3 or 4
228	Not scored	260	3 or 4
229	Not scored	261	3 or 4
230	Not scored	262	Not scored
231	Not scored	263	Not scored
232	3 or 4	264	3 or 4
233	3 or 4	265	1 or 2
234	Not scored	266	3 or 4
235	Not scored	267	3 or 4
236	1 or 2	268	Not scored
237	3 or 4	269	3 or 4
238	1 or 2	270	Not scored
239	1 or 2	271	1 or 2
240	Not scored	272	3 or 4
241	3 or 4	273	Not scored
242	3 or 4	274	Not scored
243	Not scored	275	1 or 2
244	Not scored	276	1 or 2
245	3 or 4	277	Not scored
246	3 or 4	278	Not scored
247	Not scored	279	Not scored
248	Not scored	280	Not scored

BIBLIOGRAPHY

BIBLIOGRAPHY

Allport, G. W. Personality: A Psychological Interpretation. New York: Holt, 1937.

_____. "The Historical Background of Modern Social Psychology." In G. A. Lindzey (Ed.), Handbook of Social Psychology. Cambridge, Massachusetts: Addison-Wesley, 1954, 3-56.

Astin, Helen S. "A Comparative Study of the Situational and Predictive Approaches to the Measurement of Empathy." Unpublished doctoral dissertation, University of Maryland, 1957.

Bender, I. E. and A. H. Hastorf. "The Perceptions of Persons: Forecasting Another Person's Responses on Three Personality Scales." Journal of Abnormal and Social Psychology, 1950, 45, 556-561.

_____. "On Measuring Generalized Empathic Ability (Social Sensitivity)." Journal of Abnormal and Social Psychology, 1953, 48, 503-506.

Berne, E. Games People Play. New York: Grove Press, Inc., 1964.

Buchheimer, Arnold. "Progress Report--N.D.E.A. Grant # 7-42-0550-1670." Unpublished, Hunter College of the City University of New York, 1964.

_____. "The Development of Ideas About Empathy." Journal of Counseling Psychology, 1963, 10, 61-71.

Buros, O. K. The Fifth Mental Measurements Yearbook. Highland Park, New Jersey: Gryphon, 1959.

Cartwright, Rosalind D. and Barbara Lerner. "Empathy, Need to Change, and Improvement with Psychotherapy." Journal of Consulting Psychology, 1963, 27, 138-144.

Cowden, R. C. "Empathy or Projection." Journal of Clinical Psychology, 1955, 11, 188-190.

Cronbach, L. J. "Processes Affecting Scores on 'Understanding of Others' and 'Assumed Similarity'." Psychological Bulletin, 1955, 52, 177-193.

Crow, W. J. "A Methodological Study of Social Perceptiveness." Unpublished doctoral dissertation, University of Colorado, 1954.

Dymond, Rosalind F. "A Scale for the Measurement of Empathic Ability." Journal of Consulting Psychology, 1949, 13, 127-133.

_____. "Personality and Empathy." Journal of Consulting Psychology, 1950, 14, 343-350.

English, N. B. and Ava C. English. A Comprehensive Dictionary of Psychological and Psychoanalytical Terms. New York: Longmans, Green and Company, 1958.

Fenichel, O. The Psychoanalytic Theory of Neurosis. New York: W. W. Norton, 1945.

Fiedler, F. E. "Quantitative Studies on the Role of Therapists' Feelings Toward Their Patients." Psychotherapy Theory and Research. New York: Ronald Press, 1953, 296-316.

Gage, N. L. and L. J. Cronbach. "Conceptual and Methodological Problems in Interpersonal Perception." Psychological Review, 1955, 62, 411-421.

_____. "Accuracy of Social Perception and Effectiveness in Interpersonal Relationships." Journal of Personnel, 1953, 22, 128-141.

Gordon, Kate. "A Device for Demonstrating Empathy." Journal of Experimental Psychology, 1934, 17, 892-893.

Hastorf, A. H., D. E. Bender, and D. J. Weintraub. "The Influence of Response Patterns on the Refined Empathy Score." Journal of Abnormal and Social Psychology, 1955, 51.

Kagan, Norman, D. R. Krathwohl, and R. Miller. "Stimulated Recall in Therapy Using Videotape--A Case Study." Journal of Experimental Psychology, 1963, 10, 237-243.

Katz, R. L. Empathy: Its Nature and Uses. London: Free Press of Glencoe, 1963.

Kelly, E. L. and D. W. Fiske. The Prediction of Performance in Clinical Psychology. Ann Arbor: University of Michigan Press, 1951.

Kerr, W. A. and B. J. Speroff. "Validation and Evaluation of the Empathy Test." Journal of General Psychology, 1954, 50, 269-276.

Lesser, W. M. "The Relationship between Counseling Progress and Empathic Understanding." Unpublished doctoral dissertation, Michigan State University, 1958.

Lifton, W. M. "The Role of Empathy and Aesthetic Sensitivity in Counseling." Journal of Counseling Psychology, 1958, 5, 267-275.

Lindgren, H. E. and J. Robinson. "Evaluation of Dymond's Test of Insight and Empathy." Journal of Consulting Psychology, 1953, 17, 172-176.

Norman, R. D. and Patricia Ainsworth. "The Relationships Among Projection, Empathy, Reality, and Adjustment, Operationally Defined." Journal of Consulting Psychology, 1954, 18.

O'Hern, Jane S. and D. S. Arbuckle. "Sensitivity: A Measurable Concept?" Personnel and Guidance Journal, 1964, 42, 572-576.

Osgood, C. E. Method and Theory in Experimental Psychology. New York, Oxford University Press, 1953.

Rank, R. C. "The Assessment of Counselor-Trainee Perceptions of Interview Protocols Before and After an Intensive Practicum Experience." Unpublished doctoral dissertation, University of Minnesota, 1964.

Reid, Dorothy K. and W. V. Snyder. "Experiment on 'Recognition of Feeling' in Non-directive Psychotherapy." Journal of Clinical Psychology, 1947, 3, 128-135.

Rogers, C. R. "A Theory of Therapy, Personality, and Interpersonal Relationships, as Developed in the Client-Centered Framwork." In S. Koch (Ed.), Psychology: A Study of a Science. Vol. III: Formulations of the person in the social context. New York: McGraw-Hill Book Company, 1959, 184-258.

_____. Client-Centered Therapy. Boston: Houghton Mifflin Inc., 1951.

Sarbin, T. R., R. Taft, and D. E. Bailey. Clinical Inference and Cognitive Theory. New York: Holt, Rinehart and Winston, 1960.

Siegel, Sidney. Non-Parametric Statistics for the Behavioral Sciences. New York: McGraw-Hill Book Company, 1956, 107.

Stefflre, Buford. "Client Understanding Test (first version)." Unpublished, Michigan State University, 1962.

Stone, G. C. and G. S. Leavitt. "Generality of Accuracy in Perceiving Standard Persons." Paper read at Midwest Psychological Association Conference; Columbus, Ohio, May, 1954.

Sullivan, H. S. The Psychiatric Interview. New York: W. W. Norton, 1954.

Taft, Ronald. "The Ability to Judge People," Psychological Bulletin, 1955, 52, 1-23.

Thorndike, R. L. Review of Ken and Speroffs' "The Empathy Test" in O. K. Buros, The Fifth Mental Measurements Yearbook, Highland Park, New Jersey: Gryphon, 1959.

Walz, G. R. and J. A. Johnson. "Counselors Look at Themselves on Videotape," Journal of Counseling Psychology, 1963, 10, 232-236.

Weinstein, M. and F. D. McCandless. "Empathic Communication and Anxiety in Medical Students," Mental Hygiene, 1959, 43, 526-531.

Weitz, H. "Counseling as a Function of the Counselor's Personality." Personnel and Guidance Journal, 1957, 5, 277.

Wert, J. E., C. O. Neidt, and J. S. Ahmann. Statistical Methods in Educational and Psychological Research. New York: Appleton-Century-Crofts, Inc., 1954.