AN EVALUATION OF THE OBJECTIVITY OF AN AUDIO-VISUAL COUNSELOR SCALE WHICH COMPARES THE BEHAVIORS OF M. A. CANDIDATE COUNSELORS WITH THE BEHAVIORS OF PH. D. CANDIDATE COUNSELORS

Thesis for the Degree of Ph. D.

MICHIGAN STATE UNIVERSITY

Gerald G. Griffin

1966

This is to certify that the

thesis entitled

AN EVALUATION OF THE OBJECTIVITY OF AN AUDIO-VISUAL COUNSELOR SCALE WHICH COMPARES THE BEHAVIORS OF M.A. CANDIDATE COUNSELORS WITH THE BEHAVIORS OF PH.D. CANDIDATE COUNSELORS

presented by

Gerald G. Griffin

has been accepted towards fulfillment
of the requirements for

Doctoral degree in Philosophy

Major professor

Date 3-4-66

O-169

ABSTRACT

AN EVALUATION OF THE OBJECTIVITY OF AN
AUDIO-VISUAL COUNSELOR SCALE WHICH COMPARES
THE BEHAVIORS OF M.A. CANDIDATE COUNSELORS
WITH THE BEHAVIORS OF PH.D. CANDIDATE COUNSELORS

by Gerald G. Griffin

This research examined three hypotheses associated with the development of an Audio-Visual Counselor Behavior Scale (A-V Scale). This A-V Scale was developed for the purpose of making available an objective instrument which would compare the response patterns of Master's degree candidates in Counseling and Guidance (M.A. Candidates) with the response patterns of Doctoral degree candidates in Counseling and Guidance (Ph.D. candidates). The A-V Scale consists of five forced choice, dichotomous dimensions: the Affect-Cognitive, Understanding-Nonunderstanding, Specific-Nonspecific, Exploratory-Nonexploratory, and Effective-Noneffective dimensions.

The research consisted of three separate studies. The Scale Discrimination Study tested the hypothesis that the A-V Scale would significantly discriminate the response patterns of Ph.D. candidates from those of M.A. candidates on its five dimensions. The Objectivity Study tested the hypothesis that the A-V Scale would obtain high interjudge agreement when used for the rating of counselor behaviors on audio-video tapes. The Sensitivity Study tested the hypothesis that the A-V Scale is sensitive to change in the counselor behaviors measured by its dimensions as a result of supervision which focuses on behaviors within these

dimensions.

For the Scale Discrimination Study, three trained counselor judges used the A-V Scale to independently rate 20 responses from each of ten 30 minute audio-video taped counseling sessions. Five of the counselors engaged in these interviews were M.A. candidates. The other five counselors engaged in these interviews were Ph.D. candidates. Each of the ten interviews was an initial personal-social problem interview with a client.

The rating scores obtained by the two classes of counselors on each of the five dimensions of the A-V Scale were statistically compared using "t"-tests. There were significant differences in the response scores of the two classes of counselors on all five dimensions of the A-V Scale. The responses of the Ph.D. candidates were more frequently rated affect, understanding, specific, exploratory, and effective than were the responses of the M.A. candidates.

For the Objectivity Study, three trained counselor judges used the A-V Scale to independently rate four minute segments from each of 53 audio-video counseling tapes. Forty-five of these tapes involved M.A. candidates and eight involved Ph.D. candidates.

Hoyt's analysis of variance method was applied to the rating scores. Two types of interjudge reliability and internal consistency coefficients were obtained. One was an index of item objectivity and internal consistency, and the other was an index of a global rater objectivity and internal consistency. Item interjudge reliability coefficients of .89, .79, .76, .59, and .81, and global rater interjudge reliability coefficients

of .95, .91, .89, .70, and .93, were computed for the Affect-Cognitive, Understanding-Nonunderstanding, Specific-Nonspecific, Exploratory-Nonexploratory, and Effective-Noneffective A-V Scale dimensions respectively. For these same scale dimensions, item internal consistency coefficients of .96, .92, .91, .81, and .93 respectively, and global rater internal consistency coefficients of .98, .97, .96, .87, and .98 respectively were computed.

For the Sensitivity Study, ten NDEA Guidance and Counseling Institute trainee volunteers were randomly assigned to two groups of five; one group comprised a supervised counselor-interrogation group (Supervised Group), and the other group comprised a non-supervised client-interrogation group (Nonsupervised Group). Each member of both groups underwent six 30 minute video taped counseling sessions with a college or high school student. The Supervised Group focused on affective, understanding, specific, and exploratory responses during interrogation while the Nonsupervised Group had no particular focus during interrogation.

The middle twenty responses of the first (initial) and sixth (terminal) counseling sessions for each of the ten counselors comprising the two groups were rated by three trained counselor judges. These initial and terminal ratings within each of the Supervised and Nonsupervised Groups were statistically compared using dependent "t"-tests. The results of these tests indicate that for the Supervised Group there was a significant increase in the rating scores, from the initial to the terminal ratings, on the affect, understanding, specific, exploratory, and effective dimension categories of the Scale's five dimensions. For the

Nonsupervised Group there was no significant score increase, from the initial to the terminal ratings, on these five Scale dimension categories.

AN EVALUATION OF THE OBJECTIVITY OF AN AUDIO-VISUAL COUNSELOR SCALE WHICH COMPARES THE BEHAVIORS OF M.A. CANDIDATE COUNSELORS WITH THE BEHAVIORS OF PH.D. CANDIDATE COUNSELORS

By

Gerald G. Griffin

A THESIS

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

DOCTOR OF PHILOSOPHY

Department of Counseling, Personnel Services
and Educational Psychology

1966

ACKNOWLEDGMENTS

I wish to express my sincere gratitude and appreciation to my Major Professor and Program Chairman, Dr. Norman Kagan, not only for providing me the opportunity to pursue this research endeavor but also for the very helpful support and consultation he provided towards its fulfillment. In addition, I am especially grateful to Dr. Richard Rank, Dr. Bill Kell, and Dr. Bernard Corman for the scientific refinement they added to the research through their critical evaluations and individual contributions.

My sincere thanks is also extended to those many counselors who helped in the development of the A-V Scale, who participated in the studies of this research, and who devoted their time and skills as judges. Particularly do I wish to acknowledge my gratitude to Robert Ward, Barbara Fuller, John Hartzell, Joe Mezanno, Bill DeRoo, Patricia Parisian, Robert Campbell, and Al Goldberg. I would like to make note, too, of my appreciation to Lynn Mendzel for the many secretarial services she rendered in association with this research.

Finally, to my wife, Patricia, I wish to express my gratitude for her patient cooperation and understanding during the period of this research, and to extend my sympathetic thanks to her for enduring the various typing struggles she underwent in arriving at this final dissertation document.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF APPENDICES

CHAPTER I

THE PROBLEM

This research is primarily an objectivity evaluation of an Audio-Visual Counselor Behavior Scale (A-V Scale) developed to compare two classes of counselors at Michigan State University. Essentially, Master's degree candidates in Counseling and Guidance (M.A. candidates) constitute one class of counselors while Doctoral degree candidates in Counseling and Guidance (Ph.D. candidates) constitute the other class of counselors. The A-V Scale consists of five forced-choice, dichotomous dimensions which measure the extent to which each of these two classes is characterized by affective, understanding, specific, exploratory, and effective responses in initial personal-social problem interviews.

In the remainder of this chapter the problem will be more explicitly defined under the headings of Specific Problems, Definition of Terms, Delimitations, Basic Assumptions, Hypotheses, Need for Study, and Organization of the Study.

Specific Problems

The specific problems of this research are:

1. Can the two defined classes of counselors be significantly differentiated on any of the dimensions of the A-V Scale?
2. How objective is the A-V Scale for measuring these dimensions?

3. How sensitive is the A-V Scale for indicating change in counselor behaviors on its five dimensions as a result of supervision which focuses on behaviors within these dimensions?

Definition of Terms

The terms validity, objectivity, and reliability are defined in the following manner:

1. Validity. The extent to which the A-V Scale does the job it is employed to do. That is, the property of the scale to differentiate the video-verbal behaviors of M.A. candidates from the video-verbal behaviors of Ph.D. candidates.
2. Objectivity. Interjudge reliability based upon Hoyt's analysis of variance. This is the average reliability of ratings based upon the average of interjudge correlations.
3. Reliability. Reliability based upon the internal analysis of the average ratings of judges employing Hoyt's analysis of variance method.

Delimitations

The major delimitations of this study are:

1. The video taped counseling sessions employed in the individual studies making up this research, with the exception of one objectivity study, were based upon relatively small numbers of counselors (with diverse theoretical orientations) at Michigan State University.

This was necessary because of the special problems, expense, and time involved in video taping, and because of the limited number of counselors available for the type of video taping required.

2. The counselor verbal behaviors studied were laboratory behaviors in which both the counselors and clients were aware that their counseling interviews were being video taped.

Basic Assumptions

The basic assumptions upon which this research is based are:

1. Counselor responses which are affective or cognitive, understanding or nonunderstanding, specific or nonspecific, exploratory or nonexploratory, and effective or non-effective, are available on audio-video tapes for judges to observe and measure by a rating scale.
2. Both M.A. and Ph.D. candidates in Counseling and Guidance have a pattern and manner of counseling associated with these behavior dimensions which are consistent throughout their counseling interviews when such interviews pertain to a client's personal-social concerns.
3. Those A-V Scale dimension behaviors which are characteristic of the counseling practice of Ph.D. candidates, but are not characteristic of the counseling practice of M.A. candidates, are assumed to be learned by the Ph.D candidates as a result of their training.

Hypotheses

The general expectations which this research examined are:

First Hypothesis. The frequency of affective, understanding, specific, exploratory, and effective responses in initial personal-social problem counseling interviews, as measured by the A-V Scale, will be significantly greater for Ph.D. candidates in Counseling and Guidance than for M.A. candidates beginning their practicum experiences in Counseling and Guidance.

Second Hypothesis. Based upon Hoyt's analysis of variance method, the A-V Scale will have adequate objectivity for differentiating between Ph.D. and M.A. candidate responses on each of its five dimensions.

Third Hypothesis. Counselor trainees who receive professional supervision which focuses on examining affective, understanding, specific, and exploratory responses will receive higher frequency ratings for these categories on the A-V Scale than will counselor trainees who do not receive this professional supervision.

Need for Study

Most instruments which have been developed to measure counselor performance have had poor objectivity, and therefore, poor reliability. In recent years the demands of an ever expanding, service oriented society have created a tremendous press for counselors. These demands have not only necessitated an increased need for adequate counselor education programs, but they have necessitated the need

for adequate counselor education programs which could accelerate counselor training. However, the educational and counseling professions have been seriously hampered in their ability to experiment with various counselor accelerated programs because they have lacked an objective counselor behavior measuring instrument.

Organization of the Study

A review of relevant rating measurement methodology, related research, and common counseling elements are presented in the following chapter. The development of the A-V Scale is presented in Chapter III. The methodology for testing the three hypotheses set forth in the present chapter is presented in Chapter IV. The results of this research are presented in Chapter V. The summary, conclusions, and implications are presented in Chapter VI.

CHAPTER II

REVIEW OF LITERATURE

The review of literature presented here is divided into four areas, each with a summary: relevant general principles in the construction and use of rating scales; methodological considerations in the development of rating scales; previous counselor behavior measurement; and common elements of counseling practice that are important in their evaluation. The terms "therapist" and "counselor" are used interchangeably.

Construction and Use of Rating Scales

A general review of the value and use of rating scales, and the important considerations associated with their development, are presented in this section.

"...A rating instrument generally consists of two components; a description of the behaviors to be considered and a set of alternative responses from which the rater is to choose for each person rated" (Helmstadter, 1964). A rating scale is one of the most popular techniques for describing human characteristics (Helmstadter, 1964: Torgerson, 1958). A rating scale, states Helmstadter, is the most direct [illegible] for determining the way in which others react to an individual

in question, and for this situation they are valid by definition. Helmstadter (1964, p. 199) remarks, "...if sufficient care and trouble are taken, rating scales can be used to obtain highly satisfactory quantitative descriptions of many behaviors which cannot be studied by other measuring devices."

Test construction is a never ending process (Tyler, 1964). Rating scale development is a special form of this process. In the initial stages of test construction there are no specific rules or techniques which can tell a person exactly what to measure in a particular situation. The researcher ultimately decides what the measure will include based upon his experience and judgmental skill (Helmstadter, 1964). However, Helmstadter presents the following steps which a test constructor must follow if his instrument is to be based upon sound construction.

1. The first and most essential step is a careful delimitation and breakdown of the area(s) or trait(s) involved.
2. Once the areas of knowledge and/or particular traits which need to be included have thus been laid out, the next step is to list the behaviors which the examinee should exhibit if he possesses the knowledge or characteristics under consideration.
3. When the test items have been written, the next major step is that of trying them out on an analysis group of subjects who are representative of the population for which the total test is being prepared.

4. Next, the items are to be organized into a resonably administered format and an appropriate system for scoring the test is to be developed.
5. The last logical step in development is that of making a final evaluation of the test, as revised and reorganized, on a new analysis group.

Helmstadter (1964) points out that even with the best possible psychological measuring instruments, errors will be made. He mentions that what test constructors should do is, "...to learn what sorts of error can occur; to develop procedures for estimating this error; and, insofar as possible, to utilize methods of test construction and administration which minimize these errors."

When rating scales are used, errors arise from rater characteristics and from the behavior selected for measurement. Rater biases and constant errors of leniency (tendency to be overly generous in descriptions), halo effect (rating a given individual in the same way on all behaviors), and error of central tendency (the avoidance of using extreme positions on a rating scale) are errors which arise from rater characteristics (Helmstadter, 1964).

One of the most recent innovations in the rating scale area has been the development of a forced choice technique which has been specifically designed to overcome these difficulties (Highland and Berkshire, 1951). In a forced choice rating the judge is required to consider not just one attribute, but several

characteristics all at one time. The general format requires that only one or a few of the behaviors listed in each item be selected as applicable.

To reduce error arising from the behavior selected, Helmstadter suggests selecting behaviors which can be dependably rated. He points out that the interobserver agreement (objectivity) is highest for those characteristics which are specific entities. The major logical objection to the rating scale approach is that ratings are often based on only broad impressions about the person being rated. Therefore, for rating scales to be useful to researchers and workers in applied fields, numerical ratings must in some way relate to relevant observable behaviors. Recent studies (Gebhart and Hoyt, 1958; Guilford, 1959; and Thorton, 1943 and 1944) illustrate that ratings of specifically defined characteristics obtained on carefully developed scales can have validity, while judgments of broadly defined traits are often subject to irrelevant influences.

Helmstadter (1964) offers four major techniques for the improvement of ratings. First, rate only relatively specific characteristics which "...are readily discernible and have a highly consistent meaning for a wide variety of judges." Second, be careful in the preparation of the rating scale itself. Select behavior labels which "...are short enough to fit on the scale yet sufficiently unambiguous to result in general uniformity of meaning for all potential raters. Include a concise description and a few examples along with each behavior label." Decide upon the best number of steps to be used. Five and eleven seem

to be the optimal number of steps above and below which it does not pay to proceed (Champney and Marshall, 1939). After the individual behavior scales have been developed, organize them into a complete rating instrument. Third, "...use careful selection and training of the judges who will do the rating." Forth, "...have the judges make their ratings independently and then pool their results."

Summary

A rating scale consists of a description of the behaviors to be considered and a set of alternative responses from which the rater is to choose for each person rated. It is the most direct measure currently available for determining the way in which others react. Common errors of the rating approach, such as rater biases, errors of leniency, halo effect, and errors of central tendency, are generally avoided by the forced choice technique. If sufficient care and trouble are taken in the development and use of rating instruments they can be used to obtain highly satisfactory quantitative descriptions of many behaviors which cannot be studied by other measuring devices. Ratings are improved by selection of traits for measurement which relate to specific and relevant observable behaviors; careful preparation of the rating scale itself, including appropriate selection and description of behavior levels and proper organization of individual behavior scales into a complete rating instrument; careful selection and training of the judges who will do the rating; and the pooling of the results of independent judge

ratings.

Rating Scales and Methodological Considerations

A discussion of the pertinent methodological considerations related to rating scales is presented in this section. These considerations refer to reliability, objectivity, and validity, each of which is associated with a different kind of error, and each of which represents a separate criterion for evaluating a test (Helmstadter, 1964).

With any test, reliability of the technique and validity of the results are two highly significant methodological considerations (Tyler, 1964; Helmstadter, 1964; Ruch, 1963; Lindquest, 1961). In the case of rating scales, which are used by a group of judges to rate others, a reliability consideration of importance is that of interjudge agreement (Helmstadter, 1964; Ebel, 1951; Hoyt, 1941). The central methodological question for the rating scale, then, may be stated: how useful (valid) and objective (degree of interjudge reliability) is the scale for the purpose for which it is being used?

Following is a general overview of reliability, objectivity, and validity considerations and their relation to measurement by rating scales.

Reliability

Ruch (1963) emphasizes that a measuring device cannot be valid unless it is first of all reliable. The 1959 Glossary of Measurement Terms," published by the California Test Bureau,

defines reliability (in common with others such as Buch, 1963 and Tyler, 1963) as the degree to which an individual would obtain the same score on a test if the test were readministered to the same individual. This "Glossary of Measurement Terms" distinguishes three types of reliability coefficients:

1. Coefficient of stability

   A correlation between test and retest with some period of time intervening. The test-retest situation may be with two forms of the same test.

2. Coefficient of equivalence

   A correlation between scores from two forms given at essentially the same time.

3. Coefficient of internal consistency

   A measure based on internal analysis of data obtained on a single trial of a test. More prominent of these are the split-half method and the Kuder-Richardson and Hoyt analysis of variance methods.

Reliability, in each of these cases, is defined as "...the ratio of true score variance to observed score variance and is an index of the amount of variable error (chance errors varying from person to person and from one time to the next) in a test" (Helmstadter, 1964).

There are four classical ways of estimating reliability (test-retest, parallel form, split-half, and Kuder-Richardson), all of which may be expressed in a form analogous to the quotient of the observed variance minus the error variance to the observed variance. Each of these procedures is basically different in that "...it defines what is meant by error in a slightly different way. Helmstadter (1964) points out that the procedure for

determining the test-retest correlation ...defines as variable error any time to time fluctuations;" the procedure for determining parallel form reliability "...defines as variable error both time to time and form to form fluctuations;" the procedure for determining split-half reliability "...defines variable error only as fluctuations from one half of the test to the next;" the procedure for determining Kuder-Richardson reliability (Kuder and Richardson, 1937) "...defines variable error in terms of consistency of performance from item to item."

Hoyt (1941) has presented a procedure for estimating reliability which defines a variable error in a somewhat different way from these other approaches, although the results are identical with Kuder-Richardson reliability. Hoyt's final basic way of defining what is meant by variable error involves a slightly different breakdown of observed score variance from what is usually presented. According to Hoyt's formulation, "...variation in the performance of an individual from item to item is not considered error. Rather, it is a real (non-error) difference, an intraindividual difference, and one which should not be involved in the estimation of reliability." Total variation observed, then, is conceived to be made up of three components: true interindividual differences, intraindividual differences, and error interindividual differences. According to Hoyt a better definition of reliability is expressed by the quotient of the true interindividual variance (error variance subtracted from the observed variance minus the intraindividual variance) to the observed variance minus the intraindividual variance. Using an analysis of variance procedure, then, Hoyt's reliability can

be estimated by

$$r_{xx'} = \frac{\text{MS individuals} - \text{MS residuals}}{\text{MS individuals}}$$

where:

MS individuals = mean square of deviations from the individual's means

MS residuals = mean square of the deviations left over after individual and item variation have been removed.

The reliability of rating scales is affected by the major considerations for formulating an observational measurement. The number of scale categories and the effect of the length of the judged material appear to influence reliability (Helmstadter, 1964; Bendig, 1954; Bendig and Sprague, 1954; Symonds, 1922). Reliability of rating scales is enhanced by clear and unambiguous directions (Helmstadter, 1964; Bendig, 1953), preciseness of rating criteria (Helmstadter, 1964; Rose, 1947), uniform scoring procedures (Hamlin, 1954), use of adequate samples of observed behavior (Helmstadter, 1964; Hamlin, 1954), verbal definitions of scale categories and illustrations of behaviors which fit various scale categories (Bendig, 1953; Albee and Hamlin, 1950), and the pooling of independent judge ratings (Helmstadter, 1964; Gordon, 1924). The factor which influences scale reliability the most, however, is interjudge agreement (Helmstadter, 1964; Ruch, 1963; Bendig, 1955; Bendig and Sprague, 1954).

Objectivity

Helmstadter (1964) defines objectivity as the extent to which

personal errors have been avoided in rating. Thus, a measure which is relatively free from personal errors is referred to as objective. Personal errors refer to inconsistencies among several different persons scoring the same instrument or individuals. Helmstadter states, "...any two persons observing exactly the same responses are likely to record different scores, because they can see the performance only from their own position or bias." Since such fluctuation in scores is directly attributable to the judge doing the rating, they are called personal errors. Interjudge reliability is the index which reflects the extent to which a measure is free from this type of error. Hoyt's analysis of variance method (Hoyt, 1941) is one which provides such an index. It provides a reliability coefficient which indicates the degree of consistency which is obtained when several different judges rate the same test or individual.

Helmstadter (1964) indicates that objectivity is a special form of reliability. "...Because the personal biases of judges making observations fluctuate from time to time and from subject to subject, the resulting personal errors represent a particular type of variable error. Because of this, no new index other than the index of reliability is normally needed to assess an instrument in this respect." However, Helmstadter stresses that when judges are used to rate responses, personal errors become of sufficient magnitude to require special attention and therefore deserve a special label. This is the index of objectivity.

Lack of objectivity is directly reflected as a lack of test reliability (Helmstadter, 1964; Ruch, 1963). Ruch (1963)

states "...in order to be reliable, a measuring instrument must be objective, so that two or more individuals can score the subject's responses and obtain the same result." Helmstadter (1964) states, "...a separate index of objectivity...may prove valuable to a test constructor by suggesting a specific cause of unreliability in an instrument under development." Other things being equal, the greater the objectivity of an instrument, the greater its reliability will be (Helmstadter, 1964).

Ruch (1963) suggests that the best way to insure objectivity in a measuring instrument is not to include items which must be scored on the basis of subjective judgment. The fewer the responses which must be scored by subjective judgment, the greater the probability two or more persons can score the same responses and get the same results. Ruch points out, though, that in the measurement of human behavior people must act as the measuring instrument, making it nearly impossible to obtain objective measurement uninfluenced by personal perception. However, one may assume that through the use of professional judges, those who are somewhat expert at what they are measuring, this subjectivity may be minimized. This assumption is supported by Cattell's original statement in 1903 that, "...there is, however, no other criterion for a man's work than the estimation in which it is held by those most competent to judge."

Shapiro and Eberhart (1947) recognize that perfect accuracy is neither necessary nor usually possible in observational judgments. They make recommendations that are believed to facilitate reliability, e.g., techniques to eliminate "apparent

interviewer bias." These types of techniques to insure similarity between the interviewers or judges have been applied and analyzed in various social settings by Feldman, Hyman, and Hart (1951). The use of an interview or pretest in the formulation of a rating scale has proven to be especially valuable by Rose (1947).

The personal characteristics of the judges appear to be significant in observational analysis. Bendig and Sprague (1954) found that the individual rater reliability increased with higher levels of education, and the variance depended greatly upon the length of the scale and the experience of the rater. In a follow-up study using the same rating scale, Bendig (1955) again found that the educational level of the judge affected the reliability. More bias measures were found in the undergraduate "lower education" judges.

Helmstadter (1964) declares that the careful selection and training of those judges who will do the rating is one of the most effective ways of insuring adequate results from the rating scale approach to measurement. He indicates that the first rule it to select raters who will have the opportunity to make the necessary observations of the subject in pertinent situations. Next, it is important to try to secure raters who are willing to make judgments free of distortion. One of the first tasks in the training of raters to use a scale appropriately, notes Helmstadter, is to convince them of the value of honest and accurate ratings. It also helps to point out the types of errors which raters are likely to make and provide some suggestions for avoiding them.

Validity

Validity is the most important consideration in the construction and use of all types of tests (Helmstadter, 1964; Tyler, 1963). Reliability is merely a means to the end of validity (Ruch, 1963). A measuring device cannot be valid unless it is first of all reliable.

Helmstadter (1964) states that "...measurement of behavioral characteristics is indirect. Because of this, one cannot be completely certain that a test measures the precise characteristic for which it was designed. Thus, an error can be made in using tests which is the same for every individual to which the test is applied and the same every time it is used. Such errors are called constant errors. In test measurement, the problem of these constant errors is the problem of validity." Validity is an indication of the extent to which a test is not influenced by constant error.

Cronbach (1946) and others point out that validity may have a variety of meanings. The two basic approaches to the determination of validity are logical analysis and empirical analysis ("Glossary of Measurement Terms", 1959).

Content validity and item structure are types of validity determined by logical analysis. Content validity refers to "...how well the content of the test samples the subject matter or situation about which conclusions are to be drawn." This may be determined by the opinions of a jury of experts. Content validity includes face validity, which involves a simple evaluation of content adequacy; logical validity, in which it is

carefully determined whether the test behavior is a representative sample of the behavior in question; and factorial validity, in which the extent to which certain factors contribute to total test scores are determined by examining the many relationships among test-taking behaviors (Helmstadter, 1964). Item structure includes corroborative evidence from item analysis supporting the other characteristics of the test and item composition emphasizing conceptual clarity in the expression of items.

Data showing the relationship between test and some later (criterion) behavior is called empirical validity. Concurrent validity and predictive validity are types of validity determined by empirical analysis ("Glossary of Measurement Terms", 1959). Concurrent validity refers to how well test scores match measures of contemporary criterion performance. Predictive validity relates to how well predictions made from the test are confirmed by the data collected at a later time.

A type of validity in which both logical and empirical methods may be used in its determination is construct validity, validity in which the theory underlying a test is validated. Cronbach (1957) defines construct validity as an analysis of the meaning of test scores in terms of psychological constructs. Rank (1964) points out that this definition requires that in order to establish construct validity, the test-maker must measure a construct which is tied into an appropriate theoretical framework. Helmstadter (1964) states that construct validity "...is evidence which seems to establish the test as a measure of some hypothesized trait which may affect performance in a

variety of ways in many different kinds of situations." He adds that "...construct validity involves such diverse evidences as group differences, performance change as a result of experimental manipulation of certain variables, multitrait-multimethod matrices of correlations, internal consistency information, and studies of the test-taking process." Meehl and MacCorquodale (55) suggest that no test has approached the ideal level of construct validity, indicating that this is a difficult type of validity to establish for a test.

Helmstadter (1964) feels that it is rather naive to ask whether a test is valid. Cronbach and Meehl (1955) point out that a test is never really "validated." Rather, a principle for making certain kinds of inferences about persons who obtain given test scores is verified or refuted. Information that is gathered to determine exactly what kind of inferences can be drawn from test scores is referred to as validity evidence. The question should not be "Is the test valid?" but "Is the test valid for such and such?"

Buch (1962) states that "...the validity of a measuring instrument is the extent to which it measures that which it was designed to measure." Lindquest (1961) points out that validity is always validity for a particular purpose with the essential question being how well does a test do the job it is employed to do. Lindquest (1961, p. 622) states:

> The validity of any test is its validity as an indicator of individual differences in some particular function among the members of some specified group.

Tyler (1963) feels in examining validity that the proper question to ask is "...just what is it that this test does measure?" rather that the old validity question of "...to what extent does this test measure what it purports to measure?" He feels that the content of the test must be analyzed and many correlations examined with different criteria in various groups before the answer can be known. Tyler's view is that it is not necessary for a psychologist to have an exact definition of what he hopes his test will measure at the beginning of his test development. Tyler (1963, p. 29) comments:

> ...if he has a general idea about the characteristic and its relationship to either theoretical concepts or practical situations, his precision in defining it will increase as he tries out his test in series of separate research studies.

Bayroff, Haggerty, and Rundquist (1954) point out that the use of rating scales and judges for the assessment of segments of behavior carries with it certain problems and drawbacks. One of the most important problems seems to be the obtaining of maximum rating validity. These writers note that ratings are subject both to rater contamination (agreement of ratings due to the fact that the same person makes two evaluations) and technique contamination (extent to which validity coefficients are artificially inflated because of resemblance of technique). They found that the most effective way of increasing the rating validity was to:

1. Average the ratings made by the raters.
2. Have raters identify their ratings.
3. Have the crucial ratings completed early in the series.

In 1921 Rugg stated that the rating of human character was practicable "...if the rating is done under conditions as rigorous as the following:"

1. If each final rating given a person is the average of three independent ratings, each one made on a scale as objectified as the man-to-man comparison type scale.

2. If the scales on which the ratings are made are comparable and equivalent, having been made in conferences under the instruction of one skilled in rating scale work.

3. If the three raters are so thoroughly acquainted with the person rated that they are competent to rate.

If validation evidence identifies a number of scale items which significantly distinguish between the criterion groups obtained in one sampling, the test constructor may wish to try them out on a second sample. This is known as cross-validation. If the same results occur on the second sample, the researcher can have a fair degree of confidence that his finding will hold up in subsequent studies (Helmstadter, 1964). Cross-validation is accomplished by trying out a previously developed and refined scale on a new group that resembles closely the one for which the scale will ultimately be used.

Summary

Methodological problems are the most prevelant questions regarding rating. Two of the most highly significant methodological considerations concern the reliability of the rating technique and the validity of the rating results. The reliability of a rating technique refers to its consistency as a measuring instrument.

This indicates its relative freedom from variable error. Personal errors, which are due to the impressions and biases of the raters, are a special class of variable error. They take on sufficient magnitude when rating scales are used. A special index of reliability, known as objectivity, reflects the extent to which personal errors have been avoided in ratings. Lack of objectivity is one of the most common causes of unreliability. The greater the objectivity of an instrument, the greater its reliability will be.

Using pretests in the formulation of rating scales, developing precise rating criteria, providing clear and unambiguous directions and category definitions, providing examples of scale categories, following uniform scoring procedures, providing adequate samples of behavior to be rated, and pooling of independent judge ratings are all techniques which reduce variable error and thus increase scale reliability. Personal errors are reduced, and thus objectivity and reliability increased, when careful attention is given to the selection and training of the judges who are to do the rating. The educational level and knowledge of the judges seems to be of particular importance.

The validity of test results is the most important consideration in the construction and use of a rating instrument. The problem of constant error is the problem of validity. Information that is gathered to determine exactly what kind of inferences can be made from the test scores is referred to as validity evidence. The major validity criterion of any instrument is how well does the instrument do the job it was employed to do. The validity of

rating results is increased, according to expert opinion supported by some experimental studies, by averaging judge ratings, having judges identify their ratings, and having the crucial ratings completed early in the series.

Counselor Behavior Measurement

Some of the most relevant research efforts over the past years to objectively measure various aspects of counselor behavior are reviewed in this section. A brief historical overview of counselor measurement efforts preceeds a discussion of the significant counselor behavior scales related to the present research.

History

Laswell (1929) reported the need for more adequate recording of the data of the analytic hour if such interviews were to be used for research. In subsequent articles (1935, 1936) he reported the usefulness that electrically recorded verbatim data could have for this purpose. Symonds (1939) suggested to psychologists that phonographic recordings might provide the objective data for the study of the interview process. Rogers (1942) described the ways in which such data might be used for training and research, and Covner, one of his students, wrote a series of articles (1942, 1944) spelling out the details of the rationale and technique of phonographic recording. Subsequent research under the general supervision of Carl Rogers validated the assumptions regarding the use that could be made of verbatim data.

It was recognized by early investigators that in order to conform to the rigors of the more objective approach characteristic of psychology, it was deemed necessary to translate the raw data of the interview into a form that could be amenable to statistical analyses. Lasswell (1938) devised a system of categorizing what the patient said during the psychoanalytic interview by designing codifications for the patient's positive and negative attitudes towards himself, towards others, and towards the therapist. A system was developed by Porter (1943) for the codification of what the therapist said. Porter's system of categories was refined by Snyder (1945), who added categories regarding the statements made by the patient and, in part, studied their interaction.

Other attempts to objectively describe the activity of the therapist followed the early efforts of Lasswell, Porter, and Snyder. In 1948 Keet codified the statements of the therapist to differentiate between expressive and interpretive activity. Finesinger (1948) classified therapeutic activity into three levels of activity: low, moderate, and marked, while Dittman (1952) classified therapist activity into questions reflecting feeling, clarifying feelings already expressed, clarifying feelings somewhat less surface, and depth interpretations. Similarly, Carnes and Robinson (1955) categorized therapeutic activity into clarifications, tentative analyses, interpretations, and urgings. In 1953, Collier focused on a continuum of activity level, while Harway, et. al. (1955) and Rausch, et. al. (1956) focused on a dimensional analysis of interpretive activity.

Significant Counselor Behavior Scales

The earliest validated counselor-rating scales of any significance are Rogers' Relationship Inventory Scale (1962) and Andersons' Interview Rating Scale (1962). Rogers' scale is one which permits the client to rate the counselor. It consists of three sections which focus on the type of relationship provided by the counselor. The first section consists of 72 items which describe possible positive and negative feelings and attitudes which the counselor may have toward the client; e.g., "at times he feels contempt for me." The client is asked to rate on a four point scale, ranging from, "I strongly feel this is true," to "I strongly feel that it is not true," the extent to which he feels each of these 72 items apply to his counselor as he relates to the client. The second section of Rogers' scale consists of two questions asking the client to rate, along a seven point scale ranging from extremely dissatisfied to completely satisfied, how he feels about the relationship with the counselor. The third section consists of eight sentence completions concerning the counselor and his counseling; e.g., "the counseling seemed very _____." Although the Rogers' Relationship Inventory Scale has demonstrated some utility for counselor evaluation by clients, its major limitation is that it is tied too closely to the theoretical concepts of Rogers' own counseling viewpoint. It focuses upon Empathy, Positive Regard, Congruence, and Unconditionality. Thus, it contains no questions related to effective or ineffective use of counselor questions, confrontations, interpretations, and

other counselor behaviors not emphasized as important by Rogers. Also, the scale depends entirely upon the subjective impression of the client doing the rating rather than upon objective measurement of counselor behaviors.

Andersons' Interview Rating Scale is an attempt to provide an operational definition of ideal rapport, defined in terms of effective communication in counseling interviews. It consists of 50 items that are based upon a consensus among a select group of counselors and clients concerning the meaning of rapport. Eighteen of the items refer to client behaviors and attitudes; e.g., "the client distrusts the counselor," while the remainder refer to counselor behavior; e.g., "the counselor is very patient." The items are described in terms of specific behaviors and attitudes, each of which are rated on a five point scale ranging from "never existing" to "always existing." The 50 items are evenly divided between items representing high and low rapport. The Pearson correlation between the ratings of the counselors and those of the clients, based upon median scores, was 0.98.

Anderson and Anderson have demonstrated that a core of agreement among clients and counselors can be reached concerning the nature of ideal rapport, using the Anderson Interview Rating Scale. The effectiveness of the scale as a measure of rapport in counseling has also been supported by other independent investigators. However, Andersons' scale limits itself to an effectiveness of communication concept and does not draw broadly upon other counselor behaviors such as aspects

based upon visual perception and confrontation. Like Rogers' scale, it relies more upon subjective impressions than upon objective observations.

As part of a National Institute of Mental Health project, Rioch, et. al. (1,13), developed a scale for rating counselors in interviews which used outside professionals rather than clients as judges. Rioch's scale was much simpler than either the Rogers or Anderson scales. It consisted of two small sections: one for rating the counselor on his general counseling competency during the interview and one for rating the counselor on autocriticism. The interview section consisted of nine global ratings; e.g., "respect for the patient," and the section on autocriticism consisted of five global ratings; e.g., "shows awareness of major weakness or weaknesses," both of which were marked on a five point scale from poor through excellent. The Rioch Scale was applied to eight 40 year old mothers who had one year of training in the project's two year Mental Health Program of individual and group procedures designed to train them to do psychotherapy under limited conditions. Four professional judges rated these counselors on each global dimension using the five point scale. As a group, the eight counselors received a mean score above 3.0 on each dimension. In addition, the clients of these counselors were clinically evaluated to be functioning better. These results were considered positive.

Even though these tentative results may suggest some utility for the Rioch scale, its use seems limited for three reasons. First, no reliability studies were reported for the scale.

Second, the scale is too general and too global, and thus does not permit rating of specific counselor verbal behaviors. And third, like the other scales reviewed, it relies upon subjective impressions.

During the 1963-64 academic year the Interpersonal Process Recall (IPR) Project members developed the Counseling Process Effectiveness Scale to measure counselor trainee effectiveness. IPR is a new technique in stimulated recall methodology designed to secure the maximum effect from the replay of video tapes in counseling through emphasis on the visual element during recall. This technique was developed during the course of investigation into the dynamics of empathy at Michigan State University's College of Education by Kagan, Krathwohl, and Miller (1963). The Counseling Process Effectiveness Scale was developed by IPR project members after an unsuccessful search for a suitable scale. It consisted of 33 behavioral and feeling tone items, a global rating, and a question asking the judges to write adjectives or phrases which described the trainee.

The 33 behavior items, e.g., "the counselor paces the interview inappropriately," were each rated on a four point scale with one end of the four point continuum corresponding to ineffective and the other end corresponding to effective. A few of the items were taken and modified from Andersons' Interview Rating Scale. The global rating was adapted from Rioch's scale and was scored on a representation of the normal curve with baseline divided into eight equal segments. The judges were asked to rate each counselor in comparison with counselors of comparable experience

in order to establish an absolute rather than a relative judgment so the general quality of the counselor as a product could be evaluated. On the third part of the IPR project's scale the judges were asked to indicate the main factors, in order of importance, which led to their particular rating of each counselor.

As part of the development and progress of the IPR project, Ward (1965) used the IPR technique to compare the effects of three different methods of education on counselor trainee effectiveness as measured by the project's scale. Fifty-four Michigan State University Counseling and Guidance trainees were divided equally into three counselor education treatment groups: a video recall, an audio recall, and a control group (standard counselor training). Each of the groups had three thirty minute counseling sessions with high school students. The video recall group had each of its counseling sessions video taped and replayed for recall purposes with an interrogator. The audio recall group had each of its counseling sessions audio taped and replayed for recall purposes with an interrogator. The control group had its counseling sessions audio taped but received no stimulated recall sessions with an interrogator. At the end of this training each of the 54 counselors held a counseling session with the same coached client (a college student actress). This session was audio-video taped and used as the criterion for evaluating treatment effects.

Ward employed three professional counselor educators as judges to rate each of the 54 counselor's criterion tapes for counselor trainee effectiveness using the Counseling Process

Effectiveness Scale. For each criterion tape they rated two separate four minute segments. One four minute segment was rated early in the interview session (the first to the fifth minutes) and the other was rated later in the interview session (the ninth to the thirteenth minutes). Using this scale Ward found no significant differences between the three groups in counseling ability as rated by the three professional counselor educator judges. Ward found that the three professional judges did not rate the 54 counselors participating in the study in the same way; their ratings often differed markedly from each other, and these differences were found to be statistically significant.

Based upon a two-way analysis of variance, the range of interjudge agreement between any two judges ratings for the first four minute segments was .31 to .46 with a mean of .36. The range of interjudge agreement between any two judges ratings for the second four minute segments was .52 to .61 with a mean of .56. These low objectivity findings suggest that the Counseling Process Effectiveness Scale may have questionable utility.

Kelz (1966) constructed the Counselor Effectivenes Rating Instrument, a measure for assessing counselor trainee effectiveness. This rating instrument consists of eight major categories to appraise counselor effectiveness in interview situations. The eight categories of the scale are: (1) Appearance, (2) Expression, (3) Relationship, (4) Communication, (5) Knowledge, (6) Perception, (7) Interpretation, and (8) Termination. The subdivisions of these categories represent personal (relationship) and technical (professional) elements of counselor effectiveness which are

rated on a five point scale ranging from unsatisfactory to outstanding. According to Kelz, this scale was developed by collecting effective counselor characteristics from several references and from discussions with experts in the field of counselor education. Kelz selected those characteristics for rating which he considered to be relatively independent of the nature of the problem and the method of counseling.

The counselor subjects used by Kelz to test his scale were thirty members of a National Defense Education Act Guidance and Counseling Institute conducted at the Pennsylvania State University during the academic year 1960-61. Standardized interview situations were developed using student actors who, it was felt, could meet the standards considered necessary for the portrayal of high school counselees. Two interview situations, each emphasizing problems commonly faced by secondary school counselors, were developed. One situation focused on personal factors facilitating counseling relationships (interview S) and the other emphasized technical proficiencies such as the integration and application of data (interview P).

During the study a panel of six trained professional judges from the Pennsylvania State University observed, via closed-circuit television, two 15 minute interviews (interviews S and P) for each of the thirty counselor trainees. Ratings of counselor effectiveness using Kelz's scale were made independently by the judges during and directly following their observations of the interviews. The correlations of interjudge agreement on interview S ranged from .23 to .77 with a mean of

.53. This compared to a range of .21 to .65 on interview P, with a mean of .45. The average correlation between ratings given by pairs of independent raters was .50. The average agreement of individual judges with the pooled judge ratings was .78 for interview S and .73 for interview P. The judgments of counselor effectiveness were correlated with the rankings of the counselor trainees made by counseling practicum supervisors to establish the validity for the ratings. The correlations obtained were .42 and .28 for interviews S and P respectively.

Although Kelz's Counselor Effectiveness Rating Instrument appears to be useful for assessing the counseling proficiency of counselor trainees, it lacks favorable test validation and is characterized by low objectivity, and thus, low reliability.

Switching from the global and four to five point continuum counselor rating scales, Rank and DeRoo (1965) developed a forced choice, check-off Counselor Response System (CRS) Scale for the measurement of each counselor statement along six dichotomous dimensions as follows: (1) Affective--Cognitive Content, which indicates whether or not reference to affect is present in a counselor response; (2) Affective--Cognitive Change, which deals with gross changes in feeling level between a counselor response and a preceding client statement; (3) Content--Follow-Shift, which deals with whether or not there are changes in the general topic of discussion between the client's preceding statement and the counselor's response; (4) Present vs. Past or Future, which indicates whether the counselor refers to or focuses upon something in the past, the present, or the future; (5) Restrictive-

Expansive, which deals with the extent to which the counselor limits or permits freedom of expression by the client; and (6) Client--Other Referent, which indicates whether the client or someone else is the primary referent of the counselor's response. The six dimensions are theoretically derived, but they have not been based exclusively upon any one counseling theory.

The CRS Scale was designed to describe, but not evaluate, the verbal statements of counselors during interviews, and is part of a larger system currently being developed for examining both counselor and client responses and their inter-relationships. The rating evaluation of each counselor's verbal response involves making six dichotomous judgments, one for each dimension. For adequate evaluation, any given judge is designated to rate only two of the six dimensions. Thus, a minimum of three judges are used to rate a counselor's verbal statements.

DeRoo (1965) applied the CRS Scale to audio tapes of eight doctoral candidate counselors at the end of their practicum and to a group of 30 advanced NDEA Guidance and Counseling Institute candidates at Michigan State University at the end of the 1964-65 academic year. Using trained judges, DeRoo obtained a range of reliability ratings on the scale's six dimensions from .75 to .99 with marginal reliability resulting on the Restrictive--Expansive dimension. Thus, in initial application to audio tapes, Rank and DeRoo have demonstrated very favorable interjudge objectivity and reliability across the scale's six dimensions.

Of the counselor rating scales developed to date, Rank and

DeBoe's CRS Scale appears to be the most objective; it is the most specific in the counselor behaviors it measures; it is one of the least tied to any one counseling theory; and it offers the most comparative simplicity and ease of use by raters. Because the CRS Scale is a descriptive rather than an evaluative scale, that is, because its dimensions do not necessarily suggest positive or negative characteristics for the counseling verbal behaviors it measures, it may be limited at this stage of development as an instrument for evaluating counselor education programs and counselor trainee progress, and may have greater usefulness in research as an instrument to identify counselor and client behavior.

Westman (1965) has demonstrated that the condition of stimulus presentation of material to a rater has a significant effect upon his ratings. He presented audio, visual, and audio-visual portions of a number of Hard's criterion tape segments to three separate groups of four judges. Each segment was rated globally by the judges of each group on each of fourteen counselor variables; e.g., "Empathic," along a four point continuum ranging from "imperceptible" (scored 1), to "extremely" (scored 4). The average inter-rater agreement value for the audio-visual condition was .70, the average inter-rater agreement for the visual consition was .60, and the average value for the audio condition was .48. Westman concluded that the visual mode of presentation leads to higher inter-rater agreement than the audio mode because it has higher stimulus value.

Summary

A review of the literature indicates that the scale de-

velopmental studies of the past have failed to produce an objective, and thus reliable, instrument for evaluating counselor behaviors. Although a majority of these scales appear to have good face validity, they lack objectivity. Agreement among judges using these scales is poor.

Although attempts to objectively measure counseling interview data date back to 1929, counselor rating scales did not become validated until the development of the Rogers' Relationship Inventory Scale and the Anderson Interview Rating Scale of 1962. Both of these scales had some research validation, but Rogers' scale was too closely tied to his own counseling viewpoint, and the Anderson scale did not draw broadly upon various types of counselor behaviors, nor was it designed to make use of visual cues available now through video tapes. Both scales presented rather cumbersome rating assignments based upon four or five point scale continuums that were rated according to the subjective impressions of clients. These two scales, like the Rioch Scale, the IPR Counseling Process Effectiveness Scale, and Kolz's Counselor Effectiveness Rating Instrument, which followed in development, are, for the most part, global type rating scales.

The Rioch scale, developed after the Rogers' and Anderson scales, uses professional judges rather than clients as raters. This is also true of the Counseling Process Effectiveness Scale and the Counselor Effectiveness Rating Instrument which followed the Rioch scale in development. Following the format of the Rogers' and Anderson scales, these latter two scales present

rather cumbersome rating assignments based upon four or five point scale continuums. In addition, they are not very objective and reliable scales, and they suffer from very weak validation studies.

Of all the counselor rating scales reviewed from the counselor measurement literature, Rank and DeRoo's Counselor Response System Scale appears to be the most objective and usuable scale for measuring counselor behaviors. Unlike the other scales examined, it is least biased toward any one counseling theory; it is not value laden; it deals with specific and observable behaviors rather than global and inferred constructs; and it presents a rating assignment characterized by simplicity and ease of execution. However, the CRS Scale is designed to describe, rather than evaluate, counselor verbal behaviors. It measures certain counselor behaviors and attempts to avoid placing a value as to their desirability in counseling.

With the possible exception of the IPR Counseling Process Effectiveness Scale, and Kelz's Counselor Effectiveness Rating Instrument, all the scales reviewed were developed primarily to measure audio taped counseling behaviors and are unable to take advantage of visual cues available on video tapes. Visual cues appear to enhance interjudge agreement and thus add to the objectivity of a counselor scale.

Counseling Elements Common to Various Counseling Practices

This part of the review of the research was done only after extensive lists of video taped verbal counseling behaviors were

collected and tested in an effort to determine the common counseling elements or dimensions characterizing these behaviors. This review attempts to specify those elements of counseling which are common or relevant to all counseling practice, regardless of counseling orientation. Summarizing the research data on counseling, the "Review of Educational Research" (April, 1963, pp. 184-185) concludes that the trends in the data of counseling research indicate that experienced counselors probably produce better counseling results than novice counselors and that most experienced counselors choose their methods to fit the counseling task rather than a stereotyped counseling orientation. Fiedler's findings (1950, 28) suggest that the better trained therapists of different schools agree more with each other than they agree with less trained therapists within their own school. Fiedler's findings (1950, 29) also suggest that expert therapists of differing orientations formed similar relationships with patients. In discussing counseling process Cartwright and Lerner (1963) point out that frequently studies which compare the respective percentages of successfully treated cases suggest that some common elements exist among the various approaches to counseling, regardless of theoretical orientation.

Truax and Carkhuff (1964) mention that an examination of the many formulations designed to describe successful and effective counseling reveals such common elements as the following:

1. The ability of the counselor to accurately understand the client.

2. The ability of the counselor to communicate this understanding to the client.

3. The ability of the counselor to be able to accept the client while being integrated, mature and genuine within the counseling or therapeutic relationship.

They note that these elements of the psychotherapeutic relationship are aspects of the therapist's behavior and are common elements in a wide variety of psychoanalytic, client-centered, and eclectic approaches to psychotherapy and counseling. These recurring themes are emphasized by psychoanalytic theorists such as Alexander (1948), Ferenczi (1930), Schafer (1959) and Halpern and Lesser (1960); client-centered theorists such as Dymond (1949), Jourard (1959), Rogers (1951) and Snyder (1961); and eclectic theorists such as Rausch and Bordin (1957), Strunk (1957), and Strupp (1960).

Truax and Carkhuff (1964) indicate that another aspect of the counseling process commonly pointed to by these approaches has been the client's exploration of his feelings, his values, his relationships, his fears, and other concerns. Most of these approaches, they add, consider the client's self-exploration to be one of the central happenings resulting from the process of psychotherapy. They conclude this self-exploration is most likely to take place when the counselor communicates acceptance and understanding to the client. Based upon their own research and that of others, Truax and Carkhuff (1964) feel that the findings of importance to a theoretical model for evaluating psychotherapy were that patients who explored their feelings and concerns more and more throughout

psychotherapy also demonstrated the greatest constructive personality change, while patients who underwent very little self-exploration throughout the course of psychotherapy showed little constructive personality change.

Another crucial variable for successful psychotherapy implied by the previous practitioners of analytic, client-centered, and eclectic theory has been that of concreteness. They all make references to specific problem areas and affect and regard abstract interactions as non-therapeutic and even defensive. Truax and Carkhuff (1964) note that, "...perhaps one of the most useful learnings is that the cases high in understanding, high in patient intrapersonal exploration, and high in case outcome seem heavily loaded with specificity or concreteness of interaction." They conclude that this seems to serve three important functions:

> First, by ensuring that the therapist's response does not become abstract and intellectual and, thus, more emotionally removed from the patient's feeling and experiences.
>
> Second, by forcing the therapist to be more accurate in his understanding of the client.
>
> Third, by influencing the client to attend with specificity to problem areas and emotional conflicts.

Summary

A review of significant counselor behavior literature strongly suggests that the counseling practice of experienced and trained counselors, regardless of theoretical orientation, is characterized by common elements. Specifically, these common elements are that

experienced and successful counselors (1) sensitively and accurately understand their clients and communicate this understanding to them; (2) communicate acceptance to their clients in a genuine and integrated manner; (3) permit and encourage self-exploration by their clients; and (4) deal specifically with the important concerns of their clients.

CHAPTER III

SCALE DEVELOPMENT

The six major steps followed in the development of the A-V Scale (Figure 1, page 43) are discussed here. These steps were:

1. Observation and compilation of audio-video taped counselor behaviors.

2. Categorization of the observed audio-video taped counselor behaviors.

3. Organization of categories into a scale format.

4. Application of the scale to the rating of audio-video counseling tapes.

5. Definition of scale dimensions.

6. Development of a scale scoring system.

The A-V Scale was developed for the purpose of comparing the response patterns of M.A. candidates with those of Ph.D. candidates.[1] Because of the difference in training, a reasonable expectation is that the counseling patterns of the two classes of counselors will differ. The scale developed by this research is an attempt to identify and objectively measure specific and relevant counseling variables which seem to discriminate between the counseling behavior of M.A. and Ph.D. candidates.

Observation and Compilation of Audio-Video Taped Counselor Behaviors

The first step in the development of the A-V Scale was to observe the audio-video taped counseling behaviors of both M.A.

[1]The development of this scale was based primarily upon clinical judgment.

A-V COUNSELOR BEHAVIOR RESPONSE RATING SCALE

Judge: _____ Subject: _____ Date: _____

	DIMENSIONS								Counselor Response Evaluation		
									Effective		Non-Effective
Responses	Affect	Cognitive	Under-standing	Non-under-standing	Specific	Non-Specific	Explor-atory	Non-Explor-atory	4	3 2	1
1											
2											
3											
4											
5											
6											
7											
8											
9											
10											
11											
12											
13											
14											
15											
16											
17											
18											
19											
20											
% of Re-sponses											
TOTAL									TOTAL		

Figure 1 -- A-V COUNSELOR BEHAVIOR SCALE

candidates and Ph.D. candidates. Audio-video counseling tapes of these two classes of counselors, available at the Michigan State University College of Education, were observed. The audio-video counseling tapes were of counselors carrying out personal-social problem interviews with clients at Michigan State University.

Two lists of verbal counseling behaviors were compiled from these observations.[1] One list was made for the behaviors of the Ph.D. candidates and one list was made for the behaviors of the M.A. candidates (Appendix A, p. 101). For each list, the behaviors were compiled under either an Affective heading or a Cognitive heading. Counselor statements which referred to the client's feelings or attitudes were listed under the Affective heading. Counselor statements which seemed to lack any reference to or recognition of the client's feelings were listed under the Cognitive heading. The two lists were comprised of clinically descriptive classifications of behaviors most frequently observed to be characteristic of each class of counselors.

Categorization of the Observed Audio-Video Taped Counselor Behaviors

The behavior descriptions compiled for the M.A. candidates were compared to those compiled for the Ph.D. candidates. One noticeable difference became evident from this comparison. The types of behaviors describing the M.A. candidates were those which were listed most frequently under the Cognitive heading. In contrast, the types of behaviors describing the Ph.D. candidates

[1] The compilation of these observations was made by the investigator.

were those which were most frequently listed under the Affective heading. It appeared that the two classes of counselors were apparently differentiated on the basis of affective and cognitive responses.

Because the Affective and Cognitive categories encompassed a broad range of counseling behaviors, additional categories were sought for finer categorization of the compiled counselor behaviors. The review of the literature (Chapter II, pp. 37-41) suggested that responses which were "understanding," " specific," and "exploratory" were common and important counseling behaviors characterizing the practice of many counselors of diversified orientations. These three types of behavior designations appeared useful and relevant as categories for classifying the compiled counselor behaviors.

Based primarily upon the clinical definitions of those counselor behavior designations from the literature, an effort was made to classify the compiled counselor behaviors as to whether they were understanding, exploratory, and/or specific. Briefly, understanding is defined as the counselor's ability to communicate to the client the fact that he knows what the client is feeling or talking about; exploratory is defined as the counselor's assistance and skill in helping the client to explore his feelings, values, and other concerns; and specific is defined as the counselor characteristic of being concrete and getting to the core of the client's problem areas.

A majority of the behaviors compiled for Ph.D. candidates seemed to appropriately fit under one or more of these categories

(Appendix E, pp. 105-107). A majority of the behaviors compiled for the M.A. candidates did not seem to appropriately fit under any of these three categories. However, by using "nonunderstanding, nonspecific," and "nonexploratory" behavior designations, the majority of the behaviors of the M.A. [illegible] seemed to be more appropriately classified (Appendix E, pp. 108-110).[1] Briefly, nonunderstanding is defined as the counselor's failure to recognize or communicate to the client that he knows what the client is feeling; nonexploratory is defined as the counselor's failure to help the client to explore his feelings and concerns; and nonspecific is defined as the counselor characteristic of being vague and keeping clear of the central concerns of the client.

Organization of Categories Into A Scale Format

The eight categories of Affective, Cognitive, Understanding, Exploratory, Specific, Nonunderstanding, Nonexploratory, and Nonspecific appeared to be satisfactory for incorporation into a scale form for three reasons: First, they seemed to discriminate the responses of M.A. candidates from those of Ph.D. candidates. Second, they seemed to encompass the great majority of counselor descriptive behaviors compiled. And third, they seemed to refer to behaviors which could be rather specifically defined and objectively identified.

The next problem was the selection of a scale format. A format was needed which could present these categories in the most simple

[1] It was expected beforehand that Ph.D. candidates would be characterized by affect, understanding, specific and exploratory responses while M.A. candidates would be characterized by cognitive, nonunderstanding, nonspecific, and nonexploratory responses.

but effective rating form possible. Rank and DeRoo's (1965) dichotomous dimension, forced choice CRS Scale format offered the most promise for meeting this need because it was simple and seemed capable of objective judge ratings.

It appeared that the eight categories could be satisfactorily made into four forced choice dimensions. A counselor response could be either affective or cognitive, understanding or nonunderstanding, specific or nonspecific, or exploratory or nonexploratory. The eight categories, therefore, were made into the Affect-Cognitive, Understanding-Nonunderstanding, Specific-Nonspecific, and Exploratory-Nonexploratory counselor behavior dimensions. These four dimensions were adapted to a forced choice scale format similar to that in Figure 1, page 43.

Application of the Scale to the Rating of Audio-Video Counseling Tapes

In a series of rating sessions, the four counselor behavior dimensions were used for the rating of audio-video tapes involving M.A. candidates, Ph.D. candidates, and, in some cases, counselors with the Ph.D. degree. The counselors on these tapes were engaged in personal-social problem interviews with clients. The judges used for rating the tapes were Ph.D. candidates in Counseling and Guidance. They were provided with category definitions and illustrations as presented in Appendix B, pp. 105-[illegible].

During these rating sessions some of the tapes were used for clarification, modification, and expansion of category definitions based upon judge feedback. The judges employed for

these rating sessions were in agreement that the categories of the four dimensions appeared to differentiate the Ph.D. candidates from the M.A. candidates on a rather consistent basis. Their frequency ratings for the affect, understanding, specific, and exploratory dimension categories were consistently higher for the Ph.D. candidates than they were for the M.A. candidates. These counselor categories, as measurement items for this type of scale, seemed most meaningful to these judges as counselors rating other counselors. They felt the dimension categories referred to important aspects that were readily discernible and consistent in the counseling behaviors of M.A. and Ph.D. candidates.

During the ratings it was sometimes observed that verbally a counselor could be rated affective, specific, and/or exploratory--presenting on paper a positive counseling impression--whereas in fact his nonverbal manner could be so rigid and mechanical as to have little, if any, positive effect on the client at all. It was judged that this should be taken into account by the ratings. Thus, the definition and use of a more global effective-noneffective dimension was discussed.

It was then decided to incorporate such an Effective-Noneffective dimension into the scale. The Effective-Noneffective dimension is defined in terms of counseling progress (see the next section). Effective rated responses are scored either 3 or 4 and the ineffective responses are scored either 1 or 2.

The scale with the new dimension was applied to an analysis group. Five judges were used as raters. Again, the results suggested that the scale had good discrimination power for

differentiating between the responses of M.A. and Ph.D. candidates. Also, the high interjudge agreement obtained suggested that the scale had favorable objectivity.

Definition of Scale Dimensions

The final form of the A-V Scale consisted of the Affect-Cognitive, Understanding-Nonunderstanding, Specific-Nonspecific, Exploratory-Nonexploratory, and Effective-Noneffective forced choice dichotomous dimensions as illustrated by the scale format in Figure 1, page 43. The definitions for these dimensions are based primarily on those given in the counseling literature reviewed (Chapter II, pp. 38-40), clinical judgment, and the feedback received from judges who had used the scale to rate counseling behaviors on audio-video tapes.

Affect-Cognitive Dimension

The Affect-Cognitive dimension indicates whether a counselor's response is in reference to some affective or feeling aspect of a client's verbalization (an affective response) or whether his response omits any such reference to affect or feeling in the client's verbal communication (a cognitive response). Basically, an affective response refers back to client emotions, feelings, fears, moods, desires, urges, impulses, fantasies, dreams; and any ideas, attitudes, beliefs, and convictions which are based upon such referents as these. A cognitive response, in contrast, is devoid of any basic concern, reference, or recognition of client emotions, feelings, moods, etc., which may be implied in the

client's communication. The counselor responds at a non-feeling level or asks a question or makes a statement concerning cognitive or content material. Such cognitive responses are more formal, informative, matter of fact, or on the order of a general cognitive discussion, and are altogether on a different plane from responses which make affective contact with the client.

Understanding-Nonunderstanding Dimension

The Understanding-Nonunderstanding dimension deals with how well the counselor's response communicates to the client the fact that he understands, or is seeking to understand, the client's basic concerns. Understanding responses indicate that the counselor understands the basic underlying affective or cognitive concerns of what the client is talking about, or is trying to obtain from the client sufficient information to enable him to develop an understanding of the client's underlying concerns. Nonunderstanding responses are those which indicate that the counselor lacks an understanding of the basic underlying affective or cognitive concerns of what the client is talking about, or are responses which indicate that the counselor is making no attempt to obtain appropriate information from the client which would enable him to develop an understanding of the client's underlying concerns.

Specific-Nonspecific Dimension

The Specific-Nonspecific dimension indicates whether the counselor is dealing directly with the client's basic concerns by confining his response to a concrete (specific) aspect quite

relevant to the presented problem (a core response), or whether the counselor passes over dealing directly with the client's basic concern and gives a more general, or nonspecific, response not to relevant to the presented problem (a peripheral response). A specific response, then, indicates that the counselor is handling directly and honestly the basic problem presented explicitly or implicitly by the client. The counselor's remarks are keyed to the central aspects of the problem which the client is feeling and communicating. A nonspecific response, in contrast, is a peripheral type response which indicates that the counselor is not dealing directly and honestly with the basic problem presented explicitly or implicitly by the client. Rather, the counselor's response stays away from the client and his problem, or at best handles the problem at a superficial, surface level.

Exploratory-Nonexploratory Dimension

The Exploratory-Nonexploratory dimension indicates whether the counselor's response permits, encourages, or elicits further exploration on the part of the client of any aspect of his problem areas or affective concerns. An exploratory response is one which first, indicates that the counselor has recognized some portion of the client's basic problem or concern and, second, has made some verbal reference of this back to the client in such a way that the client is free from any defined or "limited" structure in his response, and is thus free to respond to any degree and depth to the counselor's reference that he may choose. The counselor's whole manner creates an atmosphere

for further exploration. A nonexploratory response, in contrast, indicates no recognition of the client's basic problem or concern and no attempt to search them out further. This is a response which also structures or limits the client's response.

Effective-Noneffective Dimension

The Effective-Noneffective dimension is based upon the judge's professional impression as to how appropriately, from a counseling vantage point, the counselor's response deals with the counselor's verbalization and thus contributes to the development and maintenance of counseling progress. Counseling progress is generally defined as concentration of discussion on basic problems and feelings of immediate concern to the client such that there develops for the client an increasing undistorted awareness of their influence and consequence on his behavior and relationships with others. In this regard, the judge decides globally--taking into consideration the counselor's total manner and behavior, including nonverbal--the effect the counselor's response will have, or could have, toward permitting or encouraging the client to move in a direction of counseling progress. If the judge evaluates the counselor's response as effective he can either check a 4 or a 3 under the heading "effective." A rating of 4 indicates that the judge's professional impression is that the response is about as "good" or the "best" possible in terms of counseling progress. A rating of 3 indicates that the judge's professional impression is that the response is effective towards some counseling progress but it was not the "best", that is,

a more effective response or communication could have been given. If the judge evaluates the counselor's response as noneffective he can either check a 2 or a 1 under the heading "noneffective." A rating of 2 indicates a response which contributes in no way to counseling progress (sort of neutral) while a rating of 1, in addition to this, indicates also a complete lack of understanding or concern regarding the client's problem situation, or is a response which is definitely detrimental to counseling progress.

Further elaboration and illustrations of these dimension category definitions are presented in Appendix C, pp. 111-125.

Scoring System For The A-V Counselor Behavior Scale

The unit for evaluation by this scale is the counselor-client unit. A client verbalization is emitted or elicited and the counselor's response to it is then rated on all five dimensions of the scale. One forced choice rating is made independently for each of the four specific counselor behavior dimensions and a fifth rating is made evaluating the effectiveness or noneffectiveness of the counselor's response. Generally, each client-counselor unit is judged independently of preceding units, although the congruence between the direction and level of counseling movement by the counselor with the direction and level of counseling concerns the client is attempting to present must continually be kept in awareness when making rating judgments.

The manner of scoring the effective-noneffective dimension has already been described (page52). The scoring for the other

four dimensions is rather simple. For each counselor response rated, a score of 1 is assigned to each cognitive, nonunderstanding, nonspecific, and nonexploratory category marked, and a score of 2 is assigned to each affect, understanding, specific, and exploratory category marked. When all the responses for a particular counselor have been rated, total dimension scores are obtained by summing the individual scores of the two categories comprising each dimension. When the test administrator is interested in the frequency ratings for particular categories (e.g., the frequency of affect, understanding, specific, and exploratory responses within a group, or between groups, of counselors) rather than total dimension scores, then only a score of 1 is assigned to the category each time it is marked.

Summary

The A-V Counselor Behavior Scale was developed in six steps. These steps were:

1. The observation and compilation of the audio-video taped counseling behaviors of Ph.D. and M.A. candidates in Counseling and Guidance at Michigan State University.
2. The search for appropriate categories by which to categorize these compiled counselor behaviors. Affect, Understanding, Specific, and Exploratory categories seemed to appropriately classify the compiled behaviors of the Ph.D. candidates. Cognitive, Nonunderstanding, Nonspecific, and Nonexploratory categories seemed to appropriately classify the compiled behaviors of the M.A.

candidates.

3. The organization of these categories into a dichotomous dimension, forced choice scale format. The dimensions of the scale were the Affect-Cognitive, Understanding-Nonunderstanding, Specific-Nonspecific, and Exploratory-Nonexploratory counselor behavior dimensions.
4. The application of the scale to the rating of audio-video counseling tapes. This was to test how well judges could differentiate Ph.D. candidates from M.A. candidates using the scale's four dimensions. The judges were able to differentiate these two classes of counselors on the scale's four dimensions. Based upon these rating sessions a fifth Effective-Noneffective dimension was added to the scale. Category definitions were also clarified, modified, and expanded.
5. The definition of scale dimensions.
6. The development of a scale scoring system.

CHAPTER IV

METHODOLOGY

A discussion of the methodologies of the three studies designed to test the hypotheses set forth in Chapter I is presented here. This chapter is divided into three sections; a discussion of the methodology of the Scale Discrimination Study, a discussion of the methodology of the Objectivity Study, and a discussion of the methodology of the Sensitivity Study.

Methodology of the Scale Discrimination Study

This study tested the hypothesis that the frequency of affective, understanding, specific, exploratory, and effective responses in initial personal-social problem counseling interviews will be significantly greater for Ph.D. candidates than for M.A. candidates beginning their practicum experiences. The methodology designed to test this hypothesis is presented here under the headings of Counselor Subjects, Judges, Judge Training, Procedure, and Design.

Counselor Subjects

The audio-video tapes of five M.A. candidates (B, E, F, H, and I) and five Ph.D. candidates (A, C, D, G, and J) were used in this study. The M.A. candidates were undergoing their first counseling practicum experiences. The Ph.D. candidates had completed their graduate practicum training. All ten counselors underwent an initial personal-social problem interview with a high school student of approximately 30 minutes in length. These

ten counseling interviews were recorded on audio-video tapes.

The five counseling sessions involving the Ph.D. candidates were specially audio-video taped for this study. Neither these counselors nor their clients knew the purpose for which these tapes were to be used. They were simply asked if they would volunteer for a regular personal-social problem counseling session which would be audio-video taped for general counselor training research purposes. The clients were all volunteers and were assigned to the counselors on a random basis.

The five counseling sessions involving the M.A. candidates were taped earlier as part of another research project. However, none of these tapes had ever before been used in the development and evaluation of the A-V Scale.

Judges

Three judges rated the audio-video tapes of these ten counseling sessions. Two of the judges were 1965 Michigan State University NDEA Guidance and Counseling Institute graduates. The third was a Master's graduate in Guidance and Counseling. These three judges had not seen any of the tapes prior to the rating session, nor did they know any of the ten counselors they were asked to rate.

Judge Training

All three judges had been trained in the use of the A-V Scale. Each judge was given a copy of the scale, the dimension definitions and illustrations (Appendix C, p. 111), and the counselor audio-video behavior categories and illustrations (Appendix B, p. 10)

for study. Following this, a training session was arranged in which the judges were allowed to use the A-V Scale for practice ratings of counselor responses on audio-video tape. Before beginning the practice ratings, the use and rating assignment of the A-V Scale, the dimension categories, and the technicalities of the audio-video operation and the audio-video playback were discussed. The value of honest and accurate ratings was emphasized, as well as the independent scoring of each dimension according to category definitions.

The pilot use of the A-V Scale for practice ratings of audio-video counseling tapes was executed under supervision, and the results of these practice ratings were discussed. This clarified the meaning of category definitions and was very helpful in developing a uniform standard among all the judges. Counselor responses illustrative of all the scale category behaviors were observed and rated independently by the judges. Their ratings were then compared, followed by a discussion over differences in ratings, dimension definitions, scale use, and important cues used in scoring. A tabulation of the answers to some of the most frequently occuring rating problems that required clarification by the judges, including some of the most commonly asked questions about the scoring of dimension categories, is presented in Appendix D, page 126.

With this kind of training it was expected that judges would tend to develop a common frame of reference and understanding of the scale dimensions and use of the scale. This training, provided the judges have a counseling background and have thoroughly studied

the dimension category definitions and illustrations, requires approximately one to two hours. It is expected that this judge training is an integral part of the use of the A-V Scale.

Procedure

The audio-video tapes of the ten counselors were presented to the three judges, in mixed order, by audio-video replay. Twenty responses from the latter part of each tape were rated. Each counselor response was identified by number and then rated independently by the judges. The replay of the tapes was controlled by a stop-start switch, which was used to stop the replay of any tape whenever a judge felt that the counselor responses were too rapid for adequate handling.

Design

For each counselor response rated cognitive, nonunderstanding, nonspecific, or nonexploratory, a score of 1 was assigned to the respective category. For each response rated affect, understanding, specific, or exploratory, a score of 2 was assigned to the respective category. For each of the Affect-Cognitive, Understanding-Nonunderstanding, Specific-Nonspecific, and Exploratory-Nonexploratory dimensions, the total dimension score was the sum of its two category scores. Each counselor response was rated either 1, 2, 3, or 4 on the Effective-Noneffective dimension. The total score for this dimension was the sum of the 1, 2, 3, and 4 ratings. These were the scores used for making statistical comparisons.

All rating scores were averaged across the three judges to give one rating per response for each of the five dimensions of

the scale. The total score on each dimension for the M.A. candidates was compared to the corresponding total dimension score for the Ph.D. candidates. Significant differences, based upon "t"-tests, were determined at the .05 level of confidence.

Indices of item interjudge reliability and internal consistency were also established by applying Hoyt's analysis of variance method to the unaveraged rating scores for each individual counselor response.

Methodology of the Objectivity Study

This study tested the hypothesis that the A-V Scale will have adequate objectivity for differentiating between M.A. and Ph.D. candidate responses on each of its five dimensions. The methodology designed to test this hypothesis is presented here under the headings of Counselor Subjects, Judges, Procedure, and Design.

Counselor Subjects

Fifty-three counselors were observed in this study. Forty-five of the counselors were M.A. candidates (1 through 45) undergoing their initial counseling practicum experiences. Eight were Ph.D. candidates (46 through 53) undergoing their initial doctoral practicum work. These 53 counselors had an initial personal-social type interview with the same coached client. These interviews were previously audio-video taped under controlled conditions as part of another research study (Ward, 1965).[1]

[1]This study is reviewed in Chapter II, pages 30-31.

Judges

Three judges rated the video tapes of these 53 counselors. One judge was near completion of his Ph.D. in Counseling and Guidance; one was a 1965 Michigan State University NDEA Guidance and Counseling Institute graduate; and one was a Master's graduate in Guidance and Counseling. These judges had not seen any of the 53 counselor tapes prior to the rating session, nor did they know to which of Ward's three experimental conditions these counselors had been assigned. The three judges were thoroughly familiar with the dimensions of the A-V Scale. They had received training in the use of the Scale with audio-video tapes as described under judge training in the previous section, pages 57-59.

Procedure

The second half four minute selected segments of Ward's 53 video counseling session tapes were presented by video-replay to the three judges in the same random order which they had first been presented to the three judges for rating in Ward's study. All counselor verbal responses of each four minute segment were rated independently by the judges using the Scale. The video replay of the 53 counseling tapes was controlled by a stop-start switch, which was used to stop the replay of any tape whenever a judge felt that the audio-video counselor responses were too rapid for adequate handling.

Design

For each response rated cognitive, nonunderstanding, nonspecific,

or nonexploratory, a score of 1 was assigned to the respective category. For each response rated affect, understanding, specific, or exploratory, a score of 2 was assigned to the respective category. Each counselor response was rated either 1, 2, 3, or 4 on the Effective-Noneffective dimension.

Hoyt's analysis of variance method was applied to the rating scores for computation of interjudge reliability (objectivity) and internal consistency (maximum likelihood of reliability).[1] Two different analyses were made. One was an item analysis of the individual counselor responses as rated by the three judges on each of the A-V Scale's five dimensions. The other was a global rater analysis based upon the total score obtained by each counselor on each of the scale's five dimensions. The item objectivity was computed from the actual scores for each counselor response, while the global objectivity was computed from proportionate scores based upon the total cognitive, nonunderstanding, nonspecific, and nonexploratory category scores for each counselor.[2] The scores for each of these categories were totaled, then divided by the number of counselor responses rated, and then multiplied by ten to give the proportionate score for each category. Proportionate scores for the Effective-Noneffective dimensions were computed in a similar manner.

[1]Formulae for Hoyt's method are presented in Appendix F, page 139.

[2]Because the counselors in Ward's study differed in their rate of verbal responding to the client, the number of responses available on each counselor's four minute tape segment was not necessarily equal. There was a small range. Therefore, to equate for this discrepancy so that the total scores for each counselor could be compared statistically, proportionate scores were used.

The total proportionate scores for the Effective-Noneffective dimension and the four categories mentioned were also averaged across the three judges. These scores were used to compare the M.A. candidates with the Ph.D. candidates. Significant differences, based upon "t"-tests, were determined at the .05 level of confidence.

Methodology of the Sensitivity Study

This study tested the hypothesis that counselor trainees who receive professional supervision which focuses on examining affective, understanding, specific, and exploratory responses will receive higher frequency ratings for these categories than will counselor trainees who do not receive this professional supervision. The methodology for testing this hypothesis is presented here under the headings of Counselor Subjects, Judges, Procedure, and Design.

Counselor Subjects

The counselor subjects were ten of the thirty National Defense Education Act (NDEA) Counseling and Guidance Institute enrollees attending Michigan State University for the 1964-65 year who had volunteered for this study. The subjects were randomly assigned to two groups; five to a professionally Supervised Group (A, B, C, D, and E) and five to a Nonsupervised Group (F, G, H, I, and J). These NDEA Institute enrollees were assumed to have had similar intelligence, education, and high school counseling experience, as well as identical NDEA Institute instruction and group therapy

class experience during the 1964-65 year of Institute training at the University.

Judges

Three judges were employed to rate the video tapes of these ten counselor enrollees. One judge had a Master's degree in Counseling and Guidance and the other two were near completion of the Ph.D. degree in Counseling and Guidance. The three judges had no knowledge as to which of the two groups--Supervised or Nonsupervised--the counselors they were rating belonged. The three judges had received training in the use of the A-V Scale with audio-video tapes as described under judge training in the first section of this chapter, pages 57-59.

Procedure

Each counselor member in both groups underwent six separate, thirty-minute audio-video taped counseling interviews with volunteer high school and college student clients assigned to them on a random basis from an available pool. Members of the Non-supervised Group worked in two-man counselor-client interrogator teams. During any given counseling session, one team member would serve as the counselor with the client and the other team member would later serve as the interrogator with the client. During the next counseling session the two team members would exchange roles. Over the six sessions each team member served three times as the counselor and three times as the client-interrogator. The fifth member of the Nonsupervised Group alternated with an Institute member not participating in the study. Each

Nonsupervised Group counseling interview was followed by a one hour client-interrogation session.

Members of the Supervised Group worked independently of one another and had each of their counseling interviews followed by a one hour counselor interrogation session. One staff member, with a clinical background, was the counselor-interrogator (supervisor) for all five trainees following each of their six video taped counseling interviews.

Throughout each Nonsupervised Group counseling interview, the client interrogator team member observed and listened from behind a one way mirror. Following this, the client was interrogated during the video-replay of the interview while the counselor team member observed and listened to the client-interrogation from behind the one way mirror. Throughout the audio-video recall, the client interrogator team member, by his questions and comments, conducted the recall session in a manner he felt was most appropriate for an interrogator. During this client-recall session, the audio-video replay was controlled by a stop-start switch, and either the client or the client-interrogator team member could stop the replay at any point he wished in order to make a comment or ask a question.

During the counseling interviews between a client and a Supervised Group counselor member, the supervisor observed and listened to this session from behind the one way mirror. Following the interview, the counselor, rather than the client, was interrogated by the supervisor during the audio-video replay of the counseling session. Either the counselor or supervisor

could stop the audio-video replay at any point he wished to ask a question or make a comment.

The general procedure followed by the supervisor was to pursue gradually and subtley the following with the counselor:

1. The extent to which the counselor felt he was dealing with the affective concerns of the client.
2. The extent to which the counselor felt he was dealing with the real and specific concerns of the client rather than superficial considerations.
3. The extent to which the counselor felt he was communicating to the client the fact that he understood the real concerns of the client.
4. The extent to which the counselor permitted and encouraged the client to explore his real concerns.
5. The counselor's opinion of the effect that a lack of the behaviors in 1 through 4 on his part would have upon the client and their counseling progress.
6. The counselor's opinion of the effect that an emphasis of the behaviors in 1 through 4 by him would have upon the client and their counseling progress.
7. Encouragement of the counselor to consider the possible counseling use of these types of behaviors in his subsequent interviews with clients.
8. If these behaviors were emphasized in the subsequent interviews by the counselor, an examination of the counselor's evaluation of their effectiveness for the client was made. The counselor's effect on clients when he lacked these

behaviors was compared with his effect on clients when he emphasized these behaviors. The counselor could observe on audio-video tape the effect that these behaviors, or the lack of these behaviors, had on the client. Based upon these observations, he was then left to judge for himself the value these behaviors had for counseling.

Upon completion of the six counseling sessions by the members of both the Nonsupervised and Supervised Groups, the middle twenty responses of the first and sixth sessions for each counselor member were replayed and rated using the A-V Scale. There was, then, a before and after comparison for each counselor of both the Nonsupervised and the Supervised Groups. The first session rated for both groups was an initial interview with a client. The sixth session rated for both groups of counselors was the second consecutive interview with their last client.

The ratings for the initial and terminal interviews were made over two different rating sessions; one was held for the initial interviews and one was held for the terminal interviews.

Design

For each counselor response rated affect, understanding, specific, and exploratory, a score of 1 was assigned to the respective category. On the Effective-Noneffective dimension, each counselor response was rated either 1, 2, 3, or 4. These were the scores used for making statistical comparisons. Thus, the higher the score on each dimension, the greater the indication

that a counselor's responses had been rated affective, understanding, specific, exploratory, and effective.

All ratings were averaged across the three judges to give one rating per response for the Effective-Noneffective dimension and each of the four categories mentioned. The initial and terminal ratings for both the Supervised and Nonsupervised Groups were compared for significant differences on all five dimensions of the A-V Scale. The dependent "t"-test was the statistic used for making this comparison at the .05 level of confidence.

Summary

The methodologies of the Scale Discrimination Study, the Objectivity Study, and the Sensitivity Study were discussed in this chapter. These three studies tested the three hypotheses set forth in Chapter I.

The methodology of the Scale Discrimination Study was designed to test the hypothesis that the frequency of affective, understanding, specific, exploratory and effective responses in initial personal-social problem counseling interviews will be significantly greater for Ph.D. candidates than for M.A. candidates beginning their practicum experiences. Three trained counselor judges used the A-V Scale to rate 20 responses from each of ten 30 minute audio-video taped counseling interview sessions. These ten taped counseling sessions involved five M.A. degree candidates and five Ph.D. candidates engaged in personal-social problem interviews with clients. The rating scores of both groups were compared on the five dimensions of the A-V Scale by means of "t"-tests. Hoyt's

analysis of variance method was also applied to the rating scores to establish an index of item objectivity.

The methodology of the Objectivity Study was designed to test the hypothesis that the A-V Scale will have adequate objectivity for differentiating between M.A. and Ph.D. candidate responses on each of its five dimensions. Three trained counselor judges used the A-V Scale to rate four minute segments from each of 53 audio-video counseling tapes. Forty-five of these tapes involved M.A. candidates and eight involved Ph.D. candidates. All 53 counselors were engaged in an initial interview with the same coached client. Hoyt's analysis of variance was applied to the rating scores to establish two sets of objectivity and internal consistency coefficients. One set was an index of item objectivity and internal consistency and the other a global rater objectivity and internal consistency. The rating scores of the M.A. candidates were compared with those of the Ph.D. candidates on the five dimensions of the A-V Scale by means of "t"-tests.

The methodology of the Sensitivity Study was designed to test the hypothesis that counselor trainees who receive professional supervision which focuses on examining affective, understanding, specific, and exploratory responses will receive higher frequency ratings for these categories on the A-V Scale than will counselor trainees who do not receive this professional training. Ten NDEA Guidance and Counseling Institute trainees participated in the study. Five of these counselor trainees belonged to a Supervised Group which received professional interrogation by means of audio-video tape replay of their counseling

interviews. These interrogation sessions evaluated the effect that affective, understanding, specific, and exploratory responses would have on the counselor trainee's clients. This group received six of those interrogation sessions. The other five counselor trainees belonged to a Nonsupervised Group who also were exposed to the audio-video tape replay of their own counseling sessions. However, their interrogation sessions were carried out with the client, rather than themselves, by a colleague team member. These interrogation sessions pursued no particular emphasis during the audio-video replay. The pre and post tapes of the five trainees belonging to each group were rated by three counselor judges using the A-V Scale. For each group, the rating scores of the pre and post tapes were statistically compared using dependent "t"-tests.

CHAPTER V

ANALYSIS OF RESULTS

The results of the methodological procedures described in Chapter IV are reported here. This chapter is divided into three sections; a discussion of the results of the Scale Discrimination Study, a discussion of the results of the Objectivity Study, and a discussion of the results of the Sensitivity Study.

Results of the Scale Discrimination Study

The hypothesis tested by the Scale Discrimination Study was that the frequency of affect, understanding, specific, exploratory and effective responses will be greater for Ph.D. candidates than for M.A. candidates. Also, objectivity data was collected which would reflect on the second hypothesis discussed in the next section.

Analysis of Scale Discrimination Data

Table 5.1, page 72, summarizes the statistical "t" analysis of Appendix E, p. 132. The analysis presented in this appendix compares, on each of the five dimensions of the A-V Scale, the scale scores of the five Ph.D. candidates with those of the M.A. candidates. It can be seen from inspection of this Table that there is a statistically significant difference on all five dimensions between the scores of the Ph.D. candidates and those of the M.A. candidates. The responses of the Ph.D. candidates

were more frequently rated affect, understanding, specific, exploratory, and effective than were the responses of the M.A. candidates.

Table 5.1 -- Comparison of the Mean ($\overline{X}$) Scores Between the Ph.D. and M.A. Candidates on Each of the Five Dimensions of the A-V Scale (3 judges, 5 Ph.D. counselors, 5 M.A. counselors)

	Aff.-Cog.	Underst.-Nonunderst.	Spec.-Nonspec.	Expl.-Nonexpl.	Eff.-Noneff.
Ph.D. $\overline{X}$	35.2	38.2	35.4	36.4	58.8
M.A. $\overline{X}$	21.2	21.0	20.2	20.2	37.1
t - ratio	6.9**	13.4**	7.4**	7.5**	6.9**

necessary: t .05 = 2.31
necessary: t .01 = 3.36 **significant at .01 level

Aff.-Cog. = Affect-Cognitive

Underst.-Nonunderst. = Understanding-Nonunderstanding

Spec.-Nonspec. = Specific-Nonspecific

Expl.-Nonexpl. = Exploratory-Nonexploratory

Eff.-Noneff. = Effective-Noneffective

These findings support the hypothesis tested. They suggest that the dimensions of the A-V Scale are able to discriminate between the responses of these two classes of counselors. These results, then, indicate that the scale has validity for this purpose.

Analysis of Interjudge Reliability Data

Hoyt's analysis of variance method for estimating reliability was applied to the scores of the five Ph.D. candidates and the five

M.A. candidates rated in this study. The data and computations upon which this statistical analysis is based are presented in Appendix F, p. 138. The item interjudge reliability (objectivity) and item internal consistency coefficients computed by this analysis are reported in Table 5.2 for each of the five dimensions of the A-V Scale.

Table 5.2 -- Item Interjudge Reliability and Internal Consistency Coefficients Computed from the Scores of the Scale Discrimination Study (3 judges, 10 counselors)

	Aff.-Cog.	Underst.-Nonunderst.	Spec.-Nonspec.	Expl.-Nonexpl.	Eff.-Noneff.
Interjudge Reliability	.792	.816	.817	.756	.763
Internal Consistency	.919	.930	.929	.903	.906

These results suggest that perhaps the dimensions of the A-V Scale have quite favorable objectivity and reliability characteristics.

Results of the Objectivity Study

The hypothesis tested by the Objectivity Study was that the A-V Scale will have adequate objectivity for differentiating between M.A. and Ph.D. candidate responses on each of its five dimensions. The Objectivity Study also served as a type of cross-validation for the Scale Discrimination Study.

Analysis of Interjudge Reliability Data

Hoyt's analysis of variance method for estimating reliability was applied to the scores for the forty-five M.A. and eight Ph.D. candidates rated in this study. Two separate sets of interjudge reliability and internal consistency coefficients were computed by this analysis. The first set of coefficients computed were the item interjudge reliability and internal consistency coefficients based upon a comparison of individual counselor response ratings. These coefficients are reported in Table 5.3 for each of the five dimensions of the A-V Scale. The statistical analysis upon which the computation of these coefficients is based is presented in Appendix G, page 165.

Table 5.3 -- Item Interjudge Reliability and Item Internal Consistency Coefficients Computed from the Scores of the Objectivity Study (3 judges, 53 counselors)

	Aff.-Cog.	Underst.-Nonunderst.	Spec.-Nonspec.	Expl.-Nonexpl.	Eff.-Noneff.
Interjudge Reliability	.890	.792	.763	.590	.814
Internal Consistency	.961	.919	.906	.812	.929

The second set of coefficients computed were the global rater interjudge reliability and internal consistency coefficients based upon a comparison of the total dimension scores for each counselor rated. These coefficients are reported in Table 5.4, page 75, for each of the five dimensions of the A-V Scale. The statistical

analysis upon which the computation of these coefficients is based is presented in Appendix H, page 206. The internal consistency coefficients reported refer to the maximum likelihood estimate of reliability.

Table 5.4 -- Global Rater Interjudge Reliability and Internal Consistency Coefficients Computed from the Scores of the Objectivity Study (3 judges, 53 counselors)

	Aff.- Cog.	Underst.- Nonunderst.	Spec.- Nonspec.	Expl.- Nonexpl.	Eff.- Noneff.
Interjudge Reliability	.953	.911	.885	.697	.934
Internal Consistency	.964	.968	.959	.873	.977

It can be seen upon inspection of Table 5.3 that the item interjudge reliability and the maximum estimate of item reliability (internal consistency) coefficients are rather high for a counselor rating scale. Although lower, these coefficients compare favorably with those obtained from the ratings of the Scale Discrimination Study reported in Table 5.2. The results reported in Table 5.4 indicate that these coefficients are higher when the total or global scores for each dimension are compared.

These findings support the hypothesis tested. They suggest that the dimensions of the A-V Scale have favorable objectivity for the rating of Ph.D. and M.A. candidate responses. This high objectivity is accompanied, as would be expected, by a high maximum estimate of reliability.

Analysis of Data Comparing Ph.D. with M.A. Candidates

Table 5.5 summarizes the statistical "t" analysis of Appendix I, page 217. The analysis presented in this appendix compares the proportionate scores of the eight Ph.D. candidates with those of the forty-five M.A. candidates on the Cognitive, Nonunderstanding, Nonspecific, and Nonexploratory categories of the A-V Scale. A comparison of the proportionate scores of the Effective-Noneffective dimension is also made between the two classes of counselors.

Table 5.5 -- Comparison of the Mean ($\overline{X}$) Scores Between the Ph.D. and M.A. Candidates on the Cognitive, Nonunderstanding, Nonspecific, and Nonexploratory Categories, and Effective-Noneffective Dimension, of the A-V Scale (3 judges, 8 Ph.D. counselors, 45 M.A. counselors)

	Cog.	Nonunderst.	Nonspec.	Nonexpl.	Eff.-Noneff.
Ph.D. $\overline{X}$	4.8	3.8	5.3	7.0	24.9
M.A. $\overline{X}$	8.3	7.9	8.6	9.2	15.2
t - ratio	3.5**	14.1**	11.8**	8.5**	4.9**

necessary: t .05 = 2.00
necessary: t .01 = 2.66 **significant at .01 level

Inspection of Table 5.5 discloses that there is a statistically significant difference at the .01 level between the scores of the Ph.D. candidates and the scores of the M.A. candidates on all five scale dimensions. The responses of the M.A. candidates were more frequently rated cognitive, nonunderstanding, nonspecific, nonexploratory, and less effective than were the responses of the

Ph.D. candidates. These results support the findings of the Scale Discrimination Study reported in Table 5.1, and again suggest that the five dimensions of the A-V Scale are able to discriminate between the responses of M.A. and Ph.D. candidates.

Results of the Sensitivity Study

The Sensitivity Study tested the hypothesis that counselor trainees who receive professional supervision which focuses on examining affective, understanding, specific, and exploratory responses will receive higher frequency ratings for these categories on the A-V Scale than will counselor trainees who do not receive this professional training.

Table 5.6 reports the initial means of the Nonsupervised Group with the initial means of the Supervised Group on the five dimensions of the A-V Scale before interrogation sessions.

Table 5.6 -- Mean ($\bar{X}$) Scores of the Initial Interview Rating for both the Nonsupervised and Supervised Groups on the Five Dimensions of the A-V Scale (3 judges, 10 counselors)

	Aff.-Cog.	Underst.-Nonunderst.	Spec.-Nonspec.	Expl.-Nonexpl.	Eff.-Noneff.
Nonsupervised $\bar{X}$	1.2	4.2	.3	3.4	27.3
Supervised $\bar{X}$	2.6	6.8	3.4	7.2	31.0

Table 5.7, page 78, summarizes the statistical comparisons of the initial mean scores of the Nonsupervised Group before non-supervised client interrogation with the terminal mean scores of

this group following the nonsupervised client interrogation sessions. The data and computations upon which this analysis is based are presented in Appendix J, page 228.

Table 5.7 -- Comparison of the Mean Scores of the Initial and Terminal Interview Ratings for the Nonsupervised (Client-Interrogation) Group on the Five Dimensions of the A-V Scale (3 judges, 5 counselors)

	Aff.-Cog.	Underst.-Nonunderst.	Spec.-Nonspec.	Expl.-Nonexpl.	Eff.-Noneff.
Initial $\bar{X}$	1.2	4.2	.8	3.4	27.2
Terminal $\bar{X}$	1.8	2.4	1.6	2.0	25.8
t - ratio	2.5		2.2		

necessary: t .05 = 2.78

The comparison of the mean scores of the initial and terminal ratings of Table 5.7 show that there were no significant changes on any of the five dimensions of the A-V Scale. Since the means for the Understanding-Nonunderstanding, Exploratory-Nonexploratory, and Effective-Noneffective dimensions decreased rather than increased from the initial to the terminal ratings, no t-tests were necessary. It was apparent that these three dimension scores had not changed in the necessary direction.

Table 5.8, page 79, summarizes the statistical comparison of the initial mean scores of the Supervised Group before supervised counselor interrogation with the terminal mean scores of this group following the supervised counselor interrogation sessions. The data and computations upon which this analysis is based are

presented in Appendix K, page 232.

Table 5.8 -- Comparison of Mean Scores of the Initial and Terminal Interview Ratings for the Supervised (Counselor-Interrogation) Group on the Five Dimensions of the A-V Scale (3 judges, 5 counselors)

	Aff.- Cog.	Underst.- Nonunderst.	Spec.- Nonspec.	Expl.- Nonexpl.	Eff.- Noneff.
Initial $\bar{X}$	2.6	6.8	3.4	7.2	31.0
Terminal $\bar{X}$	7.2	14.0	14.6	16.6	51.0
t - ratio	3.4*	3.3*	6.1**	4.6**	4.2*

necessary: t .05 = 2.78 *significant at .05 level
necessary: t .01 = 4.60 **significant at .01 level

The comparison of the mean scores of the initial and terminal ratings of Table 5.8 show that there were statistically significant changes on each of the five dimensions of the A-V Scale. A comparison of these findings with those reported in Table 5.7 suggest that the supervised counselor interrogation with counselor trainees, which focuses on eliciting affective, understanding, specific, and exploratory counselor responses, increases significantly the frequency of counselor trainee responses which are rated affective, understanding, specific, and exploratory. These findings support the hypothesis tested.

Summary

The results of the Scale Discrimination Study, the Objectivity Study, and the Sensitivity Study are reported in this chapter.

The findings of the Scale Discrimination Study suggest that the five dimensions of the A-V Scale significantly discriminated between the responses of Ph.D. candidates and those of M.A. candidates. The responses of the Ph.D. candidates were more frequently rated affect, understanding, specific, exploratory, and effective than were the responses of the M.A. candidates. The results also suggest favorable objectivity for the A-V Scale in making these discriminations. Item interjudge reliability coefficients of .72, .82, .82, .76, and .76, and item internal consistency coefficients of .92, .93, .93, .90, and .91, were obtained for the Affect-Cognitive, Understanding-Nonunderstanding, Specific-Nonspecific, Exploratory-Nonexploratory, and Effective-Noneffective dimensions respectively.

The findings of the Objectivity Study suggest that the A-V Scale has favorable objectivity, accompanied by favorable maximum reliability estimates, for the rating of Ph.D. and M.A. candidate responses. Item interjudge reliability (objectivity) coefficients of .89, .79, .76, .59, and .81, and global rater interjudge reliability coefficients of .95, .91, .89, .70, and .93 were computed for the Affect-Cognitive, Understanding-Nonunderstanding, Specific-Nonspecific, Exploratory-Nonexploratory, and Effective-Noneffective A-V Scale dimensions respectively. For these same scale dimensions, item internal consistency (maximum estimate of reliability) coefficients of .96, .92, .91, .81, and .93 respectively, and global rater internal consistency coefficients of .98, .97, .96, .87, and .98 respectively were computed. The findings of this study also support the findings

of the Scale Discrimination Study that the A-V Scale significantly discriminates between the responses of Ph.D. candidates and those of M.A. candidates.

The findings of the Sensitivity Study suggest that supervised counselor interrogation with counselor trainees, which focuses on examining affective, understanding, specific, and exploratory counselor responses, increases significantly the frequency of counselor trainee responses which are rated affective, understanding, specific, and exploratory.

CHAPTER VI

SUMMARY, CONCLUSIONS AND IMPLICATIONS

This chapter consists of three sections; a summary of this research, the conclusions of this research, and a discussion of the implications generated by this research.

Summary

This research examined three hypotheses associated with the development of an Audio-Visual Counselor Behavior Scale (A-V Scale). This A-V Scale was developed for the purpose of making available an objective instrument which would compare the response patterns of Master's degree candidates in Counseling and Guidance (M.A. candidates) with the response patterns of doctoral degree candidates in Counseling and Guidance (Ph.D. candidates). Its development was based upon the clinical observation of counseling behaviors on audio-video tapes followed by a categorization of these behaviors according to constructs clinically evaluated to be relevant and descriptive of M.A. and Ph.D. candidates. The A-V Scale consists of five forced choice, dichotomous dimensions. These dimensions are the Affect-Cognitive, Understanding-Nonunderstanding, Specific-Nonspecific, Exploratory-Nonexploratory, and Effective-Noneffective dimensions.

The three hypotheses examined by this research are:

First Hypothesis

The frequency of affective, understanding, specific, exploratory,

and effective responses in initial personal-social problem counseling interviews, as measured by the A-V Scale, will be significantly greater for Ph.D. candidates than it will be for M.A. candidates beginning their practicum experiences in Counseling and Guidance.

Second Hypothesis

The A-V Scale will have adequate objectivity for differentiating between Ph.D. and M.A. candidate responses on each of its five dimensions.

Third Hypothesis

Counselor trainees who receive professional supervision which focuses on examining affective, understanding, specific, and exploratory counselor responses will receive higher frequency ratings for these categories on the A-V Scale than will counselor trainees who do not receive this professional training.

The Scale Discrimination Study tested the first hypothesis. For this study, three trained counselor judges used the A-V Scale to independently rate 20 responses from each of ten 30 minute audio-video taped counseling sessions. Five of the counselors engaged in these interviews were M.A. candidates. The other five counselors engaged in these interviews were Ph.D. candidates. Each of the ten interviews was an initial personal-social problem interview with a client. The judges had not seen the tapes prior to the rating session, nor did they know the counselors they were to rate.

The rating scores obtained by the two classes of counselors on each of the five dimensions of the A-V Scale were statistically

compared using "t"-tests. There were significant differences in the response scores of the two classes of counselors on all five dimensions of the A-V Scale. The responses of the Ph.D. candidates were more frequently rated affect, understanding, specific, exploratory, and effective than were the responses of the M.A. candidates. Hoyt's analysis of variance method was applied to the rating scores. Item interjudge reliability coefficients of .79, .82, .82, .76, and .76 were computed for the Affect-Cognitive, Understanding-Nonunderstanding, Specific-Nonspecific, Exploratory-Nonexploratory, and Effective-Noneffective dimensions respectively. For the same dimensions, item internal consistency coefficients of .92, .93, .93, .90, and .91 respectively were obtained.

The **Objectivity** Study tested the second hypothesis. For this study, three trained counselor judges used the A-V Scale to independently rate four minute segments from each of 53 audio-video counseling tapes. Forty-five of these tapes involved M.A. candidates and eight involved Ph.D. candidates. All 53 counselors were engaged in an initial interview with the same coached client. The judges had not seen the tapes prior to the rating session, nor did they know the counselors they were to rate.

Hoyt's analysis of variance method was applied to the rating scores. Two types of interjudge reliability and internal consistency coefficients were obtained. One was an index of item objectivity and internal consistency, and the other was an index of a global rater objectivity and internal consistency.

Interjudge reliability refers to objectivity while internal consistency refers to the maximum likelihood estimate of reliability. Item interjudge reliability coefficients of .89, .79, .76, .59, and .81, and global rater interjudge reliability coefficients of .95, .91, .89, .70, and .93, were computed for the Affect-Cognitive, Understanding-Nonunderstanding, Specific-Nonspecific, Exploratory-Nonexploratory, and Effective-Noneffective A-V Scale dimensions respectively. For these same scale dimensions, item internal consistency coefficients of .96, .92, .91, .81, and .93 respectively, and global rater internal consistency coefficients of .98, .97, .96, .87, and .98 respectively were computed. The dimension scores of the forty-five M.A. candidates were statistically compared with those of the Ph.D. candidates by means of "t"-tests. There were significant differences in the response scores of the two classes of counselors on all five dimensions of the A-V Scale. These findings supported those obtained by the **Scale Discrimination** Study which tested the first hypothesis.

The Sensitivity Study tested the third hypothesis. For this study, ten NDEA Guidance and Counseling Institute trainee volunteers were randomly assigned to two groups of five; one group comprised a supervised counselor-interrogation group (Supervised Group), and the other group comprised a nonsupervised client-interrogation group (Nonsupervised Group). Each member of both groups underwent six 30 minute video taped counseling sessions with a college or high school student. Following each counseling session each member of the Supervised Group received a one hour audio-video recall counselor-interrogation which evaluated the

effect that affective, understanding, specific, and exploratory responses would have on the counselor trainee's clients during the counseling interview. Following each counseling session each member of the Nonsupervised Group either observed through a one view mirror while a colleague team member interrogated his client for one hour, or he interrogated his colleague team member's client for one hour while his team member observed the interrogation. There was no particular emphasis to this client interrogation.

The middle twenty responses of the first (initial) and sixth (terminal) counseling sessions for each of the ten counselors comprising the two groups were rated by three trained counselor judges. These judges had no knowledge as to which of the two groups each of the ten counselors belonged. These initial and terminal ratings within each of the Supervised and Nonsupervised Groups were statistically compared using dependent "t"-tests. The results of these tests indicate that for the Supervised Group there was a significant increase in the rating scores, from the initial to the terminal ratings, on the affect, understanding, specific, exploratory, and effective dimension categories of the Scale's five dimensions. For the Nonsupervised Group there was no significant score increase, from the initial to the terminal ratings, on these five Scale dimension categories.

Conclusions

In drawing conclusions based on the findings of this research, it is necessary to consider certain variables which might have

limited or influenced these findings. First, there were the major delimitations outlined for this research in Chapter I. This refers primarily to the fact that the behaviors studied in this research were laboratory behaviors and that the number of counselors observed in two of the three studies of this research were relatively small. Second, there was the fact that the counselors and judges used in this research were primarily associated with Michigan State University.[1] Third, there was the possible operation of special factors which could have influenced the outcome of the findings. This would include the operation of such variables as rater bias, the overtraining of judges, the convenience and desirability of selected tapes, and the nonverbal conveyance of rating attitudes resulting from judges rating in the same room.

Nonetheless, within these limitations, the findings of this research appear to warrant the following conclusions about the hypotheses presented in Chapter I:

1. The data of this research indicates that the frequency of affect, understanding, specific, exploratory, and effective responses in initial personal-social problem counseling interviews, as measured by the five dimensions of the A-V Scale, is significantly greater for Ph.D. candidates in Counseling and Guidance than it is for M.A. candidates beginning their practicum experiences in Counseling and Guidance. Therefore, when used for the purpose of discriminating between the responses of

[1]Their training, however, was diversified and acquired at various other universities.

these two classes of counselors, the A-V Scale appears to have adequate validity.

2. The data of this research suggests that the A-V Scale has high interjudge reliability for discriminating between the responses of M.A. and Ph.D. candidates on each of its five dimensions. Therefore, the A-V Scale seems to have adequate objectivity. Compared to the low objectivity findings reported for most other counselor evaluation scales, these findings appear most favorable.

3. The data of this research suggests that counselor trainees who receive professional supervision which focuses on examining affective, understanding, specific, and exploratory responses will receive significantly higher frequency ratings for these categories on the A-V Scale than will counselor trainees who do not receive this professionally supervised training. Therefore, the A-V Scale seems to have adequate sensitivity for measuring the subsequent change in these category behaviors which are associated with special counselor training.

Implications

This section is divided into two parts; a discussion of the implications for counselor education and the implications for future research.

Implications for Counselor Education

Affective, understanding, specific, and exploratory counselor behaviors seem to be associated with Ph.D. candidates in Counseling and Guidance but do not seem to be associated with M.A. candidates beginning their practicum experiences in Counseling and Guidance. This would suggest that these counselor behaviors are associated with, and are perhaps relevant to, the counselor training process. If such an indication proves valid, then special emphasis on these counselor behaviors during training can perhaps accelerate counselor education. The A-V Scale would then possibly become a potential evaluator of this type of accelerated counselor education.

The affect, understanding, specific, and exploratory dimension categories may be central to effective counseling or they may have nothing to do with effective counseling. This we do not know. However, if it can be satisfactorily demonstrated that these category behaviors are significantly related to effective counseling, then this would indicate that these types of behaviors are relevant to the counselor training process. If this be the case, the A-V Scale can serve as a cumulative progress report of a counselor trainee's change in counseling behaviors as he progresses through the various phases of his counseling development. Whether or not these category behaviors are significantly related to effective counseling will have to be substantiated by further research.

During the course of the rating sessions which were a part of the three studies comprising this research, a new value and

use of the A-V Scale was suggested. NDEA Institute and other Master's level counselors who served as judges said they understood more fully the importance of affective, understanding, specific, and exploratory responses on the part of the counselor because they had been required to focus on and evaluate clearly defined aspects of the counselor's verbal behavior. This may have given them a very useful and meaningful set, one which they had not acquired before, for evaluating counselor behaviors. As a result of their rating experiences with the A-V Scale, the importance of affective, understanding, specific, and exploratory responses for effective counseling became more apparent to them. As a result, some of these counselors reported more of an emphasis on these counseling behaviors in their own counseling practice. They also reported that through the use of the Scale in rating audio-video tapes they learned more about what is "good" counseling than they had ever learned in their Master level practicum experiences.

Could it be that the use of the A-V Scale itself for rating purposes by counselor trainees would be a valuable supplement to regular counselor training program procedures in developing basic counseling understanding and growth as well as serving as a means for providing a relevant model of counselor behaviors with which to identify? Is it possible, then, that by using the A-V Scale to rate experienced and inexperienced counselors, the Scale can contribute directly to the education of counselor trainees by providing exercises for focusing on essential counselor behavior as well as providing realistic opportunities

for observing the effects affective, understanding, specific, and exploratory responses, or the lack of them, has on clients in real counseling interview situations? Could it be that such experiences would accelerate counselor education and give a sound counseling perspective early in the training period? The answers to these questions can only be acquired through further research with the Scale.

Implications for Future Research

Further development and evaluation of the A-V Scale is needed before definitive conclusions can be made as to its characteristics and value. The implications for the types of research which seem most promising for the A-V Scale are:

1. Cross-validation of the A-V Scale on audio-video taped counseling interviews at other universities using other trained judges at each of the universities.
2. Equating two groups of counselors A and B. Train group A along the dimensions of the A-V Scale. Train group B along some other counseling dimensions. Rate both groups as to global effectiveness and examine whether group A is rated higher.
3. Establishment of more intermediate counselor levels of experience and training and evaluating the Scale's sensitivity to differentiate these levels.
4. Applying the Scale to the rating of counselor trainees in various counselor education programs and observing how satisfactorily it detects change in their counseling

behaviors as they progress through the programs.

5. Evaluation of the Scale's value for accelerating counselor educational growth as a result of counselor trainees using the Scale themselves in counselor rating sessions.

Bibliography

1. Albee, G. N. and Harlin, R. M. Judgment of adjustment from drawings: the applicability of rating scale methods. J. Clinical Psychol. 1950, 6: 363-365.

2. Alexander, F. Fundamentals of Psychoanalysis. New York: Norton, 1948.

3. American Educational Research Association. Review of Educational Research, Guidance, Counseling, and Personnel Services, April, 1963, pp. 184-185.

4. Anderson, R. P. and Anderson, G. U. Development of an instrument for measuring rapport. Personnel and Guidance Journal. 1962, 41: 18-24

5. Bayroof, A. G., Haggerty, H. R., and Rundquist, E. A. Validity of ratings as related to rating techniques and conditions. Personnel Psychology. 1954, 7: 93-113.

6. Bendig, A. W. The reliability of self-ratings as a function of the amount of verbal anchoring and of the number of categories on the scale. J. Applied Psychol. 1953, 37: 38-41.

7. Bendig, A. W. Reliability and the number of rating scale categories. J. Applied Psychol. 1954, 38: 38-40

8. Bendig, A. W. Rater experience and case history judgments of adjustment. J. Clinical Psychol. 1955, 11: 127-132.

9. Bendig, A. W. and Sprague, Janine. Rater experience and the reliability of case history ratings of adjustment. J. Consult. Psychol. 1954, 18: 207-211.

10. California Test Bureau. A Glossary of Measurement Terms, Monterey, California, 1959.

11. Carnes, E. F. and Robinson, F. P. The role of the client talk in the counseling interview. Educ. and Psychol. Measmt. 1948, 8: 635-644.

12. Cartwright, R. D. and Lerner, B. Empathy need to change and improvement in psychotherapy. J. Consult. Psychol. 1963, 27: 138-145.

13. Cattell, J. M. Cited by Woodworth, Robert S. and Harold Schlosberg. Experimental Psychology, Revised Edition. New York: Henry Holt and Company, 1954.

14. Chaplin, H. and Marshall, H. Optimal refinement of the rating scale. J. Applied Psychol. 1939, 23: 323-331.

15. Collier, R. M. A scale for rating the responses of the therapist. J. Consult. Psychol. 1953, 17: 321-326.

16. Covner, B. J. Studies in phonographic recording of verbal material: I. The use of phonographic recordings in counseling practice and research. J. Consult. Psychol. 1942, 6: 104-113.

17. Covner, B. J. Studies in phonographic recording of verbal material: III. The completeness and accuracy of counseling interview reports. J. Gen. Psychol. 1944, 30: 181-203.

18. Covner, B. J. Studies in phonographic recordings of verbal material: IV. Written reports of interviews. J. Applied Psychol. 1944, 28: 89-98.

19. Cronbach, L. J. Response sets and test validity. Educ. and Psychol. Measmt. 1946, 6: 475-494.

20. Cronbach, L. J. Essentials of Pshchological Testing. New York: Harper Bros., 1960.

21. Cronbach, L. and Meehl, P. Construct validity in psychological tests. Psychol. Bull. 1955, 52: 281-302.

22. DeRoo, M. M. A study of the relationship between selected verbal behaviors of counselors and variables of the Holtzman ink blot test and the Rokeach dogmatism scale. Unpublished doctoral dissertation, 1965, Michigan State University.

23. Dittman, A. T. The interpersonal process in psychotherapy: development of a research method. J. Abnorm. and Soc. Psychol. 1952, 47: 236-244.

24. Dymond, R. A scale for the measurement of empathic ability. J. Consult. Psychol. 1949, 13: 127-133.

25. Ebel, R. L. Estimation of the reliability of ratings. Psychometrika. 1951, 16: 407-424.

26. Feldman, J. J., Hyman, H., and Hart, C. W. A field study of interviewer effects on the quality of survey data. Public Opinion Quarterly. 1951, 55: 388-398.

27. Ferenczi, S. The principle of relaxation and neocartharsis. Internat. J. Psycho-Anal. 1930, 11: 428-443.

28. Fiedler, F. E. The concept of an ideal therapeutic relationship. J. Consult. Psychol. 1950, 14: 239-245.

29. Fiedler, F. E. A comparison of therapeutic relationships in psychoanalytic, nondirective, and Adlerian therapy. J. Consult. Psychol. 1950, 14: 436-445.

30. Finesinger, J. E. Psychiatric interviewing: I. Some principles and procedures in insight therapy. Amer. J. Psychiat. 1948, 105: 187-195.

31. Frank, G. H. On the history of the objective investigation of the process of psychotherapy. J. Psychol. 1961, 51: 89-95.

32. Frank, G. H. and Sweetland, A. A study of the process of psychotherapy. J. Consult. Psychol. 1952, 26: 135-138.

33. Gebhart, G. and Hoyt, D. Personality needs of under and over achievers. J. Applied Psychol. 1958, 42: 125-128.

34. Guilford J. Personality. New York: McGraw-Hill, 1959.

35. Gordon, Kate. Group judgments in the field of lifted weights. J. Exper. Psychol. 1924, 7: 398-400.

36. Halpern, H. and Lesser, L. Empathy in infants, adults, and psychotherapists. Psychoanal. Rev. 1960, 27: 32-42.

37. Hamlin, R. M. The clinician as judge: implications of a series of studies. J. Consult. Psychol. 1954, 18: 233-238.

38. Harway, N. I., Dittman, A. T., Raush, H. L., Bordin, E. S., and Rigler, D. The measurement of depth of interpretation. J. Consult. Psychol. 1955, 19: 247-253.

39. Helmstadter, G. C. Principles of Psychological Measurement. New York: Appleton-Century-Crofts, 1964.

40. Highland, R. and Berkshire, J. A methodological study of forced-choice performance ratings. V Res. Bull. 51-9 HRRC. San Antonio, Texas, 1951.

41. Hoyt, C. Test reliability obtained by analysis of variance. Psychometrika. VI, 1941, 153-160.

42. Hoyt, C. Cited by Ebel, R. Measuring Educational Achievement. Englewood Cliffs, New Jersey: Prentice Hall, 1965.

43. Jourard, S. I-thou relationship versus manipulation in counseling and psychotherapy. J. Ind. Psychol. 1959, 15: 174-179.

44. Kagan, N., Krathwohl, D. R., and Miller, R. Stimulated recall in therapy using video tape--a case study. J. Counseling Psychol. 1963, 10: 237-243.

45. Kelly, C. D. Two verbal techniques in a miniature counseling situation. Psychol. Monogr. 1948, 62: no. 294.

46. Kelly, J. M. The development and evaluation of a measure of counselor effectiveness. APGA Journal, 1966 in print.

47. Kessler, D., Mathieu, P. and Klein, M. Sampling from the recorded therapy interview: a comparative study of different segment lengths. J. Consult. Psychol. 1965, 29: no.4, 349-357.

48. Kuder, G. and Richardson, M. The theory of estimation of test reliability. Psychometrika, 1937, 2: 151-160.

49. Lasswell, H. D. The problem of adequate personality records: a proposal. Amer. J. Psychiat. 1929, 85: 1057-1066.

50. Lasswell, H. D. Verbal references and physiological changes during the psychoanalytic interview: A preliminary communication. Psychoanal. Rev. 1935, 22: 10-24.

51. Lasswell, H. D. Certain prognostic changes during trial (psychoanalytic) interviews. Psychoanal. Rev. 1936, 23: 229-293.

52. Lasswell, H. D. A provisional classification of symbol data. Psychiatry. 1938, 1: 197-204.

53. Lindquist, E. F. Design and Analysis of Experiments in Psychology and Education. Boston: Houghton Mifflin Company, 1953, 292.

54. Lindquist, E. F. Educational Measurement. Menasha, Wisconsin: George Banta Publishing Co., 1951.

55. Meehl, P. E. and MacCorquodale, K. Operational validity of intervening constructs. In: M. H. Marx. Psychological Theory. New York, Macmillan.

56. Porter, E. H., Jr. The development and evaluation of a measure of counseling interview procedures: I. The development. Educ. and Psychol. Measmt. 1943, 3: 105-126.

57. Porter, E. H., Jr. The development and evaluation of a measure of counseling and interview procedures. II. The evaluation. Educ. and Psychol. Measmt. 1943, 3: 215-238.

58. Rosenbaum, R. C. The assessment of counselor-trainee perceptions of interview protocols before and after an intensive practicum experience. Doctoral dissertation, 1964, University of Minnesota.

59. Rosenbaum, R. C. and DeRoos, M. M. Counselor response system of the behavior interaction description system. Paper read at the APGA Convention, Minneapolis, Minnesota, 1965.

60. Rosenbaum, M. L., Sperber, A., Rigler, D. Williams, J., Harway, N. I., Bordin, E. S., Dittman, A. T. and Hays, W. L. A dimensional analysis of depth of interpretation. J. Consult. Psychol. 1956, 20: 89-97.

61. Rosenbaum, M. and Bordin, E. S. Warmth in personality development and in psychotherapy. Psychiat. 1957, 20: no. 4, 351-363.

62. Rubloch, M., Elkes, C., Flint, A. A., Usdansky, B. S., Newman, R. G., and Silber, E. National Institute of Mental Health pilo and study in training mental health counselors. The Amer. J. Orthopsychiatry. 1963, 33: 670-689.

63. Rogers, C. R. Counseling and Psychotherapy. Boston, Houghton Mifflin Company, 1942.

64. Rogers, C. R. Electrically recorded interviews in improving psychotherapeutic interviews. Amer. J. Orthopsychiat. 1942, 12: 429-435.

65. Rogers, C. R. Client-centered Therapy. Cambridge, Mass: Riverside Press, 1951.

66. Rogers, C. R. Sponsor: Mathieu, P. and Mourgand, J. P. Relationship of group counseling to subsequent academic performance at the college level. Cooperative research. Project no. OE1417, Wisconsin Psychiatric Institute, University of Wisconsin, Madison, Wisconsin, 1962.

67. Rose, A. M. Interviewing to test for validity and reliability. Intern. J. of Opinion and Attitude Research. 1947, 1: 100-101.

68. Rich, F. S. Psychology and Life. Chicago: Scott, Foresman, and Co. Sixth edition, 1963.

69. Rugg, H. O. Is the rating of human character practicable? J. Educ. Psychol. 1921, 12: 425-438 and 485-501.

70. Schafer, R. Generative empathy in the treatment situation. Psychoanal. Quart. 1959, 28: 342-373.

71. Shapiro, S. and Eberhart, J. C. Interviewer differences in an intensive interview survey. Intern. J. of Opinion and Attitude Research. 1947, 1: 1-17.

72. Snyder, W. U. An investigation of the nature of non-directive psychotherapy. J. Gen. Psychol. 1945, 33: 193-223.

73. Snyder, W. U. The Psychotherapy Relationship. New York: Macmillan, 1961.

74. Stoph, Joseph. Responses to hypothetical counseling situation as a prediction of relationship orientation in school counselors. Unpublished doctoral dissertation, University of Wisconsin, 1963.

75. Strunk, O., Jr. Empathy: a review of theory and research. Psychol. Newsletter. 1957, 9: 47-57.

76. Strupp, H. H. Nature of psychotherapist's contribution to the treatment process. Arch. Gen. Psychiat. 1960, 3: 219-231.

77. Symonds, P. M. On the loss of reliability in ratings due to coarseness of the scale. J. Exper. Psychol. 1922, 7: 456-461.

78. Symonds, P. M. Notes on rating. J. Applied Psychol. 1925, 9: 188-195.

79. Symonds, P. M. Research on the interviewing process. J. Educ. Psychol. 1939, 30: 356-353.

80. Thornton, G. The effect upon judgments of personality traits of varying a single factor in a photograph. J. Soc. Psychol. 1943, 18: 127-148.

81. Thornton, G. The effect of wearing glasses upon judgments of personality traits of persons seen briefly. J. Applied Psychol. 1944, 28: 203-207.

82. Torgerson, W. Theory and Methods of Scaling. New York: John Wiley and Sons, 1958.

83. Truax, C. B. and Carkhuff, R. R. Theory and research in counseling and psychotherapy. APGA Journal. May, 1964, 880-885.

84. Tyler, L. E. Tests and Measurements. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1963.

85. Walker, H. M. and Lev, J. Statistical Inference. New York: Holt, Rinehart and Winston, 1953.

86. Walls, G. R. and Johnston, J. A. Counselors look at themselves on video tape. J. Counseling Psychol. 1963, 10: 232-236.

87. Ward, R. Educating counselors by video-tape using the interpersonal process recall technique. Unpublished doctoral dissertation, 1965, Michigan State University.

88. Watson, P. D. and Kanter, S. S. Some influences of an experimental situation on the psychotherapeutic process: a report based on 44 treatment interviews, of the reactions of a patient and therapist to observation, recording, and physiological measurement. Psychosom. Med., 1956, 18: 457-470.

89. Westman, W. C. A comparison of three modes of rating counseling interaction. Unpublished doctoral dissertation, 1965, University of Wisconsin.

90. Winder, C., Ahmad, F., Bandura, A., and Rau, L. Dependency of patients psychotherapists' responses and aspects of psychotherapy. J. Consult. Psychol. 1962, 26: 129-134.

91. Wolf, R. and Murray, H. A. An experiment in judging personalities. J. of Psychol. 1937, 3: 345-365.

APPENDICES

APPENDIX A

A COMPILATION OF COUNSELOR VIDEO VERBAL BEHAVIORS AS OBSERVED ON AUDIO-VIDEO TAPES

VIDEO VERBAL BEHAVIORS CHARACTERIZING MASTER'S CANDIDATE COUNSELORS

Cognitive

1. Superficial
2. Clumsy structuring of interview
3. Nonverbal behavior indicates anxiety and discomfort with client
4. Prestructures client's reply
5. Distant from the client emotionally
6. Avoids client problem areas
7. Reflects or rephrases cognitive aspects
8. Expresses no feeling about client
9. Vague
10. Moves client away from problem area
11. Questions cognitive aspects of client's verbalization
12. Makes general comment in response to client's verbalization
13. Agrees with client's cognitive verbalization
14. Repeats client's comments verbatum
15. "Yeah," "um huh," "I see," "right,"
16. Comments on room
17. Makes judgmental comment concerning client or client's verbalization
18. Adds additional cognitive content or meaning to client's verbalization in restatement effort
19. Leads client in specific cognitive direction through suggestive comments
20. Lectures or sermonizes to client
21. Philosophizes to client
22. Conducts monopolizing conversation
23. Advises
24. Offers cognitive evaluation of client's situation
25. Pleads for information
26. Excessive verbalization response to client cognitive content
27. Anxiety laughter

Affective

1. Inaccurately reflects feelings
2. Inaccurately interprets feelings
3. Reflects ambiguity
4. Leads client in irrelevant affective direction through suggestive comments
5. Adds additional affective meaning to clients verbalization in reflection effort
6. Excessive verbalization response to client affective response content
7. "Yeah," "um huh," "I see," "right,"

Video Verbal Behaviors Characterizing Ph.D. Candidate Counselors

Cognitive

1. Questions cognitive aspect of client's verbalization
2. "Yeah," "um huh," "I see," "right,"
3. Reflects cognitive aspect of client's verbalization
4. Reconstructs or simplifies cognitive presentation

Affective

1. Responds at the same emotional level as the client
2. Free from discomfort with client
3. Questions client's feeling
4. Appropriately reflects feelings
5. Appropriately clarifies feelings
6. Appropriately interprets feelings
7. Asks personal-intimate questions
8. Discusses client's problem areas
9. Reassures client concerning affective aspect of verbalization
10. Offers affective evaluation of client's situation
11. Questions to clarify more specifically general feelings expressed by the client
12. Expresses to client the feeling which client's verbalization elicits from him
13. Expresses his feeling about client
14. Suggests related affective area for discussion
15. Points out to client his reluctance to discuss an affective area
16. Points out to client his refluctance to discuss more deeply his feelings
17. "Yeah," "un huh," "I see," "right,"

APPENDIX B

CATEGORIZATION OF COUNSELOR AUDIO-VIDEO BEHAVIORS ACCORDING TO BEHAVIOR CATEGORIES

COUNSELOR VIDEO VERBAL BEHAVIOR CATEGORIES AND ILLUSTRATIONS

A. Affective

This category includes all counselor responses which make reference to some affective or feeling aspect of a client's verbal communication.

Illustrations

1. Responses which enter the feeling frame of reference of the client, which deals with expressed concerns, needs, and wants of the client as expressed or implied by the client.

2. At the same emotional level as the client. Emoting with the client on his feeling level.

3. Responses which reflect, deal with, clarify, or interpret the feeling of what the client is saying.

B. Understanding

The counselor's ability to communicate to the client the fact that he knows what the client is feeling or talking about. Counselor communicates an understanding of client's feelings to the client.

Illustrations

1. Counselor reflects client feelings accurately

   Counselor attempts to express in fresh words the essential attitudes (not so much the content) expressed explicitly by the client. The counselor attempts to mirror the client's attitudes for his better self-understanding and to show the client that he is being understood by the counselor.

2. Counselor clarifies feelings in a specific, non-vague, manner.

   The counselor clarifies what was implicit in what the client said; he reflects relationships or meanings for the client which are suggested or implied in the feelings or ideas brought about by the client's actual responses.

3. Counselor interprets feelings appropriately.

   Interpretation refers to any remark which interprets, analyzes, explains, or draws inferences. Counselor presents to the client an hypothesis about reliationships or meanings of his attitude behaviors, which he had not considered before, for his consideration.

4. Counselor's verbal responses and accompanying behavior demonstrate primarily an interest in and comfort with the client and thereby conveys counselor acceptance to the client.

   a. Counselor's attention and concentration are on the client.

   b. Counselor's verbal responses demonstrate that he is following what the client says (e.g. "yes," "HmHm," "un huh," "I see," etc.)

   c. Counselor gives indications of approval which seem to give reassurance and encouragement (e.g. "good," "I feel that's a fine idea.") These may be positive confirmations by the counselor of the adequacy of client behavior, attitude, feelings, etc., as alluded to in the verbalizations of the client (e.g. "you are a capable person;" "you can be consistent;" "you can be reasonable;" "you can be organized;" "you can feel better.").

   d. Counselor's nonverbal manner is free of discomfort and anxiety (there is an absence of continuous shifting in the chair, looking away from the client, or continuous movement of arms, hands, and legs).

5. Counselor questions client's feelings generally or questions to have client clarify more specifically the general feelings which he has expressed. This is a request for general elaboration and expression of feelings.

6. Counselor expresses his feelings about the client or the feelings which the client's verbalizations have elicited from him. This includes:

   a. Statements of counselor's feelings about the client as a person.

   b. Statement of counselor's feelings about something the client has said, done, or thought.

   c. Statements in which the counselor's feelings are expressed as being his own.

C. Specific

Counselor's interaction with client is characterized by concrete (specific) responses concerning the client's problem areas and emotional conflicts.

Illustrations

1. Counselor proposes or suggests that the interview be devoted primarily to the objective of exploring the client's problem areas and emotional conflicts.

2. Counselor's responses reflect, or are directed toward, the core of the client's remarks rather than their peripheral aspects.

3. Counselor's responses move the interview discussion in the direction of the client's problem areas and emotional conflicts or bring the client's discussion back to these topics.

4. Counselor's responses are relevant to the client's problem areas and emotional conflicts rather than being irrelevant as characterized by abstract, intellectual, lectural, extraneous, etc., type of responses.

D. Exploratory

The counselor's assistance and skill in helping the client to explore his feelings, values, and other concerns.

Illustrations

1. Counselor suggests possible exploration of affective area referred to either explicitly or implicitly by client.

2. Counselor points out to the client his reluctance to explore an affective area.

3. Counselor asks personal-intimate question of client.

4. Counselor is willing to explore with the client, in a nondefensive manner, the import of his (i.e., the counselor's) behavior traits which the client observes in their relationship.

E. Cognitive

This category includes all counselor responses which omit any reference to affective or feeling aspects of a client's verbal communication.

Illustrations

1. Response is distant from the client emotionally.

2. Superficial response which passes over emotional implications of client's verbalization.

3. Not with the client emotionally.

4. A response which is not on the feeling level of the client.

5. A response which does not enter the client's emotional frame of reference--it is at a different level from the client's feelings.

6. Own feelings or apprehension preventing counselor from making emotional contact with the client.

7. Clumsy structuring of interview.

8. Answering factual questions.

9. Repeat of client's remarks verbatum.

10. Responses devoid of any reference to client affect--feeling, emotion, mood, attitude, etc.

11. Matter of fact responses, informative responses, general and superficial discussion responses.

F. Nonunderstanding

Counselor does not communicate an understanding of the client's feelings to the client.

Illustrations

1. Counselor expresses confusion as to cognitive or affective meaning of client's verbalization.

2. Counselor's responses show evidence that he has prejudged the client and thereby has classified, stereotyped, or categorized the client.

3. Counselor's responses inaccurately reflect the client's feelings.

4. Counselor's responses of clarification of client's feelings are attempted in a vague, non-specific manner.

5. Counselor's responses encourage or reflect ambiguity.

6. Counselor's responses in restatement or reflection add unwarranted additional affective or cognitive content meaning to client's verbalization.

7. Counselor responses reflect the cognitive aspects of the client's verbalization when affective aspects are central.

8. Counselor repeats client's comments verbatum.

9. Counselor pleads for information.

10. Counselor verbal and nonverbal responses indicate a lack of interest in, and discomfort with, the client.

    a. Counselor looks away from the client rather than concentrating on him.

    b. Counselor does not appear to be following what the client says.

G. Nonspecific

Counselor's interaction with client is characterized by nonconcrete (nonspecific) responses concerning the client's problem areas and emotional conflicts.

<u>Illustrations</u>

1. Counselor's responses reflect, or are directed toward, the peripheral aspects of the client's remarks rather than toward their core meanings.

2. Counselor responses are related primarily to the cognitive content of the client's verbalizations and deal only superficially with the affective content.

3. Counselor responses are irrelevant to the client's problem areas and emotional conflicts. They are characterized by abstract, intellectual, and extraneous responses.

    a. Counselor encourages expansion of superficial cognitive discussion.

b. Counselor intellectually debates cognitive or affective aspects with client.

c. Counselor remains silent and simply permits client to wander in his verbalizations.

d. Counselor is excessive in his response to client cognitive or affective content.

H. Nonexploratory

The counselor's failure to help or permit the client to explore his feelings, values, or concerns.

Illustrations

1. Pre-structures clients reply

2. Counselor monopolizes conversation, philosophizes with, or lectures to client.

3. When client presents affective content, counselor response does not encourage further exploration of this content.

4. Counselor's responses move the interview discussion in a direction away from the client's problem areas or emotional conflicts

5. Counselor makes judgmental or evaluative responses to client's verbalizations.

6. Counselor's responses are characterized by advice-giving and question-answering.

7. Counselor's responses are structured to specifically lead the client in definite cognitive or affective directions rather than permitting the client to initiate these directions.

APPENDIX C

DEFINITIONS AND ILLUSTRATIONS OF THE A-V SCALE'S DIMENSION CATEGORIES

UNIT OF EVALUATION FOR THE A-V COUNSELOR BEHAVIOR SCALE

The unit for evaluation for the A-V Counselor Behavior Scale is the counselor-client unit. A client verbalization is emitted or elicited and the counselor's response to it is then rated on all five dimensions of the scale according to the dimension category definitions. One forced choice rating is made independently for each of the four specific counselor behavior dimensions and a fifth rating is made evaluating the effectiveness of the counselor's response. Generally, each client-counselor unit is judged independently of preceding units, although the congruence between the direction and level of counseling movement by the counselor with the direction and level of counseling concerns the client is attempting to present must continually be kept in awareness when making rating judgments. The scoring of all dimension categories must be based upon concrete and observable behavior and is not to be based upon what a judge may believe could possibly be occurring within the counselor.

DIMENSIONS OF THE A-V COUNSELOR BEHAVIOR SCALE

I. The Cognitive-Affective Dimension

This dimension indicates whether a counselor's response is in reference to some affective or feeling aspect of a client's verbal communication (an affective response) or whether his response omits any such reference to affect or feeling in the client's verbal communication (a cognitive response). These categories are more clearly defined as follows:

A. Affective Responses

Affect refers to emotions, feelings, fears, moods, desires, urges, impulses, fantasy, dreams and any ideas, attitudes, beliefs, convictions, etc., which are based upon such referents as these. A counselor affective response is one which makes reference to any of this type of affective material on a feeling level regardless of how congruent it is with that being expressed by the client. An affective response is not determined by the extent to which the counselor's manner or communication is affective, but rather, it is determined solely by the fact that the counselor responds through comment (reflection, clarification, interpretation, etc.) to underlying affective elements, or to what he implies are underlying affective elements, in the client's verbal and nonverbal expressions as they are revealed in his interactional discussion with the counselor.

Examples:

1. "You feel quite inadequate in these situations."
2. "You act that way because you are angry."
3. "You feel better when she mothers you."
4. "How do you feel when your parents argue?"
5. "Maybe that's not what you want at all."

B. Cognitive Responses

A counselor cognitive response is one which is on a non-affective plane in that it is devoid of any basic concern, reference, or recognition of client feelings, emotions, moods, attitudes, or any other affective aspects which may be implied in the client's communication. Such a response passes over any affective implications inherent in the client's remarks. The counselor responds at a cognitive, non-feeling level or asks a question or makes a statement concerning cognitive or content material. Such cognitive responses are more formal, informative, matter of fact, of a general discussion order, and socially acceptable oriented, and are altogether on a different plane from responses which make emotional contact with the client through discussion of feelings, attitudes, etc. Cognitive responses include answering factual questions, "busy" talk, discussing trivialities, repeating client remarks verbatum, exclusive talk about third parties without affective reference back to the client, any any general educational, intellectual, vocational, conversational etc., type discussions.

Examples:

1. "Yes, that seems to be a logical step to take."
2. "Your father was an attorney, huh. Well that seems to give you a good background."
3. "How are you feeling today?" (If intended in a general sense)
4. "I think your performance in your science course would support a decision to enter medicine."
5. "So you want to enter graduate study. I think that's nice."

II. The Understanding--Non-Understanding Dimension

This dimension relates to how well the counselor's response communicates to the client the fact that he understands, or is seeking to understand, the client's basic problem. These categories are more clearly defined as follows:

A. Understanding Responses

These are responses which imply that the counselor understands the basic underlying affective or cognitive concerns of what the client is talking about, or is trying to get enough information from the client (through appropriate reference to what the client is feeling and expressing and are appropriately timed to the significant cues given by the client.

Examples:

1. "These feelings make it difficult for you to get along with others."

2. "In other words, you feel inadequate and self-conscious in these situations."

3. "You really want to be a big man."

4. "Could you tell me more specifically what you mean by this odd feeling about your brother?"

5. "Could you tell me more about that?"

B. Non-Understanding Responses

These are responses which imply that the counselor lacks an understanding of the basic underlying affective or cognitive concerns of what the client is talking about, or are responses which indicate the counselor is making no attempt to obtain appropriate information from the client from which he may derive an understanding of the client's underlying concerns. Non-understanding responses are those which make inappropriate reference to what the client is feeling or expressing, are inappropriately timed to significant cues, or are responses which are made after several significant client clues have either been ignored or entirely missed.

Examples:

1. CL: "When he said that, I just turned red and clutched my fists."

   CO: "Some people don't say nice things."

2. CL: "When I showed my mother my grades and she said they were acceptable, but I should be

doing better, I felt like quitting college and telling her to go to hell."

CO: "What would you do if you quit?"

3. CL: "Sometimes I have queer feelings about my parents."

CO: "I see."

4. CL: "When Ivan put his arm about me, I started crying."

CO: "You mentioned that Ivan likes the classics."

5. CL: "I'm so self-conscious in a crowd. (CO: Silence). I just can't speak to groups because of my nose (CO: Silence). My brothers always made fun of my nose (CO: Silence). My aunt said once I looked like an eagle. Boy, I hated her guts (CO: Silence). I just don't like my family."

CO: "When you go to college you can get away from your family.

III. The Specific--Non-Specific Dimension

This dimension indicates whether the counselor is dealing directly with the client's basic problem by confining his response to a concrete (specific) one quite relevant to the presented problem (a core response), or whether the counselor passes over dealing directly with the client's basic problem and gives a more general, or non-specific response not too relevant to the presented problem (a peripheral response). These categories are more clearly defined as follows:

A. Specific Responses

These are core type responses which indicate that the counselor is handling directly and honestly the basic problem presented explicitly or implicitly by the client. The counselor's remarks are keyed to the central aspects of the problem which the client is feeling and communicating. The counselor thus responds to the underlying concern, feeling, or meaning--the primary message--of the client's remarks such that his responses zero in and are congruent with the level of affect or concern being expressed by the client.

Specific responses are those which influence the client to attend with specificity to his problem areas or emotional conflicts and concerns. They are relevant to these problem areas or emotional conflicts and concerns, and move the interview discussion in this direction or bring the client's discussion back to these topics. Specific responses are usually directed at underlying affect, but may consist of cognitive responses if awareness of affect is not essential to the working through of the counseling problem.

Examples:

1. CL: "I've gotten all A's this year and I still feel lousy."

   CO: "You're sensing that obtaining such high grades is not such a satisfactory solution to your feelings of inadequacy as you thought."

2. CL: "It's a common thing for men to have mistresses in Europe. Why do they have to make such a fuss about it over here?"

   CO: "What's that got to do with your sexual problem?"

3. CL: "Hell, what do I care about being on time, my boss never says a good word about my work and he didn't do a thing about getting me that raise."

   CO: "When your boss doesn't continually praise and take care of you, you become angry and take it out on him by being late."

4. CL: "When I try to make love to my wife anymore, I just can't perform."

   CO: "You still look on your wife as your mother."

5. CL: "It's my twenty-second birthday tomorrow. I hate growing old."

   CO: "It's more fun being a little girl, then you don't have to face these adult responsibilities."

B. <u>Non-Specific Responses</u>

These are peripheral type responses which indicate that the counselor is not dealing directly and honestly with the basic problem presented explicitly or implicitly by the client. Rather, the counselor responses stay away from the client and his problem, or at best handles the problem at a superficial, surface level. These responses indicate that the counselor has missed or ignored cues which reveal the

client's main problem areas. In such responses the counselor concentrates on aspects of the client's verbalizations that are of less importance than the main message which the client has attempted to communicate. The counselor's responses refer mostly to cognitive generalizations or superficial cognitive specifics rather than their underlying concerns, and thus such responses are not congruent with the level of affect being expressed by the client.

Non-specific responses are those in which the counselor becomes abstract, intellectual, or extraneous rather than confining his remarks to the specifics of the client's problem areas or emotional conflicts. Such responses tend to be irrelevant to, or deal inappropriately with, these problem areas or emotional conflicts and thus move the interview discussion in a direction which avoids them. Such responses encourage expansion of superficial cognitive content; leads to debate with the client over cognitive or affective concerns; are typified by excessive response verbalization and verbal monopolizing on the part of the counselor; and are illustrated by all sorts of lecturing, sermonizing, and philosophizing to the client by the counselor.

Examples:

1. CO: "I've gotten all A's this year and I still feel lousy."

   CO: "I wouldn't feel so bad if I were you. There are a lot of kids who would give their eye-teeth

for those grades. Feel thankful and count your blessings."

2. CL: "It's a common thing for men to have mistresses in Europe. Why do they have to make such a fuss about it over here?"

   CO: "Well you know it's really not the Christian thing to do. Anyway, it's wise not to develop relationships which are frowned upon by our society. Have you considered locating in Europe?"

3. CL: "Hell, what do I care about being on time, my boss never says a good word about my work and he didn't do a thing about getting me that raise!"

   CO: "Don't you think though, that consistently being on time will be to your advantage in the long run? It looks much better on your record and shows a sense of maturity and responsibility. Your boss is probably a very busy man. I'm sure he will make some comment on your work in the future. And, don't be so disappointed about the raise. A lot of people have to get by on a lot less than you're making."

4. CL: "When I try to make love to my wife anymore, I just can't perform."

   CO: "I imagine that's very frustrating to your wife."

5. CL: "It's my twenty-second birthday tomorrow. I hate growing old."

CO: "My lands, child, that's not old. You're real young yet. You have your whole life ahead of you."

IV. Explorative - Non-Explorative Dimension

This dimension concerns the extent to which the counselor's response permits, encourages, or elicits further exploration on the part of the client of any aspect of his problem areas or affective concerns. These categories are more clearly defined as follows:

A. Exploratory Responses

Counselor exploratory responses are responses which first, indicate that the counselor has recognized some portion of the client's basic problem and, second, has made some verbal reference of this back to the client in such a way that the client is free from any defined or "limited" structure in his response, and is thus free to respond to any degree and depth to the counselor's reference that he may choose. In addition to the verbal response, the counselor's whole manner creates an atmosphere for further exploration. An exploratory response, then, is one which concerns some aspect of the client's basic problem and at the same time permits and encourages the client flexibility and freedom in his response. Such responses are often open ended and

allow the client to explore his own feelings and to expand upon them.

Examples:

1. "It seems that your anger is really directed toward your father."
2. "Perhaps your need for such high esteem has something to do with your choice of medicine as a career?"
3. "Maybe we should discuss a little more these strong attitudes about sex."

B. Non-Exploratory Responses

Counselor non-exploratory responses are responses which indicate no recognition of the client's basic problem, and thus no attempt to search them out further, or responses which structure or limit the client's response. A "yes" or "no" answer or a "pat" answer is often implied by such responses. Non-exploratory responses are those which give the client little opportunity to explore or expand, or to express himself freely.

Examples:

1. "What is your average in English so far this year?"
2. "That's a nice place to study, isn't it?"
3. "What was your uncle's middle name?"

COUNSELOR RESPONSE EVALUATION

This evaluation is made independently of the four counseling dimension ratings and is a dichotomous rating whereby the counselor's response is marked either effective or non-effective. This rating is not based upon an evaluation of the client's response to the counselor's remarks but, rather, is based solely upon the judge's professional impression as to how appropriately, from a counseling vantage point, the counselor's response deals with the counselor's verbalization and thus contributes to the development and maintenance of counseling progress. Counseling progress is generally defined as concentration of discussion on basic problems and feelings of immediate concern to the client such that there develops for the client an increasing undistorted awareness of their influence and consequence on his behavior and relationships with others. In this regard, the judge decides globally--taking into consideration the counselor's total manner and behavior, including nonverbal--the effect the counselor's response will have or could have, toward permitting or encouraging the client to move in a direction of counseling progress.

If the judge evaluates the counselor's response as effective he can either check a 4 or a 3 under the heading "effective." A rating of 4 indicates that the judge's professional impression is that the response is about as "good" or the "best" possible in terms of counseling progress. A rating of 3 indicates that the judge's

professional impression is that the response is effective toward some counseling progress but it was not the "best," that is, a more effective response or communication could have been given. If the judge evaluates the counselor's response as non-effective he can either check a 2 or a 1 under the heading "non-effective." A rating of 2 indicates a response which contributes in no way to counseling progress (sort of neutral) while a rating of 1, in addition to this, indicates also a complete lack of understanding or concern regarding the client's problem situation, or is a response which is definitely detrimental to counseling progress.

APPENDIX D

CLARIFICATION OF FREQUENTLY OCCURRING RATING PROBLEMS AND QUESTIONS

Right Side and Left Side Scoring Tendencies

For a given counselor response, judges should be particularly cautioned against rating the first dimension and then marking all the remaining dimensions to follow suit: e.g., marking the first dimension cognitive, and then the remaining dimensions nonunderstanding, nonspecific, nonexploratory, and noneffective, or marking the first dimension affective, and then the remaining dimensions understanding, specific, exploratory, and effective. Each dimension is to be scored independently according to category definitions.

Possible Counselor Dimension Scoring Patterns

According to the dimension category definitions, the following scoring patterns are possible:

1. A response can be cognitive and still be understanding, specific, and/or exploratory.

2. A response can be affective without being understanding, specific, and exploratory.

3. A response can be affective and understanding without being specific and exploratory.

4. A response can be understanding and specific without being affective or exploratory.

5. A response can be affective, understanding, and exploratory without being specific.

A response, however, can never be exploratory without being understanding.

Recurring Judge Questions Concerning Dimension Category Definitions

Although it is not possible to anticipate the nature of all judge questions concerning category definitions, the important

recurring ones are presented here along with the answers of clarification:

Q: "If the counselor demonstrates affect himself in his response, is this scored affect?"

A: "No, not necessarily. The main criterion is whether the counselor is referring back to some affective aspect presented by the client. The important thing is what he is referring back to, not his manner in referring."

Q: "What if the counselor makes reference back to affective aspects or implications in the client's remarks which are not justified by the client's remarks or manner. Is this scored affect?"

A: "Yes. However, if you feel the affective reference is inappropriate you would also score the response nonunderstanding, nonspecific, nonexploratory, and probably noneffective."

Q: "What if the counselor refers to or discusses the feelings and concerns of third parties which are in no way related to the concerns or involvement of the client. Is this scored affect?"

A: "No, such a response is scored cognitive. To be scored affect, discussion of feelings and concerns of third parties must be directly related to the client and tied in with the feelings and concerns of the client."

Q: "Are all questions by the counselor to be considered as seeking understanding?"

A: "No. The question has to be related to clarifying for the counselor some aspect of major concern, explicitly or implicitly implied by the client. Random questions asked by the counselor from the 'top of his head' to alleviate anxiety, to continue a cognitive discussion unrelated or leading away from the client's main concerns, or which in any way seem unrelated or inappropriate to what the client is trying to express, are scored nonunderstanding."

Q: "If the counselor fails to respond to several important affective client leads but finally responds to a lesser type of client lead, is this scored nonunderstanding?"

A: "yes."

Q: "If the counselor maintains a cognitive discussion with the client when the client's remarks are implying more affective concerns, is this scored nonunderstanding?"

A: "Yes. Whenever a counselor maintains a cognitive discussion as a means of keeping away from a discussion of the client's affective concerns, his responses are scored nonunderstanding.

Q: "When is a cognitive response on the part of the counselor scored understanding?"

A: "Whenever the cognitive content to which the counselor is responding seems to be the important concern of the client and is not a cover-up for underlying affective problems or concerns. This also applies to the specific and exploratory categories. Responses are not scored understanding if it is obvious that affective problems are of more concern to the client than cognitive content, but the counselor continues on a cognitive plane. In such cases the counselor's response is also nonspecific and nonexploratory."

Q: "Can a counselor's response be specific without being understanding?"

A: "Yes. A client can present point blank to the counselor a central concern of his and the counselor may comment on it without demonstrating any understanding of its significance or may comment on it is such a way that there is no further pursuit of it."

Q: "What is the main criterion for specific?"

A: "When the counselor's response stays at the same emotional level as the client and it is able to zero through the client's verbage right to his central concerns without being side-tracked by less important affective or cognitive concerns."

Q: "If a counselor's response is such that it is possible that the client could respond in any way that he chooses, is that response scored exploratory?"

A: "No. For a counselor's response to be scored exploratory it is essential that the counselor's response not only permit the client to respond to any length or any depth that he chooses, but it must also reflect that the counselor has some understanding that there is some aspect of real concern to the client and his response is thus encouraging further elaboration of this concern. This concern of the client may be vague and still not specifically identified by the counselor, but he

understands there is some kind of concern there and is seeking further exploration or clarification of it.

Q: "I'm not quite clear as to the difference between an effective response scored 4 and one scored 3."

A: "An effective response scored 3 is one in which the counselor is appropriately with the client and is helping him to express more clearly his feelings. An effective response scored 4 is this also, but in addition it is one which demonstrates rather exact appropriateness and timing such that the response vividly reflects back to the client the essence of what he is feeling."

APPENDIX E

STATISTICAL "t"-TEST ANALYSIS COMPARING THE A-V SCALE DIMENSION SCORES OF THE SCALE DISCRIMINATION STUDY'S FIVE PH.D. CANDIDATES WITH THOSE OF THE FIVE M.A. CANDIDATES

Affect-Cognitive Dimension

Ph.D. Candidate Counselors	Scores	Deviation From Mean	Squared Deviation
(A)	29	-6.2	38.44
(C)	40	4.8	23.04
(D)	33	-2.2	4.84
(G)	36	.8	.64
(J)	38	2.8	7.84
Sum	176		74.80
N	5		
$\bar{X}_1$	35.2	Variance $(S^2{}_1) = \frac{74.8}{5} = 14.6$	

M.A. candidate Counselors	Scores	Deviation From Mean	Squared Deviation
(B)	26	4.8	23.04
(E)	20	-1.2	1.44
(F)	20	-1.2	1.44
(H)	20	-1.2	1.44
(I)	20	-1.2	1.44
Sum	106		28.80
N	5		
$\bar{X}_2$	21.2	Variance $(S^2{}_2) = \frac{28.8}{5} = 5.8$	

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S^2{}_1 + S^2{}_2}{2}\left(\frac{1}{N_1} + \frac{1}{N_2}\right)}} = \frac{35.2 - 21.2}{\sqrt{\frac{14.6 + 5.8}{2}\left(\frac{2}{5}\right)}} = \frac{14.0}{\sqrt{4.1}} = \frac{14.0}{2.04} = \underline{6.9}$$

degrees of freedom (df) = $N_1 + N_2 - 2 = 10 - 2 = 8$

significance at .05 level = 2.31
significance at .01 level = 3.36

Understanding-Nonunderstanding Dimension

Ph.D. Candidate Counselors	Scores	Deviation From Mean	Squared Deviation
(A)	34	-4.2	17.64
(C)	40	1.8	3.24
(D)	40	1.8	3.24
(G)	37	-1.2	1.44
(J)	40	1.8	3.24
Sum	191		28.80
N	5		
$\bar{X}_1$	38.2	Variance $(S^2_1) = \frac{28.8}{5} = 5.8$	

M.A. Candidate Counselors	Scores	Deviation From Mean	Squared Deviation
(B)	24	3.0	9.0
(E)	20	-1.0	1.0
(F)	21	0.0	0.0
(H)	20	-1.0	1.0
(I)	20	-1.0	1.0
Sum	105		12.0
N	5		
$\bar{X}_2$	21.0	Variance $(S^2_2) = \frac{12.0}{5} = 2.4$	

$$t = \frac{38.2 - 21.0}{\sqrt{\frac{5.8 + 2.4}{2}\left(\frac{2}{5}\right)}} = \frac{17.2}{\sqrt{1.6}} = \frac{17.2}{1.28} = 13.4$$

df = 8

significance at .05 level = 2.31
significance at .01 level = 3.36

Specific-Nonspecific Dimension

Ph.D. Candidate Counselors	Scores	Deviation From Mean	Squared Deviation
(A)	27	-8.4	70.56
(C)	40	4.6	21.16
(D)	35	- .4	.16
(G)	36	.6	.36
(J)	39	3.6	12.96
Sum	177		105.20
N	5		
$\bar{X}_1$	35.4	Variance $(S^2_1) = \frac{105.2}{5} = 21.0$	

M.A. Candidate Counselors	Scores	Deviation From Mean	Squared Deviation
(B)	21	.8	.64
(E)	20	-.2	.04
(F)	20	-.2	.04
(H)	20	-.2	.04
(I)	20	-.2	.04
Sum	101		.80
N	5		
$\bar{X}_2$	20.2	Variance $(S^2_2) = \frac{.8}{5} = .2$	

$$t = \frac{35.4 - 20.2}{\sqrt{\frac{21.0 + .2}{2}\left(\frac{2}{5}\right)}} = \frac{15.2}{\sqrt{4.2}} = \frac{15.2}{2.05} = 7.4$$

df = 8

significance at .05 level = 2.31
significance at .01 level = 3.36

Exploratory-Nonexploratory Dimension

Ph.D. Candidate Counselors	Scores	Deviation From Mean	Squared Deviations
(A)	27	-9.4	88.36
(C)	40	3.6	12.96
(D)	37	.6	.36
(G)	38	1.6	2.56
(J)	40	3.6	12.96
Sum	182		117.20
$\bar{X}_1$	36.4	Variance $(S^2{}_1) = \frac{117.2}{5} = 23.4$	

M.A. Candidate Counselors	Scores	Deviation From Mean	Squared Deviations
(B)	21	.8	.64
(E)	20	-.2	.04
(F)	20	-.2	.04
(H)	20	-.2	.04
(I)	20	-.2	.04
Sum	101		.80
N	5		
$\bar{X}_2$	20.2	Variance $(S^2{}_2) = \frac{.8}{5} = .2$	

$$t = \frac{36.4 - 20.2}{\sqrt{\frac{23.4 + .2}{2}\left(\frac{2}{5}\right)}} = \frac{16.2}{\sqrt{4.7}} = \frac{16.2}{2.17} = 7.5$$

df = 8

significance at .05 level = 2.31

significance at .01 level = 3.36

Effective-Noneffective Dimension

Ph.D. Candidate Counselors	Scores	Deviation From Mean	Squared Deviation
(A)	48	-10.8	116.64
(C)	62	3.2	10.24
(D)	58	- .8	.64
(G)	60	1.2	1.44
(J)	66	7.2	51.84
Sum	294		180.80
N	5		
$\bar{X}_1$	58.8	Variance (S^2_1) = $\frac{180.8}{5}$ = 36.1	

M. A. Candidate Counselors	Scores	Deviation From Mean	Squared Deviation
(B)	42	4.9	24.01
(E)	37	- .1	.01
(F)	31	-6.1	37.21
(H)	38	.9	.81
(I)	38	.9	.81
Sum	186		62.85
N	5		
$\bar{X}_2$	37.1	Variance (S^2_2) = $\frac{62.85}{5}$ = 12.4	

$$t = \frac{58.8 - 37.1}{\sqrt{\frac{36.1 + 12.4}{2}\left(\frac{2}{5}\right)}} = \frac{21.7}{\sqrt{9.7}} = \frac{21.7}{3.12} = 6.9$$

df = 8

significance at .05 level = 2.31
significance at .01 level = 3.36

APPENDIX F

STATISTICAL ITEM ANALYSIS OF THE SCORES OF THE SCALE DISCRIMINATION STUDY BASED UPON HOYT'S COMPUTATION OF RELIABILITY BY ANALYSIS OF VARIANCE

Hoyt's Statistical Analysis for Determining Maximum Likelihood Estimate of Reliability (Internal Consistency) and Interjudge Reliability

	df	Sum of Squares	Variance
Counselors (Rows)	R-1	$SS_R = \frac{\sum^{R}(\sum^{C} X)^2}{C} - \frac{(\sum^{R}\sum^{C} X)^2}{RC}$	$S^2_R = SS_{R/R-1}$
Raters (Columns)	C-1	$SS_C = \frac{\sum^{C}(\sum^{R} X)^2}{R} - \frac{(\sum^{R}\sum^{C} X)^2}{RC}$	$S^2_C = SS_{C/C-1}$
Error	(R-1)(C-1)	$SS_E = SS_T - SS_R - SS_C$	$S^2_E = SS_{E/(R-1)(C-1)}$
Total	RC-1	$SS_T = \sum^{R}\sum^{C} X^2 - \frac{(\sum^{R}\sum^{C} X)^2}{RC}$	

(1) Maximum Likelihood Estimate of Reliability (Internal Consistency)

The reliability of the average ratings or the total score on the test.

$$r_{tt} = \frac{SS^2_R - SS^2_E}{SS^2_R}$$

(2) Interjudge Reliability

The average reliability of all the raters or items when between-rater variance is to be removed. This is done when decisions are based on averages of complete sets of ratings from all observers or ratings.

$$r_{tt} = \frac{SS_R - \frac{SS_E}{C-1}}{SS_R + SS_E}$$

(3) Interjudge Reliability

The average reliability of all the raters or items leaving in the between-rater variance. This should be done when decisions about pupils are based on a single score or rating which may be given by different raters.

$$r_{tt} = \frac{SS_R - \frac{SS_C + SS_E}{C-1}}{SS_R + (SS_C + SS_E)}$$

Affect-Cognitive Dimension

Counselor	Response	Raters 1	2	3	Sum
A	1	2	1	1	4
	2	1	1	1	3
	3	2	2	2	6
	4	1	1	1	3
	5	1	1	1	3
	6	2	2	2	6
	7	2	2	2	6
	8	1	1	1	3
	9	1	1	1	3
	10	2	2	2	6
	11	2	2	2	6
	12	2	2	1	5
	13	1	1	1	3
	14	1	1	1	3
	15	2	2	2	6
	16	1	1	1	3
	17	2	2	2	6
	18	2	1	1	4
	19	1	2	1	4
	20	2	2	2	6
B	1	1	2	1	4
	2	1	1	1	3
	3	2	1	1	4
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	2	4
	7	2	2	2	6
	8	2	2	1	5
	9	1	1	1	3
	10	2	2	2	6
	11	1	1	1	3
	12	2	1	1	4
	13	2	1	1	4
	14	2	2	1	5
	15	2	2	1	5
	16	1	2	1	4
	17	2	2	1	5
	18	1	1	1	3
	19	1	1	1	3
	20	1	1	1	3
C	1	2	2	2	6
	2	2	2	2	6
	3	2	2	2	6
	4	2	2	2	6
	5	2	2	2	6

(Aff.-Cog. continued)

Counselor	Response	Raters 1	2	3	Sum
	6	2	2	2	6
	7	2	2	2	6
	8	2	2	2	6
	9	2	2	2	6
	10	2	2	2	6
	11	2	2	2	6
	12	2	2	2	6
	13	2	2	2	6
	14	2	2	2	6
	15	2	2	2	6
	16	2	2	2	6
	17	2	2	2	6
	18	2	2	2	6
	19	2	2	2	6
	20	2	2	2	6
D	1	2	2	2	6
	2	1	1	1	3
	3	2	2	2	6
	4	2	2	2	6
	5	2	2	2	6
	6	1	1	1	3
	7	1	1	1	3
	8	2	2	2	6
	9	2	2	2	6
	10	2	2	2	6
	11	1	2	1	4
	12	1	2	2	5
	13	2	2	2	6
	14	2	2	2	6
	15	2	2	2	6
	16	2	2	2	6
	17	1	2	1	4
	18	1	2	1	4
	19	1	1	1	3
	20	2	2	2	6
E	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	2	1	4
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
	9	1	1	1	3
	10	1	1	1	3

(Aff.=Cog. continued)

Counselor	Response	Raters 1	2	3	Sum
	11	1	1	1	3
	12	1	1	1	3
	13	1	1	1	3
	14	1	1	1	3
	15	1	1	1	3
	16	1	1	1	3
	17	1	1	1	3
	18	1	1	1	3
	19	1	1	1	3
	20	1	1	1	3
F	1	1	1	1	3
	2	1	2	1	4
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
	9	1	1	1	3
	10	1	1	1	3
	11	1	1	1	3
	12	1	1	1	3
	13	1	1	1	3
	14	1	1	1	3
	15	1	1	1	3
	16	1	1	1	3
	17	1	1	1	3
	18	1	1	1	3
	19	1	1	1	3
	20	1	1	1	3
G	1	2	2	2	6
	2	2	2	1	5
	3	1	2	1	4
	4	2	2	2	6
	5	2	2	2	6
	6	2	2	2	6
	7	1	1	1	3
	8	1	2	1	4
	9	1	2	2	5
	10	1	1	1	3
	11	2	2	2	6
	12	2	2	2	6
	13	2	2	2	6
	14	2	2	2	6
	15	2	2	2	6

(Aff.-Cog. continued)

Counselor	Response	Raters 1	2	3	Sum
	16	1	2	2	5
	17	1	2	2	5
	18	2	2	2	6
	19	2	2	2	6
	20	2	2	2	6
H	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
	9	1	1	1	3
	10	1	1	1	3
	11	1	1	1	3
	12	1	1	1	3
	13	1	1	1	3
	14	1	1	1	3
	15	1	1	1	3
	16	1	1	1	3
	17	1	1	1	3
	18	1	1	1	3
	19	1	1	1	3
	20	1	1	1	3
I	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
	9	1	1	1	3
	10	1	1	1	3
	11	1	1	1	3
	12	1	1	1	3
	13	1	1	1	3
	14	1	1	1	3
	15	1	1	1	3
	16	1	1	1	3
	17	1	1	1	3
	18	1	1	1	3
	19	1	1	1	3
	20	1	1	1	3

(Aff.-Cog. continued)

Counselor	Response	Raters 1	2	3	Sum
J	1	2	2	2	6
	2	2	2	2	6
	3	1	2	2	5
	4	2	2	2	6
	5	2	2	2	6
	6	2	2	2	6
	7	2	2	2	6
	8	2	2	2	6
	9	1	2	1	4
	10	2	2	1	5
	11	2	2	2	6
	12	2	2	2	6
	13	2	2	2	6
	14	1	2	2	5
	15	1	1	2	4
	16	2	1	2	5
	17	2	2	2	6
	18	2	2	2	6
	19	2	2	2	6
	20	2	2	1	5
Sum		281	292	276	849
Sum^2		443	476	428	
$(Sum)^2$					240,401

$Sum^2 = 3975$ $(849)^2 = 720,801$

$\Sigma X^2 = 1347$

Source of Variation	df	Sum of Squares (SS)	Variance (S^2)
Counselor Responses	199	$\frac{3975}{3} - \frac{720,801}{600} = 123.7 (SS_R)$	$\frac{123.7}{199} = .62 (S^2_R)$
Raters	2	$\frac{240,401}{200} - \frac{720,801}{600} = .7 (SS_C)$	$\frac{.7}{2} = .35 (S^2_C)$
Error	398	$145.7 - (123.7 + .7) = 21.3 (SS_E)$	$\frac{21.3}{398} = .05$
Total	599	$1347 - \frac{720,801}{600} = 145.7 (SS_T)$	

(1) Internal Consistency $= \frac{.62 - .05}{.62} = .919$

(2) Interjudge Reliability (between-rater variance removed)

$$= \frac{.62 - .05}{.62 + (3-1)(.05)} = .792$$

(3) Interjudge Reliability (between-rater variance left in)

$$= \frac{127.7 - \left(\frac{.7 + 21.3}{2}\right)}{127.7 + (.7 + 21.3)} = .778$$

Understanding-Nonunderstanding Dimension

Counselor	Response	Raters 1	2	3	Sum
A	1	2	1	1	4
	2	2	2	2	6
	3	2	2	2	6
	4	2	1	2	5
	5	1	2	2	5
	6	2	2	2	6
	7	2	2	2	6
	8	2	1	2	5
	9	1	1	1	3
	10	2	2	2	6
	11	2	2	2	6
	12	2	2	1	5
	13	1	1	1	3
	14	1	1	1	3
	15	2	2	2	6
	16	1	1	2	4
	17	2	1	2	5
	18	2	1	2	5
	19	1	2	1	4
	20	1	2	2	5
B	1	2	2	2	6
	2	2	1	1	4
	3	2	1	1	4
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	2	1	4
	8	2	2	1	5
	9	2	1	2	5
	10	2	1	2	5
	11	1	1	1	3
	12	1	1	1	3
	13	1	1	1	3
	14	1	1	1	3
	15	1	1	1	3
	16	1	1	2	4
	17	1	1	1	3
	18	1	1	1	3
	19	1	1	1	3
	20	1	1	1	3
C	1	2	2	2	6
	2	2	2	2	6
	3	2	2	2	6
	4	2	2	2	6
	5	2	2	2	6

(Underst.-Nonunderst. continued)

Counselor	Response	Raters 1	2	3	Sum
	6	2	2	2	6
	7	2	2	2	6
	8	2	2	2	6
	9	2	2	2	6
	10	2	2	2	6
	11	2	2	2	6
	12	2	2	2	6
	13	2	2	2	6
	14	2	2	2	6
	15	2	2	2	6
	16	2	2	2	6
	17	2	2	2	6
	18	2	2	2	6
	19	2	2	2	6
	20	2	2	2	6
D	1	2	2	2	6
	2	2	2	2	6
	3	2	2	2	6
	4	2	2	2	6
	5	2	2	2	6
	6	2	2	2	6
	7	2	1	2	5
	8	2	2	2	6
	9	2	2	2	6
	10	2	2	2	6
	11	2	2	2	6
	12	2	2	2	6
	13	2	2	2	6
	14	2	2	2	6
	15	2	2	2	6
	16	2	2	2	6
	17	1	2	2	5
	18	2	2	2	6
	19	2	1	2	5
	20	2	2	2	6
E	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	2	4
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
	9	1	1	1	3
	10	1	1	1	3

(Underst.-Nonunderst. continued)

Counselor	Response	Raters 1	2	3	Sum
	11	1	1	1	3
	12	1	1	1	3
	13	1	1	1	3
	14	1	1	1	3
	15	1	1	1	3
	16	1	1	1	3
	17	1	1	1	3
	18	1	1	1	3
	19	1	1	1	3
	20	1	1	1	3
F	1	1	1	1	3
	2	2	2	2	6
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
	9	1	1	1	3
	10	1	1	1	3
	11	1	1	1	3
	12	1	1	1	3
	13	1	1	1	3
	14	1	1	1	3
	15	1	1	1	3
	16	1	1	1	3
	17	1	1	1	3
	18	1	1	1	3
	19	1	1	1	3
	20	1	1	1	3
G	1	2	2	2	6
	2	2	2	2	6
	3	2	2	2	6
	4	1	2	1	4
	5	2	2	2	6
	6	2	2	2	6
	7	2	1	2	5
	8	1	2	2	5
	9	2	2	2	6
	10	2	1	2	5
	11	1	2	2	5
	12	2	2	2	6
	13	2	2	2	6
	14	2	2	2	6
	15	2	2	2	6

(Underst.-Nonunderst. continued)

Counselor	Response	Raters 1	2	3	Sum
	16	2	2	2	6
	17	2	2	2	6
	18	2	2	2	6
	19	2	2	2	6
	20	2	2	2	6
II	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	2	1	1	4
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
	9	1	1	1	3
	10	1	1	1	3
	11	1	1	1	3
	12	1	1	1	3
	13	1	1	1	3
	14	1	1	1	3
	15	1	1	1	3
	16	1	1	1	3
	17	1	1	1	3
	18	1	1	1	3
	19	1	1	1	3
	20	1	1	1	3
I	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
	9	1	1	1	3
	10	1	1	1	3
	11	1	1	1	3
	12	1	1	1	3
	13	1	1	1	3
	14	1	1	1	3
	15	1	1	1	3
	16	1	1	1	3
	17	1	1	1	3
	18	1	1	1	3
	19	1	1	1	3
	20	1	1	1	3

(Understd.-Nonunderst. continued)

Counselor	Response	Raters 1	2	3	Sum
J	1	2	2	2	6
	2	2	2	2	6
	3	2	2	2	6
	4	2	2	2	6
	5	2	2	2	6
	6	2	2	2	6
	7	2	2	2	6
	8	2	2	2	6
	9	2	2	2	6
	10	2	2	2	6
	11	2	2	2	6
	12	2	2	2	6
	13	2	2	2	6
	14	2	2	2	6
	15	2	1	2	5
	16	2	2	2	6
	17	2	2	2	6
	18	2	2	2	6
	19	2	2	2	6
	20	2	2	2	6
Sum		297	290	299	886
Sum^2		491	470	497	
$(Sum)^2$					261,710

$Sum^2 = 4318$ $(886)^2 = 784,996$

$\Sigma X^2 = 1458$

Source of Variation	df	Sum of Squares (SS)	Variance (S^2)
Counselor Responses	199	$\frac{4318}{3} - \frac{784,996}{600} = 131.0$	$\frac{131.0}{199} = .658$
Raters	2	$\frac{261,710}{200} - \frac{784,996}{600} = .3$	$\frac{.3}{2} = .15$
Error	398	$1497 - (131.0 + .3) = 18.4$	$\frac{18.4}{398} = .046$
Total	599	$1458 - \frac{784,996}{600} = 149.7$	

(1) Internal Consistency $= \frac{.658 - .046}{.658} = .930$

(2) Interjudge Reliability (between-rater variance removed)

$$= \frac{.658 - .046}{.658 + (3-1)(.046)} = .816$$

(3) Interjudge Reliability (between-rater variance left in)

$$= \frac{131.0 - \left(\frac{.3 + 18.4}{2}\right)}{131.0 + (.3 + 18.4)} = .811$$

Specific-Nonspecific Dimension

Counselor	Response	Raters 1	2	3	Sum
A	1	1	1	1	3
	2	2	2	1	5
	3	2	2	2	6
	4	1	1	2	4
	5	1	1	1	3
	6	2	2	2	6
	7	2	2	2	6
	8	1	1	1	3
	9	1	1	1	3
	10	2	2	2	6
	11	2	2	2	6
	12	1	2	1	4
	13	1	1	1	3
	14	1	1	1	3
	15	2	1	2	5
	16	1	1	1	3
	17	1	1	2	4
	18	1	1	2	4
	19	1	1	1	3
	20	1	1	2	4
B	1	2	1	1	4
	2	2	1	1	4
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
	9	2	1	2	5
	10	1	1	1	3
	11	1	1	1	3
	12	1	1	1	3
	13	1	1	1	3
	14	1	1	1	3
	15	1	1	1	3
	16	1	1	1	3
	17	1	1	1	3
	18	1	1	1	3
	19	1	1	1	3
	20	1	1	1	3
C	1	2	2	2	6
	2	2	2	2	6
	3	2	2	2	6
	4	2	2	2	6
	5	2	2	2	6

(Spec.-Nonspec. continued)

Counselor	Response	Raters 1	2	3	Sum
	6	2	2	2	6
	7	2	2	2	6
	8	2	2	2	6
	9	2	2	2	6
	10	2	2	2	6
	11	2	2	2	6
	12	2	2	2	6
	13	2	2	2	6
	14	2	2	2	6
	15	2	2	2	6
	16	2	2	2	6
	17	2	2	2	6
	18	2	2	2	6
	19	2	2	2	6
	20	2	2	2	6
D	1	2	2	2	6
	2	2	2	1	5
	3	2	2	2	6
	4	2	2	2	6
	5	2	2	2	6
	6	1	2	1	4
	7	1	1	1	3
	8	2	1	1	4
	9	2	2	2	6
	10	2	2	2	6
	11	2	2	2	6
	12	2	2	2	6
	13	2	2	2	6
	14	2	2	2	6
	15	2	2	2	6
	16	2	2	2	6
	17	1	2	1	4
	18	1	2	1	4
	19	2	1	2	5
	20	2	2	2	6
E	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
	9	1	1	1	3
	10	1	1	1	3

(Spec.-Nonspec. continued)

Counselor	Response	Raters 1	2	3	Sum
	11	1	1	1	3
	12	1	1	1	3
	13	1	1	1	3
	14	1	1	1	3
	15	1	1	1	3
	16	1	1	1	3
	17	1	1	1	3
	18	1	1	1	3
	19	1	1	1	3
	20	1	1	1	3
F	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
	9	1	1	1	3
	10	1	1	1	3
	11	1	1	1	3
	12	1	1	1	3
	13	1	1	1	3
	14	1	1	1	3
	15	1	1	1	3
	16	1	1	1	3
	17	1	1	1	3
	18	1	1	1	3
	19	1	1	1	3
	20	1	1	1	3
G	1	2	2	2	6
	2	2	2	1	5
	3	2	2	2	6
	4	1	2	1	4
	5	2	2	2	6
	6	2	1	1	4
	7	1	1	2	4
	8	1	2	2	5
	9	1	2	2	5
	10	2	1	1	4
	11	1	2	2	5
	12	2	2	2	6
	13	1	2	2	5
	14	2	2	2	6
	15	2	2	2	6

(Spec.-Nonspec. continued)

Counselor	Response	Raters 1	2	3	Sum
	16	2	2	2	6
	17	2	2	2	6
	18	2	2	2	6
	19	2	2	2	6
	20	2	2	2	6
H	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
	9	1	1	1	3
	10	1	1	1	3
	11	1	1	1	3
	12	1	1	1	3
	13	1	1	1	3
	14	1	1	1	3
	15	1	1	1	3
	16	1	1	1	3
	17	1	1	1	3
	18	1	1	1	3
	19	1	1	1	3
	20	1	1	1	3
I	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
	9	1	1	1	3
	10	1	1	1	3
	11	1	1	1	3
	12	1	1	1	3
	13	1	1	1	3
	14	1	1	1	3
	15	1	1	1	3
	16	1	1	1	3
	17	1	1	1	3
	18	1	1	1	3
	19	1	1	1	3
	20	1	1	1	3

(Spec.-Nonspec. continued)

Counselor	Response	Raters 1	2	3	Sum
J	1	2	2	2	6
	2	2	2	2	6
	3	2	2	2	6
	4	2	2	2	6
	5	2	2	2	6
	6	2	2	2	6
	7	2	2	2	6
	8	2	2	2	6
	9	1	2	2	5
	10	2	2	2	6
	11	2	2	2	6
	12	2	2	2	6
	13	2	2	2	6
	14	2	2	2	6
	15	1	1	2	4
	16	2	2	2	6
	17	2	2	2	6
	18	2	2	2	6
	19	2	2	2	6
	20	2	2	2	6
Sum_2		278	280	281	839
Sum^2		434	440	443	
$(Sum)^2$					234,645

$Sum^2 = 3897$ $(839)^2 = 703,921$

$\sum X^2 = 1317$

Source of Variation	df	Sum of Squares (SS)	Variance (S^2)
Counselor Responses	199	$\frac{3897}{3} - \frac{703,921}{600} = 125.80$	$\frac{125.8}{199} = .632$
Raters	2	$\frac{234,645}{200} - \frac{703,921}{600} = .03$	$\frac{.03}{2} = .015$
Error	398	$143.8 - (125.8 + .03) = 17.97$	$\frac{17.97}{398} = .045$
Total	599	$1317 - \frac{703,921}{600} = 143.80$	

(1) Internal Consistency $= \frac{.632 - .045}{.632} = .929$

(2) Interjudge Reliability (between-rater variance removed)

$$= \frac{.632 - .045}{.632 + (3-1)(.045)} = .817$$

(3) Interjudge Reliability (between-rater variance left in)

$$= \frac{125.80 - \left(\frac{.03 + 17.97}{2}\right)}{125.80 + (.03 + 17.97)} = .812$$

Exploratory-Nonexploratory Dimension

Counselor	Response	Raters 1	2	3	Sum
A	1	2	1	1	4
	2	2	2	1	5
	3	2	2	2	6
	4	1	1	2	4
	5	1	1	2	4
	6	2	2	2	6
	7	2	2	2	6
	8	1	1	1	3
	9	1	1	1	3
	10	2	2	2	6
	11	2	2	2	6
	12	1	2	1	4
	13	1	1	1	3
	14	1	1	1	3
	15	2	1	2	5
	16	1	1	2	4
	17	1	1	2	4
	18	1	1	2	4
	19	1	1	1	3
	20	1	1	2	4
B	1	2	1	1	4
	2	2	1	1	4
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
	9	2	1	2	5
	10	1	1	2	4
	11	1	1	1	3
	12	1	1	1	3
	13	1	1	1	3
	14	1	1	1	3
	15	1	1	1	3
	16	1	1	2	4
	17	1	1	1	3
	18	1	1	1	3
	19	1	1	1	3
	20	1	1	1	3
C	1	2	1	2	5
	2	2	2	2	6
	3	2	2	2	6
	4	1	2	2	5
	5	2	2	2	6

(Empl.-Nonempl. continued)

Counselor	Response	Raters 1	2	3	Sum
	6	2	2	2	6
	7	2	2	2	6
	8	2	2	2	6
	9	2	2	2	6
	10	2	2	2	6
	11	2	2	2	6
	12	2	2	2	6
	13	2	2	2	6
	14	2	2	2	6
	15	2	2	2	6
	16	2	2	2	6
	17	2	2	2	6
	18	2	2	2	6
	19	2	2	2	6
	20	2	2	2	6
D	1	2	2	2	6
	2	2	2	2	6
	3	2	2	2	6
	4	2	2	2	6
	5	2	2	2	6
	6	1	2	1	4
	7	2	1	1	4
	8	1	1	2	4
	9	2	2	2	6
	10	1	2	2	5
	11	2	2	2	6
	12	2	2	2	6
	13	2	2	2	6
	14	2	2	2	6
	15	2	2	2	6
	16	2	2	2	6
	17	1	2	1	4
	18	1	2	2	5
	19	2	1	2	5
	20	2	2	2	6
E	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
	9	1	1	1	3
	10	1	1	1	3

(Expl.-Nonexpl. continued)

Counselor	Response	Raters 1	2	3	Sum
	11	1	1	1	3
	12	1	1	1	3
	13	1	1	1	3
	14	1	1	1	3
	15	1	1	1	3
	16	1	1	1	3
	17	1	1	1	3
	18	1	1	1	3
	19	1	1	1	3
	20	1	1	1	3
F	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
	9	1	1	1	3
	10	1	1	1	3
	11	1	1	1	3
	12	1	1	1	3
	13	1	1	1	3
	14	1	1	1	3
	15	1	1	1	3
	16	1	1	1	3
	17	1	1	1	3
	18	1	1	1	3
	19	1	1	1	3
	20	1	1	1	3
G	1	2	2	2	6
	2	2	2	2	6
	3	2	2	2	6
	4	1	2	1	4
	5	2	2	2	6
	6	2	1	2	5
	7	1	1	2	4
	8	1	2	2	5
	9	1	2	2	5
	10	2	1	2	5
	11	1	2	2	5
	12	2	2	2	6
	13	1	2	2	5
	14	2	2	2	6
	15	2	2	2	6

(Expl.-Nonexpl. continued)

Counselor	Response	Raters 1	2	3	Sum
	16	2	2	2	6
	17	2	2	2	6
	18	2	2	2	6
	19	2	2	2	6
	20	2	2	2	6
H	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
	9	1	1	1	3
	10	1	1	1	3
	11	1	1	1	3
	12	1	1	1	3
	13	1	1	1	3
	14	1	1	1	3
	15	1	1	1	3
	16	1	1	1	3
	17	1	1	1	3
	18	1	1	1	3
	19	1	1	1	3
	20	1	1	1	3
I	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
	9	1	1	1	3
	10	1	1	1	3
	11	1	1	1	3
	12	1	1	1	3
	13	1	1	1	3
	14	1	1	1	3
	15	1	1	1	3
	16	1	1	1	3
	17	1	1	1	3
	18	1	1	1	3
	19	1	1	1	3
	20	1	1	1	3

(Expl.-Nonexpl. continued)

Counselor	Response	Raters 1	2	3	Sum
J	1	2	2	2	6
	2	2	2	2	6
	3	2	1	2	5
	4	2	2	2	6
	5	2	2	2	6
	6	2	2	2	6
	7	2	2	2	6
	8	2	2	2	6
	9	1	2	2	5
	10	2	2	2	6
	11	2	2	2	6
	12	1	2	2	5
	13	2	2	2	6
	14	2	2	2	6
	15	1	1	2	4
	16	2	2	2	6
	17	2	2	2	6
	18	2	2	2	6
	19	2	2	2	6
	20	2	2	2	6
Sum		276	278	291	845
Sum^2		428	434	473	
$(Sum)^2$					238,141

$Sum^2 = 3933$ $(845)^2 = 714,025$

$\Sigma X^2 = 1335$

Source of Variation	df	Sum of Squares (SS)	Variance (S^2)
Counselor Responses	199	$\frac{3933}{3} - \frac{714,025}{600} = 121.0$	$\frac{121.0}{199} = .608$
Raters	2	$\frac{238,141}{200} - \frac{714,025}{600} = .70$	$\frac{.70}{2} = .350$
Error	398	$145 - (121.0 - .7) = 23.3$	$\frac{23.3}{398} = .059$
Total	599	$\frac{1335 - 714,025}{600} = 145$	

(1) Internal Consistency $= \frac{.608 - .059}{.608} = .903$

(2) Interjudge Reliability (between rater variance removed)

$$= \frac{.608 - .059}{.608 + (3-1)(.059)} = .756$$

(3) Interjudge Reliability (between rater variance left in)

$$= \frac{121 - \left(\frac{.7 + 23.3}{2}\right)}{121 + (.7 + 23.3)} = .753$$

Effective-Noneffective Dimension

Counselor	Response	Raters 1	2	3	Sum
A	1	2	2	2	6
	2	3	2	2	7
	3	3	2	3	8
	4	2	2	3	7
	5	2	2	3	7
	6	3	3	4	10
	7	3	3	3	9
	8	2	2	2	6
	9	2	2	2	6
	10	3	3	3	9
	11	3	3	3	9
	12	2	2	2	6
	13	2	1	2	5
	14	2	1	2	5
	15	3	3	3	9
	16	2	2	3	7
	17	3	2	3	8
	18	2	2	3	7
	19	2	2	2	6
	20	2	2	3	7
B	1	3	3	2	8
	2	3	2	2	7
	3	2	2	2	6
	4	2	2	2	6
	5	2	2	2	6
	6	2	2	2	6
	7	2	3	2	7
	8	2	3	1	6
	9	3	2	3	8
	10	3	2	3	8
	11	2	2	2	6
	12	2	2	2	6
	13	2	2	2	6
	14	2	2	2	6
	15	2	2	2	6
	16	2	2	3	7
	17	2	2	2	6
	18	1	2	2	5
	19	1	2	2	5
	20	1	2	2	5
C	1	3	3	4	10
	2	3	3	4	10
	3	3	3	4	10
	4	3	3	3	9
	5	3	3	4	10

(Eff.-Noneff. continued)

Counselor	Response	Raters 1	2	3	Sum
	6	3	4	3	10
	7	4	4	4	12
	8	3	3	3	9
	9	4	3	4	11
	10	3	3	3	9
	11	4	3	4	11
	12	4	4	4	12
	13	3	3	4	10
	14	3	3	3	9
	15	3	3	4	10
	16	3	3	3	9
	17	3	3	4	10
	18	3	3	4	10
	19	3	3	4	10
	20	3	4	3	10
D	1	3	3	3	9
	2	3	3	3	9
	3	3	3	3	9
	4	3	4	3	10
	5	3	4	3	10
	6	2	2	2	6
	7	3	2	2	7
	8	2	2	3	7
	9	3	3	3	9
	10	3	3	3	9
	11	3	3	3	9
	12	3	3	4	10
	13	4	4	4	12
	14	4	4	3	11
	15	3	3	3	9
	16	3	3	3	9
	17	2	2	2	6
	18	2	3	3	8
	19	3	2	3	8
	20	3	3	3	9
E	1	1	2	2	5
	2	2	2	2	6
	3	2	2	2	6
	4	2	2	2	6
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	2	2	2	6
	9	2	2	2	6
	10	2	2	2	6

(Eff.-Noneff. continued)

Counselor	Response	Raters 1	2	3	Sum
	11	2	2	2	6
	12	2	2	2	6
	13	2	2	2	6
	14	2	2	2	6
	15	2	2	2	6
	16	2	2	2	6
	17	2	2	2	6
	18	2	2	2	6
	19	2	2	2	6
	20	2	2	2	6
F	1	1	1	2	4
	2	1	2	2	5
	3	1	2	1	4
	4	1	2	1	4
	5	1	2	1	4
	6	1	2	1	4
	7	1	2	1	4
	8	1	2	2	5
	9	2	2	2	6
	10	2	1	2	5
	11	2	1	1	4
	12	2	1	1	4
	13	2	2	1	5
	14	2	2	1	5
	15	2	2	2	6
	16	2	1	2	5
	17	2	1	2	5
	18	2	2	2	6
	19	2	2	1	5
	20	1	2	1	4
G	1	3	3	3	9
	2	3	3	3	9
	3	3	3	3	9
	4	2	3	2	7
	5	3	3	3	9
	6	3	2	3	8
	7	2	2	3	7
	8	2	3	3	8
	9	3	3	3	9
	10	3	2	3	8
	11	2	3	3	8
	12	4	4	4	12
	13	3	4	3	10
	14	4	3	3	10
	15	3	3	3	9

(Eff.-Noneff. continued)

Counselor	Response	Raters 1	2	3	Sum
	16	3	3	3	9
	17	3	3	3	9
	18	3	3	3	9
	19	3	3	3	9
	20	3	3	3	9
H	1	2	2	2	6
	2	2	1	2	5
	3	2	1	2	5
	4	2	2	2	6
	5	2	2	2	6
	6	2	2	1	5
	7	1	2	1	4
	8	2	2	1	5
	9	1	1	1	3
	10	2	2	2	6
	11	2	2	2	6
	12	2	2	2	6
	13	2	2	2	6
	14	2	2	2	6
	15	2	2	2	6
	16	2	2	2	6
	17	2	2	2	6
	18	2	2	2	6
	19	2	2	2	6
	20	2	1	2	5
I	1	2	2	2	6
	2	2	1	1	4
	3	1	1	2	4
	4	1	1	1	3
	5	1	1	1	3
	6	2	2	2	6
	7	2	2	1	5
	8	1	2	2	5
	9	2	1	1	4
	10	2	2	2	6
	11	2	2	2	6
	12	2	2	2	6
	13	2	2	2	6
	14	2	1	2	5
	15	2	1	2	5
	16	2	2	2	6
	17	2	2	2	6
	18	2	2	2	6
	19	2	2	2	6
	20	2	1	2	5

(Eff.-Noneff. continued)

Counselor	Response	Raters 1	2	3	Sum
J	1	4	3	3	10
	2	4	4	3	11
	3	3	3	3	9
	4	4	3	3	10
	5	4	4	4	12
	6	3	3	4	10
	7	3	4	4	11
	8	3	4	4	11
	9	3	3	3	9
	10	3	3	3	9
	11	3	4	3	10
	12	3	3	3	9
	13	4	3	4	11
	14	3	3	3	9
	15	2	2	3	7
	16	3	3	3	9
	17	3	4	4	11
	18	4	4	3	11
	19	4	4	4	12
	20	3	3	3	9
Sum		476	474	491	1441
Sum^2		1254	1252	1351	
$(Sum)^2$					692,333

$Sum^2 = 11381$ $(1441)^2 = 2076481$

$\Sigma X^2 = 3857$

Source of Variation	df	Sum of Squares (SS)	Variance (S^2)
Counselor Responses	199	$\frac{11381}{3} - \frac{2,076,481}{600} = 332.9$	$\frac{332.9}{199} = 1.672$
Raters	2	$\frac{692,333}{200} - \frac{2,076,481}{600} = .9$	$\frac{.9}{2} = .450$
Error	398	$396.2 - (332.9 + .9) = 62.4$	$\frac{62.4}{398} = .157$
Total	599	$3857 - \frac{2,076,481}{600} = 396.2$	

(1) Internal Consistency $= \frac{1.672 - .157}{1.672} = .906$

(2) Interjudge Reliability (between-rater variance removed)

$$= \frac{1.672 - .157}{1.672 + (3-1)(.157)} = .763$$

(3) Interjudge Reliability (between-rater variance left in)

$$= \frac{332.9 - \frac{(.9 + 62.4)}{2}}{332.9 + (.9 + 62.4)} = .760$$

APPENDIX G

STATISTICAL ITEM ANALYSIS OF THE SCORES OF THE OBJECTIVITY STUDY BASED UPON HOYT'S COMPUTATION OF RELIABILITY BY ANALYSIS OF VARIANCE

Affect-Cognitive Dimension

Counselor	Response	Raters A	B	C	Sum
1	1	2	2	2	6
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
2	1	2	2	2	6
	2	2	2	2	6
	3	2	2	2	6
	4	2	2	2	6
	5	2	1	2	5
	6	2	2	2	6
	7	2	2	2	6
3	1	1	1	1	3
	2	1	1	1	3
	3	1	1	2	4
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
	9	1	1	1	3
4	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
5	1	1	1	1	3
	2	1	1	1	3
	3	2	2	2	6
	4	2	2	2	6
	5	2	1	2	5
6	1	2	1	2	5
	2	1	1	2	4
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
7	1	1	1	1	3
	2	1	1	1	3
	3	2	2	2	6
	4	2	2	2	6
	5	2	2	2	6
8	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3

(Aff.-Cog. continued)

Counselor	Response	Raters A	B	C	Sum
9	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
10	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
11	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
12	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
13	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
14	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
15	1	1	1	1	3
	2	2	1	2	5
	3	2	2	2	6
	4	1	2	2	5
16	1	1	1	1	3
	2	1	1	1	3
	3	1	1	2	4
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
17	1	1	1	1	3
	2	1	1	1	3
	3	2	2	2	6

(Aff.-Cog. continued)

Counselor	Response	Raters A	B	C	Sum
	4	1	1	2	4
	5	2	1	2	5
	6	1	1	1	3
	7	1	1	1	3
	8	2	2	2	6
18	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
19	1	2	2	2	6
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	2	2	2	6
20	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
	9	1	1	1	3
	10	1	1	1	3
21	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	2	2	2	6
22	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
23	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
	9	1	1	1	3
	10	1	1	1	3
24	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3

(Aff.-Cog. continued)

Counselor	Response	Raters A	B	C	Sum
25	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
26	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
27	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
28	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
29	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
30	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
31	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
32	1	2	2	2	6
	2	2	2	2	6
	3	2	2	2	6
	4	2	2	2	6
33	1	2	1	2	5
	2	2	2	2	6
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3

(Aff.-Cog. continued)

Counselor	Response	Raters A	B	C	Sum
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
34	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
35	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
	9	1	1	1	3
	10	1	1	1	3
36	1	2	2	2	6
	2	2	1	2	5
	3	1	1	1	3
37	1	1	1	2	4
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	2	4
38	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	2	2	2	6
	5	1	1	1	3
	6	1	1	1	3
39	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	2	4
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
40	1	1	1	1	3
	2	1	1	1	3
	3	2	2	1	5
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3

(Aff.-Cog. continued)

Counselor	Response	Raters A	B	C	Sum
41	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
42	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
43	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
44	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	2	2	2	6
	6	2	2	2	6
45	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
46	1	1	1	1	3
	2	2	2	2	6
	3	2	2	2	6
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	2	4
47	1	1	2	2	5
	2	2	2	2	6
	3	2	2	2	6
	4	1	1	1	3
48	1	1	1	1	3
	2	1	1	1	3
	3	2	2	2	6
	4	2	2	2	6
	5	2	2	2	6

(Aff.-Cog. continued)

Counselor	Response	Raters A	B	C	Sum		
49	1	1	1	1	3		
	2	1	1	1	3		
	3	1	1	1	3		
	4	1	1	1	3		
	5	2	2	2	6		
50	1	2	2	2	6		
	2	2	2	2	6		
	3	1	2	2	5		
	4	2	1	2	5		
	5	2	2	2	6		
	6	1	1	2	4		
	7	2	2	2	6		
51	1	1	2	2	5		
	2	2	2	2	6		
	3	2	2	2	6		
	4	1	1	1	3		
	5	2	2	2	6		
	6	2	2	2	6		
	7	1	1	1	3		
52	1	1	1	1	3		
	2	1	1	1	3		
	3	2	2	2	6		
	4	2	2	2	6		
	5	2	2	2	6		
53	1	1	1	1	3		
	2	1	1	1	3		
	3	1	1	1	3		
	4	2	2	2	6		
	5	2	2	2	6		
	6	2	2	2	6		
Sum		371	367	383	1121	$Sum^2=4321$	$(1121)^2=1256641$
Sum^2		487	475	523			$\sum X^2=1485$
$(Sum)^2$					419,019		

(Aff.-Cog. continued)

Analysis of Variance

Source of Variation	df	Sum of Squares (SS)	Variance (S^2)
Counselor Responses	312	$\frac{4321}{3} - \frac{1,256,641}{939} = 102.0(SS_R)$	$\frac{102.0}{312} = .330(S^2{}_R)$
Raters	2	$\frac{419019}{313} - \frac{1,256,641}{939} = 6.8(SS_C)$	$\frac{6.8}{2} = 3.4(S^2{}_C)$
Error	624	$146.7 - (132+6.8) = 7.9(SS_E)$	$\frac{7.9}{624} = .013$
Total	938	$1485 - \frac{1,256,641}{939} = 146.7(SS_T)$	

(1) Internal Consistency $= \frac{.330 - .013}{.330} = \frac{.317}{.330} = .961$

(2) Interjudge Reliability (between-rater variance removed)

$$= \frac{.330 - .013}{.330 + (3-1)(.013)} = .890$$

(3) Interjudge Reliability (between-rater variance left in)

$$= \frac{132.0 - \frac{(6.8 + 7.9)}{2}}{132.0 + (6.8 + 7.9)} = .850$$

Understanding-Nonunderstanding Dimension

Counselor	Response	Raters A	B	C	Sum
1	1	2	2	2	6
	2	1	1	1	3
	3	2	2	2	6
	4	1	1	1	3
	5	2	1	1	4
	6	1	1	1	3
2	1	2	2	2	6
	2	2	2	2	6
	3	2	2	2	6
	4	2	2	2	6
	5	2	1	2	5
	6	2	2	2	6
	7	2	2	2	6
3	1	1	1	2	4
	2	1	1	1	3
	3	2	1	2	5
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
	9	1	1	1	3
4	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
5	1	1	1	1	3
	2	1	1	1	3
	3	2	2	2	6
	4	2	2	2	6
	5	2	1	2	5
6	1	2	2	2	6
	2	1	1	2	4
	3	1	1	1	3
	4	1	1	2	4
	5	1	1	1	3
	6	1	1	1	3
7	1	1	1	1	3
	2	1	1	1	3
	3	2	2	2	6
	4	2	1	1	4
	5	2	2	2	6
8	1	1	1	1	3
	2	1	1	1	3
	3	2	2	2	6
	4	1	1	1	3
	5	1	1	1	3
	6	2	1	1	4

(Underst.-Nonunderst. continued)

Counsellor	Response	Raters A	B	C	Sum
9	1	1	1	1	3
	2	2	1	1	4
	3	1	1	1	3
	4	1	1	1	3
10	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
11	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
12	1	2	2	2	6
	2	2	2	2	6
	3	2	2	2	6
13	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
14	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
15	1	1	1	1	3
	2	2	2	2	6
	3	2	2	2	6
	4	1	2	2	5
16	1	1	1	1	3
	2	1	1	1	3
	3	1	2	2	5
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
17	1	1	1	2	4
	2	2	2	1	5
	3	2	2	2	6

(Underst.-Nonunderst. continued)

Counselor	Response	Raters A	B	C	Sum
	4	1	1	2	4
	5	2	1	2	5
	6	1	1	1	3
	7	1	1	1	3
	8	2	2	2	6
18	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
19	1	2	2	2	6
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	2	2	2	6
20	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
	9	1	1	1	3
	10	1	1	1	3
21	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	2	2	2	6
22	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
23	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
	9	1	1	1	3
	10	1	1	1	3
24	1	1	1	1	3
	2	1	1	1	3
	3	2	2	2	6
	4	2	1	1	4

(Underst.-Nonunderst. continued)

Counselor	Response	Raters A	B	C	Sum
25	1	1	1	2	4
	2	2	1	1	4
	3	2	2	2	6
	4	1	1	1	3
	5	2	1	1	4
26	1	2	1	2	5
	2	1	1	1	3
	3	1	1	2	4
27	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
28	1	1	1	2	4
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
29	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
30	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
31	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
32	1	2	2	2	6
	2	2	2	2	6
	3	2	2	2	6
	4	2	2	2	6
33	1	2	2	2	6
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3

(Underst.-Nonunderst. continued)

Counselor	Response	Raters A	B	C	Sum
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	2	4
34	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
35	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
	9	1	1	1	3
	10	1	1	1	3
36	1	2	2	2	6
	2	2	2	2	6
	3	1	1	1	3
37	1	1	1	2	4
	2	1	1	2	4
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
38	1	1	1	1	3
	2	1	1	1	3
	3	2	2	2	6
	4	2	2	2	6
	5	1	1	1	3
	6	1	1	1	3
39	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	2	2	5
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
40	1	1	2	1	4
	2	1	1	1	3
	3	2	2	1	5
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3

(Underst.-Nonunderst. continued)

Counselor	Response	Raters A	B	C	Sum
41	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
42	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
43	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
44	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	2	2	2	6
	5	2	2	2	6
	6	2	2	2	6
45	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
46	1	1	1	1	3
	2	2	2	2	6
	3	2	2	2	6
	4	2	1	1	4
	5	1	1	1	3
	6	1	1	2	4
47	1	2	2	2	6
	2	2	2	2	6
	3	2	2	2	6
	4	1	1	2	4
48	1	1	1	1	3
	2	1	1	1	3
	3	2	2	2	6
	4	2	2	2	6
	5	2	2	2	6

(Underst.-Nonunderst. continued)

Counselor	Response	Raters A	B	C	Sum		
49	1	1	1	1	3		
	2	2	1	1	4		
	3	2	2	2	6		
	4	1	1	1	3		
	5	2	2	2	6		
50	1	2	2	2	6		
	2	2	2	2	6		
	3	2	2	2	6		
	4	2	2	2	6		
	5	2	2	2	6		
	6	2	2	2	6		
	7	2	2	2	6		
51	1	1	2	2	5		
	2	2	2	2	6		
	3	2	2	2	6		
	4	1	1	1	3		
	5	2	2	2	6		
	6	2	2	2	6		
	7	1	1	2	4		
52	1	1	1	2	4		
	2	1	1	2	4		
	3	2	2	2	6		
	4	2	2	2	6		
	5	2	2	2	6		
53	1	1	1	2	4		
	2	1	1	1	3		
	3	1	1	1	3		
	4	2	2	2	6		
	5	2	2	2	6		
	6	2	2	2	6		
Sum		394	385	404	1183	$Sum^2=4937$	$(1183)^2=1399489$
Sum^2		556	529	586			$\Sigma X^2=1671$
$(Sum)^2$					466,677		

(Underst.-Nonunderst. continued)

Analysis of Variance

Source of Variation	df	Sum of Squares (SS)	Variance (S^2)
Counselor Responses	312	$\frac{4937}{3} - \frac{1,399,489}{939} = 155.27$	$\frac{155.27}{312} = .498$
Raters	2	$\frac{466,677}{313} - \frac{1,399,489}{939} = .56$	$\frac{.56}{2} = .280$
Error	624	$180.60 - (155.27 + .56) = 24.77$	$\frac{24.77}{624} = .040$
Total	938	$1671 - \frac{1,399,489}{939} = 180.60$	

(1) Internal Consistency $= \frac{.498 - .040}{.498} = \frac{.458}{.498} = .919$

(2) Interjudge Reliability (between-rater variane removed)

$$= \frac{.498 - .040}{.498 + (3-1)(.040)} = .792$$

(3) Interjudge Reliability (between-rater variance left in)

$$= \frac{155.27 - \frac{(.56 + 24.7)}{2}}{155.27 + (.56 + 24.7)} = .790$$

Specific-Nonspecific Dimension

Counselor	Response	Raters A	B	C	Sum
1	1	1	1	1	3
	2	1	1	2	4
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
2	1	1	1	2	5
	2	2	2	2	6
	3	2	2	1	5
	4	2	2	2	6
	5	1	1	1	3
	6	2	2	2	6
	7	2	2	2	6
3	1	1	1	2	4
	2	1	1	1	3
	3	2	1	2	5
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
	9	1	1	1	3
4	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
5	1	1	1	1	3
	2	1	1	1	3
	3	2	2	2	6
	4	2	2	2	6
	5	1	1	2	4
6	1	2	2	2	6
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
7	1	1	1	1	3
	2	1	1	1	3
	3	2	2	2	6
	4	1	1	1	3
	5	1	2	2	5
8	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3

(Spec.-Nonspec. continued)

Counselor	Response	Raters A	B	C	Sum
9	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
10	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
11	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
12	1	2	2	2	6
	2	2	2	2	6
	3	2	2	2	6
13	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
14	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
15	1	1	1	1	3
	2	2	1	2	5
	3	2	2	2	6
	4	1	1	2	4
16	1	1	1	1	3
	2	1	1	1	3
	3	1	1	2	4
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
17	1	1	1	1	3
	2	1	1	1	3
	3	2	2	2	6

(Spec.-Nonspec. continued)

Counselor	Response	Raters A	B	C	Sum
	4	1	1	2	4
	5	2	1	2	5
	6	1	1	1	3
	7	1	1	1	3
	8	2	2	2	6
18	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
19	1	2	2	2	6
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	2	2	2	6
20	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
	9	1	1	1	3
	10	1	1	1	3
21	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
22	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
23	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
	9	1	1	1	3
	10	1	1	1	3
24	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3

(Spec.-Nonspec. continued)

Counselor	Response	Raters A	B	C	Sum
25	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
26	1	1	1	1	3
	2	1	1	1	3
	3	1	1	2	4
27	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
28	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
29	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
30	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
31	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
32	1	2	2	2	6
	2	2	2	2	6
	3	2	2	2	6
	4	2	2	2	6
33	1	2	2	2	6
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3

(Spec.-Nonspec. continued)

Counselor	Response	Raters A	B	C	Sum
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
34	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
35	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
	9	1	1	1	3
	10	1	1	1	3
36	1	1	2	2	5
	2	1	2	2	5
	3	1	1	1	3
37	1	1	1	2	4
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
38	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	2	2	2	6
	5	1	1	1	3
	6	1	1	1	3
39	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	2	4
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
40	1	1	1	1	3
	2	1	1	1	3
	3	1	2	1	4
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3

(Spec.-Nonspec. continued)

Counselor	Response	Raters A	B	C	Sum
41	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
42	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
43	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
44	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	2	4
	6	1	1	2	4
45	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
46	1	1	1	1	3
	2	2	2	2	6
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
47	1	1	2	2	5
	2	2	2	2	6
	3	2	2	2	6
	4	1	1	1	3
48	1	1	1	1	3
	2	1	1	1	3
	3	2	2	2	6
	4	2	2	2	6
	5	1	1	1	3

(Spec.-Nonspec. continued)

Counselor	Response	Raters A	B	C	Sum		
49	1	1	1	1	3		
	2	1	1	1	3		
	3	1	1	1	3		
	4	1	1	1	3		
	5	1	2	2	5		
50	1	2	1	2	5		
	2	1	1	2	4		
	3	1	1	2	4		
	4	1	1	2	4		
	5	2	2	2	6		
	6	2	2	2	6		
	7	2	2	2	6		
51	1	1	2	2	5		
	2	2	1	2	5		
	3	2	2	2	6		
	4	1	1	1	3		
	5	2	2	2	6		
	6	2	2	2	6		
	7	1	1	1	3		
52	1	1	1	2	4		
	2	1	1	2	4		
	3	2	2	2	6		
	4	2	2	2	6		
	5	2	2	2	6		
53	1	1	1	2	4		
	2	1	1	1	3		
	3	1	1	1	3		
	4	1	2	2	5		
	5	2	2	2	6		
	6	1	1	2	4		
Sum		356	360	381	1097	$Sum^2=4173$	$(1097)^2=1203409$
Sum^2		442	454	517			$\sum X^2=1413$
$(Sum)^2$					401,497		

(Spec.-Nonspec. continued)

Analysis of Variance

Source of Variation	df	Sum of Squares (SS)	Variance
Counselor Responses	312	$\frac{4173}{3} - \frac{1,203,409}{939} = 109.41$	$\frac{109.41}{312} = .351$
Raters	2	$\frac{401,497}{313} - \frac{1,203,409}{939} = 1.15$	$\frac{1.15}{2} = .575$
Error	624	$131.41 - (109.41 + 1.15) = 20.85$	$\frac{20.85}{624} = .033$
Total	938	$1413 - \frac{1,203,409}{939} = 131.41$	

(1) Internal Consistency $= \frac{.351 - .033}{.351} = .906$

(2) Interjudge Reliability (between-rater variance removed)

$$= \frac{.351 - .033}{.351 + (3-1)(.033)} = .763$$

(3) Interjudge Reliability (between-rater variance left in)

$$= \frac{109.41 - \left(\frac{1.15 + 20.85}{2}\right)}{109.41 + (1.15 + 20.85)} = .749$$

Exploratory-Nonexploratory Dimension

Counselor	Response	Raters A	B	C	Sum
1	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	2	4
	6	1	1	1	3
2	1	2	2	2	6
	2	2	2	2	6
	3	2	1	2	5
	4	2	2	2	6
	5	1	1	1	3
	6	1	1	1	3
	7	1	2	2	5
3	1	1	2	2	5
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
	9	1	1	1	3
4	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
5	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	2	2	5
	5	1	1	2	3
6	1	2	2	2	6
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
7	1	1	1	1	3
	2	1	1	1	3
	3	2	2	2	6
	4	1	1	1	3
	5	2	2	2	6
8	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	2	1	1	4

(Expl.-Nonexpl. continued)

Counselor	Response	Raters A	B	C	Sum
9	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	2	4
10	1	1	1	2	4
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
11	1	1	1	2	4
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
12	1	1	1	1	3
	2	2	2	2	6
	3	1	1	1	3
13	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
14	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
15	1	1	1	1	3
	2	1	1	1	3
	3	2	2	2	6
	4	1	1	1	3
16	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	2	4
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
17	1	1	1	1	3
	2	2	2	2	6
	3	2	2	2	6

(Expl.-Nonexpl. continued)

Counselor	Response	Raters A	B	C	Sum
	4	1	1	2	4
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	2	1	2	5
18	1	1	1	2	4
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
19	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
20	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
	9	1	1	1	3
	10	1	1	1	3
21	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
22	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
23	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
	9	1	1	1	3
	10	1	1	1	3
24	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3

(Expl.-Nonexpl. continued)

Counselor	Response	Raters A	B	C	Sum
25	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
26	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
27	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
28	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
29	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
30	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
31	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
32	1	2	2	2	6
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	2	4
33	1	1	2	2	5
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3

(Expl.-Nonexpl. continued)

Counselor	Response	Raters A	B	C	Sum
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
34	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
35	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
	9	1	1	1	3
	10	1	1	1	3
36	1	1	2	2	5
	2	1	1	2	4
	3	1	1	1	3
37	1	1	1	2	4
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	2	4
38	1	1	1	1	3
	2	1	1	1	3
	3	2	2	2	6
	4	2	2	2	6
	5	1	1	1	3
	6	1	1	1	3
39	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
40	1	1	1	1	3
	2	1	1	1	3
	3	2	2	1	5
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3

(Empl.-Nonempl. continued)

Counselor	Response	Raters A	B	C	Sum
41	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
42	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
43	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
44	1	1	1	2	4
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	2	4
	6	1	1	1	3
45	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
46	1	1	1	1	3
	2	2	1	1	4
	3	2	2	2	6
	4	2	1	1	4
	5	1	1	1	3
	6	1	1	2	4
47	1	2	2	2	6
	2	1	2	2	5
	3	2	2	2	6
	4	1	1	2	4
48	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3

(Expl.-Nonexpl. continued)

Counselor	Response	Raters A	B	C	Sum		
49	1	1	1	1	3		
	2	1	1	1	3		
	3	1	1	2	4		
	4	1	1	1	3		
	5	1	1	1	3		
50	1	1	1	2	4		
	2	1	1	2	4		
	3	1	1	2	4		
	4	1	1	1	3		
	5	1	1	2	4		
	6	1	1	2	4		
	7	1	2	2	5		
51	1	1	2	2	5		
	2	2	2	2	6		
	3	1	1	2	4		
	4	1	1	1	3		
	5	2	2	2	6		
	6	2	2	2	6		
	7	1	1	2	4		
52	1	1	1	2	4		
	2	1	1	2	4		
	3	2	2	2	6		
	4	2	2	2	6		
	5	1	1	2	4		
53	1	1	1	1	3		
	2	1	1	1	3		
	3	1	1	1	3		
	4	1	2	2	5		
	5	1	1	1	3		
	6	1	1	1	3		
Sum		340	344	372	1056	$Sum^2=3782$	$(1056)^2=1115136$
Sum^2		394	406	490			$\sum X^2=1290$
$(Sum)^2$					372,320		

(Expl.-Nonexpl. continued)

Analysis of Variance

Source of Variation	df	Sum of Squares (SS)	Variance (S^2)
Counselor Responses	312	$\frac{3782}{3} - \frac{1,115,136}{939} = 73.09$	$\frac{73.09}{312} = .234$
Raters	2	$\frac{372,320}{313} - \frac{1,115,136}{939} = 1.94$	$\frac{1.94}{2} = .970$
Error	624	$102.42 - (73.09 + 1.94) = 27.39$	$\frac{27.39}{624} = .044$
Total	938	$1290 - \frac{1,115,136}{939} = 102.42$	

(1) Internal Consistency $= \frac{.234 - .044}{.234} = .812$

(2) Interjudge Reliability (between-rater variance removed)

$$= \frac{.234 - .044}{.234 + (3-1)(.044)} = .590$$

(3) Interjudge Reliability (between-rater variance left in)

$$= \frac{73.09 - \left(\frac{1.94 + 27.39}{2}\right)}{73.09 + (1.94 + 27.39)} = .570$$

Effective-Noneffective Dimension

Counselor	Response	Raters A	B	C	Sum
1	1	3	3	3	9
	2	2	2	2	6
	3	2	2	2	6
	4	2	2	2	6
	5	2	1	2	5
	6	1	1	1	3
2	1	3	3	3	9
	2	3	3	3	9
	3	3	2	3	8
	4	3	3	4	10
	5	2	1	2	5
	6	3	3	3	9
	7	3	3	3	9
3	1	2	2	2	6
	2	1	2	1	4
	3	2	1	3	6
	4	1	1	1	3
	5	1	1	1	3
	6	2	1	2	5
	7	1	1	1	3
	8	1	1	1	3
	9	1	1	1	3
4	1	1	1	1	3
	2	1	1	1	3
	3	2	2	2	6
	4	2	1	1	4
5	1	2	1	2	5
	2	1	1	2	4
	3	3	3	3	9
	4	3	3	3	9
	5	2	2	3	7
6	1	3	3	3	9
	2	2	2	2	6
	3	2	2	2	6
	4	2	2	2	6
	5	2	1	1	4
	6	2	1	1	4
7	1	2	2	2	6
	2	1	1	1	3
	3	3	3	3	9
	4	2	2	2	6
	5	3	3	3	9
8	1	1	1	1	3
	2	1	1	1	3
	3	2	2	1	5
	4	1	1	1	3
	5	1	1	1	3
	6	2	1	1	4

(Eff.-Noneff. continued)

Counselor	Response	Raters A	B	C	Sum
9	1	2	2	2	6
	2	2	1	1	4
	3	2	2	2	6
	4	2	1	1	4
10	1	2	2	2	6
	2	1	1	1	3
	3	2	2	1	5
	4	2	1	1	4
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
11	1	2	1	2	5
	2	2	1	1	4
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
12	1	3	3	4	10
	2	4	4	3	11
	3	3	3	3	9
13	1	2	1	1	4
	2	1	1	1	3
	3	1	1	1	3
	4	2	1	1	4
14	1	1	1	1	3
	2	1	1	1	3
	3	2	1	1	4
	4	2	1	2	5
	5	1	1	1	3
	6	1	1	1	3
15	1	2	1	1	4
	2	3	2	3	8
	3	3	3	3	9
	4	2	2	3	7
16	1	1	1	2	4
	2	1	1	1	3
	3	2	2	3	7
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
17	1	2	2	2	6
	2	2	2	2	6
	3	3	3	3	9

(Eff.-Noneff. continued)

Counselor	Response	Raters A	B	C	Sum
	4	2	2	3	7
	5	3	2	3	8
	6	2	1	1	4
	7	1	1	1	3
	8	3	2	3	8
18	1	1	1	2	4
	2	1	1	1	3
	3	2	1	1	4
	4	1	1	1	3
	5	1	1	1	3
19	1	3	3	3	9
	2	2	1	1	4
	3	2	1	2	5
	4	2	2	1	5
	5	3	3	3	9
20	1	1	1	1	3
	2	2	1	1	4
	3	2	2	1	5
	4	2	2	1	5
	5	2	1	1	4
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
	9	1	1	1	3
	10	1	1	1	3
21	1	1	1	1	3
	2	1	1	1	3
	3	2	1	1	4
	4	2	2	2	6
22	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
23	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
	9	1	1	1	3
	10	1	1	1	3
24	1	1	1	1	3
	2	1	1	1	3
	3	2	2	2	6
	4	2	1	2	5

(Eff.-Noneff. continued)

Counselor	Response	Raters A	B	C	Sum
25	1	2	2	2	6
	2	2	1	1	4
	3	1	2	1	4
	4	2	1	1	4
	5	2	1	2	5
26	1	2	1	1	4
	2	1	2	1	4
	3	2	2	2	6
27	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
28	1	2	1	2	5
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
29	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
30	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	2	4
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
31	1	1	2	2	5
	2	2	2	1	5
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
32	1	3	3	3	9
	2	3	2	3	8
	3	3	3	3	9
	4	3	3	3	9
33	1	3	3	3	9
	2	1	1	2	4
	3	2	1	2	5
	4	2	1	1	4
	5	1	1	1	3

(Eff.-Honeff. continued)

Counselor	Response	Raters A	B	C	Sum
	6	1	1	2	4
	7	1	1	1	3
	8	1	1	2	4
34	1	1	1	1	3
	2	1	1	1	3
	3	1	2	1	4
	4	1	1	1	3
	5	1	1	2	4
35	1	1	1	1	3
	2	1	1	1	3
	3	2	1	1	4
	4	2	1	1	4
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
	9	1	1	1	3
	10	1	1	1	3
36	1	2	3	3	8
	2	3	3	3	9
	3	1	1	1	3
37	1	1	2	3	6
	2	1	1	2	4
	3	1	1	2	4
	4	1	1	1	3
	5	1	2	2	5
38	1	2	2	1	5
	2	1	1	1	3
	3	3	3	2	8
	4	3	3	3	9
	5	2	1	1	4
	6	2	1	1	4
39	1	1	1	1	3
	2	1	1	1	3
	3	2	1	1	4
	4	1	1	1	3
	5	2	2	3	7
	6	1	1	1	3
	7	1	1	1	3
	8	2	1	1	4
40	1	2	1	1	4
	2	1	1	1	3
	3	3	3	1	7
	4	2	1	1	4
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3

(Eff.-Noneff. continued)

Counselor	Response	Raters A	B	C	Sum
41	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
42	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
	8	1	1	1	3
43	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
	6	1	1	1	3
	7	1	1	1	3
44	1	1	1	2	4
	2	1	1	1	3
	3	1	1	1	3
	4	3	2	2	7
	5	2	2	3	7
	6	3	2	3	8
45	1	1	1	1	3
	2	1	1	1	3
	3	1	1	1	3
	4	1	1	1	3
	5	1	1	1	3
46	1	1	1	1	3
	2	3	3	3	9
	3	3	3	3	9
	4	2	1	1	4
	5	2	1	1	4
	6	2	1	2	5
47	1	2	3	3	8
	2	3	3	3	9
	3	3	3	3	9
	4	2	1	1	4
48	1	1	1	2	4
	2	1	1	1	3
	3	3	2	3	8
	4	3	3	3	9
	5	2	2	2	6

(Eff.-Noneff. continued)

Counselor	Response	Raters A	B	C	Sum		
49	1	2	1	2	5		
	2	2	1	2	5		
	3	2	2	2	6		
	4	2	1	2	5		
	5	3	3	3	9		
50	1	3	3	4	10		
	2	3	3	3	9		
	3	2	3	3	8		
	4	2	2	2	6		
	5	4	3	3	10		
	6	2	3	4	9		
	7	3	3	4	10		
51	1	2	3	3	8		
	2	3	3	3	9		
	3	3	3	3	9		
	4	2	2	2	6		
	5	3	3	3	9		
	6	3	3	3	9		
	7	2	2	2	6		
52	1	2	2	3	7		
	2	2	2	3	7		
	3	3	3	4	10		
	4	4	3	4	11		
	5	3	3	3	9		
53	1	2	1	2	5		
	2	2	2	2	6		
	3	1	1	1	3		
	4	3	3	3	9		
	5	3	3	3	9		
	6	2	2	2	6		
Sum		518	470	508	1496	$Sum^2=8738$	$(1496)^2=2238016$
Sum^2		1046	888	1056			$\Sigma X^2=2990$
$(Sum)^2$					747,288		

(Eff.-Noneff. continued)

Analysis of Variance

Source of Variation	df	Sum of Squares (SS)	Variance (S^2)
Counselor Responses	312	$\frac{8738}{3} - \frac{2,238,016}{939} = 529.40$	$\frac{529.40}{312} = 1.70$
Raters	2	$\frac{747,288}{313} - \frac{2,237,016}{939} = 4.23$	$\frac{4.23}{2} = 2.12$
Error	624	$606.73 - (529.40+4.23) = 73.10$	$\frac{73.10}{624} = .12$
Total	938	$2990 - \frac{2,237,016}{939} = 606.73$	

(1) Internal Consistency $= \frac{1.70 - .12}{1.70} = \frac{1.58}{1.70} = .929$

(2) Interjudge Reliability (between-rater variance removed)

$$= \frac{1.70 - 1.2}{1.70 + (3-1)(.12)} = .814$$

(3) Interjudge Reliability (between-rater variance left in)

$$= \frac{529.40 - \frac{(4.23 + 73.10)}{2}}{529.40 + (4.23 + 73.10)} = .809$$

APPENDIX H

STATISTICAL GLOBAL RATER ANALYSIS OF THE SCORES OF THE OBJECTIVITY STUDY BASED UPON HOYT'S COMPUTATION OF RELIABILITY BY ANALYSIS OF VARIANCE

Affect-Cognitive Dimension Scores

Counselors	Raters A	B	C	SUM
1	8	8	8	24
2	0	1	0	1
3	10	10	9	29
4	10	10	10	30
5	4	6	4	14
6	8	10	7	25
7	4	4	4	12
8	10	10	10	30
9	10	10	10	30
10	10	10	10	30
11	10	10	10	30
12	0	0	0	0
13	10	10	10	30
14	10	10	10	30
15	5	5	3	13
16	10	10	9	29
17	6	8	5	19
18	10	10	10	30
19	6	6	6	18
20	10	10	10	30
21	7	7	7	21
22	10	10	10	30
23	10	10	10	30
24	10	10	10	30
25	10	10	10	30
26	10	10	10	30
27	10	10	10	30
28	10	10	10	30
29	10	10	10	30
30	10	10	10	30
31	10	10	10	30
32	0	0	0	0
33	8	9	7	24
34	10	10	10	30
35	10	10	10	30
36	3	7	3	13
37	10	10	6	26
38	8	8	8	24
39	10	10	9	29
40	9	9	10	28
41	10	10	10	30
42	10	10	10	30
43	10	10	10	30
44	7	7	7	21

Affect-Cognitive (con't)

Counselors	A	B	C	SUM		
45	10	10	10	30		
46	7	7	5	19		
47	5	3	3	11		
48	4	4	4	12		
49	8	8	8	24		
50	3	3	0	6		
51	4	3	3	10		
52	4	4	4	12		
53	5	5	5	15		
Sum	413	422	394	1229	$Sum^2=32,789$	$1229^2=1,510,441$
(Sum^2)	3693	3812	3478			$\Sigma X^2=10,983$
$(Sum)^2$	170569	178084	155236	503889		

Analysis of Variance

Source of Variation	df	Sum of Squares (SS)	Variance (S^2)
Counselors	52	$\frac{32789}{3} - \frac{1,510,441}{159} = 1430$	27.50
Raters	2	$\frac{503889}{53} - \frac{1,510,441}{159} = 7$	3.50
Error	104	$1483 - (1430 + 7) = 46$	.44
Total	158	$10983 - 9500 = 1483$	

(1) Internal consistency $= \frac{27.50 - .44}{27.50} = \frac{27.06}{27.50} = .984$

(2) Interjudge reliability (between-rater variance removed)

$$= \frac{27.50 - .44}{27.50 + (3-1)(.44)} = \frac{27.06}{28.38} = .953$$

(3) Interjudge reliability (between-rater variance left in)

$$= \frac{1430 - \frac{(7 + 46)}{3-1}}{1430 + (7 + 46)} = \frac{1403.5}{1483.0} = .946$$

Understanding-Nonunderstanding Dimension Scores

Counselors	Raters A	B	C	SUM
1	5	7	7	19
2	0	1	0	1
3	9	10	8	27
4	10	10	10	30
5	4	6	4	14
6	8	8	5	21
7	4	6	6	16
8	7	8	8	23
9	7	10	10	27
10	10	10	10	30
11	10	10	10	30
12	0	0	0	0
13	10	10	10	30
14	10	10	10	30
15	5	3	3	11
16	10	9	9	28
17	5	6	4	15
18	10	10	10	30
19	6	6	6	18
20	10	10	10	30
21	7	7	7	21
22	10	10	10	30
23	10	10	10	30
24	5	7	7	19
25	4	8	6	18
26	7	10	3	20
27	10	10	10	30
28	10	10	10	30
29	10	10	10	30
30	10	10	10	30
31	10	10	10	30
32	0	0	0	0
33	9	9	7	25
34	10	10	10	30
35	10	10	10	30
36	3	3	3	9
37	10	10	6	26
38	7	7	7	21
39	10	9	9	28
40	9	7	10	26
41	10	10	10	30
42	10	10	10	30
43	10	10	10	30
44	5	5	5	15

Understanding-Nonunderstanding (con't)

Counselors	A	B	C	SUM		
45	10	10	10	30		
46	5	7	5	17		
47	3	3	0	6		
48	4	4	4	12		
49	4	6	6	16		
50	0	0	0	0		
51	4	3	1	8		
52	4	4	0	8		
53	5	5	3	13		
Sum	375	394	359	1128	$Sum^2=28752$	$1128^2=1,272,384$
(Sum^2)	3193	3428	3075			$\Sigma X^2=9696$
$(Sum)^2$	140625	155236	128881	424742		

Analysis of Variance

Source of Variation	df	Sum of Squares (SS)	Variance (S^2)
Counselors	52	$\frac{28752}{3} - \frac{1,272,384}{159} = 1582$	30.40
Raters	2	$\frac{424742}{53} - \frac{1,272,384}{159} = 12$	6.00
Error	104	$1694 - (1582 + 12) = 100$	.96
Total	158	$9696 - 8002 = 1694$	

(1) Internal consistency $= \frac{30.40 - .96}{30.40} = \frac{29.44}{30.40} = .968$

(2) Interjudge reliability (between-rater variance removed)

$$= \frac{30.40 - .96}{30.40 + (3-1)(.96)} = \frac{29.44}{32.32} = .911$$

(3) Interjudge reliability (between-rater variance left in)

$$= \frac{1582 - \frac{(12 + 100)}{2}}{1582 + (12 + 100)} = \frac{1526}{1694} = .901$$

Specific - Nonspecific Dimension Scores

Counselors	Raters A	B	C	SUM
1	10	10	8	28
2	3	1	3	7
3	9	10	8	27
4	10	10	10	30
5	6	6	4	16
6	8	8	8	24
7	8	6	6	20
8	10	10	10	30
9	10	10	10	30
10	10	10	10	30
11	10	10	10	30
12	0	0	0	0
13	10	10	10	30
14	10	10	10	30
15	5	7	3	15
16	10	10	9	29
17	6	8	5	19
18	10	10	10	30
19	6	6	6	18
20	10	10	10	30
21	10	10	10	30
22	10	10	10	30
23	10	10	10	30
24	10	10	10	30
25	10	10	10	30
26	10	10	7	27
27	10	10	10	30
28	10	10	10	30
29	10	10	10	30
30	10	10	10	30
31	10	10	10	30
32	0	0	0	0
33	9	9	9	27
34	10	10	10	30
35	10	10	10	30
36	10	3	3	16
37	10	10	8	28
38	8	8	8	24
39	10	10	9	29
40	10	9	10	29
41	10	10	10	30
42	10	10	10	30
43	10	10	10	30
44	10	10	7	27

Specific - Nonspecific (con't)

Counselors	A	B	C	SUM
45	10	10	10	30
46	8	8	8	24
47	5	3	3	11
48	6	6	6	18
49	10	8	8	26
50	4	6	0	10
51	4	4	3	11
52	4	4	0	8
53	8	7	3	18
Sum	447	437	402	1286
(Sum^2)	4133	4011	3592	
$(Sum)^2$	199809	190969	161604	552382

$Sum^2=34836$ $1286^2=1,653,796$ $\Sigma X^2=11736$

Analysis of Variance

Source of Variation	df	Sum of Squares (SS)	Variance (S^2)
Counselors	52	$\frac{34836}{3} - \frac{1,653,796}{159} = 1211$	23.90
Raters	2	$\frac{552382}{53} - \frac{1,653,796}{159} = 21$	10.50
Error	104	$1335 - (1211 + 21) = 103$	.99
Total	158	$11736 - 10401 = 1335$	

(1) Internal consistency $= \frac{23.90 - .99}{23.90} = .959$

(2) Interjudge reliability (between-rater variance removed)

$$= \frac{23.90 - .99}{23.90 + (3-1)(.99)} = \frac{22.91}{25.88} = .885$$

(3) Interjudge reliability (between-rater variance left in)

$$= \frac{1211 - \left(\frac{21 + 103}{2}\right)}{1211 + (21 + 103)} = \frac{1149}{1335} = .860$$

Exploratory - Nonexploratory Dimension

Counselors	Raters A	B	C	SUM
1	10	10	8	28
2	4	4	3	11
3	10	9	9	28
4	10	10	10	30
5	10	8	6	24
6	8	8	8	24
7	6	6	6	18
8	8	10	10	28
9	10	10	7	27
10	10	10	9	29
11	10	10	9	29
12	7	7	7	21
13	10	10	10	30
14	10	10	10	30
15	7	7	7	21
16	10	10	9	29
17	6	8	6	20
18	10	10	8	28
19	10	10	10	30
20	10	10	10	30
21	10	10	10	30
22	10	10	10	30
23	10	10	10	30
24	10	10	10	30
25	10	10	10	30
26	10	10	10	30
27	10	10	10	30
28	10	10	10	30
29	10	10	10	30
30	10	10	10	30
31	10	10	10	30
32	7	7	5	19
33	10	9	9	28
34	10	10	10	30
35	10	10	10	30
36	10	7	3	20
37	10	10	6	26
38	7	7	7	21
39	10	10	9	29
40	9	9	10	28
41	10	10	10	30
42	10	10	10	30
43	10	10	10	30
44	10	10	7	27

Exploratory - Nonexploratory (con't)

Counselors	A	B	C	SUM		
45	10	10	10	30		
46	5	8	8	21		
47	5	3	0	8		
48	10	10	10	30		
49	10	10	8	28		
50	10	9	1	20		
51	6	4	1	11		
52	6	6	0	12		
53	10	8	8	26		
Sum	481	474	424	1379	$Sum^2=37613$	$1379^2=1,901,641$
(Sum^2)	4515	4402	3804			$\Sigma X^2=12721$
$(Sum)^2$	231361	224676	179776	635813		

Analysis of Variance

Source of Variation	df	Sum of Squares (SS)		Variance (S^2)
Counselors	52	$\frac{37613}{3} - \frac{1,901,641}{159}$	= 578	11.1
Raters	2	$\frac{635813}{53} - \frac{1,901,641}{159}$	= 36	18.0
Error	104	761 - (578 + 36)	= 147	1.4
Total	158	12721 - 11960	= 761	

(1) Internal consistency $= \frac{11.1 - 1.4}{11.1} = \frac{9.7}{11.1} = .873$

(2) Interjudge reliability (between-rater variance removed)

$$= \frac{11.1 - 1.4}{11.1 + (3-1)(1.4)} = \frac{9.7}{13.9} = .697$$

(3) Interjudge reliability (between rater variance left in)

$$= \frac{578 - \frac{(36 + 147)}{2}}{578 + (36 + 147)} = \frac{486.5}{761} = .639$$

Effective - Noneffective Dimension Scores

Counselors	A	B	C	SUM
1	12	11	12	35
2	20	18	21	59
3	12	11	13	36
4	6	5	5	16
5	11	10	13	34
6	13	11	11	35
7	11	11	11	33
8	8	7	6	21
9	8	6	6	20
10	11	10	9	30
11	9	7	8	24
12	10	10	10	30
13	6	4	4	14
14	8	7	8	23
15	10	8	10	28
16	9	9	11	29
17	18	15	18	51
18	6	5	6	17
19	12	10	10	32
20	14	12	10	36
21	6	5	5	16
22	4	4	4	12
23	10	10	10	30
24	6	5	6	17
25	9	7	7	23
26	5	5	4	14
27	7	7	7	21
28	5	4	5	14
29	6	6	6	18
30	7	7	8	22
31	7	8	7	22
32	12	11	12	35
33	12	10	14	36
34	5	6	6	17
35	12	10	10	32
36	6	7	7	20
37	5	7	10	22
38	13	11	9	33
39	11	9	10	30
40	11	9	7	27
41	6	6	6	18
42	8	8	8	24
43	7	7	7	21
44	11	9	12	32

Effective - Noneffective (con't)

Counselors	A	B	C	SUM		
45	5	5	5	15		
46	13	10	11	34		
47	10	10	10	30		
48	12	9	11	32		
49	11	8	11	30		
50	19	20	23	62		
51	18	19	19	56		
52	14	13	17	44		
53	13	12	13	38		
Sum	520	471	509	1500	$Sum^2=49114$	$1500^2=2,250,000$
(Sum^2)	5856	4837	5817			$\sum X^2=16510$
$(Sum)^2$	270400	221841	259081	751322		

Analysis of Variance

Source of Variation	df	Sum of Squares (SS)	Variance (S^2)
Counselors	52	$\frac{49114}{3} - \frac{2,250,000}{159} = 2220$	48.0
Raters	2	$\frac{751322}{53} - \frac{2,250,000}{159} + 25$	12.5
Error	104	$2359 - (2220 + 25) = 114$	1.1
Total	158	$16510 - 14151 = 2359$	

(1) Internal consistency $= \frac{48.0 - 1.1}{48.0} = \frac{46.9}{48.0} = .977$

(2) Interjudge reliability (between-rater variance removed)

$$= \frac{48.0 - 1.1}{48.0 + (3-1)(1.1)} = \frac{46.9}{50.2} = .934$$

(3) Interjudge reliability (between-rater variance left in)

$$= \frac{2220 - \frac{(25 + 114)}{2}}{2220 + (25 + 114)} = \frac{2150.5}{2359} = .911$$

APPENDIX I

STATISTICAL "t"-TEST ANALYSIS COMPARING THE A-V SCALE SCORES OF THE OBJECTIVITY STUDY'S EIGHT PH.D. CANDIDATES WITH THOSE OF THE FORTY-FIVE M.A. CANDIDATES

Affect-Cognitive Dimension

COGNITIVE CATEGORY

M.A. Candidates	Cognitive Scores	Deviation From Mean	Squared Deviation
1	8	- .3	.9
2	0	-8.3	68.9
3	10	1.7	2.9
4	10	1.7	2.9
5	4	-4.3	18.5
6	8	- .3	.9
7	4	-4.3	18.5
8	10	1.7	2.9
9	10	1.7	2.9
10	10	1.7	2.9
11	10	1.7	2.9
12	0	-8.3	68.9
13	10	1.7	2.9
14	10	1.7	2.9
15	5	-3.3	10.9
16	10	1.7	2.9
17	6	-2.3	5.3
18	10	1.7	2.9
19	6	-2.3	5.3
20	10	1.7	2.9
21	7	-1.3	1.7
22	10	1.7	2.9
23	10	1.7	2.9
24	10	1.7	2.9
25	10	1.7	2.9
26	10	1.7	2.9
27	10	1.7	2.9
28	10	1.7	2.9
29	10	1.7	2.9
30	10	1.7	2.9
31	10	1.7	2.9
32	0	-8.3	68.9
33	8	- .3	.9
34	10	1.7	2.9
35	10	1.7	2.9
36	3	-5.3	28.1
37	10	1.7	2.9
38	8	- .3	.9
39	10	1.7	2.9
40	9	.7	.5

(Aff.-Cog. continued)

M.A. Candidates	Cognitive Scores	Deviation From Mean	Squared Deviation
41	10	1.7	2.9
42	10	1.7	2.9
43	10	1.7	2.9
44	7	-1.3	1.7
45	10	1.7	2.9
Sum	373		329.8
N	45		
$\bar{X}_1$	8.3	Variance (S^2_1) = $\frac{329.8}{45}$ = 7.3	

Ph.D. Candidates	Cognitive Scores	Deviation From Mean	Squared Deviation
46	7	2.2	4.8
47	3	-1.8	3.2
48	4	- .8	.6
49	8	3.2	10.2
50	3	-1.8	3.2
51	3	-1.8	3.2
52	4	- .8	.6
53	5	.2	.4
Sum	37		26.2
N	8		
$\bar{X}_2$	4.8	Variance (S^2_2) = $\frac{26.2}{8}$ = 6.3	

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S^2_1 + S^2_2}{2}\left(\frac{1}{N_1} + \frac{1}{N_2}\right)}} = \frac{8.3 - 4.8}{\sqrt{\frac{7.3 + 6.3}{2}\left(\frac{1}{45} + \frac{1}{8}\right)}} = \frac{3.5}{\sqrt{1}} = \frac{3.5}{1} = 3.5$$

degrees of freedom (df) = $N_1 + N_2 - 2 = 45 + 8 - 2 = 51$

significance at .05 level = 2.00
significance at .01 level = 2.66

Understanding-Nonunderstanding Dimension

NONUNDERSTANDING CATEGORY

M.A. Candidates	Nonunderstanding Scores	Deviation From Mean	Squared Deviation
1	7	- .9	.8
2	0	-7.9	62.4
3	9	1.1	1.2
4	10	2.1	4.4
5	4	-3.9	15.2
6	8	.1	.0
7	6	-1.9	3.6
8	8	.1	.0
9	10	2.1	4.4
10	10	2.1	4.4
11	10	2.1	4.4
12	0	-7.9	62.4
13	10	2.1	4.4
14	10	2.1	4.4
15	3	-4.9	24.0
16	9	1.1	1.2
17	5	-2.9	8.4
18	10	2.1	4.4
19	6	-1.9	3.6
20	10	2.1	4.4
21	7	- .9	.8
22	10	2.1	4.4
23	10	2.1	4.4
24	7	- .9	.8
25	6	-1.9	3.6
26	7	- .9	.8
27	10	2.1	4.4
28	10	2.1	4.4
29	10	2.1	4.4
30	10	2.1	4.4
31	10	2.1	4.4
32	0	-7.9	62.4
33	9	1.1	1.2
34	10	2.1	4.4
35	10	2.1	4.4
36	3	-4.9	24.0
37	10	2.1	4.4
38	7	- .9	.8
39	9	1.1	1.2
40	9	1.1	1.2

(Underst.-Nonunderst. continued)

M.A. Candidates	Nonunderstanding Scores	Deviation From Mean	Squared Deviation
41	10	2.1	4.4
42	10	2.1	4.4
43	10	2.1	4.4
44	5	-2.9	8.4
45	10	2.1	4.4
Sum	354		384.8
N	45		
$\bar{X}_1$	7.9	Variance $(S^2_1) = \frac{384.8}{45} = 8.6$	

Ph.D. Candidates	Nonunderstanding Scores	Deviation From Mean	Squared Deviation
46	5	1.2	1.4
47	3	- .8	.6
48	4	.2	.4
49	6	2.2	4.8
50	0	-3.8	14.4
51	3	- .8	.6
52	4	.2	.4
53	5	1.2	1.4
Sum	30		24.0
N	8		
$\bar{X}_2$	3.8	Variance $(S^2_2) = \frac{24.0}{8} = 3.0$	

$$t = \frac{7.9 - 3.8}{\sqrt{\frac{8.6 + 3.0}{2}\left(\frac{1}{45} + \frac{1}{8}\right)}} = \frac{4.1}{\sqrt{.85}} = \frac{4.1}{.29} = 14.1$$

df = 51

significance at .05 level = 2.00
significance at .01 level - 2.66

Specific-Nonspecific Dimension

NONSPECIFIC CATEGORY

M.A. Candidates	Nonspecific Scores	Deviation From Mean	Squared Deviation
1	10	1.4	2.0
2	3	-5.6	31.4
3	9	.4	.2
4	10	1.4	2.0
5	6	-2.6	6.8
6	8	- .6	.4
7	6	-2.6	6.8
8	10	1.4	2.0
9	10	1.4	2.0
10	10	1.4	2.0
11	10	1.4	2.0
12	0	-8.6	74.0
13	10	1.4	2.0
14	10	1.4	2.0
15	5	-3.6	13.0
16	10	1.4	2.0
17	6	-2.6	6.8
18	10	1.4	2.0
19	6	-2.6	6.8
20	10	1.4	2.0
21	10	1.4	2.0
22	10	1.4	2.0
23	10	1.4	2.0
24	10	1.4	2.0
25	10	1.4	2.0
26	10	1.4	2.0
27	10	1.4	2.0
28	10	1.4	2.0
29	10	1.4	2.0
30	10	1.4	2.0
31	10	1.4	2.0
32	0	-8.6	74.0
33	9	.4	.2
34	10	1.4	2.0
35	10	1.4	2.0
36	3	-5.6	31.4
37	10	1.4	2.0
38	8	- .6	.4
39	10	1.4	2.0
40	10	1.4	2.0

(Spec.-Nonspec. continued)

M.A. Candidates	Nonspecific Scores	Deviation From Mean	Squared Deviation
41	10	1.4	2.0
42	10	1.4	2.0
43	10	1.4	2.0
44	10	1.4	2.0
45	10	1.4	2.0
Sum	389		316.2
N	45		
$\bar{X}_1$	8.6	Variance $(S^2{}_1) = \frac{316.2}{45} = 7.0$	

Ph.D. Candidates	Nonspecific Scores	Deviation From Mean	Squared Deviation
46	8	2.7	7.3
47	3	-2.3	5.3
48	6	.7	.5
49	8	2.7	7.3
50	3	-2.3	5.3
51	4	-1.3	1.7
52	4	-1.3	1.7
53	6	.7	.5
Sum	42		29.6
N	8		
$\bar{X}_2$	5.3	Variance $(S^2{}_2) = \frac{29.6}{8} = 3.7$	

$$t = \frac{8.6 - 5.3}{\sqrt{\frac{7.0 + 3.7}{2}\left(\frac{1}{45} + \frac{1}{8}\right)}} = \frac{3.3}{\sqrt{.79}} = \frac{3.3}{.28} = \underline{11.8}$$

df = 51

significance at .05 level = 2.00
significance at .01 level = 2.66

Exploratory-Nonexploratory Dimension

NONEXPLORATORY CATEGORY

M.A. Candidates	Nonexploratory Scores	Deviation From Mean	Squared Deviation
1	10	.8	.64
2	4	-5.2	27.04
3	9	- .2	.04
4	10	.8	.64
5	8	-1.2	1.44
6	8	-1.2	1.44
7	6	-3.2	10.24
8	10	.8	.64
9	10	.8	.64
10	10	.8	.64
11	10	.8	.64
12	7	-2.2	4.84
13	10	.8	.64
14	10	.8	.64
15	7	-2.2	4.84
16	10	.8	.64
17	6	-3.2	10.24
18	10	.8	.64
19	10	.8	.64
20	10	.8	.64
21	10	.8	.64
22	10	.8	.64
23	10	.8	.64
24	10	.8	.64
25	10	.8	.64
26	10	.8	.64
27	10	.8	.64
28	10	.8	.64
29	10	.8	.64
30	10	.8	.64
31	10	.8	.64
32	7	-2.2	4.84
33	9	- .2	.04
34	10	.8	.64
35	10	.8	.64
36	7	-2.2	4.84
37	10	.8	.64
38	7	-2.2	4.84
39	10	.8	.64
40	9	- .2	.04

(Expl.-Nonexpl. continued)

M.A. Candidates	Nonexploratory Scores	Deviation From Mean	Squared Deviation
41	10	.8	.64
42	10	.8	.64
43	10	.8	.64
44	10	.8	.64
45	10	.8	.64
Sum	414		95.20
N	45		
$\bar{X}_1$	9.2	Variance $(S^2_1) = \frac{95.20}{45} = 2.1$	

Ph.D. Candidates	Nonexploratory Scores	Deviation From Mean	Squared Deviation
46	8	1.0	1.0
47	3	-4.0	16.0
48	10	3.0	9.0
49	10	3.0	9.0
50	7	0	0.0
51	4	-3.0	9.0
52	6	-1.0	1.0
53	8	1.0	1.0
Sum	56		46.0
N	8		
$\bar{X}_2$	7.0	Variance $(S^2_2) = \frac{46.0}{8.0} = 5.8$	

$$t = \frac{9.2 - 7.0}{\sqrt{\frac{2.1 + 5.8}{2}\left(\frac{1}{45} + \frac{1}{8}\right)}} = \frac{2.2}{\sqrt{6.8}} = \frac{2.2}{.26} = 8.5$$

df = 51

significance at .05 level = 2.00
significance at .01 level = 2.66

Effective-Noneffective Dimension

M.A. Candidates	Scores	Deviation From Mean	Squared Deviation
1	17	1.8	3.2
2	29	13.8	190.4
3	13	- 2.2	4.8
4	13	- 2.2	4.8
5	22	6.8	46.2
6	18	2.2	4.8
7	22	6.8	46.2
8	12	- 3.8	14.4
9	15	- .2	0.0
10	13	- 2.2	4.8
11	11	- 4.2	17.6
12	33	17.8	316.8
13	10	- 5.2	27.0
14	13	- 2.2	4.8
15	25	9.8	96.0
16	11	- 4.2	17.6
17	23	7.8	60.8
18	12	- 3.2	10.2
19	20	4.8	23.0
20	12	- 3.2	10.2
21	13	- 2.2	4.8
22	10	- 5.2	27.0
23	10	- 5.2	27.0
24	15	- .2	0.0
25	13	- 2.2	4.8
26	17	1.8	3.2
27	10	- 5.2	27.0
28	13	- 2.2	4.8
29	10	- 5.2	27.0
30	10	- 5.2	27.0
31	12	- 3.2	10.2
32	30	14.8	219.0
33	15	- .2	0.0
34	12	- 3.2	10.2
35	10	- 5.2	27.0
36	23	7.8	60.8
37	13	- 2.2	4.8
38	18	2.8	7.8
39	13	- 2.2	4.8
40	13	- 2.2	4.8

(Eff.-Noneff. continued)

M.A. Candidates	Scores	Deviation From Mean	Squared Deviation
41	10	-5.2	27.0
42	10	-5.2	27.0
43	10	-5.2	27.0
44	18	2.8	7.8
45	10	-5.2	27.0
Sum	682		1511.2
N	45		
$\bar{X}_2$	15.2		

Variance $(S^2{}_2) = \frac{1511.2}{45} = 33.6$

Ph.D. Candidates	Scores	Deviation From Mean	Squared Deviation
46	18	-6.9	47.6
47	25	.1	0.0
48	22	-2.9	8.4
49	22	-2.9	8.4
50	33	8.1	65.6
51	29	4.1	16.8
52	28	3.1	9.6
53	22	-2.9	8.4
Sum	199		164.8
N	8		
$\bar{X}_1$	24.9		

Variance $(S^2{}_1) = \frac{164.8}{8} = 20.6$

$$t = \frac{24.9 - 15.2}{\sqrt{\frac{20.6 + 33.6}{2}\left(\frac{1}{8} + \frac{1}{45}\right)}} = \frac{9.7}{\sqrt{4.0}} = \frac{9.7}{2} = \underline{4.9}$$

df = 51

significance at .05 level = 2.00
significance at .01 level - 2.66

APPENDIX J

DEPENDENT -"t" STATISTICAL ANALYSIS COMPUTATIONS COMPARING THE NONSUPERVISED GROUP'S INITIAL AND TERMINAL INTERVIEW SCORES

Affect Category

Nonsupervised Group (Client-Interrogation)

Counselor Enrollees	Affect Score Initial Interview	Affect Score Terminal Interview	Difference (D)	Difference^2 (D^2)
F	1	1	0	0
G	2	3	1	1
H	0	0	0	0
I	2	3	1	1
J	1	2	1	1
Sum	6	9	$\Sigma D = 3$	$\Sigma D^2 = 3$
N	5			
$\bar{X}$	1.2	1.8	$(\Sigma D)^2 = (3)^2 = 9$	

$$\text{dependent "t"} = \frac{\frac{\Sigma D}{N}}{\sqrt{\frac{N\Sigma D^2 - (\Sigma D)^2}{N^2(N-1)}}} = \frac{\bar{D}}{S^2_D} = \frac{\frac{3}{5}}{\sqrt{\frac{5(3) - 9}{25(4)}}} = \frac{.6}{\sqrt{.06}} = \frac{.6}{.24} = 2.5$$

degrees of freedom = N-1 = 5-1 = 4
significance at .05 level = 2.78

Understanding Category

Nonsupervised Group (Client-Interrogation)

Counselor Enrollees	Understanding Score Initial Interview	Understanding Score Terminal Interview
F	8	3
G	2	0
H	1	0
I	4	2
J	6	2
Sum	21	12
N	5	
$\bar{X}$	4.2	2.4

It can be seen that there was a decrease rather than an increase from the initial to the terminal ratings.

Specific Category

Nonsupervised Group (Client-Interrogation)

Counselor Enrollees	Specific Score Initial Interview	Specific Score Terminal Interview	Difference (D)	Difference (D^2)
F	2	4	2	4
G	0	1	1	1
H	0	0	0	0
I	1	1	0	0
J	1	2	1	1
Sum	4	8	4	6
N	5			
$\bar{X}$	.8	1.6	$(4)^2 = 16$	

$$\text{dependent "t"} = \frac{\frac{4}{5}}{\sqrt{\frac{5(6) - 16}{25(4)}}} = \frac{.8}{\sqrt{.14}} = \frac{.8}{.47} = \underline{2.2}$$

df = 4
significance at .05 level = 2.78

Exploratory Category

Nonsupervised Group (Client-Interrogation

Counselor Enrollees	Exploratory Score Initial Interview	Exploratory Terminal Interview
F	9	6
G	0	2
H	0	0
I	5	2
J	3	0
Sum	17	10
N	5	
$\bar{X}$	3.4	2.0

It can be seen that there was a decrease rather than an increase from the initial to the terminal ratings.

Effective Category

Nonsupervised Group (Client-Interrogation)

Counselor Enrollees	Effective Score Initial Interview	Effective Score Terminal Interview
F	36	37
G	22	24
H	21	21
I	31	25
J	26	22
Sum	136	129
N	5	
$\bar{X}$	27.2	25.8

It can be seen that there was a decrease rather than an increase from the initial to the terminal ratings.

APPENDIX K

DEPENDENT - "t" STATISTICAL ANALYSIS COMPUTATIONS COMPARING THE SUPERVISED GROUP'S INITIAL AND TERMINAL INTERVIEW SCORES

Affect Category

Supervised Group (Counselor Interrogation)

Counselor Enrollees	Affect Score Initial Interview	Affect Score Terminal Interview	Difference (D)	Difference^2 (D^2)
A	7	8	1	1
B	0	4	4	16
C	4	10	6	36
D	1	4	3	9
E	1	10	9	81
Sum	13	36	$\sum D = 23$	$\sum D^2 = 143$
N	5			
$\bar{X}$	2.6	7.2	$(\sum D)^2 = (23)^2 = 529$	

$$\text{dependent "t"} = \frac{\frac{\sum D}{N}}{\sqrt{\frac{N\sum D^2 - (\sum D)^2}{N^2(N-1)}}} = \frac{\bar{D}}{S^2_D} = \frac{\frac{23}{5}}{\sqrt{\frac{5(143) - 529}{25(4)}}} = \frac{4.6}{\sqrt{1.86}} = \frac{4.6}{1.36} = 3.4$$

degrees of freedom = N - 1 = 5 - 1 = 4
significance at .05 level = 2.75
significance at .01 level = 4.60

Understanding Category

Supervised Group (Counselor Interrogation)

Counselor Enrollees	Understanding Score Initial Interview	Understanding Score Terminal Interview	Difference (D)	Difference^2 (D^2)
A	14	17	3	9
B	3	15	12	144
C	5	14	9	81
D	7	5	2	4
E	5	19	14	196
Sum	34	70	40	434
N	5			
$\bar{X}$	6.8	14.0	$(40)^2 = 1600$	

$$\text{dependent "t"} = \frac{\frac{.40}{5}}{\sqrt{\frac{5(434) - 1600}{25(4)}}} = \frac{8.0}{\sqrt{5.70}} = \frac{8.0}{2.38} = 3.3$$

df = 4
significance at .05 level = 2.75
significance at .01 level = 4.60

Specific Category

Supervised Group (Counselor Interrogation)

Counselor Enrollees	Specific Score Initial Interview	Specific Score Terminal Interview	Difference (D)	Difference^2 (D^2)
A	7	17	7	49
B	1	14	13	169
C	3	10	7	49
D	4	14	10	100
E	2	18	16	256
Sum	17	73	53	623
N	5			
$\bar{X}$	3.4	14.6	$(53)^2 = 2809$	

$$\text{dependent "t"} = \frac{\frac{53}{5}}{\sqrt{\frac{5(623) - 2809}{25(4)}}} = \frac{10.6}{\sqrt{3.06}} = \frac{10.6}{1.73} = 6.1$$

df = 4
significance at .05 level = 2.75
significance at .01 level = 4.60

Exploratory Category

Supervised Group (Counselor Interrogation)

Counselor Enrollees	Exploratory Score Initial Interview	Exploratory Score Terminal Interview	Difference (D)	Difference^2 (D^2)
A	15	19	4	8
B	1	16	15	225
C	6	14	8	64
D	8	14	6	36
E	6	20	14	196
Sum	36	83	47	529
N	5			
$\bar{X}$	7.2	16.6	$(47)^2 = 2209$	

$$\text{dependent "t"} = \frac{\frac{47}{5}}{\sqrt{\frac{5(529) - 2209}{25(4)}}} = \frac{9.4}{\sqrt{4.36}} = \frac{9.4}{2.09} = 4.6$$

df = 4
significance at .05 level = 2.75
significance at .01 level = 4.60

Effective Category

Supervised Group (Counselor Interrogation)

Counselor Enrollees	Effective Score Initial Interview	Effective Score Initial Interview	Difference (D)	Difference^2 (D^2)
A	44	54	10	100
B	25	49	24	276
C	29	52	23	529
D	33	42	9	81
E	26	60	34	1156
Sum	157	257	100	2442
N	5			
$\bar{X}$	31	51	$(100)^2 = 10,000$	

$$\text{dependent "t"} = \frac{100}{\sqrt{\frac{5(2442) - 10,000}{25(4)}}} = \frac{20}{\sqrt{22.1}} = \frac{20}{4.7} = 4.2$$

df = 4
significance at .05 level = 2.75
significance at .01 level = 4.60