THE RELATIVE EFFECTIVENESS OF LARGER UNITS
USED IN INTERVIEW ANALYSIS

Dissertation

Presented in Partial Fulfillment of the Requirements
for the Degree Doctor of Philosophy in the
Graduate School of The Ohio State
University

By

John Edward Muthard, B.S., A.M.
The Ohio State University

1952

THE OHIO STATE
UNIVERSITY

Approved by:

Adviser

TABLE OF CONTENTS

Chapter		Page
I	Introduction.....	1
II	History of the Problem.....	9
III	Description of Data, Dimensions and Methodology.....	81
IV	Results and Discussion.....	132
V	Summary and Conclusions.....	201
	Bibliography.....	220
	Appendix 1.....	226
	Autobiography.....	263

LIST OF TABLES

Table		Page
I	Interview Sample Used in This Research..	84
II	Mean, Standard Deviation and Range of Ratings on Four Dimensions Obtained When 36 Interviews Were Rated by Discussion Topic, Problem Area and Fractional Units	133
III	Intercorrelations of the Mean Values of Four Dimensions for Each of Three Units of the Interview.....	135
IV	Reliability Coefficients for the Combined Ratings of the Several Judges for the Four Dimensions on Three Different Type Units.....	137
V	Inter-rater Reliability Coefficients for Four Variables Rated in Three Different Interview Units.....	138
VI	Significance of the Differences Between Reliability Coefficients of Combined Ratings of Four Dimensions on Three Types of Units.....	142
VII	Significance of the Difference Between Correlation Coefficients for Mean Ratings of Within-Unit Halves Obtained for Discussion Topic, Problem Area and Fractional Units of the Interview.....	144
VIII	Significance of the Differences Between Correlations of Amount of Lead Ratings for Within-Unit Halves and Between-Unit Halves When Three Types of Units are Used in Rating Interview Typescripts....	149
IX	Significance of the Difference Between Correlations of Working Relationship Ratings for Within-Unit Halves and Between-Unit Halves When Three Types of Units are Used in Rating Interview Typescripts.....	150

Table		Page
X	Significance of the Difference Between Correlation Coefficients of Three Raters for Within-Unit Halves on Working Relationship.....	153
XI	Significance of the Difference Between Sets of Difference Scores for Three Types of Interview Units.....	155
XII	Mean Ratings on Four Dimensions for Four Unit Categories in Discussion Topic Units.....	159
XIII	Mean of Between-Unit Difference Scores for Ratings of Four Dimensions in Three Types of Units.....	161
XIV	Distribution of Unit Classifications for the 174 Discussion Topic Units Used in This Investigation According to Each of the Twelve Cases Studied.....	186
XV	Distribution of Unit Classifications for the 174 Discussion Topic Units Used in This Investigation According to the Counseling Center Where They Were Obtained.....	190
XVI	Mean and Standard Deviation Values of Amount of Lead Ratings for Discussion Topics Within Cases and by School.....	192
XVII	Mean and Standard Deviation Values of Counselor Assumption of Responsibility Ratings for Discussion Topics within Cases and by School.....	193
XVIII	Mean and Standard Deviation Values of Client Assumption of Responsibility Ratings for Discussion Topics within Cases and by School.....	194
XIX	Mean and Standard Deviation Values of Working Relationship Ratings for Discussion Topics within Cases and by School.....	195

LIST OF CHARTS

Chart		Page
I	Amount of Lead Ratings by Rater for Discussion Topics of a "Typical" Interview.....	169
II	Counselor Assumption of Responsibility Ratings by Rater for Discussion Topics of a "Typical" Interview.....	170
III	Client Assumption of Responsibility Ratings by Rater for Discussion Topics of a "Typical" Interview.....	171
IV	Working Relationship Ratings by Rater for Discussion Topics of a "Typical" Interview.....	172
V	Remark-by-Remark Ratings for Amount of Lead in the First Interview of Case III of Chicago.....	176

ACKNOWLEDGMENT

The writer is grateful for the understanding guidance received from Dr. Francis P. Robinson throughout the course of this investigation. The help of Dr. John R. Kinzer and Dr. Harold B. Pepinsky who served as members of the reading committee is appreciated.

This research was supported in part from funds granted to The Ohio State University by the Research Foundation for aid in fundamental research.

The assistance of Dr. D. R. Whitney and Mrs. Lydia Kinzer of the Statistics Laboratory, Department of Mathematics, The Ohio State University and Dr. Robert J. Wherry in the planning of statistical procedures is acknowledged.

The author is thankful for the help of David Danskin, Jane Hamilton, Philip Pendleton, James E. Smith and Roy Warman in making the many ratings and judgments involved. For guidance and assistance in the use of IBM punch cards the writer is indebted to L. Murray Baughman, Robert Dugan and George Vening.

For her considerable aid and assurance during the progress of this study, the writer is deeply grateful to his wife, Ella.

THE RELATIVE EFFECTIVENESS OF LARGER UNITS USED IN INTERVIEW ANALYSIS

CHAPTER I

INTRODUCTION

The main purpose of this study is to investigate the relative effectiveness for studying client and counselor behavior of several larger units of analysis used in typescript research. This problem is regarded as one of the basic methodological questions which needs to be resolved before it will be possible to compare and integrate interview research findings. In addition, we are interested in this study in contributing knowledge about the extent to which various units enable one to determine the different roles assumed by the client and counselor during their relationship and whether these roles affect their behavior. These problems seem to warrant intensive investigation. Before elaborating on these problems, it seems worthwhile to define what is meant by the field of counseling and indicate the importance of the field.

An examination of recent texts on counseling gives the impression, at first glance, that each author has a unique conception of the nature of counseling. As Wrenn (63) suggests, "counseling" has probably been more

misused and misunderstood than any educational term. There are those who equate it with the whole field of personnel work or guidance and others who see it solely as advice giving or receiving. Even counselors and teachers of counselors find it difficult to concur on a statement covering the counseling function. Among the main points of difference in defining counseling are: the kinds of discussion which should be included under that name; the nature of the relationship between the counselor and the client and the need for diagnosis in helping the client.

There are, however, several points of agreement on the definition of counseling. Many writers (26, 45, 48, 60, 63) agree that counseling is a two-person relationship. They also concur in stating that the counselor is qualified by both training and experience to aid the client in dealing with his difficulties. There is also the implication that the client, as a result of the counseling experience, learns how to deal more effectively with his difficulties. Since it incorporates the points of agreement among the diverse definitions, that of Robinson (45, p. 3) was chosen as the definition of counseling for this study. He says, "Counseling...covers all types of two person situations in which one person, the client, is helped to adjust

more effectively to himself and to his environment."

Counseling, as defined here, is an important activity carried on in many settings. Survey evidence, such as that of Froehlich (21) and Kamm and Wrenn (29), indicates that it is an increasingly important function of schools and colleges. Counseling has also become a more important function of the Federal, state and municipal governments, community agencies and business and industry. Some other indications of the importance of counseling in our society is the continued demand for counseling services, the increased membership of professional groups whose primary interests are counseling and the larger number of publications devoted in whole or part to discussing various facets of counseling.

The above discussion indicates that there has been a considerable increase in the numbers of counseling facilities and personnel and that there is a continued demand for their services. To meet this demand those in the counseling field are obligated to improve the efficacy of their procedures. Such knowledge needs to be based upon research findings; otherwise, it is doubtful whether the field of counseling can effectively serve either the best interests of its clients or the counseling profession itself.

An important factor in the continued growth of

counseling has been the increasing experimentation in this area. However, it is only 11 years ago that the counseling process was first experimentally studied by use of verbatim descriptions of the interview (40). Since then there has been an increasing amount of research directed toward testing the hypotheses which counselors use as a basis for their practices. The merit of different approaches and methods in counseling continues to need experimental verification.

Protocol research has used a wide range of units to group client and counselor statements for the purpose of analysis. Early investigators in this area were principally concerned with studying single speech units. Both Porter (40) and Royer (50) used the single remark as their principal unit while Snyder (56) and Covner (13) used idea units. Later researchers have also used these briefer units, but have become more concerned with client and counselor behaviors in larger segments of the interview. Curran (14) and others studied the whole interview. Sherman (55) and Carnes (10) were among those who applied the discussion topic unit to typescript research. Fractions of the interview or interview series have been used by Curran (14), Sheerer (54) and others. Comparing and integrating findings arising from this multiplicity of interview

units presents an almost insuperable task. Research contributing to determining which unit is most effective for studying client and counselor behavior and their interaction seems needed.

The investigations just cited have been primarily concerned with studying factors affecting client or counselor behavior that were observable in single speeches. It now seems that, in addition to these limited factors, it is desirable to attempt to tease out some aspects related to larger divisions of the interview. For example, Robinson (47) suggests that the roles that the counselor and client assume as the interview progresses may be a worthwhile approach to furthering the study of counseling interviews. Ascertaining the degree to which some of the larger units facilitate study of such factors as roles would also be of interest to investigators of interview behavior.

One major problem which has arisen in research work on counseling is: What can be done to permit the comparison and integration of the experimental findings obtained in the many studies of the counseling interview? Some of the factors which interfere with any efforts to deal with this problem are the use of: (1) non-comparable units to segment the interview for purposes of analysis,

(2) counseling interviews from a single counseling orientation or center and (3) non-comparable dimensions to conceptualize client and counselor behavior. Investigations which deal with or take into account these factors would contribute to the development of a more uniform data-language in typescript research.

Elton (19) has indicated in discussing the problem of non-comparable units that, "...In counseling research, it may be that the varying results reported by different investigators may be due, to some degree, to the differing interview units which have been selected for analysis. At least the development of a research methodology in the area of counseling is not apt to make additional progress until a comparative study is made of each of the already proposed units of analysis."

Bordin (6) has also pointed out the significance of the comparability problem. In 1950 he said, "One of the big problems of research in this field is our lack of knowledge sufficient to evaluate the comparability of samples of clients and counselors."

It is thus evident that one problem that research in the area of counseling should explore is, Which unit of those presently used or what new type of unit represents the most meaningful organization of statements of the interview? That is, meaningful in the sense that

the unit brings together material which is psychologically related and partitions that which is different. Determination of this unit would facilitate the study of client and counselor behavior and would, if applied to a wide variety of researches, permit comparisons not possible today. Such a unit may also prove useful for getting at complex factors of the counseling interview such as roles. To move toward these goals it seemed desirable to study the relative efficiency of units since it is unlikely that procedures can be developed to transform findings obtained by using one unit into what would have been obtained by use of another unit.

Although this study was primarily concerned with the unit aspect of the non-comparability problem, the other two factors mentioned as barriers to the comparison and integration of research findings, i.e., the use of interviews from a single counseling center or orientation and non-comparable dimensions, were also considered.

In the three chapters which follow there is a detailed discussion of how this problem was studied. The next chapter examines the history pertinent to the problem. Chapter III presents (1) a description of the data used, (2) the procedures used for selecting and

training judges, (3) some characteristics of the units studied, (4) some characteristics of the client and counselor dimensions used in this study, (5) the experimental design for this research and (6) two secondary hypotheses. Chapter IV discusses the results. First, it presents findings regarding: (a) some characteristics of the ratings, (b) the reliability of the ratings and (c) the hypotheses concerning the distinguishing characteristics of units. The discussion section considers: (a) factors associated with the discussion topic and problem area unit, (b) some implications of this research for subsequent study and (c) some limitations of this study. The final section gives the findings pertinent to the secondary hypotheses investigated.

CHAPTER II

HISTORY OF THE PROBLEM

In the preceding chapter the significance of the role of counseling in American society was cited. It was pointed out that the great need existing for this type of service imposed an obligation upon its practitioners to use the methods of science to improve the effectiveness of their function. Experimentation in this area of knowledge was seen as moving rapidly ahead in the last decade, but such research is beginning to be faced with the problem of the comparability of its findings. The different types of units used for studying the counseling conference was seen as one of the barriers interfering with the comparison of findings obtained at different research centers. This investigation was proposed as a beginning attack on that problem. In this chapter the previous research pertinent to this study is considered.

This chapter reviews in four sections the research on interview typescript analysis which is pertinent to this study. The first section presents the high spots in the short history of interview typescript research. The second and major section discusses the units used in typescript analysis in the past. Finally, the historical basis for selecting the unit classification schema and

counseling dimensions used in this study is reviewed.

Brief History of Interview Typescript Research

Early Studies

Little more than a decade ago investigators of the interview began to use verbatim transcriptions of interviews which had been electrically recorded. Until that time, researchers either despaired of the possibility of studying the interview process (and many did) or used admittedly incomplete case reports. The best of these were based on the combination of notes taken during the interview and immediate recall. Lewis' research (32), frequently cited as one of the earliest attempts at the categorization of client responses and the study of these responses within a series of interviews, was based on such protocols.

In 1941, Porter (40) initiated the use of what has come to be known as typescripts in interview research. He developed a checklist of categories to evaluate the directiveness of the counselor's statements. His major contribution was demonstrating that interview behavior, as seen in a typescript, could be conceptualized into useful categories and that these categories could be judged reliably. Soon after Porter's study, Covner (13) compared counselor prepared interview reports with

typed transcriptions of electrical recordings and found that: (1) Less than one-third of the material appearing in a counseling interview is included in the counselor's report. The mean percentages of inclusion of items for four groups of interviewers ranged from 9.7 percent to 35.3 percent. (2) From three-fourths to nine-tenths of the material included in counselor reports is accurate.¹

The Rogerian orientation of many of the early researchers in this field is apparent in the choice of typescripts studied and the questions posed by the following group of investigators. Snyder's work (56) went a step further than Porter's to explore the relationship between counselor statements and immediately succeeding client statements. His research was limited to non-directive counselors, but represented a significant step in interview analysis because he developed methodology which has served as the base for much of the research since that time. Royer's (50) study offered supplementary proof that the process of therapy can be studied by categorizing the responses in an interview,

¹In 1934 Betts (4), as a result of his evaluation of techniques for the study of oral composition, concluded, "For case study research the electrical recording apparatus is the most economical and accurate method of recording the oral language activities."

and that more than 55 percent of the counselor responses during non-directive counseling were responses that followed the characteristic steps in a non-directive counseling situation. Curran (14) attempted further refinement by making an intensive analysis of a single case of twenty interviews which previously had been rated as high in counselor non-directiveness.

Raimy's work (41) is significant primarily because it heralded a rebirth of the self-concept in counseling and applied the typescript analysis approach to the development of personality theory. He used six categories of self reference to study 14 cases (seven successful, four unsuccessful, and three doubtfully successful). In his graphical analysis, he pointed out that, "In successful cases the Positive Self-Reference lines ascended while the Negative and Ambivalent lines descended. This was not true of the unsuccessful cases."

More Recent Studies

Since World War II there has been a resurgence of interest in typescript research. The two centers for this type of investigation have been Ohio State and the University of Chicago. At the former, the emphasis has been primarily on the study of the counseling process with special interest in exploring client and counselor interaction as well as the effectiveness of counselor

techniques. From Rogers' recent text (48) it appears that much of Chicago work up to 1949 was primarily interested in the changes in the client's attitudes and self percepts during counseling. In Chicago Interim Report #3 (53) an extensive research program dealing with the relationships of client dimensions and behaviors both within and outside of the interview is described.

The Ohio State interview research up to 1950 is summarized and discussed by Robinson (45). As part of its discussion of the entire field of student counseling, this text presents the research pertinent to the efficacy of various counseling procedures. For example, after exploring the role of a particular counseling concept such as client talk or resistance, studies were then made to determine what counselor techniques will result in successful conferences according to the internal criteria used. The following brief resume of this research shows that, although the above example is typical, it has not been the only approach used.

The work of Carnes (10) on the extent of counselor flexibility and its relationship to immediate outcomes is typical of Ohio State studies. He used typescripts of counseling sessions involving both professional and advanced graduate student counselors and clients who were students enrolled in a remedial course, Psychology

of Effective Study and Individual Adjustment. Three raters made over-all ratings of the dimension counselor assigning of responsibility for 126 discussion topics of 32 interviews. They also judged the dimension, amount of lead, but on a remark-by-remark basis. Carnes found that, although his counselors varied a great deal on a mean leading, i.e., a directive, non-directive continuum, they were quite flexible in their use of leading. On a scale having a range of 50 points, over half the discussion topics had a standard deviation of 10 points. He also found a wide range of variability in amount of lead between different types of discussion topics. This variability was even more marked for the counselor dimensions, assigning of responsibility. He analyzed the relationship of his three measures of flexibility to immediate outcomes² in the discussion unit. In the over-all series of interviews, he found that flexibility was unrelated to such outcomes. Since the least flexible counselor was really quite flexible, he concludes that his findings merely show that beyond a certain level these measures of flexibility are unrelated

²The immediate criteria of the counseling interview as discussed by Robinson (45, Ch. V). In this instance the variables included are the client's growth in insight, working relationship and the client's acceptance of responsibility.

to outcomes in the interview.

Sherman (55) was the first to make a study using the discussion topic as a unit. She not only demonstrated that this new unit could be used, but also contributed to an understanding of the relationship between counselor techniques and outcomes in the counseling relationships. Allen (1) studied the characteristic development of the topical unit and found that topics have discernible steps or sections of sequential development, and that successful units usually were better developed than less successful units.

Several studies investigated the relations of specific counselor techniques to immediate outcomes within the interview or interview series. Tindall and Robinson (59) demonstrated that silence can be a useful technique in counseling. In many instances the client did add new information, clarify plans or formulate a statement of his problem after the counselor had been silent. They point out that overuse of silence may cause resistance. Daulton (15) devised a checklist for studying resistance in the interview. One of her principal findings was that resistance, as determined by her checklist, is a good measure of working relationship. In a complementary study, Davis and Robinson (16) analyzed the relationship of resistance-reducing

techniques to counselor technique, topic of the unit and effectiveness of the unit. Nelson (37), in her study of "optimum leading" found, "That as the deviations on either side of the optimum become less, (within-interview) outcomes improve...the data indicate that it is as bad (in terms of within-interview outcomes) for a counselor to lead too little as it is for him to lead too much." Optimum lead was defined as, "The degree of leading which should have been present to bring about the most favorable outcomes in terms of the ultimate goals of counseling." In his study of client responsibility, Elton (20) concluded that the primary counselor technique influences the amount of responsibility that the client assumes, but that, even with this effect controlled, differences in responsibility taking still exist and are related to other outcomes.

Studies on the within-interview-series outcomes of counseling were made by Carnes and Robinson (11), Good and Robinson (23) and McCormick (34). The first of these studies found that client talk ratio³ is not justified as the sole criterion of the effectiveness of

³"The proportion of client talk to total talk in the interview." Robinson (45, p. 102)

the interview. Carnes and Robinson felt that it may reflect rather than lead to insight, working relationship and client responsibility. Good and Robinson found that the correlation between either of two measures of feeling and a composite immediate criterion they used were approximately zero. In his exploratory study of immediate and delayed criteria in study skill interviews, McCormick discovered a high relationship between immediate and delayed behaviors within an interview, but found, on the other hand, that immediate conference behavior was not related, in his cases, to the out-of-conference behavior which was measured by reading and work-study skills test results.

Bugenthal (9) and Kauffman and Raimy (30) extended the research on the self concept begun by Raimy. The first study used the Conceptual Matrix Method, a variation of the PNAV Method of Raimy, to investigate varied subjects and types of protocols. The latter study found a strong positive relationship ($r=.838$) between the Discomfort Relief Quotient (DRQ) of Mowrer and Dollard and the PNAV findings.

The counseling interview research of investigators at the University of Chicago is described in the recent Rogers text (48). Its emphasis has been, in line with client-centered counseling theory, on the study of the

changes in the client's attitudes toward himself and his associates.

Typical of Chicago research are the six studies described and analyzed by Raskin (42, 43). Seeman (52) determined that counselor responses in Rogerian type counseling in 1947 were more frequently non-directive (85 percent) than in 1940 (62.6 percent). He also noted that negative attitudes predominate at the beginning of counseling and that positive client attitudes, particularly those expressed in the present tense, are stated with increasing frequency at the end of the process. Hoffman (28) derived an index of behavioral maturity from his analysis of client's report of his own behavior. Successful cases, he discovered, increased on this index, while unsuccessful cases showed no change. Stock (58) found positive correlations between client's attitudes toward self and attitudes toward others. Acceptance of self and acceptance of others were positively correlated and increased with the progress of counseling, Sheerer (54) concluded. Haigh (27) found that those clients showing decreased defensiveness during the course of counseling were increasingly aware of defensive behavior, while those showing increased defensiveness showed the opposite trend.

Summary and Discussion

This brief review shows that typescript analysis has been a fertile field of research since its inception a decade ago. There has been a continued interest in the study of client responses and changes in client behavior; this, in recent years, has been studied most ardently by those hoping to develop a personality theory from such research findings. The early interest in counseling techniques continues to be the subject of investigation. There has been some refinement in the formulation and application of behavior concepts to such research. However, with the exception of the Allen (1) and Robinson (47) studies there are no investigations devoted to studying the units used in typescript research.

A reviewer of the research based on typescript analysis could hardly help wonder about the reasons for the diversity of units used. They range in length and complexity from the simple idea unit to the whole interview or even a series of interviews. The reason for this variety of bases for interview analysis is the variety of purposes and the differences in orientation which investigators have had in studying the recorded interview. There is considerable agreement that statements made by the counseling participants can be rated

and that this rating can contribute to analyzing the counseling process. However, there is less agreement upon what grouping of remarks should comprise the unit to be rated for studying the dynamics of segments of the conference larger than the immediate responses of one participant to those of the other.

Although there exists a variety of ways to combine counselor and client expressions for research purposes, analysis of the research done shows that two of the principal ways have been discussion topics or fractions of the interview or series of interviews. These two types of units are considered in detail in the next section, but it seems appropriate to note here that they parallel each other in terms of what they are used for. Despite the fact that both these units have been used frequently and for comparable purposes, there has been no investigation of their relative effectiveness for studying the counseling process. This research was undertaken to contribute knowledge toward an answer of this problem.

The Unit in Interview Analysis

Before starting the discussion of the various units, it might be well to define the concept "interview unit." A unit includes those combinations of words made by the

client and/or the counselor which serve as the basis for rating or analyzing the behavior of either participant, or the result of their interaction. The principal units used have been idea or thought units, single remarks, the whole interview, fractions of the interview or interview series, and the discussion topic. The following historical discussion of each of these interview units presents its definition, its rationale, its previous use and its advantages and disadvantages.

The Single Remark Unit

The single remark unit includes all the consecutive talk of one of the interview participants between the two speeches of the other participant. An occasional exception to this definition may occur in long speeches when the transcriber does not catch that a distinct shift in thinking has occurred (the two parts are usually separated by a pause).

In his pioneer study, Porter (40) used each separate counselor speech as his unit of analysis. He constructed a checklist which was used to categorize the procedures used by counselors in 19 recorded interviews. The 24 categories he used were subsumed by him under five major counselor functions including: (a) defining the interview situation, (b) developing

the problem situation, (c) developing the client's insight and understanding, (d) information giving and explanations, (e) sponsoring client activity or fostering decision making. The rater agreement among three trained raters in categorizing responses was 63% for placing responses in the same broad functional category and 45% for classifying in exactly the same subcategory. Another measure of reliability, the Profile Comparability Index, was $\neq .86$.

In another early study, Royer (50) used the single remark to determine the characteristics of non-directive counseling. In her analysis, "Each speech was considered a response, except in those cases where there appeared a definite break from the type of response given in the earlier part of the speech." Such divisions within remarks were arbitrarily decided by the raters.

Researchers at Ohio State such as Sherman (55), Carnes (10), Elton (19,20) and Nelson (37) have all used the single remark as a basis for categorizing counselor remarks in terms of amount of lead. Besides this, Nelson also had judges determine the optimum amount of lead on a remark-by-remark basis.

Others have used the single remark as a basis for studying a wide variety of conference phenomena. Davis and Robinson (16) evaluated the resistance-reducing

devices of counselors by classifying counselor remarks used in that fashion. Reid and Snyder (44) required seven judges to classify client responses in terms of "feeling" indicated. Assum and Levy (2) used a whole counselee response as the basic unit of their study of the Discomfort-Relief Quotient in a single non-directively counseled case. Stock (58) used the single remark as her basic unit in studying the relationship between the self concept and feelings directed toward other persons. A variant of the single remark unit was the basis for Sheerer's (54) investigation of the relationship between the client's self acceptance and his acceptance of others. She divided the client statements into units for rating, being careful not to distort meaning when separating the units, and using one complete client response whenever possible. Blau (5) categorized client remarks during the initial interview to make a preliminary test of the efficiency of his index for predicting success in psychotherapy.

This unit was used early in typescript research and continues to be used for a number of reasons: (1) It stands out structurally in the transcript of an interview and is thus readily identifiable. (2) It has psychological unity in the sense that it is the expression of one of the participants in response to the other.

(3) It permits a direct study of and provides descriptive information about the kinds of counselor or client remarks used and their immediate effect on the following speech.

It is apparent that the single remark probably will continue to be a useful unit for certain purposes. However, as McCormick (34) has pointed out, it is limited in certain respects: (1) "When a single speech expresses several ideas, classification of them under a single category may be impossible or meaningless. (2) The brevity of single remarks makes it difficult to rate them reliably. (3) The long-range effects of a single remark, or the cumulative effects of a series of remarks are not apparent in the single remark." (4) Also, it should be added that while it can show sequential changes from one remark to the next, it does not show or permit the study of the larger structural organization of conferences.

The Idea Unit

Various investigators have used the idea or thought unit as a means of eliminating one type of classification difficulty arising from the use of the single remark. The thought or idea unit can be described as a consecutive series of client or counselor words which expresses

a complete thought. Dollard and Mowrer (18), Good and Robinson (23), and Kauffman and Raimy (30) found it useful to identify this unit as the clause unit. Independent clauses thus represent separate idea or thought units.

Snyder (56) was the first typescript researcher to use the idea unit as his basic unit. In his study of the nature of non-directive therapy, he extended Porter's approach to study the relationship of counselor remarks to client responses and used the idea rather than the single remark unit. In the rating of the interview typescripts he himself divided the interviews into ideas which were then rated by his judges. Judges categorized both counselor and client ideas in terms of content and feeling expressed. Snyder regarded the problem of determining the boundaries of the unit of material as a significant one. He saw it as determining the boundaries of "ideas" and accepted Covner's (13) statement that these are places which "clearly indicated change in subject matter or attitude of the client's thinking." Since Snyder alone decided on the point of dividing remarks into ideas his study offers no check on the reliability with which idea units can be determined.

The thought unit has been used to study "tension"

or "feeling" in case records and interviews. Dollard and Mowrer (18) found that the thought unit was more reliable than single words or complete sentences. That is, judges could rate the DRQ (Discomfort-Relief Quotient) more reliably with it as a unit than the others. In their study of feeling as a criterion of success in different types of counseling interviews, Good and Robinson (23) also used the thought unit as a basis for determining the DRQ. They found that two judges dividing typescript remarks into thought units disagreed only once in 300 times.

Researchers of client-centered orientation have used the thought unit to study the characteristics of client responses. Curran (14), in his exhaustive analysis of a single series of 22 interviews, placed thought responses into 19 categories. These data were then analyzed for trends from one interview to another and from the first half of an interview to its second-half. Bugenthal (9) used the thought unit in his investigation of the relationship of the conceptual matrix to the self concept. He found that independent judges correlated $r = .937$ in setting the unit bounds. He also found that the evaluation of categories and attitudes was highly reliable. Seeman's (52) replication of the early Snyder study of the nature of non-

directive therapy also used this unit. It was also the basic unit of Barry's study (3) of the relations of verbal reactions to adjustment level.

A number of other investigators used the thought unit as the basic unit, but selected for analysis only those statements pertinent to the questions they were studying. Hoffman's (28) basic unit of behavior reference was "a group of words reported by the client which describes his overt or planned behavior." The unit used by Haigh (27), in his analysis of defensive behavior in the interview, seems most nearly classifiable as an idea unit since he examined typescripts, statement by statement, in terms of the value, concept, or experience expressed. Daulton (15) and Tindall and Robinson (59) also used thought unit variations. The former first identified statements in which the client said anything that seemed to show that rapport was changed and then classified these comments as to type and severity of resistance. This, at times, extended beyond the usual idea unit in size, but in conceptualization seems related to it. Tindall and Robinson in their study of the use of silence, used as their unit "as much of those elements of conversation that directly preceded and followed the pause situation as we felt led up to the occurrence of the pause and were affected by it." In an

attempt to develop a system of classification for the verbalized data in psychoanalytic therapy, Lasswell (31) has suggested the "reference". This is also called a unit of meaning by him, and consists of a sentence which presents the self or not-self as indulged or deprived.

The advantage of the idea unit is that it divides interview remarks in a way which includes only a single expression of feeling or emotion regardless of the length or diversity of statements within the remark. This increases the reliability of rating and makes it useful for studying client expression of feeling or emotion or other aspects of interview behavior that can be isolated in that unit.

As Robinson (47) points out, idea units are not without difficulty. "...Difficulty has frequently been encountered in their determination in actual interview analysis; each worker has usually found it necessary to develop an elaborate Manual of Directions for adjudicating the many points of decision." And, as in the use of the complete single remark, "It is difficult to judge reliably what is happening within a single speech, and while the client's next response indicates his immediate reaction it does not furnish evidence of later effects. Thus a client may at first resist a counselor's

suggestion but later accept it or treat it as his own idea." When one studies the relationship of a given thought unit with what precedes or follows it, as Snyder (56) did, it is almost impossible to determine what idea in the previous remark or whole series of statements the response is being made to. Finally, its very structure makes it unsuited for determining the presence of factors affecting larger segments of the interview.

The Whole Interview

Another obvious possibility as a unit in typescript research is the entire interview. This unit includes all the client and counselor statements during the course of a single conference.

Few researchers, using recorded interviews, have rated the whole interview as the primary unit. In his early study, Porter (40) had his raters make an over-all rating of the interview for directiveness in addition to judging individual counselor remarks. Grant and Grant (24) made independent ratings of the entire initial interview to determine the feasibility of making judgments about therapy readiness.

Quite a few, however, have combined or calculated mean values for the ratings secured from all of the single speech units within a conference. In his study

of the accuracy and completeness of the counseling interview report, Covner (13) compared the phonographic record of the interview with counselor reports. Raimy (41) in his previously cited study of the self concept, tabulated the frequency of each type of statement in each interview. He then graphically studied the trends of the self-reference categories from interview to interview in the series. In a more recent study, Bugenthal (9) made his trend analysis on the same basis. In attempting to devise an index predictive of therapy success, Blau (5) used the combined percentages of three categories of client responses made in the initial interview of counseling cases. Assum and Levy (2) in the study of a single non-directive case rated each response and then calculated the DRQ for the entire interview. Curran (14) used the entire interview to study trends in the series of 21 interviews he analyzed. His data, for the interview as a whole, were a summation of the idea units within each interview. Daulton (15) summated the weighted values of individual resistance ratings for single speeches within each interview to determine its total resistance weighting. For each interview she used, Royer (50) determined the total number of counselor and client ideas classified in each of the categories she used.

The principal advantage of this unit is that it serves as an easily available and readily determined basis for studying the trends of both client and counselor behavior during the course of an extended series of interviews. Another thing which recommends it for this purpose is that it tends to have a definite pattern of organization and seems to show certain developmental trends.

The weaknesses of this unit of analysis are: (1) Its size makes it possible to observe only gross relationships and obscures the significance of many client-counselor interactions within a conference. As Robinson (47) put it, "Over-all ratings of the interview can indicate the general counseling plan used and whether the interview was effective or ineffective, but this approach cannot very well get at the dynamics of how it happened." (2) It does not permit the investigation of the relationship between various counseling dimensions (see below) and discussion topic units. A conference frequently involves a discussion of several problems, e.g., study skills and feelings of inferiority, which have already been shown to involve different patterns of both client and counselor behavior. (3) Finally, it assumes that the single remarks making up the interview are similar enough on the various dimensions of the

counseling situation to be summated. As indicated in (1) and (2) above such an assumption is unwarranted.

Fractions of the Interview

Investigators, almost from the beginning of type-script research, have been interested in exploring the general trends in the counseling process. Although it was of interest to learn about the relationship between counselor remarks and client responses as Snyder (56) did, there was and continues to be interest in problems which such an approach could not attack. That is, interest in: (1) the nature of therapy and remediation process, (2) the changes in the kinds of client statements during the course of counselling sessions, (3) the significance of immediate behavior, e.g., immediate resistance and later acceptance, (4) the development of more reliable measures of interview behavior and (5) the analysis of factors in interview dynamics represented by larger units than single remarks but smaller units than the whole interview, has stimulated researchers to consider larger units of the counseling process.

A rather obvious procedure for analyzing the first two above would be to use the whole interview, but it is rather unlikely that the various interview series one wishes to investigate would each have the same number of interviews. One solution to this problem of

obtaining comparable data from series of unequal length, i.e., cases varying in the number of interviews held, is dividing each series into fractions. This approach was first used by Lewis (32) who divided each series into tenths. Subsequent researchers have used other fractions, but all have made data from interview series of varied length comparable in order to permit assessment of the general trends within the various cases.

Lewis (32) used deciles of the interview to study the changes in the responses of delinquent girls during a series of interviews. After categorizing her protocols she went through a series of steps which enabled her to construct tables and charts based on the average of six cases by decile intervals. In discussing her choice of unit she said, "A tenth of the total course of interviews was accepted as comparable to the same tenth of the course of the interviews in another case. A decile interval was defined as one-tenth of the total course of interviews." In addition she added, "This division into units gave a basis for comparison, but was not regarded as a functional concept derived from the material."

In the series of studies described by Raskin (43), quintiles of the interview series were used by several of the investigators as the basis for combining data

from the series of conferences for ten cases. No particular reason was given for using fifths rather than some other fraction. This grouping procedure may have been used in order to permit the combining of a sufficient number of single speech ratings to obtain reliable ratings for the fractional unit used.

In order to test certain hypotheses about changes in the client's responses during the course of a single interview, Curran (14) divided each interview of the series which he studied into halves and compared the first and second halves.

Two advantages are gained by using fractions as units. Its main value is that it provides a means for combining data from a group of interview series of unequal length. This procedure permits one to investigate, for a group of cases, whether there are trends in either client or counselor behavior at different points in the conference series. For example, Lewis studied the frequency of negative or positive expression in each tenth of the conference series. Another merit of this unit is the unquestionable reliability with which it can be determined.

Although this unit of analysis facilitates the combining of data in series of unequal length and has other advantages as mentioned above, it is not without

difficulties. The major criticism is that it assumes that the statements within an interview are equivalents and therefore additive. As Robinson (47) has noted, "There are episodes of small talk, of questioning the other for information, of making arrangements for next time and, within the body of the interview, of discussing various topics which have quite different dynamics." In addition there may be larger meaningful units affecting client-counselor relations, e.g., topics under discussion. The characteristics of such larger units of the interview would be difficult to study by using fractions which divide the conference or series of conferences without regard for the problems being discussed and may even join the end of one interview with the beginning of another. Further, while summation of smaller units into fractions provides some measure of delayed effects, it does not isolate the subject matter but mixes in whatever occurs within the fraction and excludes relevant material not in that fraction. These characteristics of the fraction as a unit for typescript research suggest that it would tend to obscure many important relationships and show only the most gross dynamic relationships within the interview.

The Discussion Topic

Because of the limitations of both the single speech or idea unit and of the whole interview or fractional unit, Robinson (47) and his students have proposed and used another unit in typescript analysis. This unit, the discussion topic, is defined as all of the counselor and client conversation devoted to a single topic, e.g., client and counselor may spend 10 minutes discussing examination skills and then switch to a discussion of dependence upon the family for the next 25 minutes. As the above examples suggest, this unit is a functional unit, i.e., its boundaries serve to distinguish the segments of the interview dealing with different problems. This type of division is valuable since it represents a meaningful structure which affects the conference behavior, i.e., the client and counselor have been found to vary their behaviors when the discussion topic shifts (10, 20, 45).

To determine the units within an interview, the judges read it through and make a mark whenever the major topic changes. All of the statements between these marks is treated as a unit. The following counselor statement is a typical example of the ending of one

discussion topic and the beginning of a new topic.

Co: If you'll try to outline that chapter before next time, we'll go over it together. (Pause)// Last time you mentioned you were not sure about your vocational plans.

This procedure also sets off from the main body of the interview such special situations as ending the interview and friendly discussions which are held at the beginning of many interviews. These, as Robinson's analysis (47) of earlier Snyder data indicates, stand out as distinct sections within the interview. In addition, the main body of the interview is divided at each shift of problem during the client-counselor discussion.

Sherman (55), who first used this unit in typescript analysis, found that it can be determined with a high degree of reliability. In her study three judges reading a random sample of interviews had a 95 percent agreement as to place of unit division. Though this is an adequate level of reliability, it is somewhat below what can be obtained by an arbitrary division such as fractions. It also requires the time of two or three trained judges to make judgments about the division points.

The discussion topic as a unit has been utilized in research for several purposes. Like other larger units it has been used as a means for summing ratings.

Unlike the fraction, it has frequently been used as the basis for global ratings which are not only more reliable, but permit the study of delayed effects. Research has also been directed toward determining the structure of this unit.

Since Sherman first used the discussion topic many other investigators at Ohio State have made it their primary unit for interview research. Among those using it as a basis for summing ratings of single remarks or thought units have been Carnes (10), Carnes and Robinson (11), Elton (19,20), Good and Robinson (23), and Nelson (37). In analyzing the role of client talk in the counseling interview, Carnes and Robinson determined the talk ratio, i.e., the half lines of client talk to the total half lines, within each discussion topic and related it to various other variables. Amount of lead ratings were summated within the discussion topic in research by Nelson, Carnes and in both Elton studies. Good and Robinson grouped thought unit ratings of positive and negative feelings in order to determine whether differences in the DRQ for three different type discussion topics were significant.

Several investigators have made use of the discussion topic to study delayed or cumulative behavior

within the conference. Various behavior constructs such as working relationship, counselor assigning of responsibility, client insight and client acceptance of responsibility have been judged in this manner since single speech units are unsuitable for such ratings. Sherman (55), Nelson, Good and Robinson, Carnes, Elton and McCormick (34) have rated different combinations of the above variables with the discussion topic as the basic unit. Although amount of lead or the directiveness of a counselor's statements can be rated in the single remark, Sherman, Carnes and Robinson, Davis and Robinson (16) and Elton have found it useful to rate the primary technique, i.e., the most frequently used technique of leading, of the counselor within the discussion topic. Davis and Robinson analyzed the use of resistance-reducing techniques by the counselor within different kinds of discussion topics, i.e., units in which the problem was classified study skill, other scholastic problems, vocational problems and therapy.

Allen (1) has studied the statements grouped into discussion topics to determine whether they had a discernible structure; i.e., whether there is a pattern of development within such topics and the degree to which this pattern is found. On the basis of authority

and empirical analysis of interviews she established the following major steps in the development of a unit: (1) statement of the problem, (2) development of the problem, (3) making of plans, (4) summary. Rating discussion topics within interviews on the degree to which these phases were developed she found that units at the end of interviews possess the greatest degree of development, middle units are as well patterned as the typical over-all unit, and beginning units are less well patterned than either middle or end units. Compared to the least successful units of progress, the most successful units: (1) "Progress at least through the making of plans step and many complete the summary. (2) Tend to occur in the middle or latter half of the conference. (3) Clearly indicate that the counselor is following a developmental pattern. (4) Show client insight present in the most instances."

The advantages of this unit of analysis as compared to others which have been proposed are: (1) The role that the client and counselor assume tends to be consistent during the discussion of a given problem and to change with a change in the topic, e.g., counselors tend to assume more responsibility for the progress of the interview when skill problems are being discussed than when adjustment problems are being

considered. (2) It divides the main body and special parts of the interview, i.e., the introductory and ending parts usually differ from the main body of the interview; any method of division which combines these special units with the discussion proper will confound or obscure relationships that might be found if this was not done. (3) Since it includes all the discussion about a single problem, it lends itself to the analysis of counseling dynamics in terms of the type of learning required to deal with the problem or in terms of the diagnostic category (see p. 48). Such a system provides a common basis for the study of dimensions of client and counselor behavior, e.g., Carnes found that during the discussion of study skills counselors tend to use a greater amount of lead than during the talk about adjustment problems. (4) Study of this unit, like other larger units, enables investigators to observe either delayed or cumulative behavior within the interview. Unlike fractions or whole interviews which are not readily classifiable according to type of problem, the discussion topic permits study of the changes in a particular problem area at various stages of the interview series. (5) The mean rating for dimensions by the topic as a whole are more reliable than single remark ratings.

Among the advantages listed for the discussion topic was its own potentiality for categorization since it, by definition, represents all the consecutive client and counselor talk about a particular problem. Such a procedure provides a link between diagnosis (in this instance classifying the problem) and the counseling process (the relationship between and the changes of interview variables studied). However, it has been difficult to decide upon the categories to be used in the classification of topics.

When she pioneered the use of the discussion topic, Sherman (55) set up the following seven categories: (1) vocational problems, (2) personal problems dealt with on an information basis, (3) therapy, (4) other scholastic problems, (5) study skill, (6) visiting, and (7) unclassifiable. This system was subsequently used by Davis and Robinson (16), Allen (1) and Carnes and Robinson (11). In discussing these early studies, Robinson (45) states that the first four of the listed Sherman topic categories "showed quite similar behavior and might well be grouped together." Subsequently Good and Robinson (23), Nelson (37), Elton (20), and Carnes (10) did research using the topic categories of lack of skill, non-emotional adjustment problems, emotional adjustment problems and special topics.

Their research showed that the breakdown of adjustment problems into emotional and non-emotional categories was not a useful one. As a result, a three-fold classification including skill problems, adjustment problems and special units was used by Elton (19) and McCormick (34).

The problem of classifying topics is still not settled; at the present time Ohio State researchers are weighing the desirability of revising the present classification schema. A preliminary investigation of such a revised system will be made in this research. Though using the same classes and definitions for those classes, it would combine, for purposes of research, those discussion topics falling in the same class which are contiguous, i.e., if the last three discussion topics of an interview fall in the adjustment category they would be combined for the purposes of summing ratings or making global ratings. There is some feeling that there may be smaller units of meaningful structure within the present topic. Another possibility, which has been broached, is that the variations in roles played by the client and counselor during the course of the interview would be a meaningful unit for study. These variations in ways of grouping interview

remarks arises from different bases for meaningfulness, i.e., the topic is a unit of subject matter while the role played would be concerned with the variations in the counseling relationship. If these meaningful structures exist, it is important to find which is most important and in what ways.

Summary

As the above discussion shows, each of the various units of division used in typescript research have both advantages and disadvantages. There is evident, however, an increasing interest in attempting to study additional factors affecting interview dynamics, to study trends in client behavior during the counseling experience, determine the delayed effects of counselor techniques and increase the reliability of the ratings made in such research. In order to enable researchers using typescripts to obtain maximum information about the variations within the interview and the relation of these changes to the kinds of problems the client presents, it is desirable to attempt devising or selecting a unit which has those potentialities. It would also be useful to be able to make comparisons between studies; especially, when these studies have been made over the same typescripts. In the light of these

considerations it is proposed that a special study of discussion topics, problem area units and fractions be made to see: (a) if the several units differ from each other in structure, and (b) to determine which unit division is most useful or in what way each is useful for studying the dynamics of client and counselor behavior.

The next section will present a more thorough discussion of the background of the various methods of classifying client problems in counseling interview research and additional reasons for using the categories suggested by Robinson (45) in this research.

Classification of Client Problems

Since one of the units being assessed is the problem area and the division of the interview into units of that type is based on a problem analysis, it was necessary that we adopt a classification system for categorizing client problems. There are several classification systems to choose from; therefore, it becomes necessary to review those available and select that schema which is best for our purposes.

Bordin (7) has suggested three criteria for testing classification constructs. These are used here as a frame of reference for assessing the three systems

considered below. As he saw it, these constructs should: (1) enable the clinician to understand more clearly the significance of the individual's behavior, (2) be independent of each other and (3) form a basis for the choice of treatment.

Classification by Problem Location

Until recently, the most widely used basis for classifying the problems of students was the locus of the problem, e.g., educational problems and vocational problems. This seems to have been a function of the way clients described their problems and of the division of labor that developed in the field of personnel work. It continues to be the practice, in many institutions, for clients to seek counselors, and counselors to function, on such a basis. The best known example of this classification schema in student personnel work is that proposed by Williamson and Darley (62) which follows:

- (1) Educational guidance
- (2) Vocational guidance
- (3) Personal guidance (including social, emotional and leisure-time guidance)
- (4) Health guidance
- (5) Economic (i.e., financial) guidance

To the extent that the client and counselor can modify the identifiable locus of the problem, this system might be of some value to the personnel worker. When this system is tested by the above three criteria it appears woefully lacking. It helps the counselor little in understanding the client's problem to know that he is having difficulty making the necessary grades since the basis for this might range wide. The interdependence that exists among the classifications in this system was shown by Bordin (7). He found that in 2,000 cases of the Student Counseling Bureau at the University of Minnesota that there was a "high degree of patterning in the occurrence of the problem categories." In only one instance did a single category occur an appreciable number of times by itself; twenty-three percent were classified as having only vocational problems. The next highest class occurring alone, educational problems, included only 1.6 percent of the total. Since the classifications are not independent and fail to present a conceptualization of the client's problem which helps the counselor understand that problem, it follows that these categories cannot form an adequate basis for selecting procedures which will alleviate the client's problem.

Classification by Diagnostic Category

Workers in the field of remedial teaching adopted a causal categorization scheme at least twenty years ago. Instead of simply saying that the student had a reading problem they determined what deficiencies caused the reading difficulty and directed their research toward finding specific treatment content that would remedy these deficiencies, e.g., visual problem and prescribing glasses. This case study approach produces a multiplicity of causes, especially when we consider the many possible combinations, and therefore, requires that there be an equivalent number of treatment plans. Because of this, interest in counseling shifted to developing a more generalized system for categorizing cases.

In discussing the need for a set of diagnostic categories based on psychological constructs, Pepinsky said:

Bordin has urged that the causes of maladjustment be studied, that maladjustments be classified in terms of causes, and that the effects of treatment based on particular classifications be determined. This method of studying treatment hypotheses follows from Bordin's assertion that "there can be no completely definitive demonstration of the differential validity of treatment without knowledge of what we are treating."

The grouping of clients' maladjustments in terms of diagnosed causes does appear to have a

practical justification. As in physical medicine and psychiatry, the justification lies in the efficacy with which we can treat new cases, in our ability to demonstrate that by categorizing maladjustments we save time and energy and reduce the error of our work.

Our first task is to provide a logical basis for the operations involved. A category in our use of the term, may be defined as a generalized conception of how a particular kind of problem is to be identified and solved. (39, p. 10)

He then proposed and studied a set of categories which modified and expanded a classification scheme suggested by Bordin (7). It includes the following categories:

(1) Lack of information

(2) Self conflict

(a) Intrapersonal (the individual's desire to assume a role is frustrated by some aspect of the self)

(b) Interpersonal (the unwanted role is forced on the individual by others who are important to him)

(c) Cultural (the individual's desires are frustrated by cultural factors, as prejudices, mores, etc.)

(3) Choice anxiety

(4) Lack of assurance

(5) Lack of skill

(6) Dependence

To test the reliability and independence of his system, Pepinsky had three judges classify each of 115 case records under one of the above categories. Best agreement in classifying cases was obtained for the lack of assurance category. Of the 39 cases given this classification, three judges agreed on its presence in 12 cases, two judges in 13 cases, and one judge diagnosed as present in 13 cases. He hypothesized that his results did "indicate considerable accuracy of diagnosis for the categories as a whole." (39, p. 33), but also indicated that, "It would be difficult to determine the extent to which the present findings exceed chance expectation." With the exception of 2c, 3 and 6 in the above listing, it was concluded that the categories were relatively independent.

It is to be remembered that these constructs were not developed as categories to be used in classifying interview protocols or segments of such records. Rather, they have been used as a schema for classifying client's underlying problems on the basis of examination of extensive case records. No attempt has been made to apply these concepts to typescript data. Research is needed to show whether they are applicable to such records.

These categories meet the criterion proposed by Bordin more completely than either of the preceding two. Five of the eight were shown to be independent. They are also a step forward toward permitting a better understanding of the client's behavior and serve as a basis for hypotheses about the choice of treatment. The findings regarding inter-judge reliability for them is inconclusive. Despite their merit, their applicability to typescript research and especially to the categorization of interview segments is somewhat doubtful. This judgment rests on the fact that they were developed for analysis of complete case records and have yet to be applied to typescript analysis.

Classification by Type of Counseling Situation

Robinson (45) has set up a classification system which is principally concerned with determining categories of conference conditions which will be instructive in deciding the manner of helping a client. The assumption is made that from time to time something may happen which changes the behavior patterns of the client and counselor, but in between these points of change there is a consistency in behavior due to mutually accepted values. If this is true, these attitudes would represent important factors in counseling and an attempt should be made to understand

them more thoroughly.

From the research done under his supervision, Robinson has determined that the patterns of client and counselor behavior in four types of situations are different. His classification system follows:

- A. Special situations, i.e., visiting, structuring, interview arrangements, etc.
- B. Body of the interview
 - 1. Adjustment problems
 - 2. Lack of skill
 - 3. Immaturity

A comparison of the above schema with that proposed by Pepinsky readily suggests parallels. The first four of his categories are similar to the adjustment problems classification. There is a lack of skill category in both and the Pepinsky Dependence category is similar to Immaturity above. They differ, however, in that the Pepinsky categories are more concerned with the underlying bases for the client's problem while the latter categories are concerned with taking into consideration the dynamics of client and counselor interaction which relate to the client's manifest problems. In terms of treatment methods then, the Pepinsky categories lead to treatment procedures for those clients falling in a given category. The Robinson

schema differs in that it makes it feasible to investigate the social psychology of client-counselor interaction and its effect on client learning.

For a number of reasons the classification of immaturity is seldom used in typescript analysis for the categorization of discussion topics. An important reason is that the client seldom presents his difficulties as arising from immaturity. Since topic classification hinges primarily on the subject being discussed, judges seldom find clear-cut examples of immaturity problems. Instead of presenting immaturity as his difficulty, the client tends to project this into other problem areas. This requires that the counselor or judge consider behavior in addition to the client's words and also have either a long acquaintance or extensive case records to safely make this classification. A further reason for the infrequent classification of immaturity problems is that persons having such problems seldom feel the need for counseling. For these reasons this classification was not included in this research.

In the previous section on units, it was pointed out that researchers using the discussion topic had classified them according to this plan of classification. These interview typescript investigations have

suggested that this system is valuable in that it:

(1) facilitates research by enabling investigators to combine units which are similar in terms of the kind of learning the client engages in and (2) permits the determination of the characteristic properties of certain types of counseling situations, the analysis of their immediate and delayed effects, and their relation to other counseling situations. Therefore, it is proposed to use this classification system in this research, i.e., special, skill and adjustment.

Other Dimensions of the Counseling Interview

Not only must one decide in typescript research upon the kind of unit and classification system to use, but it is also necessary to select the "dimensions" of the client and counselor behavior which might be affected by those unit divisions one plans to consider. These "dimensions" are constructs devised by researchers as means of conceptualizing common characteristics of interview behavior. Most of them assume, as the term dimension suggests, that the behavior concepts are present in a more or less degree, and that judges can reliably rate their presence. Such dimensions are necessary whenever one attempts to study the structure and dynamics of the interview.

One useful way of categorizing the multitude of dimensions used is according to which participants behavior is being assessed. The discussion which follows first takes up the dimensions describing counselor behavior, next those describing client behavior and attitudes, and finally considers the few dimensions which describes behavior of both of the participants.

In this section the many dimensions which have been used in typescript research will be considered. After the concept is defined, its use in research will be discussed, its demonstrated sensitivity considered, and the reliability data concerning it cited. Thus, this section presents sufficient data about each dimension so that a decision can be made whether to use it or not in the present study.

Counselor Dimensions

Interest in and use of concepts describing the counselor's expressions or attitudes dates from Porter (40). He divided his 24 categories into five major groups which included defining the interview situation, bringing out and developing the problem situation, developing the client's insight and understanding, information giving and explanations and sponsoring client activity or fostering decision making. Many of the

checklist definitions he used are now classified under the dimension now called amount of lead. Those which would not fall under that heading are similar to counselor assigning of responsibility. A thorough discussion of amount of lead follows.

Amount of Lead: This dimension refers to the degree which a counselor's remark (a) seems to be ahead of the content of the client's last remark and/or (b) brings pressure upon the client to accept new ideas. A counselor's statement may vary on this variable from simple acceptance of the client's remark to strong urging of the client to pursue a particular course of action.

Soon after Porter's research, Snyder (56) studied the relationship between counselor and client responses. All but four of the 17 counselor categories he used were concerned with the degree of directiveness inherent in the counselor's statement. A comparison of his definitions for these levels of directiveness and the definitions for various categories for amount of lead described in Appendix 1 shows a marked similarity. Using essentially the same categories, Royer (50) studied the extent that interviews with three clients followed non-directive precepts. More recently,

Seeman (52) duplicated Snyder's work.

Before rating amount of lead assumed its present form, several researchers determined the primary technique of discussion topics. A primary technique is essentially the category of counselor technique of leading which is most frequently present in the unit being studied. In her study, Sherman (55) found that four of the 28 counselor categories on her checklist accounted for most of the discussion topics when these were rated for primary technique. These four categories were clarification, tentative analysis, interpretation and urging.⁴ Later, primary technique was a key variable in the investigations of Carnes and Robinson (11), Davis and Robinson (16) and Elton (20).

The present scale for amount of lead is essentially the one developed by Carnes (10) in his study of counselor flexibility. He had a group of 42 judges classify a set of counselor speeches on a 100 point continuum. From this data the median value for each remark was determined and an empirical basis for assigning amount of lead categories at specific points on a continuum obtained. These points served principally as orientation points since a remark falling in

⁴Definitions of these concepts are presented in the MANUAL FOR TYPESCRIPT ANALYSIS, Appendix 1.

a category on the scale might well be either more or less leading than its scale value. Carnes reduced the scale to 50 points for use in his research. Both he and Nelson (37) marked the amount of lead rating for each counselor remark within a discussion topic on a single scale. Elton (19) modified this by making a Single Remark Rating Sheet which permits the assessment and recording of each counselor remark separately.

Each of the researchers using this dimension has tested the reliability with which it is rated. Carnes and Robinson obtained a reliability coefficient for rating amount of lead of $\sqrt{.844}$ (combined ratings of three judges using the Spearman-Brown prophecy formula). Elton (19) found a reliability coefficient between two raters of $\sqrt{.87}$ which became $\sqrt{.93}$ when the combined rating reliability was computed by the Spearman-Brown prophecy formula.

The work of Porter (40), Sherman (55), Carnes (10) and others indicates that lead ratings for either single remarks or entire units vary widely. Both the profiles graphed by Porter and the data discussed by Carnes show that counselors of different orientations not only differ in mean lead, but also tend to use a wide range of techniques, whatever their orientation. The four primary techniques, which Sherman found occurred most

frequently, also ranged from near the bottom of the lead scale to near the top. This wide variation in rating, accompanied by reliability of levels described above, marks this dimension as one of the most sensitive to changes in the interview.

From the above we see that amount of lead or some variant of that concept has been used in typescript analysis ever since the initiation of that research media. It has been used in at least ten investigations using typescripts. In the course of its development, its rating has become increasingly reliable and has continued to be a sensitive measure of counselor statements.

Counselor Assumption of Responsibility: One of the dimensions which was regarded as significant by Bordin (8), Collier (12) and Robinson (45) is the division of responsibility between the client and counselor in the conference. Until Elton's study (20), investigators had assessed the division of responsibility in the interview mainly in terms of the responsibility taken by the client. His research established counselor assumption of responsibility as a significant counselor dimension. It is concerned with "the amount of responsibility which the counselor himself assumes, which he permits the client to assume,

or which he forces the client to assume." In rating this variable one disregards the client's intentions or desires or the resistance of the client to the counselor's structuring.

Although counselor assumption of responsibility is, in comparison to some concepts, a newcomer in the array of constructs used in typescript research, the importance of the behavior which it describes, has been recognized from the very beginning. Three of the five major groupings of checklist definitions of Porter (40) described aspects of counselor behavior which are determinants of ratings for counselor assumption of responsibility. They were: defining the interview situation, bringing out and developing the problem situation, and sponsoring client activity or fostering decision making. Both Snyder (56) and Royer (50) used categories such as forcing client to choose and develop topic, proposing client activity and structuring all of which are related to this dimension. Since Elton's study of client responsibility (20), both Carnes (10) and Elton (19) have used counselor assumption of responsibility as a dimension in their research. The Tindall and Robinson (59) study of the use of silence as counseling technique investigated one special technique used by the counselor in assigning responsibility.

Both investigators using this variable determined its reliability. Carnes (10) found that the reliability coefficients between two raters ranged from $\sqrt{.61}$ to $\sqrt{.75}$. When the Spearman-Brown prophecy formula was applied to take into consideration the fact that ratings used are based on the judgments of three raters, he obtained a correlation of $\sqrt{.844}$. The reliability coefficient between two judges of this dimension was found to be $\sqrt{.81}$ by Elton (19). When this was recomputed by the Spearman-Brown prophecy formula, the correlation obtained was $\sqrt{.86}$.

The sensitivity of this measure is attested to by the range of mean scores obtained for it in the studies of Carnes and Elton. In graphing the means of assumption of responsibility for eight counselors, Carnes shows the least variable of the eight ranged over two-tenths of the responsibility scale, while the most variable ranged over seven-tenths of the responsibility scale. Elton found that the mean ratings (average of three judges) for this dimension varied from 6.50 to 37.45 on a 50 point scale in the sample of discussion topics he studied. In more descriptive terms, the topics analyzed range from discussions in which the counselor assumed complete responsibility to those in which the client was assigned primary responsibility. This considerable

variability in conjunction with reliability of a useful level, indicates adequate sensitivity.

Counselor assumption of responsibility has been recognized as a significant variable since the beginning of interview research, but until recently was incorporated with client responsibility taking. Research by Elton has shown the need for a counselor responsibility dimension. In the two investigations for which it was used it was found to be both a reliable and sensitive measure.

Core: One of the basic orientations in counseling listed by Robinson (45, p. 74) was "Dealing with the core of the client's remarks." Doing this, as he saw it, involved responding to the point that was of most interest to the client. That is, a counselor speaking to the core of a client's remark, is discussing the main idea, either topical or feeling, which concerns the client. This concept is further described in Appendix 1.

In the single instance in which this dimension was used, Elton (19) tested the reliability with which core is rated by applying a test of homogeneity. His chi-square findings showed that the judgments of two raters were correlated and not independent.

The findings of Elton indicate that Core does not

meet the criterion of sensitivity. In most discussion topics, the counselor made core responses 80 to 90 percent of the time. The secondary and new categories of this dimension occurred so infrequently as to be negligible; in only three of 10 counseling relationships did their combined percentage exceed five percent and none of the exceptions exceeded 12 percent.

This counselor dimension has been shown amenable to reliable rating, but seems unsuited for this research because it does not meet the criteria of representativeness, sensitiveness, or long development.

Counselor Sensitivity to Feeling: Bordin (8), Rogers (48) and others have considered the counselor's sensitivity to the client's attitudes and feelings an important counseling dimension. That is, does the counselor respond to the client's feeling about the topic or to the subject matter in the client's statements.

To measure this concept, Elton (19) set up a three category system for classifying counselor statements. These categories provide for the extreme categories on this dimension, content and feeling and in addition has an unclassifiable category. Elton found that it could be reliably rated. His chi-square test of the ratings indicated that the judgments of

two raters were not independent, but correlated at a significant level.

This dimension suffers the same inadequacies as Core. Though there was a wider distribution of the percentage of content or feeling found in different client-counselor relationships, this was limited primarily to discussion topics dealing with adjustment problems. In most instances the number of counselor responses falling in the content and unclassifiable categories exceeded 80 percent. Unless a more precise definition of the content-feeling dimension is developed, i.e., one which differentiates the unclassifiable category, it is not likely to meet the sensitivity criterion set here.

Feeling as a counselor dimension was deemed unsuited for this research for much the same reasons as Core. Although it apparently can be rated reliably and is considered by some to be an important variable, it has not been adequately developed nor proven sensitive to changes in client-counselor behavior.

Client Dimensions

Researchers, who first studied client behavior in the counseling conference, developed category systems which have since been developed into variables which have dimensional characteristics. Lewis (32) developed the initial set of client categories. Soon afterwards Snyder (56) developed a more thorough set of categories

for client expressions. Some of these were: statement of the problem, asking for advice or information, acceptance of clarification of feeling, understanding or insight, answers to questions, rejection of clarification, discussion of plans, ending of contact, free discussion, not related to problem and unclassifiable. This great variety of categories, 19 in all, was also used by Royer (50) to study the process of non-directive counseling. Both found that the results of an objective analysis of non-directive counseling follows the steps non-directive counseling would presumably take. With the exception of Seeman's work (52), which duplicated Snyder's study, there have been no additional applications of this category system.

The above client categories are of historical interest only. Since they "pigeon-hole" client behavior rather than deal with it as involving behaviors which are present in a more-or-less degree, they are unsuited to the purposes of this investigation.

Working Relationship: Among the categories used by Snyder (56) were several, such as acceptance of clarification of feeling and rejection of clarification, which presaged the dimension working relationship. This concept is even more akin to that older interviewing construct, rapport. From these relationships it becomes

apparent that working relationship is concerned with the degree that there is mutual respect between the participants and the client feels free to present his problems to the counselor. The factors considered in rating this variable, which also further define it, are found in Appendix 1.

Sherman (55) developed a five-point scale for categorizing working relationship. She found that it correlated $r = .75$ with growth of insight and skill in the unit and was related ($r = .43$) with responsibility taking. In her study, working relationship was significantly influenced by the primary technique (see p. 57) of the counselor. This dimension has since been used in researches by Carnes (10), Carnes and Robinson (11), Elton (20), Good and Robinson (23), Nelson (37) and McCormick (34). Daulton (15) and Davis and Robinson (16) have studied resistance in the interview, an important aspect of working relationship. The rating scale for this dimension was modified by Carnes to permit ratings along a 50 point continuum. It is used in this form today.

This variable has been rated with increasing reliability since it was developed for rating purposes by Sherman. For 26 interview units she obtained perfect agreement with another judge 69 percent of the time; differed one degree, 12 percent; varied two degrees, 15 percent

and differed three degrees once or four percent. Degrees of difference refers to the number of categories, on a five category continuum, which one rating of a unit was distant from a second rating of it. The reliability correlation coefficient obtained by Carnes between pairs of raters ranged from $\sqrt{.52}$ to $\sqrt{.62}$. Application of the Spearman-Brown prophecy formula to these findings, yielded a correlation coefficient of $\sqrt{.807}$. This is justifiable since the rating used for research purposes is obtained by averaging the judgments of three raters. McCormick obtained a reliability coefficient of $\sqrt{.91}$ between himself and another rater on working relationship.

There is some indication that working relationship is sensitive to changes within the interview. Sherman's data show that the discussion topics in her study ranged from category two to five on a five point scale. In terms of her checklist this means that the relationship varied from being "somewhat resistive - rejects counselor point of view or suggestion but in a polite manner, does not talk freely..." to one in which the "counselor and counselee work together on a real problem, talk freely..." Similar variability was found in the studies of Carnes and Robinson, Good and Robinson and Elton.

Working relationship has a long history as a dimension in counseling research and has gone through several stages of development before being rated as it is at present. Since Sherman formulated a scale for rating working relationship there have been six additional studies done using this dimension. Each of these studies has found sufficient reliability to warrant use of the dimension and there is some suggestion that it is sensitive to changes within the counseling relationship.

Client Assumption of Responsibility: Among the counselor dimensions considered above was the counselor's assumption of responsibility. The client counterpart of this is client assumption of responsibility. To rate the client's statements on this construct, the judge considers the extent the client takes an active role in clarifying his problem and participates in making plans for dealing with it.

Such a variable was anticipated by such categories as planning and detachment as used by Lewis (32) and the asking for advice or information and discussion of plans categories of Snyder (56). This dimension, as it is now used, was first devised by Sherman (55). Responsibility taking, as she called it, ranged along a five category continuum from situations in which the

counselor assumed practically all responsibility through various proportions of sharing responsibility to those circumstances in which the primary responsibility for the interview's progress rested with the client. It was used in that form by Carnes and Robinson (11), Good and Robinson (23), Nelson (37), and Elton (20). As the result of Elton's work, which suggested that responsibility would be better conceptualized as two dimensions, Carnes (10) developed a counselor and a client scale for this dimension. This client scale is called Client Interest in Assuming Responsibility. As its name suggests, it is concerned with how much responsibility the client desires to assume. Rating instructions make a special note of telling the rater to discount the influence of counselor technique upon the amount of responsibility that the client actually assumes. It has been used in this form by Carnes and McCormick (34).

In testing the reliability of this dimension, Sherman obtained perfect agreement 65 percent of the time and a difference of one degree between the raters the remaining 35 percent of the time. Her ratings, it should be remembered, were done on a five category scale. In the Good and Robinson research the median reliability of the inter-correlated ratings of three judges was $r = .73$. This, when corrected by the Spearman-

Brown prophecy formula to take into account the combining of three ratings to obtain a mean value for research purposes, yielded a reliability coefficient of $\sqrt{.890}$. On a sample of the study skill units used in his research, McCormick obtained an inter-rater reliability coefficient of $\sqrt{.88}$ when using the scale devised by Carnes. These data indicate that this dimension is susceptible to reliable rating.

This dimension has proven quite sensitive to variations in the interview. Inspection of the array of responsibility ratings found by Elton (20, pp. 47-48) shows that each of the categories one through four is represented by a sizable proportion of the discussion units and that there are a few in the extreme category five. Good and Robinson do not cite separate data on client assumption of responsibility, but do report the variability of the composite criterion which included it. This measure varied from 10 to 44 in 94 units on a 50 point scale.

Client assumption of responsibility qualifies for use in this research on the basis of its frequent use, continued development as a rateable concept, sensitivity to interview changes and acceptable reliability.

Client Growth in Insight: Another client dimension, which was considered important even before present

research methods were developed for studying the counseling relationship, is client growth in insight and understanding. This variable deals with "evidences of new insights, gains in skill, self acceptance and plans." (10)

Both Lewis (32) and Snyder (56) had client categories in their research which dealt with this area. Following Snyder, Curran (14) used four insight categories in his interview content analysis. The scale, which is now used by Ohio State investigators interested in this aspect of client behavior, was first devised by Sherman (55). To study what she called Changes in Insight or Skill During a Unit, she developed a five category scale. This scale was also modified by Carnes (10) into a 50 point continuum which is shown in Appendix 1. The same scale Sherman used was incorporated into the investigations of Carnes and Robinson (11), Elton (19), Good and Robinson (23) and Nelson (37). Carnes' later variation of this scale was also used by McCormick (34).

Variations of growth of insight were used in several other studies. These include Stock (58), attitudes toward self; Sheerer (54) acceptance of and respect for self; Seeman (52), understanding and insight; Hoffman (28), maturity of behavior and Haigh (27), defensiveness. In his analysis of the effects of knowledge about the

client on judges typescript ratings, Dial (17) rated the client statements on the variable Client Progress, i.e., progress toward learning to deal with problems being considered in the topic, a concept apparently concerned with much the same behavior as growth in insight.

Although growth of insight has long been considered a desired immediate outcome of the counseling relationship, it has proven more difficult to rate reliably and accurately than many other dimensions. Robinson (45) cites several factors making it difficult to determine accurately. They are: (a) the tendency for clients to say they have been helped or to develop plans primarily to please the counselor; (b) the difficulty of assessing acceptance remarks and (c) the inability to make a judgment about the client's unverbilized insights or plans. On rerating his judgments after a two-month interval, Curran (14) found that his two sets of categorizations, including insight ratings, did not differ significantly. The insight dimension was the least reliably rated by judges in Sherman's study (55). There was perfect agreement on a five point categorization of insight 42 percent of the time. The median inter-correlation among three judges obtained by Good and Robinson (23) for this dimension was $\neq .66$. The

Spearman-Brown correction for grouping the three ratings yielded a correlation of $\sqrt{.856}$. McCormick (34) obtained a correlation coefficient of $\sqrt{.90}$ between himself and another judge on this variable. The median reliability phi coefficient obtained in Dial's work (17) was $\sqrt{.45}$.

Growth in insight has evidenced about the same degree of variability as working relationship and client assumption of responsibility, but does not meet the criteria of sensitivity because of the difficulties arising in rating it.

Growth of client insight or understanding has been used as a dimension by at least 10 researchers if we include its early variants. The reliabilities obtained have varied from low levels to those at acceptable levels. This same unreliability makes its variability a function of chance factors to a greater degree than is desirable in a sensitive measure.

Feeling: Another client dimension which has been used to study interview typescripts since the beginning of such research is feeling. Various techniques have been used to attempt to measure the emotional tone of a client's expressions. They have been used to measure such things as "percent of remarks containing expressions of feeling, average intensity of feelings expressed, ratio of negative to positive feelings, frequency that

expressions of feeling are directed toward self, others, or the environment." (45, p. 97)

Both Lewis (32) and Snyder (56) included categories assessing the client's feelings in their rating schemes. Raimy (41) rated positive, negative and ambivalent feelings toward self. Others who explored various aspects of the feeling dimension have been Curran (14), positive and negative emotional responses; Reid and Snyder (44), feeling; Raimy, Kauffman and Raimy (30) and Blau (5), positive, negative and ambivalent self concepts; Assum and Levy (2) and Good and Robinson (23), Discomfort-Relief Quotient; and Barry (3), perceived strength and perceived threat. Many investigators have rated different types of feeling reactions. However, there is little agreement among them concerning either the name of a representative construct or how it should be rated.

The levels of reliability obtained for rating feeling reactions varies with the definition used by the researcher, but in several instances is relatively low. Snyder found that categorization of responses for feeling content was the least reliable of the client responses he rated. On a sample of 356 responses Raimy found that three of four judges agreed in placing responses in one of six categories from 50 to 82 percent

of the time. The two poorest rated categories were positive and ambivalent feelings toward self. Reid and Snyder determined that when 15 judges were asked to classify remarks into one of 25 categories that a remark was placed in five different categories, on the average. The percent of feeling remarks and Discomfort-Relief quotient obtained by different raters for discussion topics was found to correlate from $r = .659$ to $r = .896$, varying with the type of problem in the Good and Robinson study. Combined ratings, which were used in the research, were calculated to possess reliability coefficients ranging from $r = .794$ to $r = .945$. These data do not reflect the reliability of rating single remarks, since both DRQ and percent of feeling remarks are measures of total incidence of two categories within a topic.

The data of Good and Robinson on the percentage of feeling remarks and mean Discomfort-Relief Quotients suggest that these measures are moderately variable. The mean percentage of feeling remarks for lack of skill topics was 14.50; for emotional adjustment topics it was 31.92. The standard deviation of this measure indicates that a considerable number of topics of all types evidenced little or no feeling remarks. Mean DRQ's for different type topics was .531 for skill topics, .636 for non-emotional adjustment topics and

-.774 for emotional adjustment topics. The difference for these findings are statistically significant at either the one or five percent level, but the marked variability of skill topics indicates that they overlap non-emotional adjustment topics considerably on this variable.

In summary it appears that though investigators have been interested in rating feeling reactions they have not been able to agree on just what they are trying to measure and have had considerable difficulty rating reliably those measures which they used. It would also be difficult to select a feeling reaction concept which would be representative.

Other Client Dimensions: There are several client dimensions and categories which have been used in the past, but which describe only occasional remarks or which have been designed for purposes unlike those of this investigation. These are mentioned only briefly below.

As part of his research evaluating brief psychotherapy, Morton (36) devised a rating procedure which required judges to make global judgments about the adjustment of the client in 14 areas, e.g., mother, father, etc. and one overall judgment of adjustment. These ratings were made on a six-point scale which

considered the intensity of the conflict and the generalization of effect shown by the client for a given area. This was shown to be a useful method for studying changes in adjustment, but is too broad in scope for use with the units used in this research.

Two other specialized dimensions, devised to predict the client's chances for success in therapy, based upon analysis of the initial interview, were developed by Grant and Grant (24) and Blau (5). The Grants propose a concept which they call "therapy readiness", i.e., the client's attitudinal set at the start of therapy which makes him more or less amenable to counseling. They demonstrate that rating interviews globally or by means of six-sub-dimensions yields essentially the same rating, but have not empirically established the usefulness of the construct therapy readiness. Blau devised a predictive index for the same purposes as the above. Each client remark was rated by the Raimy (41) PNAV formula for both valence and direction. The predictive index is the positively valenced Self percent plus the ambivalent valenced Self minus the negatively valenced Self; each of these values is in terms of percent of total client talk. Reclassification of the remarks ten days later by the investigator yielded 95 percent agreement. This

predictive index correlated with therapy success $r = .90$. It also is interesting, but unsuited to the purposes of this study.

Seeman and Dymond (53) report a number of unpublished studies carried out at the University of Chicago which factor analyzed Q sorts based on three aspects of self and others: present self, desired self and the "average person". These sorts were made at different stages in the counseling process and may yield an interesting set of client dimensions. The Chicago Interim Reports describe many other studies which are concerned with new client dimensions.

Both Dimensions

Two dimensions have been used which involve the evaluation of both client and counselor remarks at the same time. One of these, silence, has been already considered in conjunction with the discussion of the counselor's assigning of responsibility (see p. 59). The other is talk ratio.

Talk ratio, as defined by Carnes and Robinson (11), is the ratio of half lines of client talk to the total half lines within the unit. It is neither a representative nor frequently used dimension. In its only use, Carnes and Robinson found that it was positively correlated

with client responsibility taking and working relationship; it was also related to client growth in insight, but only in vocational problem units. They concluded it was not justified as the sole criterion of effectiveness of an interview and may reflect rather than lead to other outcome dimensions. Because it is somewhat related to all the dimensions selected for use and is not a representative dimension, it was rejected for use in this investigation.

Summary

From the above discussion of the many dimensions which have been used in typescript research, it is apparent that not all of them are suited for use in this investigation. Thus, it is necessary to select a sample of dimensions appropriate to the problem under study. The appropriateness of a dimension for this research is based on several criteria. A first consideration is representativeness, i.e., the degree to which the concept used, or similar concepts, is typical of the possible dimensions. Another criterion is the level of development of the concept, i.e., the extent it has been found useful in the past. A third criterion of particular import for this research is the sensitiveness of the dimension. By this we mean its potentiality for

showing changes in client and counselor expression independent of any unreliability of rating. A final criteria, which is almost assured by the previous standards, is that the concept be amenable to reliable rating.

The previous discussion of the various dimensions shows that some of them meet the criterion for use in this research better than others. In addition to selecting variables which were appropriate, according to the criteria, it was thought desirable to use both client and counselor dimensions. With these two considerations in mind, two client and two counselor concepts were chosen. The counselor dimensions deemed most suitable were amount of lead and counselor assumption of responsibility. To measure client expressions, working relationship and client assumption of responsibility were regarded as most useful for the purposes of this research.

CHAPTER III

DESCRIPTION OF DATA, DIMENSIONS AND METHODOLOGY

Students of the counseling process have had many different purposes in investigating interview behavior as seen in interview typescripts. They have been interested in studying problems which fall, for the most part, in the following categories: (1) client and counselor responses and their interaction; (2) factors influencing client and counselor responses and (3) the variations and changes occurring during the counselor-client relationship. To study these problems investigators have used different kinds of units. It seems to follow that research studying the effectiveness or meaningfulness of a unit must test the degree which such a unit facilitates the study of the problems mentioned above.

The discussion of the methodology of this study is divided into five sections. First, the data are described. Second, the selection and training of judges participating in this research is outlined. Third, the three methods of grouping remarks and the classification schema used are considered. Fourth, the dimensions of counselor and client behavior rated in this research are discussed. Finally, the experimental

design for this study is presented in detail.

DESCRIPTION OF DATA

In order to permit the widest possible generalization from this study, interviews were selected which represent not only different counselors, different clients and different type problems, but also different institutions. These interviews came from a pool of verbatim typescripts prepared from electrically recorded interviews held at the University of Chicago, University of Minnesota, University of Missouri and Ohio State University. The cases recorded are said to be a representative sample of the kinds of problems and counseling methods found at these institutions. To insure that a specific counseling orientation or type of student problem was not given undue weight in this study an equal number of interviews was used from each institution. The main purpose in having the various schools represented was to make possible the widest possible generalization from the findings and not to permit comparison of the practices from one school to another.

Three consecutive interviews for each client and counselor pair were used. Wherever possible cases were selected which represented a complete series of

interviews, i.e., the initial to the final conferences inclusive. In instances where this was not feasible because some conferences in the series were not recorded or the series extended into six or eight interviews, e.g., Minnesota and Ohio State, a group of three contiguous interviews within the same case was used. Since there is a tendency for the initial counseling interview to follow a pattern, the samples from the longer series were taken from the later interviews.

The decision to use series of interviews rather than single conferences was based on a number of considerations. First there was reason to believe that a single conference might not be representative of the particular counseling relationship. Another reason was that groups of interviews would subject the units studied to a test comparable to the previous research which had used them. That is, fractions, such as quintiles, have been used to analyze trends in interview behavior during interview series. Thirdly, the analysis of series permits some tentative exploration of the trends in counseling behavior during successive interviews. Finally, the initial interview tends to follow much the same pattern from one counselor to another.

Table I
Interview Sample Used in This Research

<u>Institution</u>	<u>Inter- viewer</u>	<u>Case No.</u>	<u>Interview Number</u>	<u>Total Int. by Inst.</u>
Minnesota	A	38700	2-4	
	C	38705	3-5	
	D	33728	1-3	<u>9</u>
Chicago	Z	Boo	1-3	
	V	Sketch	1-3	
	X	Will	1-3	<u>9</u>
Missouri	B1	1	1-3	
	Di	6	1-3	
	Sh	13	1-3	<u>9</u>
Ohio State	F2	F23	3-5	
	F3	F33	2-4	
	F5	F55	5-7	<u>9</u>

The sample cases and interviews used are identified in Table I above. The total pool from which these were selected is found in Appendix 1. The first entry in Table I indicates that Interviewer A(a code designation)of the University of Minnesota is the counselor for case number 38700 which is represented in this research by the second through the fourth interviews.

Though the interviews making up the pool from which the sample studied was drawn, are said to be representative there are several selective factors which need to be mentioned. First, it is likely that some selection entered into what material was recorded and duplicated, e.g., recordings which include a great deal of unintelligible material are usually not transcribed. Secondly, the characteristic nature of services rendered at one counseling center is frequently unlike those at others represented here. For example, most clients coming to the Chicago counseling center seek assistance in dealing with their personal problems; at Minnesota they usually seek aid in educational and vocational decisions. This difference is discussed in greater detail below. In addition to these two selection factors which are true of most typescript research, this research imposes two additional criteria which, though they have not been shown to be related to client-counselor behavior, are selective factors. The demand that cases include three consecutive transcribed conferences eliminated many series of the pool. The additional criterion of different counselors for each series further narrowed the possibilities for inclusion. The above comments indicate what selection factors may be operative in this study. They do not mean that the sample used was

selected on any of the variables studied in this research.

Since it is of some interest to know the general characteristics of counseling cases from each institution, a brief comment about each is given below. These comments are based on observations made during the reading of the typescripts in the pool. While there was wide variation in method and types of cases at each institution and much overlap, these remarks help show some of the general differences at each institution.

On the basis of the typescripts obtained from the University of Chicago, the writer got the impression that the counseling there dealt with functional, non-incapacitating problems. In their description of the work being done at the Chicago Counseling Center, Grummon and Gordon (25) state that the problems dealt with include, "...The usual maladjustments ranging from what might be called the superficial and situational to more deep seated personal disturbances. A large proportion of clients fall into the normal range." Though occasionally a client broaches a study skills problem, the counselor continues, in such instances, to deal with the feelings and attitudes of the client. There is a tendency for the counselor to spend more time defining the relationship in these conferences than in those of Minnesota or Ohio State. All the counselors

adhere closely to the client centered counseling method. Snyder's casebook (57) includes protocols typical of counseling done at Chicago.

Minnesota cases, as a result of that school's early emphasis on testing and vocational counseling, tend to center on vocational and educational problems. Typically, the client presents his problems and selects, with the counselor, the tests he thinks might help him secure information about himself. Subsequent interviews hinge on a discussion of the implications of the test results and a synthesis of the data available during a systematic discussion of the choices which the client has in mind. Counselors at Minnesota do not appear to hold to any single method. A description of the counseling program at this school can be found in Williamson (61).

The following description of the services of the University of Missouri Counseling Bureau is an excerpt from a bulletin of that institution.¹

Students may consult the counseling Bureau for help in choosing a department in which to major, for aid in selecting a vocation, or for help in solving various personal problems. The

¹Personal communication from Dr. Paul C. Polmantier, Director, University Testing & Counseling Services, University of Missouri.

counselor works with the student individually to assist him in integrating the personal, educational, and vocational aspects of his life into a pattern which is meaningful and satisfying to him.

Counseling may include personal interviews in which the particular situation of the student is discussed; psychological and educational tests, individually selected for the student; and interpretation of information about occupations.

The Ohio State typescripts were prepared from counseling sessions held with undergraduate students enrolled in the course, The Psychology of Effective Study and Individual Adjustment. The counselors were either faculty or advanced graduate students in the student personnel area. As a regular part of the course, students were assigned to work one day a week for about 30 minutes with a counselor. These students presented a wide range of problems, but most frequently discussed their study skill problems. Typically the client spent early conferences dealing with study skill problems, but as the relationship continued adjustment problems were brought up for consideration. Here also, no one counseling method was used. A thorough discussion of this counseling program is presented by Robinson (46).

Thus, the basic data for this research comes from four different large universities which differ, in varying degrees, both in terms of counselor methods and

client problems. These schools have extensive counseling programs which cannot be regarded as necessarily typical of what is done throughout the United States. However, typescripts from these institutions represent a better sampling of counseling, as it is done in American universities, than one could obtain from interviews of a single school.

SELECTION AND TRAINING OF JUDGES

Persons participating in the ratings of typescripts for this research were advanced graduate students in the student personnel program at Ohio State University, or had equivalent background and experience. With the exception of the dimension which is rated remark-by-remark, i.e., amount of lead, each of the variables is rated by three judges. Because of the greater reliability of amount of lead ratings these were done by two raters.

All judges studied appropriate sections of the MANUAL FOR TYPESCRIPT ANALYSIS (Appendix 1) prior to engaging in any rating. In addition, they were thoroughly acquainted with the Robinson text (45) which describes at length the concepts used in this research. Before doing ratings which were to be used in this

research, each judge made practice ratings over materials previously rated. After completing the rating of a sample of units or remarks, he then checked his judgments against the original ratings and noted any discrepancies. These were also checked with a more experienced rater. If this process showed that the new rater still lacked a thorough grasp of the rating principles outlined in the manual, he undertook additional practice ratings.

The task of rating the typescripts for the 12 cases was divided in the following manner. Two raters made a complete set of global judgments for the whole group of interviews. Three-fourths of the third set of ratings for the global dimensions was made by a third judge; a fourth judge assessed the remaining part of the third set. These last two raters made ratings on all three types of units for the interviews they rated. The single remark ratings of amount of lead were made by six different judges. Here again, however, each judge rated them for all three types of units.

Two procedures were used to avoid the introduction of systematic variance aside from that inherent in the material being rated. Where different judges were used an attempt was made to spread their work so no systematic error would be introduced. This seemed especially

desirable for the global ratings, but was deemed less necessary for the relatively well-defined and highly reliable amount of lead ratings. To make global judgments each rater first worked through an entire case making judgments on one type of unit, then reread the case making ratings on the second type and finally went over the typescript again to make ratings of the third type of unit. In order to minimize any advantages that the same serial order of rating might obtain for any one of the three types of units, these were rated in a counterbalanced order. That is, the three types of units, in the course of being rated three times in each case, were rated in each of the three serial positions and at varied positions in the whole set of cases.

The nature of the actual rating process is given in the following sections.

UNITS AND THEIR CLASSIFICATION

This research centers on the analysis of the relative effectiveness of the larger units which have been used in typescript research for combining data, i.e., discussion topics, problem areas and fractions. The characteristics of each of these and the classification system used in categorizing the problem type is discussed below.

Discussion Topic

The discussion topic represents all the conversation of both client and counselor about the same topic or subject. The subject matter of a conversation is, in most instances, readily identifiable. Shifts in topics are also easily noted. Thus, during the course of a conference, if a client begins talking about his inability to handle examinations, turns to his lecture note-taking procedures, and then goes on to his difficulties in reading textbooks, each of these is a separate discussion topic.

The several advantages of the discussion topic have been discussed at length in Chapter II. They are, in brief: (1) the client and counselor role tends to remain consistent during the discussion of a single topic; (2) it separates special phases from the main body of the interview; (3) it permits study of delayed or cumulative behavior within the interview and (4) the ratings obtained are more reliable than those based on single remarks.

From the above statements, it follows that the three judges who marked off discussion topics read the typescript and placed marks at those points which indicate a change in the subject under discussion. After the judges independently demarcated the discussion topics of

an interview, they met as a group to discuss any differences in assigning division points. Whenever the three judges concurred the division point was accepted without further question. When only two judges agreed, the basis for the division was briefly discussed, but, in most cases, the break agreed upon earlier by two judges was used. If none of the judges concurred, the division points were determined by discussion among the judges. Thus, all the remarks between two marks in the typescript constituted a discussion topic unit. For identification purposes, each of these units was numbered. A combination of this number with the interview notation permits singular identification of every unit in the pool, e.g., F111-3 represents the third discussion topic of the first interview between counselor F1 and client F1-1.

In the process of setting of the various topics within the interviews, it was found that judges do occasionally differ on the points of demarcation. The two major sources of these differences were: (1) the sub-division into separate topics by one judge what another judge regarded as all one topic and (2) the setting of the demarcation point when there were transitional remarks which might logically be considered part of either topic. One should note that, in

both these instances, differences arose because of different views as to what was meaningful or what was the subject of discussion. It is particularly likely that one judge may group together what another judge regards as several separate topics because they are inter-related in such a way as to suggest a higher order label, i.e., they are a meaningful combination in themselves. The relevance of this will be seen in discussing the next type of unit, "problem area".

Units which are designated as E (for extraneous) were not rated on the dimension amount of lead. They were marked as E in the typescripts which the judges rated. Although they were relatively few and these were all brief, the raters were told not to consider this material in making global judgments. Short units, i.e., those less than eight remarks or 12 lines in length, were not used in obtaining ratings for halves of a unit. It was felt that such ratings would not be reliable enough for use in this investigation.

To determine the reliability with which discussion topics are delineated the demarcations of three judges were examined. In the case of Minnesota, Missouri and Chicago interviews, the middle interview from each of the sets of three interviews in each case was used. In the case of the Ohio State interviews earlier data

reported by Carnes is used. In rating 52 discussion topics from nine interviews, the following results were obtained:

	Percentage
All three judges agreed exactly or within the same remark	31
Two judges agreed exactly or within the same remark	52
Only one judge marked a transition point	17

In the instances in which complete agreement was not obtained, 33 percent of the judgments fell within three remarks of the consensus point. An additional 36 percent of those instances involved a splitting of a discussion topic into two or more units by one or two of the judges.

Since the data on the demarcation of discussion topics in the Ohio State interviews used in this investigation are no longer available, the findings of Carnes (10, p. 49) on 11 unmarked interviews is reported. His results were:

	Percentage
All three judges agreed exactly or within the same remark	43
All three judges agreed within a range of three remarks	12
Two judges agreed within a range of three remarks	10
Only one judge marked a transition point	35

Though the reported findings show some disagreement in separating discussion topics, the group discussion which preceded the final setting of division points when there was any disagreement enhanced the actual reliability of the division. This notion of increased reliability through group discussion of ratings is based on a comparable situation in personnel rating, i.e., a jury of raters discussing the rating to be assigned a worker.

Problem Area Units

All the contiguous statements dealing with the same kind of problem make up what is called a problem area unit. These are readily ascertained after discussion topics are marked in the typescript books since they are groupings of discussion topics dealing with the same class of problem. That is, two or more adjacent discussion topics which fall in the lack of skill category combine to make a problem area unit. Thus, the number of problem area units within an interview is a function of the number of shifts in type or category of problem discussed, rather than the number of separate subjects discussed as with the discussion topic breakdown.

The use of such a unit is suggested by Robinson's

research (47). On the dimension he used, amount of lead, several types of problems, which had been categorized separately, showed similar interview behavior. On the basis of these findings Ohio State investigators adopted the threefold classification system described in Chapter II, i.e., skill and adjustment problems, and special discussions. These developments suggest that one seeking to study changes in either client or counselor behavior or their interaction might profitably study typescripts by organizing and assessing discussion falling within a given problem area. It also suggests that the client-counselor relationship may remain much the same during the discussion of each kind of problem area and tend to change when the kind of problem under discussion changes.

In the process of classifying discussion topics, the judges noted that during the interpretation of test scores the roles of the client and counselor seemed to differ markedly from the roles assumed in the above three categories. Therefore, it was considered desirable to add a fourth class, test interpretation, to the schema for setting off problem area units. Units were placed in this class when the topic was concerned primarily with the presentation of test results, i.e., those instances in which the counselor told the client

what his test scores meant. Where the test discussion involved both participants and was integrated with the client's decision-making, the topic was classified in the adjustment category. The need for this new classification arose from the inclusion of typescripts from Minnesota and Missouri in this research; typically, counselling at these schools includes the selection and, subsequently, the interpretation of tests. This research may, in part, establish whether such a classification is essentially different from the three others used in this study.

This new type unit seems to have advantages in addition to those that obtain for discussion topics. Material grouped in this manner seems more meaningful since it comprises remarks related in terms of a psychological construct rather in terms of a subject construct. Although discussion topics are classified by problem area, there remains in such a division the assumption that adjacent discussion about the same type of problem has different characteristics. If such is not the case, it adds little to our knowledge of counselling to evaluate each set of remarks dealing with a single subject. It would, in that case, be more valuable for us to explore the characteristics of problem area units in order to arrive at a better

understanding of counseling dynamics, i.e., the interaction between client and counselor and the changes in their statements during the counseling process.

One final advantage is the increased reliability with which these can be divided as compared to discussion topics. As indicated above, the two most common bases of differences in the division of discussion topics are grouping of related subjects and the setting of boundaries in transition material. Both of these would be considerably less with the problem area unit. Units dealing with different type problems would be regarded as related in relatively few cases. Furthermore, transitional statements, which occur primarily between two topics dealing with the same type problem, would not have to be partitioned.

The demarcation of problem areas involved three operations. The first of these was the setting of the limits for discussion topics which has been described in detail above. In the second step judges classified each discussion topic as to whether it dealt with skill, adjustment, test interpretation or special situations. As a third step all contiguous discussion topics of the same type were then lumped together as one larger "problem area", i.e., one that emphasized a relatively constant client-counselor relationship.

For example, an interview composed of five topics classified in the following way: mother (adjustment), chemistry (skill), note-taking (skill), girl (adjustment) and roommate (adjustment) would become three problem area units.

To find the reliability of partitioning this unit, the judgments of three raters in determining the shift in topic at the problem change points were compared. This differs from assessing the reliability of discussion topics in that it is concerned with demarcating the points where the type of problem changes rather than every change of topic. Since no previous researchers have used the problem area unit, there is no previous evidence for reliability. In the same sample used to examine the reliability of discussion topics the following results were obtained:

	Percentage
All three judges agreed exactly or within the same remark	46
Two judges agreed exactly or within the same remark	39
Only one judge marked a transition point	13

Twenty-two percent of the judgments not in complete agreement fell within three remarks of the demarcation point decided upon in conference. An additional 30 percent of this group arose from the fact that judges sometimes

split units which were subsequently adjudged single units when the judges conferred. Here again, information about the demarcation of the Ohio State interviews used is not available.

There was relatively high agreement among judges in the classification of problem area units, e.g., adjustment, skill, etc. For 87 percent of the units three judges concurred on the classification for a unit. The remaining 13 percent showed agreement between two judges.

Fractions

As a unit, the fraction comprises a definite percentage of the total remarks within a series of counseling interviews. For example, if a given series included 750 statements and the investigator wished to study each fifth of the series, he would set off units in terms of successive 150 remarks.

This procedure assures high level reliability in setting points marking off units, but has the obvious flaw of neglecting the meaningful organization of talk within an interview and assuming that the remarks made in one fraction are comparable to those in another. As indicated above, the assumption of homogeneity breaks down when one considers what Robinson (45) has called the special category. Snyder (56) and Robinson (47)

have demonstrated that these special situations involve a client and counselor relationship which differs from those units within the body of the interview. Since divisions into fractions combine the talk about different kinds of problems and incorporate special situations into the material studied, one would predict that this would obscure data which can be determined by using more meaningful units.

In this study the interview series were divided into quintiles. The only reason for using this fraction was that it has been used by investigators at the University of Chicago to study interview series similar in length to those included in this research. That is, this research was interested in comparing fractional and topic units as they have been used.

The reliability of determining these fractions is obviously so high that there is little point in actually testing the reliability. That is, it was a simple matter to count off a number of remarks and make a mark. After the quintiles were marked, later rating of remarks showed no significant errors in these divisions.

Summary

In the discussion above the characteristics of each of the units was considered. Two of them have been

used in previous research while the other, problem area units, were first used in this investigation. In terms of factors such as size and purpose they were found comparable. Since making comparisons involves studying the behavior within each unit, it was necessary to use concepts describing such activity. The four dimensions used in this study are considered in detail in the next section.

OTHER COUNSELING DIMENSIONS

As indicated in Chapter II, research on the counseling process and its dynamics has used a variety of concepts to categorize counselor and client expressions. Since this research is interested in determining whether there are structural consistencies within the units under study and to what degree adjacent units differ, it necessarily needs to assess the units in terms of some of these concepts. Because investigators have used a wide variety of dimensions, it is necessary to select from these those which are most appropriate.

On the basis of the criteria described in Chapter II, four dimensions were selected for use in this research. The two concepts describing counselor remarks used were amount of lead and counselor assigning of responsibility. To assess client behavior working

relationship and client assumption of responsibility were chosen. Each of these will be considered in turn. There will be a discussion of the nature of the concept, the procedures for rating it, and the reliability with which it has been rated.

Amount of Lead

This concept refers to the degree which a counselor's remark: (a) seems to be ahead of the content of the client's last remark and (b) brings pressure upon the client to accept new ideas. Remarks vary on this variable from simple acceptance statements which lead hardly at all, to those which urge the client to take a particular course of action. A thorough discussion of the different categories of amount of lead can be found in Appendix 1, p. 250.

In the rating of this dimension the judge examined each counselor remark in the light of the just preceding client statement. Keeping in mind the two factors mentioned above, he made a judgment which located the remark on a 50 point continuum. This graphic rating scale, developed by Carnes (10), is reproduced below. The judge used the descriptive statements as points of orientation, but since a statement may appear to lead more or less than the modal value of the category in

which it is judged to fall, his rating may also be expected to vary.

Silence	Acceptance	Clarification	General Lead	Tentative analysis	Urging
.....					

Raters placed their ratings for this dimension on the Single Remark Rating Scale (Appendix 1). When rating amount of lead on this sheet the rater placed numbers at the left edge of the form corresponding to the numbers of the remarks which he was rating. Each remark was rated; unclassifiable remarks are checked in that category. Each rater prepared a separate set of sheets for each interview rated.

The determination of the mean amount of lead value involved essentially the same steps whatever unit one studied. After the numerical value of each rating was determined, the rating sheets were marked to indicate where the units used separated parts of the interview. To determine the mean rating for a unit, the average

rating for all remarks in that unit, with the exception of those marked unclassifiable, was computed. In this research, the amount of lead rating for any given unit was obtained from the average of two mean ratings found in the manner described above.

Amount of lead proved the most reliably rated dimension of the four used in this investigation. Although six different raters participated in making judgments on amount of lead for the various sets of interviews used in this research, the lowest of the reliability coefficients obtained for the three different units was $\neq .86$. When this is corrected by the Spearman-Brown prophecy formula to take into account the combining of two ratings the correlation becomes $\neq .93$. Additional discussion of the reliability of this variable can be found under the discussion of Hypothesis 1 in the next chapter.

Counselor Assumption of Responsibility

As the name suggests, this dimension is concerned with the degree to which the counselor assumes for himself or assigns to the client responsibility for dealing with the problems presented by the client. Further definition is incorporated in the discussion of rating procedures below and can also be found in Chapter II

and Appendix 1.

In rating this dimension the judge evaluated the counselor expressions for the entire unit being rated rather than assessing each remark. He was concerned only with the degree the counselor generally took over or threw responsibility. In deciding where to place his rating on the graphic rating scale illustrated below, he considered only the intent or purposes of the counselor and disregarded the reaction of the client. This scale, one of the dimensions on the Unit Analysis Rating Sheet (see Appendix 1), also ranges on a 50 point continuum. At the midpoint and other intermediate points are brief descriptive phrases for use as guides by the rater. These descriptive statements and their positions on the continuum were determined rationally.

.....				
Counselor assumes complete responsi- bility	Counselor primarily responsible	Responsi- bility evenly divided	Client assigned primary responsi- bility	Client assigned nearly complete responsi- bility

The ratings for this variable in each of the three units used in this research were made in the following manner. Each rater assessed the statements of the unit

in an over-all fashion. From this evaluation he decided what point he should mark on his rating sheet. The checks of each of the three raters were then translated into numerical values, each dot representing a value between one and fifty. The mean of the three ratings for a unit was the value used in this research.

The reliability coefficients for ratings of this variable were found to range from $\sqrt{.63}$ to $\sqrt{.86}$ in the nine pairs of inter-rater correlations obtained. When corrected to take into account the use of three raters (by using the Spearman-Brown prophecy formula) the reliability coefficient for discussion topics was $\sqrt{.90}$, for problem areas $\sqrt{.86}$ and for fractions $\sqrt{.94}$. The reliability of rating this variable is given further consideration in the next chapter.

Client Assumption of Responsibility

This dimension is concerned with the degree of responsibility which the client is interested in assuming. It is concerned with the extent the client plays an active role in clarifying his problem and making plans for dealing with it. It is the client counterpart of Counselor Assumption of Responsibility described above. However, these responsibility dimensions should not be conceived as variables which divide the responsibility "pie" since it is possible for both rating to be

either high or low. Further discussion of this rating is found in Chapter II and Appendix 1.

As indicated in Chapter II and Appendix 1, judges rating this dimension discount the influence of counselor technique on the amount of responsibility which the client assumes. That is, they are rating the amount of responsibility the client himself wishes to assume and subtracting the influence of counselor throwing responsibility, i.e., using techniques which place responsibility with the client, or taking for himself responsibility during the conference.

In rating this dimension the counselor uses the scale below which is part of the same rating sheet used for Counselor Assigning of Responsibility. The complete Unit Analysis Rating Sheet is found in Appendix 1.

.....			
Client desires	Q ₁	Md	Q ₃ Client desires
no responsi-			to assume a
bility or even			great deal of
rejects it			responsibility

This scale has descriptive labels for its extreme points, but uses the notations Md, Q₁, Q₃ for intermediate points on the hypothesized continuum. With this scale the counselor is evaluating client assumption of responsibility of a given unit in terms of the range of

such behavior which he has observed in other interview units from counseling situations with college students. If the unit is judged about average in comparison to the rater's "norm group" it is marked at the median. Should it be regarded as better than 75% of the "norm group" units on client assumption of responsibility it is marked at the Q_3 point. The rating of other units is done similarly.

In this research, ratings on Client Interest in Assuming Responsibility varied in reliability from one type of unit to another and also according to the pair of raters. In the nine sets of rater comparisons, the correlation coefficients ranged from $r/.28$ to $r/.66$. The corrected coefficients (again using the Spearman-Brown prophecy formula to take into account the use of three raters to obtain an average rating) were $r/.72$ for topics, $r/.71$ for problem areas and $r/.73$ for fractions. The problem of rater reliability is also considered in Chapter IV.

Working Relationship

Good working relationship or good rapport is generally accepted as a necessary ingredient of an effective conference. As indicated in Chapter II, an effective working relationship involves a feeling of mutual respect

between the participants and a willingness on the part of the client to talk about the problems which he regards as important. At the other extreme of this variable, the client is resistive or apathetic.

This dimension is also rated on the Unit Analysis Rating Sheet (see Appendix 1) on a 50 point rating scale which is shown below.

.....
 Poor Q₁ Md Q₃ Excellent

The polar points on this rating tool are labelled to denote the extremes of this variable. Intermediate points which show the median and quartile points on this hypothesized continuum are further guides for raters. As in the case of Client Interest in Assuming Responsibility, the judge makes his ratings in terms of behavior observed in other interview situations with college students. With this scale the rater is evaluating the working relationship of a given unit in terms of the range of relationships which are included in his "norm group".

The reliability coefficients obtained for the

rating of whole units on this variable ranged from $r = .39$ to $r = .58$. When corrected by the Spearman-Brown prophecy formula to take into consideration the combining of three ratings, the following reliability coefficients were found: topics, $r = .71$; problem areas, $r = .69$ and fractions, $r = .77$. The ratings of halves of all three types of units was done about as reliably as for the entire unit. The inter-rater reliability correlation coefficients ranged from $r = .35$ to $r = .58$ and the corrected correlation coefficients ranged from $r = .67$ to $r = .78$.

In one special approach (testing hypotheses 1b and 2) ratings of unit halves were correlated on two dimensions. To determine this value for amount of lead the rating sheet was marked to coincide with the division of halves in the typescript book and the mean value for the half was computed in the same fashion as for the whole unit. To rate working relationship by halves, the rater read the remarks within the unit to the point marking the half division point for that unit. He then made a judgment over the first half. After reading to the end of the unit, he made a rating of the second half and at the same time made his global judgment. For both variables, the value used for this study was an average of the judges ratings.

In the above discussion of the counseling dimensions

used in this research, the dimension was defined; the method of rating it described; the procedure for combining individual ratings on it for the same unit indicated and finally, the inter-rater reliability found in this investigation presented.

Now that the data, units, classification system and counseling dimensions used in this research have been discussed at length, it is in order to consider in detail the experimental design of this study. The next section does this.

EXPERIMENTAL DESIGN

The question which suggested this investigation was: Which of the three methods of dividing interviews, fractions, discussion topics, or problem areas, gives the greatest evidence of meaningful structure being present? This section discusses in detail the methodology used in attacking this question. First there is a discussion of the assumptions made in setting forth hypotheses to test. Secondly, the hypotheses formulated are discussed. Thirdly, the statistical procedures are detailed. Finally, there is a description of the procedures used to make some qualitative observations concerning the data.

Assumptions

The following statements are the presuppositions which are basic to the several hypotheses below. As such, they are not the subject of research themselves, but rather are accepted on the basis of previous research and logical analysis. The discussion following each statement makes the assumption more explicit.

1. During the course of the counseling process the attitudes and behaviors of both client and counselor vary.

By this we mean that both the client and the counselor not only vary the character of their statements, but also assume different roles. Roles or attitudes are defined here as internalized states which are inferred from the patterns of response consistency found in the overt behavior of the participants. Since the data for this research were interview typescripts, it follows that the behavior under study, and from which inferences about attitudes were made, was the recorded speech of the participants in the counseling sessions.

2. These attitudes and behaviors can be satisfactorily rated.

Within the definitions of attitude given

above, this assumption is supported by the considerable evidence which shows that investigators have rated various constructs in typescript research with satisfactory reliability.

3. There are factors associated with these changes, some of which are yet to be determined.

This means that the variations in behavior and attitude which occur are related to other, for the most part unnamed, constructs and structural characteristics of the interview. That such relationships exist (in the sense that they also may be conceptualized) is presumed because behavioral change cannot be regarded as a spontaneous event.

4. The unit which makes these changes more discernible will be more effective, for the purpose of describing counselor and client behavior and the factors associated with it, than the unit which obscures them.

This holds that, among the different ways of ordering or grouping statements of a conference or a conference series, the method which permit investigators to differentiate inferred attitudes is more likely to lead to a determination of the factors associated with changes

which occur. Since the scientific study of phenomena has as its main purpose the prediction and control of events, it seems reasonable to the writer to assume that a methodology which opens the way to the differentiation among phenomena is more efficient than one which does not permit such analysis.

5. While changes taking place during the interview series will cause differences between widely separated units, the crucial test of the differentiating value of a particular unit is its capacity to reflect changes in attitude between contiguous units in an interview series.

Units which are widely separated in time may differ because of factors extraneous to the immediate counseling relationship. At least, that possibility cannot be ruled out. Adjacent units, in which these extraneous factors are minimal would, therefore, be a crucial test of the differentiating value of a particular method of organizing counseling speeches. The differentiating value of a unit, for our purposes, is the degree to which its organization of interview behavior results in ratings which differ more from one unit to the next

than those ratings obtained by using other types of units.

Hypotheses

The following statements are null hypotheses which this research proposes to test empirically. Following the statement of each hypothesis there is presented its alternative and a brief discussion of its significance for this research.

1. The reliability of rating the various dimensions does not differ significantly from one type of unit to another. The two testable hypotheses, derived from this more general hypothesis, which this research examines are:

- a. There is no difference between the inter-rater reliability coefficients obtained for the three different types of units.
- b. The units do not differ in the degree which the first half of each unit correlates with the second half.

As an alternative to (a) it is predicted that those units which organize interview statements into a meaningful combination are more amenable to reliable rating for interview behavior constructs than those units based on an arbitrary division of interview statements. It is

predicted that the rating reliabilities for problem area units and discussion topics will be higher than those for fractional units.

As an alternative to (b) it is predicted that discussion topic and problem area units have greater internal consistency than fractional units. That is, the correlations between the ratings of first and second halves of discussion topic and problem area units will be significantly larger than comparable correlations for fractions.

The writer believed that the alternatives suggested for hypotheses 1a and 1b would hold since there is a tendency for client and counselor to maintain a consistent pattern of behaviors during the discussion of a particular problem or type of problem. Also, that behavior pattern tends to vary when the subject or the type of problem varies. In the case of fractional units, however, the statements are grouped without regard to subject or type of problem. This appears to bring together a variety of behaviors which would fall at varying levels on the dimensions being studied. Thus, the job of the rater was seen as relatively simple and clear-cut in the rating of units where the statements all fell at pretty much the same level. However, since a fractional unit is likely to include diverse elements,

it was thought it would be rated less reliably than the other two units.

If the above alternatives hold, it can be construed as support for the major hypothesis of this research, i.e., that problem area and discussion topic units are more meaningful and useful for interview typescript research than fractional units. This conclusion is based on the inference that units which can be rated with high reliability tend to incorporate within their bounds relatively homogeneous behavior. In this context, greater homogeneity would suggest that the more reliably rated unit has greater psychological unity than those less reliably rated.

2. There is no difference between the correlations obtained between ratings of two halves of a unit and those obtained between ratings of the last half of a unit and the contiguous half of the next unit.²

The alternative predicted is that in discussion topic and problem area units the ratings of within-unit halves will correlate

²This hypothesis differs from 1a and 1b in that it is primarily concerned with the degree material within a unit differs from that of adjacent units. On the other hand, the reliability hypotheses, 1a and 1b, deal with the internal consistency and homogeneity of statements within a unit.

to a significantly greater degree than those of between-unit halves. It is also predicted that comparable findings for fractions will not be significantly different.

Should the research testing this hypothesis reveal that the ratings of two halves of a unit correlate with each other more than they correlate with contiguous halves of the adjacent unit, it could be construed as evidence for the existence of meaningful structure within the unit being studied. In the light of the previously stated assumptions (1, 4, and 5), a unit which organizes statements in the above manner will be more effective for the analysis of the interview than one which fails to do so. Such a unit not only brings together related facets of the conference, but also partitions segments of the interview in a manner which makes variations in client and counselor behavior or attitudes more readily discernible.

3. Observable shifts in client and counselor behaviors and attitudes from each unit to its next adjacent unit do not differ when the difference scores of the three units under study are compared.

The alternative is, of course, that the degree of shift from unit to unit is greater between series of discussion topic and between

series of problem area units than between fractions of the interview series. It is predicted that the difference scores, i.e., the absolute value obtained when the numerical rating of each unit is subtracted from the rating of its adjacent units, is significantly larger in problem area and discussion topic units than in fractions.

If the alternative hypothesis is warranted it suggests that the problem area and discussion topic are more effective than fractions in making variations in the client-counselor relationship discernible. It would also, in terms of our fifth assumption, indicate that units with larger difference scores are more sensitive on the crucial test of contiguous units.

Statistical Procedures

In this section, the steps taken to put the above hypotheses to test are discussed in detail. Each hypothesis is considered in turn. A step-by-step statement of the methods employed in testing the hypotheses follows a brief statement of the objective of the procedures.

Hypothesis 1a (see p. 117): The purpose of this

analysis was to determine and compare the level of agreement obtained by different judges rating each of the four dimensions within three different kinds of units. The steps followed to make this test are listed below:

1. Determine the mean amount of lead for each type unit.

Two judges have previously rated each counselor remark in each conference included in this investigation. The demarcation points of each of the three types of units is made on the completed Single Remark Rating Sheet. After the numerical values for each rating is noted, the mean rating of each type unit by each of the raters is then computed.

2. Determine the global ratings of working relationship, counselor assumption of responsibility and counselee assumption of responsibility for each of the three raters and on each of the three units.
3. Correlate the ratings of the two judges on amount of lead; also inter-correlate the ratings of the three judges on each of the global dimensions.

This yields an inter-rater reliability

coefficient for each pair of raters on each of the three units and on each of the four dimensions.

4. Calculate the significance of the difference between the reliability coefficients obtained for the same dimension on the three different types of units.

The procedure for doing this is set forth in Lindquist (33, pp. 214-218). After transforming the correlation coefficients into z's through the use of a table, the procedure is essentially the same as that for determining the significance of the difference between uncorrelated means.

Hypothesis 1b (see p. 117): The assessment of the internal consistency of ratings of client and counselor behavior within the various units via this hypothesis involves correlating the ratings obtained for the first and second halves of units.³

³Since there is a significant relationship between the counselor dimension amount of lead and the counselor's assumption of responsibility, and also a significant relationship between working relationship and client assumption of responsibility, it was felt that for testing both hypotheses (1b) and (2) amount of lead and working relationship would suffice. They meet the several criteria for use of a dimension in this research as well or better, in some instances, than the other two variables and have the additional merit of not being correlated to the same extent as the other two. Because of the above considerations it was regarded desirable and sufficient to use these as representative of the sample of dimensions.

The steps followed to test hypothesis 1b were:

1. Determine the mean ratings for amount of lead for each half of each type of unit.

This involves first computing the mean rating of each rater for each half of a unit. The average of the two means for each half-unit was then computed.

2. Determine the working relationship ratings on each half of each unit.

Since three judges make this rating, the rating value used is the average of the three ratings.

3. Then correlate the ratings for the first and second half for each of the three kinds of units.
4. Test the significance of the difference between the correlations obtained from step three.

For each of the two dimensions, the correlation coefficient obtained for discussion topics was compared with that found for problem area units and fractions. A third comparison between problem area and fractional units was also made. The procedures for making this test are the same as in step four of hypothesis 1a.

Hypothesis 2 (see p. 119): The steps followed in testing this hypothesis are as follows:

1. The first three steps of hypothesis 1b supply the basic data for testing this hypothesis.
2. Correlate the average ratings for the second half of each unit and the first half of the following unit. This was done for both working relationship and amount of lead.
3. Test the significance of the difference between the correlations obtained for within-unit halves and between unit-halves for each type of unit.

The procedures described in step four of hypothesis 1a were also used in making this test.

Hypothesis 3 (see p. 120): The following procedures were used to test the hypothesis that observable shifts in client and counselor attitude between successive units do not differ with the various kinds of units.

1. For each unit of the three different types of units determine the average rating of each dimension.
2. Compute the difference scores between successive units of each type.

Such scores are simply the numerical difference between adjacent units without regard for sign. Although it is not the main object of this

research to compare the sets of difference scores within each series, the statistical test to be used calls for this type of analysis as an intermediate step. Findings obtained from this series-by-series analysis may suggest hypotheses concerning variations of client-counselor behavior at different institutions or in dealing with different kinds of problems.

3. Test the hypothesis that the difference scores for the different types of units are equivalent by using the U test (35).

Each set of ratings of a single interview series constitutes a single experiment. For example, to test whether the difference scores for fractions and discussion topics are equivalent, one first combines the difference scores for both these units in rank order. For identification purposes one writes difference scores for topics with lead pencil and those for fractions with red pencil. Since the alternative to be tested is that topic difference scores are greater than fraction scores, one counts the number of pencilled numbers following each red number.

The sum of these is the U statistic. The significance of this is determined by procedures described by Mann and Whitney (35). If the total units involved are twenty or less, however, the significance can be read from tables.

While the 12 interview series represent 12 replications of the experiment, the general significance of the data as a whole can be determined by computing the U statistic for the total experiment. To combine the data of the 12 series it is necessary to compute the variance of each series and determine the variance for the entire 12 series. One can then determine whether the three sets of difference scores for each of the dimensions differ significantly from each other.

Secondary Questions

Since this study is probably the first interview typescript research which includes samples of interviews from several universities, it presents an opportunity to examine some questions about differences among institutions. However, since the study was not primarily designed to investigate such differences, the sample of

interviews used here, i.e., 12 series --- three each from four schools permits only a tentative statement concerning the characteristics of counseling interviews at different centers. Recognizing these limitations upon any findings obtained, it was thought desirable to test some questions about differences among the interviews from various universities. Testing these secondary questions was expected to yield additional information about the sample of interviews studied and also, possibly, suggest hypotheses which subsequent researchers might study intensively.

Distribution of Types of Problems: One of the questions which has been of some interest to students of counseling has been: Do institutions differ in terms of the types of problems their counseling service deals with? Some investigators have used checklists or summarized descriptive statements of the activities at different institutions in seeking information on this question. To the writer's knowledge, none have used typescripts as a basis for investigating the above question.

On the basis of observations made during reading typescripts from the several institutions, the following question was suggested. In counseling interview typescripts from four different universities do the

distribution of types of problems considered differ significantly from one institution to the next?

The data were tabulated by school and type of problem to permit a ready examination of the pattern of problems found at each institution. Although there are only three cases from each university, it was thought that making the tentative assumption that the cases are representative of those occurring at each school, and that the units within each interview series were randomly representative of units at such schools would permit a tentative exploration of this problem by means of a chi-square test. Since we cannot be sure that these two assumptions are true, the findings here will only be suggestive. Chi-square was applied to determine whether the distribution of problem types differed significantly. The classification schema previously described in this chapter (see P. 96) was used to categorize 174 discussion topics.

Counselor and Client Dimensions at Different Schools: There is also some interest in knowing whether the rated behaviors of counselors and clients at different centers differ. Information was available about the counselor dimensions, amount of lead and counselor assigning of responsibility, within the various

interviews making up the sample for the major problem of this study. Similar information was available for client acceptance of responsibility and working relationship, the two client variables considered. The question considered was: Do the mean and standard deviation of each of the four dimensions rated differ among the four institutions represented in the pool of typescripts studied?

Since an analysis of each of the twelve cases indicated variation among the three cases for that institution and quite a bit of overlap between institutions, it was deemed undesirable to attempt more refined statistical analysis. Therefore, the examination of this question was restricted to a presentation and brief discussion of the mean and standard deviation statistics of the four dimensions for each case and each institution.

Summary

This chapter has described in detail the data, dimensions and methodology used in this research. It presented the: (1) basis for selection and the general characteristics of the interview typescripts used, (2) procedures used in the selection and training of judges, (3) nature of the units being investigated,

(4) definitions, rating procedures and reliabilities for the four dimensions of client and counselor attitudes rated and (5) assumptions, hypotheses and statistical procedures involved in this research.

In the next chapter the experimental findings are presented and their significance discussed. In addition to systematically considering each of the hypotheses making up the main part of the investigation, certain secondary findings regarding the data are presented.

CHAPTER IV

RESULTS AND DISCUSSION

Introduction

This chapter presents and discusses the significance of the findings of this investigation into the relative effectiveness of discussion topics, problem areas and fractions as units in typescript research. It considers: (1) some characteristics of the ratings, (2) the reliability of the ratings, (3) the two reliability hypotheses, (4) the two hypotheses concerning distinguishing characteristics of units, (5) a discussion of the following three points which bear on this investigation: (a) the factors associated with discussion topic and problem area units, (b) implications of this study for subsequent research and (c) the limitations of this study and (6) the findings obtained from examination of the secondary questions.

Some Characteristics of the Ratings

Since some of the statistics used in this research involve the assumption that the data being studied are normally distributed, e.g., tests of significance between means, it was felt desirable to obtain some picture of the characteristics of the ratings made on the three types of units. These findings are shown in Table I.

Table II

Mean, Standard Deviation and Range of
Ratings on Four Dimensions Obtained When
36 Interviews Were Rated by Discussion
Topic, Problem Area and Fractional Units

<u>Dimension</u>	<u>Unit*</u> <u>Type</u>	<u>Mean</u>	<u>Standard</u> <u>Deviation</u>	<u>Range</u>
Amount of Lead	D	19.3	9.7	6 - 45
	P	19.4	9.8	6 - 43
	F	15.0	8.8	9 - 41
Counselor Assumption of Responsibility	D	21.9	11.5	4 - 48
	P	21.7	11.2	4 - 48
	F	19.7	12.6	5 - 48
Client Assumption of Responsibility	D	25.6	8.0	10 - 46
	P	25.9	8.5	10 - 46
	F	23.2	7.8	14 - 45
Working Relationship	D	24.4	7.8	11 - 45
	P	24.1	7.6	12 - 45
	F	20.5	8.7	12 - 45

* D - Discussion Topic (N = 174)
P - Problem Area (N = 106)
F - Fraction (N = 60)

This table, which shows the mean, standard deviation and range values of the ratings used in this research, reads as follows: The first row of data indicates that the dimension amount of lead when rated in discussion topics had a mean rating of 19.3, a standard deviation of 9.7 and a range of six to 45 on a 50 point continuum. The data suggest that the distribution of ratings for client assumption of responsibility and working relationship are very close to those expected of a variable which is normally distributed. Both amount of lead and counselor assumption of responsibility show some positive skewness; however, this was not regarded as so marked as to warrant modification of the statistical procedures used.

The findings shown in Table II are also of interest as gross descriptive measures of the ratings obtained. The table indicates that the ratings made with the three different type units rather closely parallel each other in terms of the mean, standard deviation and range values obtained for the four variables rated. It suggests that any differences obtained among the units studied are not a function of any considerable disparity of the distribution of ratings.

Table III

Intercorrelations of the Mean Values
of Four Dimensions for Each of Three
Units of the Interview

Discussion Topics (N = 174)

	CoR	CeR	WR
L	-.67	-.46	-.41
CoR		.68	.72
CeR			.68

Problem Areas (N = 106)

	CoR	CeR	WR
L	-.62	-.38	-.38
CoR		.69	.71
CeR			.69

Fractions (N = 60)

	CoR	CeR	WR
L	-.78	-.47	-.61
CoR		.64	.79
CeR			.77

Key to Abbreviations:

L - Amount of Lead

CoR - Counselor Assumption of Responsibility

CeR - Client Assumption of Responsibility

WR - Working Relationship

Table III shows the table of inter-correlations of the combined ratings obtained for four dimensions. There is a separate matrix for each type of unit. In the table for discussion topics, the first line is read as follows: amount of lead is correlated $-.67$ with counselor assumption of responsibility in 174 units, correlated $-.46$ with client assumption of responsibility and $-.41$ with working relationship. The rest of the table reads similarly.

From Table III two things are apparent. (1) Inter-correlations for each type of unit follow pretty much the same general pattern. That is, the raters seem to have held the same orientation toward rating the several dimensions whether the units being rated were discussion topics, problem areas or fractions. (2) The four dimensions are related. This would be expected since two of them reflect counselor methods and the other two client outcomes. Not only are these pairs linked to a certain degree, but it is also evident that counselor techniques have a relationship with client outcomes. However, the intercorrelations obtained are low enough to show the need for the four measures used.

Table IV

Reliability Coefficients for the Combined
Ratings of the Several Judges for the Four
Dimensions on Three Different Type Units

(Spearman-Brown Prophecy Formula Used to
Take into Account Combined Ratings)

<u>Dimension</u>	<u>Discussion Topic</u>	<u>Problem Area</u>	<u>Fractions</u>
Amount of Lead	.93	.93	.96
Counselor Assumption of Responsibility	.90	.86	.94
Client Assumption of Responsibility	.72	.71	.73
Working Relationship	.71	.69	.77

Table V

Inter-Rater Reliability Coefficients for Four
Variables Rated in Three Different Interview Units

Dimension	Rater 1 and 2			Rater 1 and 3*			Rater 2 and 3*		
	<u>D.T.**</u>	<u>P.A.</u>	<u>Fr.</u>	<u>D.T.</u>	<u>P.A.</u>	<u>Fr.</u>	<u>D.T.</u>	<u>P.A.</u>	<u>Fr.</u>
Amount of lead***	.87	.86	.92						
Counselor Assumption of Responsibility	.73	.63	.83	.72	.69	.79	.81	.69	.86
Client Assumption of Responsibility	.32	.28	.39	.35	.38	.47	.66	.65	.56
Working Relationship	.43	.39	.56	.44	.46	.43	.49	.44	.58

* Rater 3 represents two raters; one rated 27 interviews and the other nine.

** D.T. - Discussion topics (N - 174)

P.A. - Problem areas (N - 106)

Fr. - Fractions (N - 60)

*** On this one dimension the data listed for rater 1 and rater 2 incorporate the ratings of several pairs of raters on different sets of interviews.

Reliability of Ratings

One of the necessary facets of investigations involving ratings is a consideration of the reliability with which the ratings were made. Although reliability data obtained for ratings done in this investigation have already been cited in Chapter III, it seemed desirable to present and discuss them here as a set of findings. The tables of reliability coefficients, Tables IV and V, show that the reliability obtained in rating the four variables was adequate for use in typescript research. The first row of correlation coefficients in Table IV reads as follows: the combined ratings of two judges rating amount of lead obtained a correlation coefficient of $r = .93$ when the Spearman-Brown prophecy formula was used to take into account the use of combined ratings to obtain average ratings. In the same fashion, amount of lead ratings done on problem areas yielded corrected reliability coefficients of $r = .93$ and those done over fractions $r = .96$. The remaining data are read similarly, except that the combined ratings of three judges are involved.

This difference in number of raters also shows up in Table V. The first row of this table reads as follows: The ratings of amount of lead for discussion topic units between two sets of raters correlated $r = .87$,

for problem area units $\sqrt{.86}$ and for fractional units $\sqrt{.92}$. The remaining data differ only in that they present the intercorrelations from three raters rather than two.

Two of the dimensions, amount of lead and counselor assumption of responsibility, were rated slightly more reliably than in most previous studies (see p. 58 and p. 61). Working relationship and client assumption of responsibility were rated somewhat less reliably than in previous studies (see p. 67 and p. 69). However, the differences are not great and do not seem to represent an important difference affecting this study.

As indicated earlier, amount of lead is rated remark-by-remark and the mean of the ratings within a unit is used as the rating for that unit. Its high reliability may be a function of both this combining of large numbers of ratings and the relatively clear-cut definition given various levels of lead. The lower reliabilities obtained for the global dimensions, client assumption of responsibility and working relationship are in part due to the single global rating used. They may also be, in part, a function of the basic difference in counseling orientation of one of the raters (see Table V). Raters two and three have similar counseling orientations --- they are student personnel majors at Ohio State --- while the other rater, i.e., rater one,

has interests and background in clinical psychology. More intensive training might have developed a more similar frame of reference for these judges. However, the training for these ratings was no less intensive than for counselor assumption of responsibility and that was rated as reliably as on previous occasions.

Although some rating reliabilities were lower than previously reported, all of the correlation coefficients obtained were significant at the .01 level of confidence. Thus it appears that the ratings of problem area, discussion topic and fractional units are all reliable enough to be used in typescript research. It should be noted that the fractions used were quintiles of a three interview series, thus incorporating unusually large segments of the conference into a single unit. Shorter fractional segments of the interview might not be rated with sufficient reliability to justify their use.

Reliability Hypotheses

In developing testable hypotheses to attack the problem of the relative effectiveness of three larger units of analysis in typescript research, one of the first considered was that of the relative reliability of ratings within each unit. Translated into null

Table VI

Significance of Differences Between Reliability Coefficients
of Combined Ratings of Four Dimensions on Three Types of Units
(Differences are z transformations)

<u>Dimensions</u>	Problem Areas vs. Topics		Topics vs. Fractions		Problem Areas vs. Fractions	
	Diff. [*] <u>$z_P - z_T$</u>	C.R.	Diff. <u>$z_T - z_F$</u>	C.R.	Diff. <u>$z_P - z_F$</u>	C.R.
Amount of Lead	0	0	-.288	1.88	-.288	1.74
Counselor Assumption of Responsibility	-.179	1.43	-.266	1.74	-.445	2.70**
Client Assumption of Responsibility	-.021	.17	-.021	.14	-.042	0.25
Working Relationship	-.039	.31	-.133	.87	-.172	1.04

*Standard error of difference $z_1 - z_2$
 Problem Area vs. Topic .1247
 Topic vs. Fraction .1529
 Problem Area vs. Fractions .1621

**Significant at .01 level

hypothesis form, this became: the inter-rater reliability in rating the various dimensions does not differ significantly from one type of unit to another. The alternative predicted was that ratings of variables by discussion topic and problem area units would be significantly more reliable than those rated by fractions. The findings pertinent to this hypothesis are presented in Table VI. It shows the significance of the difference between the correlations listed in Table IV.

Table VI reads as follows: When problem areas are compared with topics the reliability coefficients transformed into Fisher's z for combined ratings on amount of lead do not differ and thus the critical ratio obtained is not significant. In only one of the twelve comparisons of correlations obtained for combined ratings was the obtained difference significant. When ratings of counselor assumption of responsibility for problem area and fractional units were compared, the reliability correlation coefficients of the latter were significantly higher. Since such a difference could occur by chance during the course of 12 comparisons, the findings are interpreted as not providing sufficient evidence to reject the null hypothesis.

Table VII

Significance of the Difference Between Correlation Coefficients for Mean Ratings of Within-Unit Halves Obtained for Discussion Topic, Problem Area and Fractional Units of the Interview

(Based on z transformations of correlation coefficients)

<u>Directionality of Prediction</u>	<u>Dimension</u>	<u>Obtained Difference</u>	<u>C.R.*</u>
Problem Area larger than Discussion Topic	Amount of Lead	.098	.70
	Working Relationship	.050	.36
Discussion Topic larger than Fraction	Amount of Lead	.065	.38
	Working Relationship	.129	.76
Problem Area larger than Fraction	Amount of Lead	.163	.87
	Working Relationship	.179	.96

*A C.R. of 1.64 would be significant at the .05 level.

Hypothesis 1b is also concerned with the reliability with which dimensions are rated in the different units. As a null hypothesis it states: The units do not differ in the degree which the first half of each unit correlates with the second half. As indicated in Chapter III only two variables, amount of lead and working relationship, were used to test this hypothesis. It is concerned with the internal consistency of the units. It aims to test whether amount of lead and working relationship ratings of both halves of the unit show that that particular unit is made up of behaviors which are more closely related, insofar as the two constructs used are concerned than are the split-halves of the other units. As a measure of reliability, it parallels the split-half test frequently used in examining the reliability of psychological tests.

None of the six comparisons of split-half reliability coefficients listed on Table VII yielded critical ratios which were significantly larger than could have occurred by chance. The null hypothesis 1b could not be rejected. The topmost row of data in Table VII reads as follows: When problem area units are predicted as showing higher split-half reliability coefficients than discussion topic units, the obtained difference between

the pair of coefficients is .098 which yields a critical ratio of .70. Since this is not significant the null hypothesis cannot be rejected for this comparison.

Although the above findings do not support the alternatives proposed in Chapter III, the results are amenable to explanation. As indicated earlier, one of the special advantages of large units over small units is their higher reliability. It has been demonstrated in both rating and test measurement that increasing the number of items upon which a rating or test score is based tends to increase the reliability of that finding. Since the 36 interviews were divided into 60 fractional, 106 problem area and 174 discussion topic units, it is apparent that fractions generally were about three times as long as discussion topics and almost twice as long as problem area units. It may be that this difference in length favoring the fractions operates to negate both hypotheses 1a and 1b.

From the above findings on the reliability of ratings in this investigation, the following conclusion may be drawn. None of the three units used can be said to be rated more reliably than any of the others. The evidence indicates that the ratings of any well-defined dimension of the counseling situation is likely to be

rated as reliably with discussion topic as with problem area units. It also is apparent that large fractional units can be rated as reliably as units partitioned according to subject matter or type of problem. However, only further research can determine whether fractional units of the same length as discussion topic or problem area units would be rated as reliably as those two units.

Hypotheses Concerning the Distinguishing Characteristics of Units

There is no question that it is important that a unit used in interview typescript research should permit the reliable rating of variables describing the statements incorporated in the unit. However, the unit should have characteristics in addition to those of reliability. Another important criteria of the effectiveness of an interview unit for studying client and counselor behavior is the degree with which it makes discernible variation from one unit to another within an interview or series of interviews. That is, to what extent does it enable one to observe the considerable variety of behaviors incorporated within an interview.

The remaining two hypotheses which this research investigated sought empirical evidence regarding the relative effectiveness of each unit studied for

investigating the variations in client and counselor behavior. The form the hypotheses took was predicted upon the previously stated assumptions (Chapter III, pp. 114-117). Those assumptions which are pertinent here are: (1) client and counselor behavior and attitudes vary during the course of counseling; (2) the unit which makes these variations more discernible will be more effective and (3) the crucial test of the differentiating value of a particular unit is its capacity to reflect changes in behavior and attitude between contiguous units in an interview series. . The results bearing on this aspect of the problem are presented below.

Hypothesis 2: Stated in the null form this hypothesis is: There is no difference between the correlations obtained between ratings of two halves of a unit and those obtained between the last half of a unit and the contiguous first half of the next unit. The alternative predicted was that in both discussion topic and problem area units, the ratings of within-unit halves will correlate to a significantly greater degree than those of between-unit halves. It was also predicted that comparable correlations of ratings for fractions would not be significantly different.

This hypothesis was tested by using the two

Table VIII

Significance of the Differences Between Correlations of
Amount of Lead Ratings for Within-Unit Halves and
Between-Unit Halves When Three Types of Units Are Used
in Rating Interview Typescripts

<u>Type Unit</u>	<u>N</u>	<u>Correlations for Within- Unit Halves</u>	<u>Correlations for Between- Unit Halves</u>	<u>Obtained* Difference Between <u>z</u>'s</u>	<u>Critical Ratio</u>
Discussion Topic	147	.75	.49	.437	3.71**
Problem Area	81	.79	.35	.705	4.40**
Fractions	48	.72	.85	-.348	1.65

*Standard error of z
 Discussion Topics .1179
 Problem Areas .1601
 Fractions .2108

**Significant at the .001 level using the one-tail test

Table IX

Significance of the Difference Between Correlations of
Working Relationship Ratings for Within-Unit Halves and
Between-Unit Halves When Three Types of Units are Used
in Rating Interview Typescripts

<u>Type Unit</u>	<u>N</u>	<u>Correlations for Within- Unit Halves</u>	<u>Correlations for Between- Unit Halves</u>	<u>Obtained* Difference Between $\bar{z}$'s</u>	<u>Critical Ratio</u>
Discussion Topic	147	.89	.66	.629	5.34**
Problem Area	81	.90	.57	.824	5.15**
Fractions	48	.86	.86	.000	0.00

*Standard error of $\bar{z}$

Discussion Topics - .1179

Problem Area - .1601

Fractions - .2108

**Significant at the .001 level of confidence using one-tail test

dimensions, amount of lead and working relationship. Ratings were made in the same fashion as for hypothesis 1b.

The findings are shown in Tables VIII and IX. Table VIII is read as follows: In 147 discussion topic units, the amount of lead ratings for the first half of each unit correlated with similar ratings for the second half $r = .75$; the lead ratings for the second half of each unit correlated with those of the first half of the next unit $r = .49$; when transformed into z 's the difference between the z for the first correlation and the z for the second correlation is .437 in favor of the correlation of within-unit halves ratings; this difference yields a critical ratio of 3.71 which could have occurred by chance fewer than one in one thousand times. The remainder of Table VIII and Table IX are read in similar fashion.

As Table VIII and IX indicate the null hypothesis can be rejected at the .001 level of confidence for both problem area and discussion topic units on both of the variables tested. In each of the comparisons involving these two units, the correlations for within-unit halves were always significantly greater than those for between-unit halves. However, the results obtained from testing the null hypothesis in the case of

fractional units did not permit rejection of the null hypothesis.

These findings indicate that discussion topic and problem area units have unifying characteristics which differentiate them from contiguous discussion topic and problem area units. The nature of some of these characteristics is considered in the next section. The results also show that, although the within-unit halves of each fractional unit are positively correlated, there is no significant difference between this correlation and that obtained for contiguous halves of adjacent fractions. This suggests that the organization of counseling statements into fractions obtains units which average the rated values for discussion over different subjects and types of problems and thus obscures the variations in client and counselor behavior which are found within the interview or series of interviews when either discussion topic or problem area units are used.

Since both Hypotheses 1b and 2 involved correlating ratings on working relationship which were made consecutively and since the writer as one of the raters knew the hypotheses being tested in this study, it was thought that the possibility of contamination of the ratings should be considered. Therefore, the correlations

Table X

Significance of the Difference Between Correlation
Coefficients of Three Raters for Within-Unit Halves
on Working Relationship

(Differences are z transformations)

<u>Type Unit</u>	$\sigma_{z_1 - z_2}$	Rater 1 vs. 2		Rater 1 vs. 3		Rater 2 vs. 3	
		<u>Obtained Difference</u>	<u>CR</u>	<u>Obtained Difference</u>	<u>CR</u>	<u>Obtained Difference</u>	<u>CR</u>
Topics	.1179	.761	6.45*	.576	4.89*	-.185	1.56
Problem Area	.1601	.733	4.58*	.539	3.37*	-.194	1.21
Fractions	.2108	.371	1.76	.340	1.61	-.031	0.15

*Significant at the .01 level of confidence

of the within-unit halves ratings of working relationship for the three raters were compared. As Table X indicates, the within-unit halves ratings of rater 1 correlated to a significantly greater degree than those of either rater 2 or 3. The judgments of rater 2, the writer, did not differ significantly from those of rater 3, but in all instances tended to correlate less than those of rater 1. This was regarded as evidence that knowledge of the hypotheses involving the rating of within-unit halves did not contaminate the ratings of the writer. Since amount of lead, the other dimension used to test these hypotheses, is based upon single-remark ratings and since these were largely made by persons unfamiliar with the hypotheses of this study, it was thought that contamination on that variable was improbable.

Hypothesis 3: This hypothesis is principally concerned with the relative effectiveness of the three units under consideration as means of organizing remarks so that variability in client and counselor behavior is fully discernible. Stated in the null form, it is: Observable shifts in client attitude from each unit to its next adjacent unit do not differ when the difference scores of the three units under study are compared. The predicted alternative was that the degree of shift

Table XI

Significance of the Difference Between Sets of
Difference Scores for Three Types of Interview Units

(Based on analysis by Mann-Whitney U test)

<u>Dimension</u>	<u>Unit Comparisons</u>					
	Problem Area vs. Fraction		Topic vs. Fraction		Problem Area vs. Topic	
	<u>x</u> *	Sign. Level**	<u>x</u>	Sign. Level	<u>x</u>	Sign. Level
Amount of Lead	5.31	.001	4.81	.001	1.50	.07
Counselor Assumption of Responsibility	3.41	.005	3.30	.005	.76	.22
Client Assumption of Responsibility	2.62	.01	1.48	.07	1.64	.05
Working Relationship	2.67	.01	1.81	.05	1.21	.12

*Similar to the critical ratio

**Level of significance - based on the one-tail test since the direction of the difference was predicted.

from unit to unit is greater between series of discussion topics and between series of problem area units than between fractions of the interview series.

All four of the client and counselor dimensions used in this investigation were studied in the test of this hypothesis. As previously indicated, sets of difference scores, i.e., the absolute value of the differences between ratings assigned adjacent units, were obtained for each type of unit. The Mann-Whitney U test (35) was used to determine whether these sets of difference scores differed from each other in the predicted fashion.

The findings from examination of this hypothesis are shown in Table XL. The data obtained by comparing each of the three units studied with the other units is shown in three columns of paired sets of data, i.e., the x's and significance level for each combination of unit comparisons. The rows summarize the findings according to each dimension.

On all four of the variables used in this investigation, problem area unit ratings yielded difference scores which were significantly larger, at better than the one percent level, than those obtained for quintiles of the interview series. The hypothesis that observable shifts in client attitudes from each unit to its next

adjacent unit do not differ must be rejected when problem area and fractional units are compared.

This hypothesis is also rejected on three of the four variables studied when discussion topic and fractional units are compared. On the dimensions of amount of lead and counselor assumption of responsibility the levels of significance were .001 and .005 respectively. The differences obtained between the two sets of difference scores on the variable working relationship was significant at the .05 level. On the fourth variable, client assumption of responsibility, the difference --- in the predicted direction --- was significant at the .07 level of confidence.

A partial explanation of the lower significance level obtained for the last two variables may be the lower reliability with which they were rated. While the combined ratings on the dimensions amount of lead and counselor assumption of responsibility yielded reliability correlation coefficients of about $\sqrt{.90}$, the combined ratings for client acceptance of responsibility and working relationship were found to have a reliability coefficient of correlation of about $\sqrt{.70}$. Since the reliability with which a variable is rated is an essential aspect of the sensitivity of that dimension, i.e., its capacity to reveal clear-cut

differences in behavior within the interview, lower level reliability would be expected to be associated with less definitive difference scores.

The hypothesis that difference scores for problem area units do not differ significantly from those for discussion topics could be rejected on only one of the four variables studied. Ratings on the variable client assumption of responsibility yielded difference scores which were significantly larger for problem areas at the .05 level of confidence. The findings obtained are not regarded as sufficient evidence to reject the general null hypothesis comparing problem area and discussion topic units.

The lack of difference between these two units may be, in part, due to the considerable overlap of problem area and discussion topic units. Problem area units, as indicated earlier (see Chapter III, p. 96), are made up of contiguous discussion topics which fall in the same problem category. In the sample of typescripts studied the problem area and discussion topic unit were identical 75 percent of the time. In the rating of these identical units, it was felt unnecessary to rate them twice since they are identical segments of the interview. Thus, when the units are the same, the ratings used in this study are the same.

Table XII

Mean Rating on Four Dimensions for Four
Unit Categories in Discussion Topic Units

<u>Dimension</u>	Unit Categories			
	<u>Adjustment (N-85)</u>	<u>Skill (N-33)</u>	<u>Special (N-46)</u>	<u>Test I.* (N-10)</u>
Amount of Lead	15.0	27.9	19.6	26.0
Counselor Assumption of Responsibility	25.9	17.1	19.8	12.0
Client Assumption of Responsibility	28.0	24.3	22.5	22.0
Working Relationship	25.2	25.0	23.0	20.0

* Test Interpretation

As a result, 58 percent of the difference scores obtained for problem area units had exact counterparts among the difference scores for discussion topics with which they were compared.

Although the above tests establish that difference scores for both problem area and discussion topic units are significantly larger than those for fractions, it is of some interest to consider the basis for this. An examination of Table XII shows that both client and counselor tend to have different behavioral patterns during the discussion of different type problems. The columns indicate the four different unit categories used in this study and the rows show the four counseling dimensions. For example, the first entry in the table indicates that the discussion topics categorized as adjustment were found to have a mean lead rating of 15.0 in the sample considered. It appears that, as the type of problem changes, the client and counselor behaviors and attitudes tend to vary. As an example of this, if the topic shifted from an adjustment unit to a test interpretation unit considerable variation would be expected on the variable amount of lead.

Another factor entering into the findings for Hypothesis 3 is the relatively limited variability of ratings within fractional units of the same case.

Table XIII

Mean of Between-Unit Difference Scores for
Ratings of Four Dimensions in Three Types
of Units

<u>Dimension</u>	<u>Discussion Topic (N=161)</u>	<u>Problem Area (N=94)</u>	<u>Fraction (N=48)</u>
Amount of Lead	6.33	7.50	2.79
Counselor Assumption of Responsibility	7.05	8.58	4.08
Client Assumption of Responsibility	4.66	5.73	3.98
Working Relationship	3.60	4.56	3.84

This can be seen by a closer examination of the difference scores obtained. Table XIII indicates that although the ratings for all fractional units range wide over the total possible ratings (see Table II, p. 133), they show a marked tendency to cluster about the same level within a single case. Table XIII presents the mean of the between-unit difference scores for ratings of the three types of units. The rows of the table indicate the dimension and the columns the type unit. The mean difference scores in Table XIII show that fractional units, in most instances, when compared to the other units, are less variable from one unit to another. The combining of ratings over diverse problems into a single rating seems to level off the variations present when the subject matter or problem serves as the demarcation point. This applies to both the single remark ratings of amount of lead and two of the global ratings. In making global ratings over material which brings together diverse sub-units, the judge apparently resorts to a rating which is an average of the component parts of the fractional unit.

The above data indicate that both problem area and discussion units are more effective than fractional units in making variations in the client-counselor relationship discernible. Although problem area units

tend to be somewhat more sensitive than discussion topics, only one of the four comparisons was significant within the fiducial limits set.

Discussion

This section discusses further ramifications of the findings presented in the previous section. The discussion which follows considers: (a) conjectures about factors associated with discussion topics and problem areas, (b) the implications of this study for subsequent research and (c) some limitations of this research.

Factors Associated with Discussion Topic and Problem Area Units: Both the systematic ratings and the unsystematic observations of the typescripts suggest that there is a tendency for the client and counselor to adopt different patterns of behavior when the subject matter or type of problem varies. This is supported both by the evidence of Table XII, which shows that the mean ratings of topics in different categories tend to vary and by the evidence found upon testing Hypothesis 3. One of the bases for these variations in behavior seems to be the roles which the counseling participants accept.

For the purposes of this discussion the definition of role given by Sargent (51, p. 360) seems the most

appropriate of the several considered. He says, "A person's role is a pattern or type of social behavior which seems situationally appropriate to him in terms of the demands and expectations of those in his group." Further consideration of three aspects of this definition might help clarify later discussion. (1) Sargent indicates that although "roles are culturally and socially defined delimitations of behavior, or patterns of permitted alternatives," much variation is possible within the bounds of the role. Newcomb (38) also emphasizes that "roles and prescribed roles do not refer to the actual behavior of any given individual. ...Individual behavior is personally motivated and characteristic of the individual himself; never consists of acting out a role prescribed in advance." (2) The "situationally appropriate" aspect of the definition indicates the importance of perception for the person taking a role. As Sargent puts it, "One does not respond to a situation as defined objectively, but rather as he perceives or interprets it." (3) The phrase, "demands and expectations" of others emphasizes that roles are learned through social experience. In this connection, Sargent says, "Most roles are reciprocal; their structure is patterned through the mutual expectations of group members." With this definition as

background, the following discussion considers the usefulness of the role concept for studying behavior in the counseling conference.

One hypothesis which seems worth considering is that the type of problem or the subject being discussed has an effect on the "role behaviors"¹ of the client and the counselor. That is, the variations within the conference which were cited above, may be a function of the demands and expectations which the client and counselor perceive in the counseling relationship. As a function of their learnings and anticipations about the counseling relationship and also their broader life roles, both the client and counselor assume "role behaviors" which they deem congruent with the subject or problem under discussion. It should be emphasized that the role concept, as here defined, does not eliminate individual differences or variability, but rather suggests limits or patterns of permitted alternatives.

It may be helpful to consider, in terms of concept of role, some of the typical situations which arise during the course of a conference. At the opening of

¹"Motive pattern on the part of a specific individual as he takes a role." T. Newcomb. Social Psychology, New York: Dryden Press, 1950.

the interview, especially for the initial session, the participants engage in a friendly discussion. Such a discussion is, of course, not peculiar to counseling situations, but can be seen as an example of where broader life roles are carried into counseling. When the discussion turns to the purpose of the conference, the client usually presents his problem and the counselor typically says only enough to keep the client moving and help the client tell his story. This too fits into the wider pattern of roles which people in our culture accept in relationships where one person is telling another his problems.

Assume that the first problem which concerns the client is his rate of reading. After some discussion of his practices, usually in response to the questioning of the counselor, both the counselor and client expect that the former will assume major responsibility for dealing with this problem in the conference. Such role behavior follows the more general pattern in which the uninformed seeking technical information sits back and expects the professional or technical person to inform him. Likewise, the technician or professional worker regards it fitting that he should assume the role of advisor or teacher under such circumstances.

If the client now presents a problem which involves

personal values and the resolution of conflicts, the counselor typically acts in a manner designed to place much of the responsibility with the client and to make it easier for the client to deal with various aspects of his problem.²

At the close of the conference, the pattern of exchanges tends to follow the general pattern that is seen when two people are parting. That is, there may be some talk about whether the participants should meet again; if so, when; what might be done until the next meeting and a parting good-bye. This also seems quite similar to many well-defined relationships found outside the counseling conference.

A more detailed consideration of a "typical" interview from among those incorporated in the present study may throw further light on the possibilities and limitations of role concepts in the study of the counseling interview. The interview used was selected because it included about an average number of shifts

²It is true that both clients and counselors frequently take roles at variance with the described relationships. Rather than a denial of the application of the role concept, this disagreement suggests what may be an interesting hypothesis. It may be that there is no generally accepted social norm for role behaviors when two people talk about matters which are similar to what has been classified adjustment or skill in this study.

in type of problem being considered and the difference scores between units (see p. 125), were about average for the group of cases. It was thought that a set of charts indicating the ratings made for each unit in an interview would facilitate the discussion of various examples of role behaviors in that interview.

Each of the four charts which follow were taken from ratings on the first interview of the Sketch case from Chicago. Each chart deals with the ratings of one of the four dimensions; each rating was for an entire unit. Chart I is read as follows: the mean amount of lead rating for unit 1, an adjustment topic, was about seven for rater 1 and eight for rater 2. The remainder of Chart I is read the same way. Charts II, III and IV differ only in that three raters were used there.

At first examination Charts I, II, III and IV suggest that the ratings of different raters for the same unit are quite variable. However, as an inspection of Table V (see p. 138) shows, the inter-rater reliability coefficients were high enough (ranging from $\sqrt{.28}$ to $\sqrt{.92}$ with a median of $\sqrt{.58}$) to be useful, especially when these ratings are combined (see Table IV, p. 137).

The discussion which follows considers each of the units in turn. It relates the data of the several

Chart I

Amount of Lead Ratings by Rater* for Discussion
Topics of a "Typical" Interview

(Case---Sketch from Chicago)

* Raters:

Rater 1
Rater 2
Rater 3

** Type Units:

A - Adjustment
Sp - Special

Chart II

Counselor Assumption of Responsibility Ratings
by Rater* for Discussion Topics of a "Typical"
Interview

(Case---Sketch from Chicago)

* Raters:

Rater 1
Rater 2

** Type Unit:

A - Adjustment
Sp - Special

Chart III

Client Assumption of Responsibility Ratings
by Rater* for Discussion Topics of a
"Typical" Interview

(Case---Sketch from Chicago)

* Raters:

Rater 1
Rater 2
Rater 3

**Type Units:

A - Adjustment
Sp - Special

Chart IV
Working Relationship Ratings by Rater*
for Discussion Topics of a "Typical"
Interview

(Case---Sketch from Chicago)

* Raters:

Rater 1
Rater 2
Rater 3

Type Units:

A - Adjustment
Sp - Special

charts to what was being discussed in the interview. The ratings cited are for the unit as a whole. In examining the data in terms of role behavior the emphasis is upon (1) the compatibility of the roles assumed by client and counselor, primarily in terms of division of responsibility and (2) the shift in roles when the type of problem or topic changed.

At the opening of the interview, the client described in detail the changes in cultural taste and personal outlook which he has noticed in himself during the past six months. As is usually the case, while the client was presenting his problem, the counselor tended to lead relatively little and assign most of the responsibility to the client (see Charts I and II). The differences in ratings for client acceptance of responsibility (see Chart III) may arise, in part from the fact that although the client does most of the talking in this unit, on several occasions he makes comments indicating he wishes the counselor to tell him what to do, e.g., "Now, what do I do?" It thus appears that the counselor is consistent in the role he is taking, but that the client is uncertain about the role he should assume.

In the second unit we find that the client presses the counselor for an explanation of "...what do you do

about something on this order," and how one can go about getting a better understanding of oneself. Here the client is seen as somewhat less willing to accept responsibility although the counselor continues to place major responsibility with the client (see Charts III and II). Two of the raters saw this unit as having a slightly less satisfactory working relationship than the first unit (see Chart IV) this may reflect the somewhat conflicting roles taken by the counselor and client so far as responsibility for the progress of the interview is concerned.

Unit 3 was primarily concerned with the counselor's inquiring of the client whether he would be willing to cooperate in a research project. Chart II indicates that here the counselor modified his usual manner of handling responsibility and the amount of lead used, to suit the character of the discussion. In this situation, the counselor assumed more responsibility and the client was interested in accepting less responsibility than in the adjustment type units 1, 5 and 6. Although there is a shift in responsibility, the ratings for working relationship remained at about the same level as before.

In unit 4, the client asked the counselor what he is expected to do. As the ratings for this unit indicate (Charts II and III) the counselor intended to place major

responsibility for the discussion with the client and the latter did not appear interested in accepting responsibility. Along with this difference in intent there was a somewhat lower rating for working relationship. In terms of role and role behaviors, it seems that the counselor behavior was not congruent with the expectancies of the client. That is, the counselor did not assume the role that the client thought he should.

At the beginning of the next unit, unit 5, the client asked the counselor what he should do. After a few responses by the counselor which indicated that the responsibility for developing the problem rested with the client, the student took the initiative by presenting paintings which he had done. He used these to point out what he saw as changes in his outlook during the period just preceding counselling. In this adjustment unit, the conferees were in accord regarding the division of responsibility and concomitantly the working relationship rating is higher than in unit 4. It seems that both assumed roles appropriate to the situation in which a person seeking help presents his problem to an accepting advisor.

At the start of unit 6, the client again asked, "What can I do now to help?" The counselor structured the relationship to indicate that the answers to the

client's problems "lay within" the client. The discussion which followed dealt with the change in values which the client was experiencing. Ratings show that the counselor assigned a great deal of responsibility to the client and the latter seemed interested in accepting it. Here again, the roles taken by the counselor and the one which the client was interested in assuming for himself seemed in accord.

The closing unit in this interview is typical of the close of many interviews. After the counselor gave the cue which indicated that the interview had to end, the conferees discussed arrangements for the next meeting. According to the ratings (see Chart II), the counselor assumed a somewhat larger share of the responsibility than in units 5 and 6. The client was (see Chart III) less willing to accept responsibility for the progress of the interview. It thus seemed that client expectancies and counselor role behavior once again did not concur. The ratings on working relationship (see Chart IV) also were somewhat lower.

Since roles are "patterns of permitted alternatives" (see p. 164), it is of some interest to see whether any such pattern might be observable in the speech-by-speech analysis of an interview. It was suggested earlier

Chart V

Remark-by-Remark** Ratings for Amount of Lead
in the First Interview of Case III of Chicago

(Ratings are those of one rater)

* Long vertical marks along abscissa indicate unit division points. The first break is at remark 59.

** Type Units:

A - Adjustment

Sp - Special

(see p. 165) that roles may vary with the subject or type of problem being discussed and that roles taken while dealing with similar problems would be expected to follow similar patterns. To further examine this hypothesis, the ratings of counselor remarks on the amount of lead dimension for the first interview of Sketch were graphed in Chart V. This chart is read as follows: on the dimension amount of lead, the first counselor remark was rated 18, or at about the level of a general lead (see Appendix 1), remark 2 was rated 5 or about at the level of an acceptance remark, etc. The end of unit 1, an adjustment topic, falls immediately after remark 59 and remark 60 begins unit 2, a special unit.

An examination of Chart V shows that the speech-by-speech behavior within units do not cluster neatly at one level during the course of one unit and shift abruptly to another level at the beginning of the next unit. However, as the discussion which follows indicates, there is some pattern in the ratings.

In the first unit of the interview, the client presented his problem and the counselor assumed the role of understanding listener. With few exceptions, amount of lead ratings in this adjustment topic fell at about the level of an acceptance or clarification

remark (see Appendix 1 for description of amount of lead techniques).

The three units which follow were all classified as special topics. In unit 2 the counselor tended to lead somewhat more than in the initial unit. His remarks structured the counseling relationship as a situation in which the client is expected to take most of the responsibility for deciding how to deal with his own problems. In the next unit, the counselor asked the client whether he would be willing to cooperate in some research being done at the counseling center. The ratings are unusually consistent and also somewhat higher than for adjustment topics in this interview. The discussion returned to the problem of defining the roles of the conferees. The client talked as if he wanted the counselor to tell him what he should do, at least what steps he had to take to tackle his problem. Here again the ratings were slightly higher than in unit I, V or VI.

The next two units, V and VI, are a further amplification of the client's problem. It should be noted that although they deal with different subjects, these two adjustment units show pretty much the same pattern of ratings. Acceptance remarks predominate with occasional clarification remarks being introduced.

The above anecdotes and charts illustrate some of the hypotheses regarding the possible significance of role and role behaviors in the counseling conference. They contribute to the more systematic data which show that client and counselor behaviors tend to vary from subject to subject or problem to problem. It also is apparent that there tends to be some patterning within a unit, e.g., the counselor tended to use techniques falling at a low level during adjustment topic discussions and use techniques which lead more when discussing special topics. However, it is also evident that the anecdotal descriptions and charts show marked variations indicating that assumption of a role does not result in rigidly consistent behavior.

The implications for subsequent research of role behaviors of the conferees, or the lack of such patterns of behavior, is considered briefly in the next section.

Implications for Subsequent Research: If the assumptions set forth in Chapter III are accepted, the greater efficacy of discussion topics or problem area units over fractional units follows from the evidence cited above. The first two units have been found to:

- (1) show greater sensitivity to variations in behavior during the course of a conference and
- (2) bring together client and counselor responses which reflect related

behaviors and attitudes. This latter characteristic of problem area and discussion topic units makes it feasible to use them to study some new factors in the dynamics of the interview; i.e., the significance of counselor and client attitudes and roles in the counseling process.

Therefore, it seems that subsequent research investigating the counseling interview process should use units which group statements according to some meaningful organizing principle. From the findings of this investigation, one must conclude that the use of units which artificially demarcates segments of the interview will result in the loss of information about significant variations in the behavior of the participants. It seems imperative that any unit used in interview typescript research take into consideration both the within-case variation and the within-unit consistency of client and counselor role behaviors. As a corollary, it seems vital that any unit used for such research be based on the principle that units should be psychologically meaningful. That is, they should unify behaviors and attitudes which are of the same pattern and partition those segments of the interview which are dissimilar psychologically.

The need to consider the principle just cited suggests another avenue of research on the methodological tools to permit comparisons in typescript research. It may be feasible to conceptualize a unit which is based upon client and/or counselor roles rather than the type of problem or subject. Such a unit would organize the statements of the interview in terms of role behaviors or some other psychosociological construct. It would facilitate the examination of hypotheses which the present units (discussion topics and fractions) are not especially suited to test. For example, one hypothesis --- suggested by the anecdotal findings in the discussion of the role concept --- might be that: When the counselor and the client assume complementary role behaviors, the progress of the conference is more satisfactory than if the roles they assume are not in accord. An approach somewhat like that used by Nelson (37) in studying optimum lead might be used. That is, ratings of the role behaviors assumed by the counselor could be compared with judgments representing the most desirable role behavior pattern for that particular situation.

The data on the inter-correlations of the variables studied and the subjective reports of the participating

raters suggest another next step in typescript research. As Table III (see p. 133) shows, the four variables used are significantly inter-correlated. A factorial analysis of the dimensions presently used might lead to the establishment of fewer but more comprehensive variables. The development of such dimensions would be another step forward in the establishment of methodology making counseling interview research comparable.

In connection with obtaining factorially developed dimensions, there also seems a need for rating scales which link various levels of rating with behavioral description. All three raters participating in this investigation found it somewhat difficult to rate the two client dimensions used. The task of placing a particular unit somewhere along a normative continuum, which the rater conceptualizes as a result of his own experience, offers several possibilities for unreliability, e.g., the differential backgrounds of the raters and even the individual variability of a rater.

One method for linking the rating procedure more closely to interview behaviors would be to adapt the Thurstone attitude scaling methods to the rating of interview typescripts. In brief, the steps for doing this would be as follows: (1) Have a group of experienced counselors or raters of counseling interviews make a

series of statements concerning specific behaviors which they consider in rating the dimension for which the scale is being developed. (2) Items are eliminated or edited to ensure that they are (a) short and specific, (b) to the point and (c) not doubled-barreled, but amenable to either acceptance or rejection. (3) A number of judges, preferably persons who possess about as much counseling experience as those raters who will use the completed scale, are asked to evaluate each of the items on a numerical scale. (4) The scale value of each item is obtained by taking the average of all the judges ratings of it. (5) Items for which judges show wide disagreement are discarded and only those with a high degree of agreement retained. (6) As a final step, items are selected from the pool of items obtained from the first five steps which represent the entire range of values for the dimension. The above procedure was adapted from Ghiselli and Brown (22, pp. 115-116).

The usefulness of this checklist might be checked in at least two ways: (1) The reliabilities obtained by use of the checklist could be compared with those obtained by previous methods. In terms of reliability, the checklist would be more efficient if the desired level of reliability could be obtained by fewer raters

or if it yielded higher reliability coefficients for the same number of raters as the earlier methods.

(2) Ratings obtained by using the checklist could be compared with those assigned by a three-judge jury which conferred about the rating to be used after each member had made an independent judgment. The jury procedure which has been found to be one of the most reliable methods of rating, would provide a type of criterion against which to test the more readily obtained checklist ratings.

Limitations of This Study: In planning this study, an attempt was made to eliminate or minimize possible weaknesses in the procedures used. However, in retrospect this study appears to have several limitations. These have usually been discussed during the course of the presentation, but it was felt that a summary discussion would be helpful. They are:

(1) In some instances more than one rater was used in obtaining a set of ratings over the 36 interviews. That is, in the rating of the global dimensions, the ratings of "Rater 3" actually involved two judges. One rated 27 interviews and the other nine. Also in the more tedious amount of lead ratings, the judgments of six different raters were paired. Although these conditions held for all three types of units, the

introduction of additional rater variability may have decreased the reliability of the data obtained.

(2) It also seems that the training of the raters could have been more intensive. The writer believes that one of the bases for lower reliability on two of the three global dimensions was that training was not carried to the point that all three raters held the same frame of reference for the two client dimensions. (3) The use of fractional units of about the same length as the average discussion topic or problem area would have made the test of the reliability hypotheses more adequate, i.e., it would have eliminated the distorting element of the greater length of fractions of the interview series. However, it was felt that the above factors did not interfere with the demonstration of the hypotheses proposed by this investigation.

Although this study considers three of the larger units used for the study of interview typescripts, there are two others which it does not consider. One of these is the whole interview. Another possibility is a new type of unit based on the role behaviors of the client and counselor. These additional units should probably be investigated in subsequent research.

Table XIV

Distribution of Unit Classifications for the 174 Discussion
Topic Units Used in This Investigation According to Each of
the Twelve Cases Studied

<u>Unit Classification</u>	Chicago			Minnesota			Missouri			Ohio State		
	<u>1*</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>	<u>7</u>	<u>8</u>	<u>9</u>	<u>10</u>	<u>11</u>	<u>12</u>
Adjustment	7	5	6	17	8	5	12	6	15	0	1	3
Test Interpretation	0	0	0	0	3	0	4	3	0	0	0	0
Skill	0	0	0	0	0	0	0	0	2	11	14	6
Special	2	8	2	4	4	3	5	5	5	3	5	0

*Case number

Secondary Questions and Findings

Kinds of Problems: One of the questions posed early in this investigation was: Do the kinds of problems counselors at different schools deal with differ significantly? Impressions gained from reading typescripts from different institutions suggested this query. In this investigation, to counteract this possible fact, the interviews used were taken from four schools.

The distribution of problem classifications for the 174 discussion topics studied are presented in Table XIV. As that table indicates, although different sets of interviews from the same institution show some differences in the distribution of problem types, the three cases from each school tend to follow much the same pattern. Table XIV is read as follows: In case one, a series of interviews from Chicago, seven topics fell in the adjustment category and two were classified special. There were neither test interpretation nor skill units in this case.

In the sample of interviews studied, the distribution of types of problems found varied to a certain degree from one institution to another. Units from Minnesota and Missouri interviews showed somewhat similar distributions among the four classes used here. Both of

these counselling centers typically deal with clients faced with decision making in educational and vocational problems. Usually the client expects to take and does take a battery of tests to gain additional information about himself. The incidence of units in which the main subject was test interpretation is higher than the data shows since units in which test interpretation was integrated with the client's decision making were classified as adjustment topics.

In the three cases from Chicago, 60 percent of the units fell in the adjustment category and the remainder in the special classification. The adjustment units of this group fit the general pattern previously cited in Chapter III. The relatively high percentage of special topics reflects both the usual closing conversations of conferences and instances in which the counselor needed to structure the counseling relationship. Although there are neither skill nor test interpretation units, it would not be correct to say that clients coming to the Chicago center never presented such problems. In the instances in which such problems were presented, the counselor dealt with them as if they were decision-making problems.

The distribution of problem types for Ohio State discussion topics was markedly different from that of

other schools sampled in the study. Thirty-one of the total 33 skill units in the total sample came from Ohio State. The proportion of adjustment type units was less than for other schools. Unlike the sample from Minnesota and Missouri, it did not include any test interpretation units.

These findings concur reasonably well with the descriptions of the types of problems by school which was presented in Chapter III. This earlier description was based on subjective impressions after reading much larger samples of interviews from each of the schools and published descriptions of the counseling done at the centers.

Although the sample from each institution is limited --- nine interviews and three clients for each school --- the general similarity of the distribution of problems for each case from a single institution suggests that the cases used here might be representative of the type of case and problems handled at each school. This tentative assumption made, it was believed that a test of the hypothesis, that there are no differences among the schools in the distribution of types of problems, would be of interest.

Table XV

Distribution of Unit Classifications for the 174 Discussion
Topic Units Used in This Investigation According to the
Counseling Center Where They Were Obtained

<u>Unit Classification</u>	<u>School</u>				<u>Total</u>
	<u>Chicago</u>	<u>Minnesota</u>	<u>Missouri</u>	<u>Ohio State</u>	
Adjustment	18	30	33	4	85
Skill and Test Interpretation	0	3	9	31	43
Special	12	11	15	8	46
	—	—	—	—	—
Totals	30	44	57	43	174

The combined data for each school are shown in Table XV. In it the rows indicate the type of problem and the columns show the institution, e.g., there were 30 adjustment topics in the Chicago sample.

To do this the data for the three cases from each school were combined as shown in Table XV. It will be noted there that it was also necessary to combine the test interpretation and skill category units³ in order to meet the requirements of a theoretical frequency of five in each cell. A chi-square analysis was then made for the entire distribution. With six degrees of freedom the obtained chi-square was significant at beyond the .01 level of confidence.

Assuming that these cases and their problems are representative, this indicates that school differences exist which affect the typescripts obtained. Whether this is due to student differences at these schools, to the kind of client attracted, or to counselor predilections, or some other factor is not clear. Obviously, more research is needed on this point.

Subsequent analysis revealed that the chief contribution to this difference among the schools was

³These units tend to be rated similarly (see p. 150). They also occurred less frequently than the other two in the total series of units.

Table XVI

Mean and Standard Deviation Values of
Amount of Lead Ratings for Discussion
Topics within Cases and by School

	<u>Mean</u>	<u>S.D.</u>
<u>Chicago</u>		
I	10.6	3.7
II	8.5	3.6
III	10.7	1.7
All Cases	9.7	3.4
<u>Minnesota</u>		
IV	13.1	5.2
V	15.7	5.1
VI	14.4	6.8
All Cases	14.2	5.6
<u>Missouri</u>		
VII	17.9	10.0
VIII	24.3	8.6
IX	26.4	7.7
All Cases	22.7	9.6
<u>Ohio State</u>		
X	32.9	5.3
XI	22.3	7.2
XII	27.2	4.8
All Cases	26.8	7.7

Table XVII

Mean and Standard Deviation Values of
Counselor Assumption of Responsibility
Ratings for Discussion Topics within
Cases and by School.

	<u>Mean</u>	<u>S.D.</u>
<u>Chicago</u>		
I	38.4	5.3
II	30.8	10.7
III	41.3	6.5
All Cases	35.8	9.5
<u>Minnesota</u>		
IV	19.8	5.2
V	14.0	5.5
VI	20.0	10.9
All Cases	17.8	7.3
<u>Missouri</u>		
VII	26.7	10.4
VIII	19.7	9.9
IX	14.6	5.6
All Cases	20.3	10.2
<u>Ohio State</u>		
X	9.0	4.0
XI	25.8	5.8
XII	22.2	6.3
All Cases	18.4	10.6

Table XVIII

Mean and Standard Deviation Values of
Client Assumption of Responsibility
Ratings for Discussion Topics within
Cases and by School.

	<u>Mean</u>	<u>S.D.</u>
<u>Chicago</u>		
I	39.5	5.5
II	28.1	8.2
III	28.2	6.6
All Cases	31.5	8.8
<u>Minnesota</u>		
IV	25.3	5.0
V	20.4	3.9
VI	29.4	7.7
All Cases	24.3	6.2
<u>Missouri</u>		
VII	24.1	6.5
VIII	31.8	4.9
IX	18.9	5.4
All Cases	23.9	7.6
<u>Ohio State</u>		
X	15.7	3.7
XI	30.5	5.2
XII	21.7	3.3
All Cases	24.9	7.9

Table XIX

Mean and Standard Deviation Values of
Working Relationship Ratings for Dis-
cussion Topics within Cases and by
School

	<u>Mean</u>	<u>S.D.</u>
<u>Chicago</u>		
I	38.4	2.4
II	28.5	6.0
III	31.3	6.0
All Cases	32.2	6.7
<u>Minnesota</u>		
IV	23.8	2.6
V	15.7	5.1
VI	26.9	5.6
All Cases	21.6	6.1
<u>Missouri</u>		
VII	25.3	3.6
VIII	21.1	6.0
IX	18.4	3.8
All Cases	21.6	5.3
<u>Ohio State</u>		
X	15.0	3.8
XI	33.3	2.9
XII	24.5	5.5
All Cases	25.5	8.9

the distribution of Ohio State units. In each comparison of the Ohio State with the other distributions of problem types, the distributions were significantly different. Comparisons among the other schools did not permit rejection of the null hypothesis, although a larger sample of cases might well show differences among these schools. The relatively larger number of skill units and proportionately fewer adjustment type units in the Ohio State sample is undoubtedly the basis for this finding. It is to be remembered that the above data are subject to the limitations that a sample of three cases from each institution imposes. Thus, these findings are primarily suggestive for further research rather than evidence themselves for generalizations about the types of problems arising at different counseling centers.

Further Characteristics of Interviews at Different Schools: The data presented in Tables XVI, XVII, XVIII and XIX indicate the nature of the ratings obtained for each case and for the three cases from each institution sampled. All the tables read as follows: The columns refer to the mean or standard deviation value while the rows refer to the case or group of cases. For example, in Table XVI the first line indicates that Case I from Chicago was rated as showing a mean amount of lead of

10.6 (this is at about the level of a clarification remark --- see Appendix 1) and a standard deviation of 3.7.

The data in the several tables are not presented as evidence indicating the general characteristics of counseling at the several institutions represented in this study, but are primarily meant to make available to future investigators the findings from a small sample of client-counselor relationships at the four schools. Only three separate client and counselor pairs are involved for each school. Such a small sampling may be atypical even though the procedure for selecting interviews was not selective. Keeping in mind the sampling limitation, it was felt desirable to present these findings since they may suggest hypotheses for subsequent research. A brief discussion of the findings is presented below.

From Table XVI it appears that mean ratings for amount of lead tend to be somewhat lower at Chicago and Minnesota than at Missouri and Ohio State. Lead ratings at the last two schools tend to be much more variable than those at Chicago and somewhat more variable than those at Minnesota. Since ratings on this variable are in part a function of the type of problem (see Table XII, p. 159), some of the differences, such as those involving

Ohio State, may be related to the sampling of units representing the particular institution.

The mean ratings for counselor assumption of responsibility as shown in Table XVII tend to be less consistent within each school than those for amount of lead. Counselors at Minnesota, Missouri and Ohio State tend to assume slightly more responsibility than the clients for the progress of the conference. With the standard deviations of ratings for these schools falling between seven and 10 points, it is apparent that the counselors varied on this dimension considerably. Counselors at Chicago tended to place somewhat more responsibility with the client, but their ratings were also quite variable.

On the dimension client assumption of responsibility, as indicated in Table XVIII, there was not only considerable variability within each case, but also a tendency for the mean ratings within each institution to cover a wider range. Though the mean for all Chicago cases was somewhat higher than the mean rating for this dimension at the other schools, the overlap was such that each of the other three institutions was represented by one case which showed a higher mean rating on this dimension than two of the Chicago cases. It appears that any conclusions about the differences

between schools for client assumption of responsibility must await more extended study.

The findings for ratings of working relationship which are shown in Table XIX also must be regarded as merely suggestive. In three instances the mean rating for one case varies considerably from the values of the other two. There is a tendency for the mean ratings on working relationship to be higher in the Chicago cases, but one should note that there is considerable overlap in the ratings actually obtained.

The statements made in this section about the characteristics of counseling interviews at several universities must be considered as quite tentative. In examining both sub-hypotheses, the sampling restriction of only three cases from each school demands that any findings be tendered cautiously. This limitation, in conjunction with the considerable variability between ratings at the same institution, necessitated limiting the discussion of differences on the four dimensions studied to tabulation and few descriptive comments.

Probably the most significant conclusion that can be drawn from the above findings bears on the design of comparative studies of the counseling process at different schools. The variability of ratings on

cases at the same institution suggests that if "random samples" are to be studied, the number of units from each institution should be reasonably large. Since, in at least one instance, the distribution of types of problems varied markedly from that obtained at other institutions, it appears desirable for comparative studies to take account of the differential behaviors associated with different kinds of problems.

CHAPTER V

SUMMARY AND CONCLUSIONS

Counseling has become an important activity in many areas of daily living. Some of the situations where an increasing number of workers engage in counseling are: schools and universities, industry and business, state and federal government agencies and various community agencies. Despite the increased numbers of professionally qualified persons available for counseling and related work, the demand for these services continues to be greater than can be accommodated.

The continuous demand for counseling services and the growth of counseling as a professional activity has made it increasingly important that workers in the counseling field improve the efficacy of their procedures. Although the last decade has seen greater interest in studying the counseling process, there is little doubt that much work remains to be done here. It has become increasingly evident during the past decade that the soundest approach for advancing counseling theory and practice is experimentation. In response to the need for knowledge about counseling and the acceptance of the experimental method as the

best way to investigate the counseling process, there has been an increased amount of experimentation.

One major problem which has arisen in research work on counseling is: What can be done to permit the comparison and integration of the many experimental findings obtained in studies of the counseling interview? Some of the factors which interfere with any efforts to deal with this problem are the use of: (1) counseling interviews from a single counseling center or representing a single counseling orientation, (2) non-comparable dimensions to conceptualize client and counselor behavior and (3) non-comparable units to segment the interview for purposes of analysis.

If counseling research is to move toward permitting comparison and integration of the studies at different research centers, it seems necessary to study the factors which interfere with progress toward that goal. This study marks the beginning of dealing with the first factor, i.e., the use of interviews from a single counseling center or orientation. It incorporates interviews from 12 client-counselor relationships in four institutions into a single study. An effort was also made to take into consideration the second factor, i.e., the non-comparability of dimensions.

This was done by selecting for use in this study, dimensions which are generally accepted as important concepts for describing client and counselor behavior. However, the problem posed by this factor obviously needs further study. It was the primary purpose of this investigation to contribute to the problem posed by the last factor, i.e., the use of non-comparable units to segment the interview for purposes of analysis.

The purpose of this study was to investigate the relative effectiveness for studying client and counselor behavior of several larger units of analysis used in typescript research. It was necessary to study the relative efficiency of the units since it is unlikely that procedures can be developed to transform findings obtained by using one unit into what would have been obtained by use of another unit. Several considerations suggested that this first study deal with larger units rather than single speeches. Larger units:

- (1) have been found to be more reliably rated than single speech units,
- (2) are comparable to each other both in terms of size and use,
- (3) have been used as the principal unit of study more frequently by recent investigators,
- (4) can be used to get at meaningful structures in the interview which would otherwise be ignored and
- (5) permits study of the delayed effects

on the progress of the interview of both client and counselor responses.

Of the three principal, larger units which have been used, i.e., the whole interview, discussion topic and fractions, it was decided to study the last two. These were more nearly comparable. They also represented sharply divergent points of view on the problem of which unit to use, i.e., grouping remarks according to an organizing principle such as the subject being discussed vs. grouping statements by dividing them into fractions. A resolution of this difference was regarded as an important sub-goal in dealing with the larger problem of unit comparability. In addition to discussion topics and fractions, the merit of a new unit, the problem area, was deemed worth investigating.

Description of Data, Dimensions and Methodology

Selection of Interviews:

In order to permit the widest possible generalization from this study, interviews were selected which represent not only different counselors, different clients and different type problems, but also different institutions. Thirty-six interviews were selected from a pool of verbatim typescripts prepared from electrically recorded interviews at four different

universities, Chicago, Minnesota, Missouri and Ohio State. The final sample included three sets of three consecutive interviews from each of the institutions. Wherever possible the cases selected included all the interviews in the series of contacts, but when this was not feasible (because of the length of the case) three consecutive interviews were selected from the case.

Three consecutive conferences for the same client-counselor were used instead of individual interviews for several reasons: (1) they would be more likely to be representative of a particular counseling relationship, (2) groups of interviews would subject the units studied to a test comparable to the previous research using them, (3) the analysis of series permits some tentative exploration of the trends in counseling behavior during successive interviews and (4) the initial interview tends to follow much the same pattern from counselor to counselor.

After selecting the interviews to be studied, three other methodological decisions had to be made. They were: the selection and description of the units, the classification system to use in conjunction with the units and the client and counselor dimensions to rate.

Units of the Counseling Interview Studied

As indicated at the beginning of this chapter, it was thought that a first study of the relative effectiveness of units would most profitably be concerned with larger units of the interview. From among the larger units which might be studied, it was decided that the discussion topic and fraction should be used and that a third unit, the problem area unit, which had heretofore not been used, also be studied. A brief discussion of the character of these three units is presented below.

The discussion topic represents all the conversation of both client and counselor about the same topic or subject. Some of its advantages are: (1) the client and counselor roles tend to remain consistent during the discussion of a single topic; (2) it separates special phases from the main body of the interview; (3) it permits study of delayed or cumulative behavior within the interview and (4) the ratings obtained are more reliable than those for single remarks. The division points for this unit were determined by the pooled judgments of three independent raters. After making divisions of the interview independently, the judges met in conference to arbitrate any differences. For a sample of nine interviews, the two or three

judges used agreed exactly or within the same remark for 83 percent of the discussion topics.

A problem area unit consists of all the contiguous statements dealing with the same kind of problem. Since they are groupings of discussion topics dealing with the same class of problem the first step for determining their boundaries was the same, but in addition it was necessary to classify each topic and place those falling in the same class together. The unit classification schema suggested by Robinson (45) was used to classify topics. To the adjustment, special and skill categories of that schema was added test interpretation. This new unit seemed to have advantages in addition to those found for discussion topics. Material grouped in this manner seems meaningful since it comprises remarks related in terms of a psychological construct rather than in terms of a subject matter construct. It was also somewhat more reliably determined. In the same sample of interviews used to examine the reliability of topics, two or three judges agreed within the same remark 85 percent of the time.

As a unit, the fraction comprises a definite percentage of the total remarks within a series of counseling interviews. For this research quintiles of each

series was the fractional unit used. The reliability of determining its bounds is obviously quite high. Since it, like the other units, is made up of many remarks, it can be more reliably rated than single speeches.

Dimensions of the Counseling Process Used

To compare the three units studied, the behavior within each unit had to be considered. The four concepts selected for use in this investigation are the dimensions amount of lead, counselor assumption of responsibility, client assumption of responsibility and working relationship. The first two are counselor and the last two client dimensions. All of these were previously used in several studies and had been found to be both reliable and sensitive measures of interview behaviors.

Judging Procedures

Persons participating in the ratings for this research were advanced graduate students in the student personnel program at Ohio State University or had equivalent background and experience. The judges were trained in making whatever judgments they participated in. Three judges participated in demarcating and classifying discussion topic and problem area units.

For each of the units studied the three global dimensions were rated by three judges. Only two raters were used to rate amount of lead, which is rated remark-by-remark, since it had been previously found highly reliable.

Precautions were taken to avoid introducing systematic variations which might favor any of the three units being studied. In the instances where a rater did not make judgments over the entire set of cases, he did make judgments for all three types of units for the interview series he rated. To make the global ratings, each judge read through a case making ratings for all units of one type. He then read them to make ratings of units of another type. Thus in all, each judge read through each case three times, usually in succession, to make all the ratings for the three types of units. Since a fixed order of reading might favor one unit over another, the sequential order was planned so that each unit was rated by one rater at each of the serial positions.

Experimental Design

In developing procedures for examining the question posed, it was necessary to develop: (1) assumptions about the behavior within the counseling interview and the characteristics of units most effective for studying such behavior; (2) a set of testable hypotheses;

(3) statistical procedures for testing the hypotheses and (4) some secondary questions regarding the data obtained.

Assumptions

The assumptions made were, in brief: (a) During the course of the counseling process the attitudes and behaviors of both client and counselor vary. (b) These attitudes and behaviors can be satisfactorily rated. (c) There are factors associated with these changes, some of which are yet to be determined. (d) The unit which makes these changes more discernible will be more effective, for the purpose of describing counselor and client behavior and factors associated with it, than the unit which obscures them. (e) While changes taking place during the interview series will cause differences between widely separated units, the crucial test of the differentiating value of a particular unit is its capacity to reflect changes in behavior between contiguous units in an interview series.

Hypotheses

In the light of the above assumptions, four testable hypotheses were proposed. Two of these dealt with the reliability with which the three units could be rated on the four dimensions used. In the null form

they were: (a) There is no difference between the interrater reliability coefficients obtained for the three different units and (b) The units do not differ in the degree which the first half of each unit correlates with the second half. In both cases, the alternative predicted was that problem areas would be rated most reliably, discussion topics somewhat less reliably and fractions least reliably.

The two other hypotheses consider a unit's capacity for distinguishing useful segments of the interview. That is, they examine the extent which the units studied make variations within the interview discernible.

Hypothesis 2 was: There is no difference between the correlations obtained between ratings of two halves of a unit and those obtained between ratings of the last half of a unit and the contiguous half of the next unit. The predicted alternative was that ratings of within-unit halves would correlate to a significantly greater degree than those of between-unit halves in the case of discussion topic and problem area units. Fractions were predicted as not showing such a difference.

The final major hypothesis proposed was based on the last two assumptions given above. The null hypothesis posed was: Observable shifts in client and

counselor behaviors from each unit to its next adjacent unit do not differ when the difference scores of the three units under study are compared. It was predicted that the shift from unit to unit would be greater in problem area and discussion topic units than in fractional units.

Statistical Procedures

The discussion which follows briefly describes the statistical procedures used. All the comparisons made here are between findings for the three types of units.

The data used in testing the above hypotheses were the combined ratings of the several raters judging each unit. For Hypotheses 1a and 3, the ratings for all four client and counselor dimensions were used. The procedures for testing Hypotheses 1b and 2 differed in that only amount of lead and working relationship ratings were used and these were for halves of the units.

To test Hypothesis 1a the ratings of each judge were correlated with each other judge who rated the same interviews. The obtained correlation coefficients were averaged --- by using the z transformation --- to secure a single reliability coefficient. The reliability coefficient for the combined ratings was then determined by applying the Spearman-Brown prophecy formula to take

into account the number of raters used. The significance of the difference between each pairing of the obtained reliability coefficients of the units was then determined.

The testing of Hypothesis 1b required the correlation of ratings of the first half of each unit with those for its second half. This was done for only one client dimension and one counselor dimension. The significance of the difference between the correlations obtained for each type of unit was obtained. For example, a test was made whether within-unit halves ratings on amount of lead correlated more in problem area units than in fractional units.

The first few steps in Hypothesis 2 were similar to those for 1b. Ratings were secured on the unit halves. The correlation between ratings of within-unit halves was used. In the next step, the ratings for the last half of each unit were correlated with the ratings for the first half of the next contiguous unit. The significance of the difference between the two sets of correlations, those for within-unit and those for between-unit halves, was then calculated.

The procedure for testing Hypothesis 3 was as follows: (1) The difference scores between combined ratings of adjacent units were determined. (2) The

Mann-Whitney U test was used to test the hypothesis that the distribution of difference scores for the different types of units were equivalent. (3) The U statistic was computed for the entire set of cases and its significance ascertained.

Secondary Questions

Since this is one of the first studies which uses counseling interview records from several institutions, it was thought that there would be some value in investigating whether the schools differ on the variables studied. It was recognized that the data secured represented a limited sampling --- three different clients in nine interviews --- from each campus, and that any findings obtained would have to be considered very tentative. Keeping this in mind, two questions were posed: (1) Do the types of problems dealt with at the several counseling centers differ? (2) Do the mean and standard deviation values for the combined ratings of the four dimensions differ among the four institutions?

In both instances, the data pertinent to the queries were tabulated. The first of these secondary questions was examined by a chi-square analysis. Since the data obtained for the second question suggested that the three interviews at each of the schools tended to differ one from the other, it was felt that an

application of other than descriptive statistical measures was not warranted. Therefore, the data regarding this problem were only tabulated.

Results

The conclusions which are suggested by the data collected in this investigation are summarized below. The findings are considered in the following order: (1) characteristics of the data, (2) major hypotheses of the study and (3) secondary questions.

Characteristics of the Data

1. The distribution of combined ratings for discussion topic units were found to fit reasonably close to the theoretically normal pattern. This was particularly true of the client dimensions, client assumption of responsibility and working relationship. Both of the counselor dimensions, amount of lead and counselor assumption of responsibility, showed some positive skewness; however, this did not appear so marked as to warrant modification of the statistical methods used.

2. Inter-correlations of the four variables when rated on the three different units showed quite similar patterns. This suggests that in rating the different units raters held quite similar frames of reference. These data also show that the four dimensions are inter-related to a significant degree. However, these

relationships were low enough to justify use of the four dimensions.

3. In each instance, the reliability coefficients obtained for the combined ratings indicated that the variables could be rated consistently enough from one rater to another to permit their use in research. For the three different type units, the median reliability coefficient for the combined ratings on amount of lead was $\nearrow .93$. Similarly, the median for counselor assumption of responsibility was $\nearrow .90$, for client assumption of responsibility $\nearrow .72$ and for working relationship $\nearrow .71$.

Major Hypotheses

1. Problem area, discussion topic and fractional (quintiles of the interview series) units were all rated with about the same degree of reliability on two counselor and two client dimensions of the counseling interview. In terms of reliability, no one of the three units was found more advantageous than another. All permitted ratings which were reliable enough to use in typescript research.

2. A test of the internal consistency of the several units showed that they do not differ in the degree to which they are internally consistent. Amount of lead ratings for the two halves of each type unit

inter-correlated $r=.75$ for discussion topic units, $r=.79$ for problem area units and $r=.72$ for fractional units. For working relationship ratings the respective inter-correlations were: $r=.89$, $r=.90$ and $r=.86$.

3. Problem area and discussion topic units not only bring together similar behaviors, but also set off parts of the interview which are less related. That is, ratings for amount of lead and working relationship correlate from $r=.75$ to $r=.90$ for within-unit halves for these two units. Comparable between-unit half ratings correlated from $r=.35$ to $r=.66$. In each instance the within-unit correlations were significantly larger than the between-unit correlations.

4. The fractional unit used did not partition the counseling interview in a manner which differentiated segments of the interview which were behaviorally different. Specifically, the amount of lead ratings for within-unit halves of fractions correlated $r=.72$ while between-unit halves ratings correlated $r=.85$. On working relationship these two correlations were the same, $r=.86$. The differences between within-unit and between-unit correlation coefficients were not significant in either case.

5. Use of problem area and discussion topic units makes shifts in client and counselor behavior --- as

measured by ratings --- more discernible than does use of the fractional unit studied. This was demonstrated by the test of Hypothesis 3 which indicated that in every instance the measures of shifts for problem area units were significantly greater than those for fractions. In like manner, the variations in ratings from one discussion topic to another were significantly greater than those for fractions with three of the four dimensions studied and was quite close to the fiducial limits set in the fourth.

6. Problem area and discussion topic units do not differ significantly in terms of observable shifts of behavior from one unit to the next. Since the variations between problem area units, as measured by difference scores, were significantly greater than for discussion topic units on only one of the four dimensions studied, it was felt that the null hypothesis could not be rejected in the case of these two units.

Secondary Questions

1. In making studies which are concerned with the differences in client and counselor behaviors at several institutions, it appears desirable to take into consideration any differential distribution of problem types at each of the schools studied. This seems advisable since some schools seem to differ, e.g., Ohio State cases

studied in this investigation had more skill and fewer adjustment topics than the other three schools sampled.

2. Studies which incorporate counseling records from several institutions need to take into consideration the client and counselor differences at each institution before attempting any comparisons between schools. The data dealing with the differential patterns of client and counselor behavior at different schools showed that there tended to be considerable variability among the client and counselor behaviors from a single institution. With few exceptions, the combined ratings for a dimension tended to have a range of from 10 to 15 rating scale points for the three cases studied from each school.

BIBLIOGRAPHY

1. Allen, K. P. An analysis of the development within a topical unit in the counseling interview. M. A. Thesis, Ohio State Univ., 1946.
2. Assum, A. L., and Levy, S. J. Analysis of a non-directive case with follow-up interview. J. abnorm. soc. Psychol., 1948, 43, 78-89.
3. Barry, J. R. The relations of verbal reactions to adjustment level. J. abnorm. soc. Psychol., 1950, 45, 647-658.
4. Betts, E. A. An evaluation of certain techniques for the study of oral composition. Univ. Iowa Studies: Studies Educ., 1934, 9, No. 2, 7-32.
5. Blau, T. H. Report on a method of predicting success in psychotherapy. J. clin. Psychol., 1950, 6, 403-406.
6. Bordin, E. S. Counseling methods: therapy. In C. P. Stone (Ed.), Annual Review of Psychology, 1950, 1, 267-276.
7. Bordin, E. S. Diagnosis in counseling and therapy. Educ. psychol. Measmt., 1946, 6, 169-184.
8. Bordin, E. S. Dimensions of the counseling process. Amer. Psychologist, 1947, 2, 415-416.
9. Bugenthal, J. F. T. An investigation of the relationship of the conceptual matrix to the self-concept. Ph. D. Dissertation, Ohio State Univ., 1948.
10. Carnes, E. F. Counselor flexibility: Its extent, and its relationship to other factors in the interview. Ph. D. Dissertation, Ohio State Univ., 1949.
11. Carnes, E. F., and Robinson, F. P. The role of client talk in the counseling interview. Educ. psychol. Measmt., 1948, 8, 635-644.

12. Collier, R. M. A basis for integration rather than fragmentation in psychotherapy. J. consult. Psychol., 1950, 14, 199-206.
13. Covner, B. J. Studies in phonographic recordings of verbal material: III. The completeness and accuracy of counseling interview reports. J. gen. Psychol., 1944, 30, 181-203.
14. Curran, C. A. Personality factors in counseling. New York: Grune & Stratton, 1945.
15. Daulton, J. A study of factors relating to resistance in the interview. M. A. Thesis, Ohio State Univ., 1947.
16. Davis, S. E., and Robinson, F. P. A study of the use of certain techniques for reducing resistance during the counseling interview. Educ. psychol. Measmt., 1949, 9, 297-306.
17. Dial, H. An investigation of the effects on certain judges ratings on interview typescripts when certain additional client information is provided. Ph. D. Dissertation, Ohio State Univ., 1950.
18. Dollard, J., and Mowrer, O. H. A method of measuring tension in written documents. J. abnorm. soc. Psychol., 1947, 42, 3-32.
19. Elton, C. F. The effect of client and topic on counselor behavior. Ph. D. Dissertation, Ohio State Univ., 1951.
20. Elton, C. F. A study of client responsibility: counselor technique or interview outcome? Educ. psychol. Measmt., 1950, 10, 728-737.
21. Froehlich, C. P. Counselors and guidance officers in public secondary schools. Occupations, 1947-48, 26, 522-527.
22. Ghiselli, E. E., and Brown, C. W. Personnel and Industrial Psychology. New York: McGraw-Hill Book Co., 1948.

23. Good, Jeanne E., and Robinson, F. P. Feeling as a criterion of success in different types of counseling interviews. Educ. and Psychol. Measmt., 1951, 11, 639-645.
24. Grant, J. D., and Grant, Marguerite Q. "Therapy readiness" as a research variable. J. consult. Psychol., 1950, 14, 156-157.
25. Grummon, D. L. and Gordon, T. The Counseling Center at the University of Chicago. Amer. Psychologist, 1948, 3, 166-171.
26. Hahn, M. E., and MacLean, M. S. General clinical counseling in educational institutions. New York: McGraw-Hill Book Co., 1950.
27. Haigh, G. Defensive behavior in client-centered therapy. J. consult. Psychol., 1949, 13, 181-189.
28. Hoffman, A. E. A study of reported behavior changes in counseling. J. consult. Psychol., 1949, 13, 190-195.
29. Kamm, R. B., and Wrenn, C. G. Current developments in student personnel programs and the needs of the veteran. Sch. and Soc., 1947, 65, 89-92.
30. Kauffman, P. E., and Raimy, V. C. Two methods of assessing therapeutic progress. J. abnorm. soc. Psychol., 1949, 44, 379-385.
31. Lasswell, H. D. A provisional classification of symbol data. Psychiatry, 1938, 1, 197-204.
32. Lewis, V. Changing the behavior of adolescent girls. Arch. Psychol., 1943, No. 279.
33. Lindquist, E. F. Statistical analysis in educational research. Houghton Mifflin, 1940.
34. McCormick, K. F. A comparison of immediate and delayed criteria in counseling evaluation. Ph. D. Dissertation, Ohio State Univ., 1951.

35. Mann, H. G., and Whitney, D. R. On a test of whether one of two random variables is stochastically larger than the other. Ann. Math. Statist., 1947, 18, 50-60.
36. Morton, R. B. An evaluation of brief psychotherapy: conflicts analyzed within the theoretical structure of Rotter's social learning theory. Amer. Psychologist, 1949, 4, 272-273.
37. Nelson, Jane A. A study of optimum leading: its relation to actual leading and the relation of the deviation of actual leading from optimum leading to outcomes in the counseling interview. M. A. Thesis, Ohio State Univ., 1949.
38. Newcomb, T. Social Psychology. New York: Dryden Press, 1950.
39. Pepinsky, H. B. The selection and use of diagnostic categories in clinical counseling. Applied Psychology Monographs, 1948, No. 15.
40. Porter, E. H., Jr. The development and evaluation of a measure of counseling interview procedures. Educ. psychol. Measmt., 1943, 3, 105-126, 215-238.
41. Raimy, V. C. Self reference in counseling interviews. J. consult. Psychol., 1948, 12, 153-163.
42. Raskin, N. J. An analysis of six parallel studies of the therapeutic process. J. consult. Psychol., 1949, 13, 206-220.
43. Raskin, N. J. The development of the "parallel studies" project. J. consult. Psychol., 1949, 13, 154-156.
44. Reid, D. K., and Snyder, W. U. Experiment on 'recognition of feeling' in non-directive psychotherapy. J. clin. Psychol., 1947, 3, 128-135.
45. Robinson, F. P. Principles and procedures in student counseling. New York: Harper and Brothers, 1950.

46. Robinson, F. P. Two quarries with a single stone. J. Higher Educ., 1945, 16, 201-206.
47. Robinson, F. P. The unit in interview analysis. Educ. psychol. Measmt., 1949, 9, 709-716.
48. Rogers, C. R. (Ed.) Client-centered therapy. New York: Houghton Mifflin Co., 1951.
49. Rogers, C. R. Counseling and psychotherapy. Boston, Mass.: Houghton-Mifflin Co., 1942.
50. Royer, Anne E. An analysis of counseling procedures in a non-directive approach. M. A. Thesis, Ohio State Univ., 1942.
51. Sargent, S. S. Role and ego in psychology. In J. H. Rohrer and M. Sherif (Ed.), Social Psychology at the Crossroads. New York: Harper and Bros., 1951, Pp. 355-370.
52. Seeman, J. A study of the process of non-directive therapy. J. consult. Psychol., 1949, 13, 157-168.
53. Seeman, J., and Dymond, Rosaline. The process and outcomes of client-centered therapy. Third Interim Report, Nov., 1951, Counseling Center, University of Chicago, Chicago. (mimeographed)
54. Sheerer, Elizabeth T. An analysis of the relationship between acceptance of and respect for self and acceptance of and respect for others in ten counseling cases. J. consult. Psychol., 1949, 13, 169-175.
55. Sherman, Dorothy M. An analysis of the dynamic relationship between counselor technique and outcomes in larger units of the interview situation. Ph. D. Dissertation, Ohio State Univ., 1945.
56. Snyder, W. U. An investigation of the nature of non-directive therapy. J. gen. Psychol., 1945, 33, 193-223.
57. Snyder, W. U. (Ed.) Case book of non-directive counseling. Boston, Mass.: Houghton-Mifflin Co., 1947.

58. Stock, Dorothy. An investigation into the inter-relations between the self concept and feelings directed toward other persons and groups. J. consult. Psychol., 1949, 13, 176-180.
59. Tindall, R. H., and Robinson, F. P. The use of silence as a technique in counseling. J. clin. Psychol., 1947, 3, 136-141.
60. Williamson, E. G. Counseling adolescents. New York: McGraw-Hill Book Co., 1949.
61. Williamson, E. G. Counseling and the Minnesota point of view. Educ. psychol. Measmt., 1947, 7, 141-155.
62. Williamson, E. G., and Darley, J. G. Student personnel work: an outline of clinical procedures. New York: McGraw-Hill Book Co., 1937.
63. Wrenn, C. G. Student personnel work in college: with emphasis on counseling and group experiences. New York: Ronald Press, 1951.

APPENDIX

MANUAL FOR TYPESCRIPT ANALYSIS

INTERVIEW IDENTIFICATION

The pool of verbatim typescripts available for study and listed in Table I include protocols from Chicago, Michigan, Minnesota, Missouri, Ohio State and Penn State. The Chicago interviews are identified by client pseudonyms. In the Ohio State typescripts, the letter identifies the quarter in which the protocol was obtained; the first number, the counselor, the second, the client; and the third, the interview in the series. Penn State is identified by a similar system, but the letter "S" stands for Snyder and not for the quarter the record was made. Minnesota and Missouri cases each have a code number; the number or letter after this code number indicates the place of the interview in the series. This is also the case with Michigan interviews. Chicago records label each interview as first, second, etc.

The grouping of interview series in Table I indicates which cases are bound together. Bound volumes of Ohio State, Missouri and Minnesota are labeled with the appropriate case numbers. Others have the name of the school only. The second column in Table I tells the inclusive pages for each series, the code designation of the counselor when that is available and the number of interviews recorded for that series.

Minnesota, Michigan, Ohio State and Penn State type-scripts have the client's speeches take the same number as the preceding counselor's speech and begin a new series at the start of each interview. Chicago and Missouri interviews also number the remarks by pairs, but the change of number depends upon which participant opened the interview series, i.e., when the client makes the first statement, the numbers of remarks change with each client statement, and likewise, when the counselor opens the conference, remarks in that series then change with the counselor's statements.

An explanation of the significance of the symbols or special notations used within the interviews follows:

- C - identifies counselor remarks.
- S - identifies client remarks.
- C 15 - the number indicates the location of the remark within the interview.
- (pause) - indicates a brief gap in the consecutive talk of either counselor or client.
- - is used to show hesitation or indicate an incomplete statement. When these intervals are prolonged they are labeled either (pause) or (silence) and the time is indicated when this break is more than a few seconds.
- u. - the recording was unintelligible at this point.

Table I. Typescripts Available

University of Minnesota

Case #	Inclusive page #	Counselor code letter	No. of interviews in the series
33661	0 - 18	D.K.	1
33728	19 - 65	D	3
38700	66 - 123	A	4
38705	124 - 184	C	6
38715	1 - 15	B	2
38728	16 - 55	D	2
38731	56 - 61	B	1
38745	62 - 85	D.K.	1
38761	86 - 121	A	3
38850	122 - 188	C	7
38740	1 - 284	A	12
38886	1 - 27	D	1
38967	28 - 46	B	2
38935	47 - 75	B	2
39020	76 - 122	D	2
39026	123 - 142	B	2
39060	143 - 157	A	1
39090	158 - 168	D	1
39103	0 - 17	B	1
39154	18 - 29	B	2
39206	30 - 54	D	1
39209	55 - 89	A	2
39260	90 - 106	D	1
39268	1 - 19	B	2
39329	20 - 24	B	1
39345	25 - 33	D	1
39368	34 - 41	B	1
39400	41.1 - 52	D	1
39405	53 - 60	C	1
39538	61 - 67	A	1
41711	68 - 106	D.K.	2

D.K. The counselor's code identity unknown

The grouping of case numbers above indicates which cases are bound together.

The Ohio State University

Case #	Inclusive page #	Coun- selor code letter	No. of inter- views in the series
A 5-5 (2,3)	1 - 23	A 5	2
A 7-7 (1,2)	24 - 45	A 7	2
A 8-8 (1)	46 - 52	A 8	1
A 9-9 (2,3)	53 - 80	A 9	2
A 10-10 (1,2)	81 - 105	A 10	2
E 1-1 (1-9)	98 - 178	E 1	9
E 2-2 (1-8)	1 - 97	E 2	8
E 3-3 (1-5a,5b-7)	1 - 92	E 3	8
E 4-4 (2-7,9,10)	1 - 64	E 4	8
F 1-1 (1-8)	63 - 189	F 1	8
F 1-2 (1-7)	190 - 272	F 1	7
F 2-2 (1,3)	1 - 46	F 2	2
F 2-3 (1-6)	47 - 175	F 2	6
F 3-3 (2-4,8)	1 - 76	F 3	4
F 3-4 (1,3-4,5)	77 - 139	F 3	4
F 4-4 (1,2)	0 - 30	F 4	2
F 4-5 (1-4)	31 - 90	F 4	4
F 5-5 (1-8)	91 - 188	F 5	8
F 5-6 (1-8)	1 - 99	F 5	8
F 5-7 (2)	100 - 104	F 5	1
F 6-6 (1-7)	105 - 198	F 6	7
X 2-2 (4)	1 - 3	X 2	1
X 3-3 (3)	4 - 6	X 3	1
M 4-4 (4)	7 - 17	M 4	1
M 5-5 (5)	18 - 31	M 5	1
B 1-1 (2)	32 - 51	B 1	1
B 2-2 (1,2)	52 - 74	B 2	2
B 3-3 (5)	75 - 82	B 3	1
B 7-7 (2,3)	83 - 108	B 7	2
B 10-10 (1-4)	109 - 170	B 10	4
B 11-11 (1-3)	171 - 190	B 11	3
B 2-2 (3)	191 - 203	B 2	1
FF	204 - 212	FF	1
GG (1)	213 - 220	GG	1

University of Michigan

Case #	Inclusive page #	Counselor code letter	No. of interviews in the series
15688 - 1	1 - 15	A	1
17285 - 3	16 - 29	C	1
17790 - 1	30 - 50	D	1
17905 - 1	51 - 74	B	1
17967 - 1	75 - 82	A	1
18045 - 1	83 - 105	B	1
18253 - 2	106 - 123	A	1

Penn State

Case #	Inclusive page #	No. of interviews in the series
S 2 - 1 - (4,5,6,7)	5 - 8 9 - 16 17 - 22 23 - 27	4
S 3 - 1 - (1)	28 - 33	1
S 4 - 1 - (1,3)	34 - 43	2
S 5 - 1 - (3,4)	45 - 51	2
S 6 - 1 - (22)	59 - 63	1
S 8 - 1 - (3,5,7)	65 - 77	3
S 9 - 1 - (3)	84 - 98	1
S 10- 1 - (1,18)	99 - 104 117 - 122	2
S 12- 2 - (6)	125 - 129	1
S 13- 1 - (1)	130 - 137	1
S 14- 1 - (1,11)	139 - 145	2

University of Chicago

Case #	Inclusive page #	No. of interviews in the series
Boo	2 - 23	3
Foh	27 - 36	2
Lee	36 - 38	1
Pir	42 - 76	9
Sketch	87 - 109	3
Wab (1,3)	119 - 134	2
Will	138 - 153 /	3

University of Missouri

Case #	Counselor code letter	Client number	No. of interviews in series
1	B1	1	3
2	B1	2	2
3	B1	3	3
4	D1	1	10
5	D1	2	4
6	D1	3	3
7	Pr	1	2
8	Pr	2	4
9	Pr	3	2
10	Ru	1	2
11	Ru	2	2
12	Ru	3	2
13	Sh	1	3
14	Sh	2	2
15	Sh	3	2
16	Ya	1	2
17	Ya	2	2
18	Ca	1	1

TRAINING OF JUDGES

The judges should be inducted into typescript rating by having them carefully read the MANUAL FOR TYPESCRIPT ANALYSIS. Chapters IV, V, VI and VII of Robinson's text, PRINCIPLES AND PROCEDURES IN STUDENT COUNSELING might well be supplementary to the manual; especially when the rater has not read appropriate parts of it recently.

Practice ratings on the dimensions which he is to evaluate should then be made. Frequent references to appropriate sections of the manual should be made, especially when the new rater is uncertain about his categorization. These practice ratings should be made over materials previously rated so that there is a ready check for the neophyte. After completing the rating of several interviews, the new rater should check his judgments against the original ratings and make a note of any discrepancies which this second look does not dissipate. These discrepancies should then be discussed with an "expert" rater. It may be advisable to go through this process once more for two or three interviews.

Judging Procedure

Each judge will be assigned specific interviews to classify and rate.

The first job is to divide the interview into units, and at the same time classify the units according to topic. A discussion of the principles to use in this process is included below under a separate heading. When inter-rater agreement is part of the study, judges should independently record on plain paper the numbers of the remarks which represent demarcation points. An example of the notation is:

1. C 1
C 12 as you see it.
2. C 12 How is
S 37

Judges participating in the unit division will confer in order to set a division point as the basis for subsequent analyses of the typescripts.

The agreed-upon division point will be marked in the typescript book on the right margin with a pencil line. When the break is within a remark, a pencil mark will indicate this point.

A Unit Analysis Summary Sheet (following) should be prepared for each discussion topic. This could be prepared from a corrected version of one of the initial records of unit division prepared by a judge. The appropriate page notation might also be made at this time. This need not be prepared by the judge.

UNIT ANALYSIS SUMMARY

Unit Identification

Topic

Growth

*.....

.....

*.....

.....

*.....

.....

Combined.....

Combined.....

Mean Lead

Acceptance of Responsibility

.....

.....

.....

.....

.....

.....

Combined.....

Combined.....

SD of Lead

Talk Ratio

.....

Composite Criterion.....
(See Carnes dissertation)

.....

.....

Combined.....

Mean Optimum Lead

Assigning Responsibility

.....

.....

.....

.....

.....

.....

Combined.....

Combined.....

DRQ

Working Relationship

.....

.....

.....

.....

.....

.....

Combined.....

Combined.....

* Judges:

To rate the dimensions based on an evaluation of individual counselor remarks the rater will fill in on the following Single Remark Rating Form: (1) the interview number, (2) the judge's name, (3) the date. Next in the appropriate line on the rating form the judge will rate each counselor remark on the appropriate dimensions, using as many rating forms as necessary for each interview. Additional forms should be numbered along the left side to correspond with the number of the remark being rated. In some instances the remark represents two or maybe more purposes on the part of the counselor or the client. An example is:

C 20. Well, you've brought up your grade in psychology in the last two weeks, you've ahm, dropped a little bit, but still maintain your above average, ah, grade in anatomy.//
Now what would you feel would be the next step in coordinating your study activities? (OSU, F114, p. 118)

This is called two separate remarks. In order to identify these parts of a remark the rater will label them 117A, 117B, as the case may be, on the rating form. This permits checking of the typescript against the rating sheets. At the end of each interview, a new rating form must be started.

When dimensions which require a global evaluation of a unit are judged, these can be made concurrently.

	Silence	Acceptance	Clarification	General lead	Tentative Analysis	Urging	Unclassifiable Counselor Assumes Complete respons.	Counselor Primarily respons	Respons. Evenly Divided	Client Assigned Primary respons.	Client Assigned Respons. Unclassifiable	Core	Secondary	New	Unclassifiable	Content	Feeling	Unclassifiable
1																		
2																		
3																		
4																		
5																		
6																		
7																		
8																		
9																		
10																		
11																		
12																		
13																		
14																		
15																		
16																		
17																		
18																		
19																		

Interview No.

Judge

Discussion Un

Date

The dimensions included under this process would be assigning responsibility, working relationship, growth, and acceptance of responsibility. The judge will record his rating on the following Unit Analysis Rating Sheet for each of the above dimensions being considered in the research being done. When these are translated into numbers on a 50 point scale they can, at the same time, be recorded on the Unit Analysis Summary Sheet.

In filling out any rating sheet be sure that all identifying data are given, i.e., interview number, discussion unit number, judge's name and the date.

DIVISION OF AN INTERVIEW INTO DISCUSSION UNITS

Definition. The discussion unit is defined as all the consecutive counselor and counselee remarks that pertain to one topic. A topic is a problem or subject around which the discussion of the counselee and counselor center. It is a lower-level concept than the general category of lack-of-skill, and concerns the content or subject matter of the interview. For example, it is possible that a counselee is having trouble getting good grades. Several ways of improving these grades could be discussed, such as improving reading ability, better examination-taking skills; each of these subdivisions is to be considered as a topic. Thus it is

Unit Analysis (See Manual of directions for complete explanation)

Judge

Date

Unit Identification

I. Topic of unit (Check one)

Lack of skill

Too short

Adjustment problem

Special situation

II. Counselor Techniques

A. Amount of lead in the counselor remark

Each separate counselor remark must be checked on the scale. If there are ten counselor speeches in the unit, there will then almost always be ten checkmarks on the rating scale. Median values of a few techniques have been worked out and are indicated on the rating scale as reference points for the rater. However, attention must be given to the total context to determine if that particular remark leads more, less or the same as the point marked on the scale.

Silence

Acceptance

Clarification

General lead

Tentative Analysis

Urging

Little or no lead

Much lead

B. Counselor intention as to division of responsibility

The counselor technique of taking over or throwing responsibility. The judge is asked to rate the amount of responsibility which the counselor himself assumed, which he permitted the client to assume, or which he forced the client to assume. (Disregard desire or intention or resistance of client to this counselor structuring) Make one check mark as an overall rating of the unit.

Counselor assumes complete responsibility

Counselor primarily responsible

Responsibility evenly divided

Client assigned primary responsibility

Client assigned nearly complete responsibility

III. Criteria

The remainder of the ratings are to be based on overall impressions of the entire unit. The scales are marked with Md, Q₁ and Q₃ (median, first quartile and third quartile). Thus a unit which is believed to be slightly below average (for typical interview units occurring in counseling situations with college students) would be checked between Md and Q₁; a unit believed to be better than three quarters or more such units on a given characteristic would be checked to the right of Q₃.

A. Counselor-client working relationship

Good working relationship is the presence of mutual respect and warmth, of respectful consideration for each others ideas, and of client willingness to explore his problems. It is also the absence of symptoms of resistance. At the other end poor working relationship is the presence of symptoms of resistance and the absence of the positive signs of warmth, respect and willingness to explore problems. In judging poor working relationship, consider both frequency and severity of resistance, and make a global rating of the unit.

Poor Q₁ Md Q₃ Excellent

B. Growth during the interview

Growth consists of evidences of new insights, gains in skill, self-acceptance and plans. It is important to rate only gains within the unit, not gains carried over from a previous unit. Again as above, consider both number and importance of gains.

Poor Q₁ Md Q₃ Excellent

C. Client interest in assuming responsibility

How much responsibility does the client want to take, discounting counselor techniques and their influence upon the amount of responsibility that the client actually assumes? That is, this is a measure of client interest in initiating behavior.

Client desires
no responsibility
or even rejects it Q₁ Md Q₃ Client desires
to assume a
great deal of
responsibility

possible for an interview to be divided into several topics or discussion units, all of which might come under the general classification of lack-of-skill. The categories of discussion units will be described later.

In dividing into discussion units, each unit is marked on the typescript or a separate sheet of paper if rating is to be independent. Then each unit is numbered consecutively beginning with one. The division point between units does not have to occur between two distinct speeches, but may occur within one remark of either the counselor or the counselee. In fact, the division may be within one sentence.

Each counselor and counselee remark, in a sense, is a slightly different aspect of a problem. Also, there sometimes is a gradual drifting from one approach to another without any clear break. As long as these various remarks are related to the topic they should be considered as one unit. In general, it is better to make unit divisions at major transition points rather than to sub-divide closely related approaches to one problem.

Divisions between topic areas are often clear in that the exploration or making of plans pertaining to one subject has been exhausted and either the counselor

or the client presents a clearly different subject he would like to talk about.

1. Ce ...So I think I understand where he gets the stuff he tests on now.
Co You think you'll do better on the next test.
Ce Yeah, I really do.
Co Uh huh.
2. Ce But another thing that's been worrying me is that I just don't seem interested in most of the courses I'm taking...

In the above example, the counselee brings up a new topic for discussion and the transition point is obvious.

In the discussion of a given problem, various aspects may be considered which are related to the central topic. Occasionally, however, what starts out as a related aspect may become so developed and diverge so far from the original approach that it must be considered as a new unit. Generalizations are difficult here in that degree of relatedness is a continuum and when and where to make a break is hard to describe. As a rule, if the new approach is tied to the last with subsequent counselor or counselee remarks, or possibly by a return to the original approach before too much time elapses, it may be considered as one unit. However, if the comments originally meant to pertain to the first topic, go further and further afield and have a development and goal of their own, it is a new unit. In these

difficult divisions, the break should be made where there is the most evident change in topic. Since all interviews are divided into units by two independent judges, any disagreements in ratings are discussed and the best division point agreed upon by committee consensus, often with the aid of a third judge.

Often purpose as well as content must be considered in making logical divisions. This is particularly pertinent to the study of many initial and final interviews. As the previous discussion indicates, the initial discussion often includes presentation by the client and inquiry by the counselor which has a single major purpose, the statement of the problem and events relevant to the problem. This process may result in talk about a number of subjects, but will still have a single purpose. In the final interview, when there is a synthesizing and summarizing of previous discussions we also have a variety of topics linked by one purpose. It is more efficient to group these remarks into one unit than to divide them into many small, discrete topics.

Sometimes, in the transition from one unit to another, there are remarks which do not seem to be too directly related to either unit. If these comments can in some way be related to the preceding or subsequent

unit they can be included in the most similar unit. If, however, they cannot be related to either surrounding discussion unit, they must be considered a discussion topic in themselves and be given a new number in regular sequence.

Customarily, at the beginning and end of an interview there are a few social remarks which are not connected in purpose or topic with the main body of the interview. If these are few in number, it does not seem logical to make a separate unit out of them, nor, on the other hand should they be considered as representative of the main body of the interview. Within the interview there also may appear a few remarks which do not seem to have any bearing on the topic under discussion. As a rule, if these client and counselor remarks are less than eight in number, they are to be given the number of the larger unit within which they occur and are also to be given the extra designation of E, for extraneous. These units are not included in the sample of units selected for research study. If, however, the discussion on any one of these special topics continues for more than eight remarks or twelve lines, it must be considered as a unit and numbered in the regular sequence of units. This length was set because units with fewer statements have

previously been found difficult to rate reliably.

TYPES OF DISCUSSION TOPICS

Discussion topics are classified under five headings. The four topic categories are skill, adjustment, lack of maturity and special. Units are given these labels on the basis of the subject of the topic. When any unit includes less than eight client and counselor remarks, it is also labeled "short". The topic category is ordinarily determined by the intent of the client. However, where the client and counselor have different intentions, it should be classified according to the purpose of the counselor.

Definitions of topics follow:

1. Skill

The unit is concerned with tutoring or coaching, or a discussion of certain skills. The client, in such topics, needs to rely on the counselor to determine just what the deficiencies are and how he and the counselor can work together to improve the deficient skills, e.g., study skills, social skills, speech, etc. It differs from an adjustment topic, in that the client is unable to resolve his difficulties through his own thinking.

Typical examples of skill topics include discussions

of: reading rate and comprehension, exam preparation, techniques for reading study materials, etc. Illustrative examples in the typescripts can be found in the Ohio State # F111 to F127 casebook beginning on pages 64, 83, and 169.

2. Adjustment

The unit is concerned with a clarification of thinking on some problem disturbing the client, including vocational, curricular, or personal problems. The counselor's task is to help clarify the client's thinking so he can see the issues in his problem more clearly and accept their implications.

Adjustment problems have at times been divided into "nonemotional" and "emotional" adjustment problems for convenience in discussion. In the former, the client usually has difficulty because he does not know or has not recalled many important aspects of his problem. Given the necessary information and help in integrating it, he can work out a satisfactory solution. In problems of "emotional" adjustment, people have difficulty in understanding and facing certain personal aspects of a problem. While such sub-divisions were once made, only the general classification of adjustment is now used.

Examples of adjustment problems would be vocational choice, financial problems, inability to adapt socially, conflict arising from family relationships, talking about anxieties, etc. Samples of units placed in this category are located in Ohio State casebook F111 to F127, beginning on pages 109 and 163.

3. Immaturity

Such topics are concerned with problems of dependence, overconscientiousness as to the opinion of others, egocentrism, etc. The efforts of the counselor are usually directed toward helping the client think through these problems so as to attain more mature orientations. There are several reasons why discussion topics which may be basically instances of immaturity are seldom identifiable as such during the course of typescript analysis. One reason is that a client seldom feels so immature or dependent that he seeks the help of a counselor for this problem. Many immature persons have no anxieties about their immaturity, and, if they do, are likely to project it into situational problems such as vocational choice, poor grades, etc. Therefore, immaturity problems are difficult to identify, from the client's speeches within a discussion topic and are ordinarily diagnosed as such only from an overall judgment based on long acquaintance with the client.

Consequently, topics which are essentially part of an immaturity problem are usually categorized as adjustment topics.

An example of this type topic can be found in case X55 on pages 9 to 13 in the 781 Training Typescript Book.

4. Special

These units most frequently occur at the beginning and end of an interview. They are not part of the main body of the interview, but tend to serve much the same purpose as the preliminary and closing formalities and statements of a letter. They include discussion which serves to get the interview started; structure the relationship; or end the interview.

Within the interview there also may appear a few remarks which do not seem to have any bearing on the topic under discussion. They are often elicited by outside stimuli such as the class bell ringing, interruptions, something seen from the window, etc.

Typical of special units at the beginning of the interview are comments on the weather, the health of the participants and statements by the counselor structuring the relationship. As suggested above, those imbedded in the body of the interview are usually responses to extraneous stimuli. Special units at the close of the

interview often involve plans for subsequent interviews and "good-by" statements. Samples of the above may be found in the Ohio State casebook # F221 to F236.

Beginning - p. 47

Within body - p. 165

End - p. 21

5. Short

Any unit with less than eight client and counselor remarks or having fewer than twelve lines falls in this category. If it is also identifiable as being in one of the four above-mentioned categories it would be designated by that category and short, e.g., Skill-Short or Special-Short. Short units are given their label solely on the basis of length.

The rationale for this additional category is that brief units have proven difficult to rate reliably. The major difficulty in rating these arises from attempts to make a global evaluation of the unit for a given dimension. Mean dimension values calculated from but a few single remark ratings are unstable and there is some difficulty placing units in one of the four major classes. These units are not ordinarily used in sampling units for research study.

Examples of units that have been judged as being in this group can be found in Ohio State casebook F111

to FL27, on pages 162 and 174.

INTERVIEW DIMENSIONS: COUNSELOR TECHNIQUES AND CLIENT OUTCOMES

The four dimensions of counselor technique which have been used include amount of lead, dealing with the core of the client's remarks, content vs. feeling, and assigning responsibility. All but the last are rated on a remark-by-remark basis on the Remark by Remark Rating Scale¹. Ratings on these dimensions involve evaluating the counselor's remark in the light of the just previous client remark. That is, the judge reads the client remark to determine the ideas or attitudes the client is attempting to express and then rates the following counselor remark with this information as his point of reference.

In addition to the rules given below for rating each dimension, there are two general rules which apply to all single remark ratings in typescript analysis: (1) Whenever the client or counselor speech on the typescript seems so incomplete as to make rating difficult, that speech is to be marked in the "unclassifiable" category, (2) Whenever a rating is made on anyone of the three dimensions and the judge is employing the context

¹Rating of assigning responsibility is done on a global basis and will be described in detail below.

to arrive at his rating rather than the immediately preceding client remark, that rating must be indicated by encircling a dot (or dots) on the rating scale. It needs to be emphasized here that the great majority of the ratings will be made on the basis of the immediately preceding client remark.

Amount of Lead

Variation in amount of lead is a technique used by counselors to further the client's thinking. The degree to which a remark leads usually depends on two characteristics: (a) how far the content of the counselor's remark seems to be ahead of the content of the client's last remark, and (b) the degree of pressure or definiteness in the counselor's remark used to bring about client acceptance of new ideas. To illustrate, an acceptance remark adds no new idea and little pressure (except, perhaps, to continue talking); tentative analysis brings in a new idea which logically follows but little pressure is exerted to accept it; an interpretation remark states a new idea with definiteness which indicates that the client is expected to accept, etc. While a counselor may not intend his remark to lead very much, if it seems to do so to the client, it does. The criterion of leading is the relation the counselor remark has to the preceding client remark in the eyes of the client.

A preliminary study indicates that typical examples of a particular type of counselor technique have certain modal values of amount of lead. These values for certain techniques are indicated on the rating scale as reference points for the rater. These techniques are defined below and further discussion and examples will be found on pages 82-95 of PRINCIPLES AND PROCEDURES IN STUDENT COUNSELING.²

²Silence. Occasionally following a client remark the counselor is silent. There is no new idea here and generally little pressure. It is conceivable, however, that silence could mean as much pressure as exerted by a rejection remark.

Acceptance. Following a client remark, the counselor may merely indicate that he understands and accepts what the client has said by a "yes" or "mm hmm" response. Such remarks usually lead very little. Care must be taken, however, to differentiate an acceptance remark from "yes" in response to a client question, or a "yes" which has a connotation of approval.

Clarification. Following a remark by the client, the counselor phrases the important ideas in more concise or clearer terms. A clarification remark leads slightly more than acceptance in that the counselor has done some selecting of the various ideas in the client's remarks.

General lead. A request for the client to select a topic for conversation or to continue talking about a topic. It is of the type, "what do you want to talk about today" or "will you tell me more about that."

Tentative analysis. The counselor presents an additional aspect or a new approach to the problem being discussed, but in a purely tentative, non-forcing manner. The client feels free to accept or reject the idea. Frequently it is in a question form.

Urging. Those instances in which the counselor puts considerable pressure on a client to accept an idea or course of action. It is characterized by the use of such devices as arguments for the action and occasionally even threats. Rejection of an idea would be an extreme form of urging.

On the other hand, it is important to note that all counselor remarks of a given type, e.g., clarification, do not represent this modal value. Some such remarks may lead more or less than this indicated value. In fact, a given type of remark may have a value beyond the next labeled point on the scale. In judging each remark individually, even of the same type listed on the scale, attention must be given to the total context to determine if that particular remarks leads more, less, or the same as the point marked on the scale. While the types of remarks listed on the scale tend to be more consistent in degree of leading than many other types of remarks, it should be noted that a particular urging remark, for example, may represent a weak lead but an urging remark of the rejection type would represent a very strong lead. Generally there is a tendency for those techniques at the right end of the scale to vary more from the median value than those at the left. Some indication of the broad areas in which highly variable and unlisted techniques tend to fall is given in footnote 3 below. Also see Table 3 on page 83 of PRINCIPLES AND PROCEDURES IN STUDENT COUNSELING for some typical score values obtained by a sampling of various kinds of remarks.

³Some techniques are not included on the rating scale either because they occur infrequently or they have such a wide range of leading in terms of context and manner. Definitions follow for help in classifying these remarks in your ratings.

Summary clarification. An organization and summary of the ideas present in a series of client remarks. Since the whole series of remarks is clarified, this technique leads a little more than simple clarification.

Approval. In response to a client remark, a counselor may express approval of selected points and increase the likelihood of their influence. Since the client has stated the point, he feels little pressure. this technique frequently falls within the left half of the scale.

Interpretation. The counselor states something which can be inferred from what the client has said or done, but which the client has not specifically discussed. It has a wide range of amount of lead, shading into (and difficult to distinguish from) clarification at one extreme and urging at the other.

Assurance. In response to a client remark, the counselor makes a response of the type, "Everything is going to be all right." This may lead a great deal if it tends to reject the importance of the client's problem.

Introducing an apparently unrelated topic. In the eyes of the client, the relatedness of a topic which a counselor introduces may vary a good deal, and hence the amount of lead varies considerably. The more unrelated it seems, the more the remark leads. It may seem to the client both a rejection of what he has been discussing and the introduction of a new topic.

Special rules. The matter of (a) counselor replies to client questions, (b) questions asked by the counselor more for his own than for furthering the client's thinking, and (c) structuring speeches represent special

situations and will be classified as follows:

(a) When a counselor answers a direct question of the client in a manner that allows the client some freedom in his reply, such a response is considered at about the level of a general lead. For example,

S - Do you have my test scores?

C - I don't have them, but they are readily available to me.

S - I forgot, what day did you set for our next meeting?

C - Next Wednesday at ten o'clock.

(b) Counselor questions, such as the following, which seem irrelevant, lead a great deal and their modal value on the scale is usually considered as leading more even than a typical urging remark.

(The previous conversation deals with when the counselor asked about the client's notes.)

S - They help me. I don't know... there's so much to remember.

C - That's a five hour course, isn't it?

S - I don't know what to do... that's my trouble. We outline every chapter every night.

C - In class?

S - No, we do that in the evening. He assigns lessons to be outlined which I like.

C - Does that count on your grade?

(c) When a counselor gives a structuring speech such as is illustrated in the example, the modal scale value is also considered around the general level of an urging remark.

C - Did the instructor explain to you the function of these interviews?

- S - Not exactly, he explained without giving details.
 C - I'm available here for one meeting a week to help you with whatever sort of problem you might have. Usually the interview will last about thirty minutes but that depends on what you want to talk about.

In using the rating scale below, which is part of the Single Remark Rating Sheet, each separate counselor remark must be rated by using an individual line and identifying number. A counselor remark is defined as all the counselor talk between two statements by the client. Exceptions to this may occur when the client interjects a remark during a counselor statement or a break in a long counselor speech indicates that more than one "remark" is being made. Notice that silence, which will be shown on the typescript as "(pause)", when it occurs alone and seems deliberately continued by the counselor is rated as a counselor remark.

Since the dimensions dealing with the core of the client's remarks and content vs. feeling were not used in this investigation, the parts of the manual describing rating procedures for them was deleted.

Responsibility Division: Counselor Assumption of Responsibility and Client Assumption of Responsibility

The progress of the interview is determined by a number of events which occur during the development of a unit, and the responsibility for initiating and continuing each of these events may be carried by either of the participants in varying degrees. Typically there is first an introduction of a topic; during the discussion there may be a diagnosis of difficulties, interpretation of behavior or test scores, demonstrations of behavior or skills, suggestions of ideas, etc. Finally there is the making of plans and sometimes a summary of the unit.

The degree of responsibility a participant takes for the progress of the unit is a function of how much he wants to take and of what the other participant forces on him or permits him to take. A dominating counselor may have a client who appears satisfied to have him take complete charge of the interview, or he may have a client who tries himself to take a great deal of the

responsibility. In like manner a counselor who does not want to direct the interview may have a client who seems willing to take primary responsibility, or he may have one who, like himself, does not. We have here a problem of differentiating between the counselor technique and the client behavior.

In order to assess the amount of responsibility of both participants in the interview, two rating scales are used. In scale A, which follows, the judge is asked to rate the amount of responsibility which the counselor himself assumed, which he permitted the client to assume, or which he forced the client to assume, i.e., the counselor technique of taking over or throwing responsibility. This is somewhat independent of how much responsibility the client actually takes; the essential point to look for is the amount of responsibility that the counselor plans for himself in this unit.

While the other counselor dimensions have been rated on a remark-by-remark basis, this counselor dimension represents a global evaluation of the general tendency to assign responsibility during the entire discussion topic. That is, what general impression would the client or an outside observer have of the general manner in which the counselor was dividing the responsibility.

A. Counselor intention as to division of responsibility:

.....				
Counselor assumes complete respons.	Counselor primarily respons.	Respons. evenly divided	Client assigned primary respons.	Client assigned nearly complete respons.

It is important to remember that we are interested in what the counselor is attempting to do with "responsibility", and that we are not concerned in the above scale with the actual division of responsibility, i.e., whether the client accepts it or not.

Acceptance of Responsibility.

The client behavior counterpart of counselor assigning of responsibility is influenced by the same factors as that dimension, but it needs an independent rating. This is the first of the client outcomes to be discussed but it seems more convenient to discuss it here in conjunction with counselor assigning of responsibility.

In the second rating scale the judge is asked to rate the amount of responsibility which the client is interested in or appears willing to take. That is, the counselor influences by his technique the amount of responsibility that the client actually assumes, but discounting counselor techniques, how much responsibility

does the client want to take? That is relative to clients in other counseling units where the counselor has assumed similar responsibility (scale A) does this client seem on his own to be taking an average amount of responsibility or more or less? (Scale B) For example in an "urging" unit, a client does not have much chance to participate but his limited behavior here may indicate a much greater desire to contribute to the conference than another client in a "clarification" unit, although the latter is forced to do more. In brief, this is a measure of client interest in initiating behavior. The Q_1 , Md, and Q_3 refer to a distribution of client's assumption of responsibility under comparable counselor assigning of responsibility. That is, is this client in the top quarter or other quarter of students in trying to take responsibility under these conditions? Here again it should be remembered that a global rating for the entire discussion topic is wanted, not a remark-by-remark rating.

B. Client interest in assuming responsibility:

Client				Client
desires	Q_1	Md	Q_3	desires
no respons.				to assume
or even				a great
rejects				deal of
it				respons.

Practically all of the remainder of the ratings are based on overall impressions of the entire unit rather than on a remark-by-remark analysis. The scales are marked with Md., Q_1 and Q_3 (median, first quartile, and third quartile). Thus a unit which is believed to be slightly below average (for the typical interview units occurring in counseling situations with college students) on a certain characteristic would be checked between Md and Q_1 ; a unit believed to be better than three quarters or more such units on a given characteristic would be checked to the right of Q_3 .

Working Relationship.

Good working relationship between the counselor and client may be defined, positively, as the presence of mutual respect and warmth, and of the respectful consideration for each other's ideas and suggestions. It is usually characterized by a pooling of efforts in the attack on the problem and by free talking. It may be defined also as the absence of indications of resistance on the part of the client.

Poor working relationship, then, is the inverse of the above, and is characterized by numerous and/or severe symptoms of resistance on the part of the client. These symptoms may consist of a belligerent rejection of the counselor's point of view, unwillingness or

refusal to talk about a real problem, apathy as evidenced by frequent silences or monosyllabic responses or polite speech which remains superficial and unenthusiastic, attempts to close the interview, or other signs of obvious lack of cooperation with the counselor.

In judging poor working relationship, it is necessary to consider both the severity and the frequency of the resistance. That is, outright rejection of the counselor may be indicative of strong resistance. On the other hand, silence and monosyllabic responses at one point may not be serious, but if present throughout a unit, may indicate fairly strong resistance. Research evidence indicates that continued client apathy may be as strong as active resistance. It should also be noted that a silent client may not be evidencing resistance to the counselor but may be confronted with a problem which is difficult to face or verbalize. This latter would not necessarily represent a poor working relationship.

The fact that a client modifies or questions a counselor idea does not necessarily indicate poor working relationship; it is sometimes a sign of good working relationship that a client feels free to respond frankly. It must be judged from the client's manner whether he seems to be showing resistance to the counselor himself

or to the ideas. Also notice that a client response of "Mm hmm" is not necessarily indicative one way or the other; other factors must be considered on making the rating.

Working relationship:

.....
 Poor Q_1 Md Q_3 Excellent

Descriptions of the dimensions growth and talk ratio were omitted from this appendix since they were not used for ratings made in this research.

AUTOBIOGRAPHY

I, John Edward Muthard, was born in Reading, Pennsylvania, April 17, 1917. I received my secondary school education in the public schools of the city of Reading, Pennsylvania. My undergraduate training was received at the State Teachers College at Kutztown, Pennsylvania, from which I received the degree of Bachelor of Science in Education in 1942. I received the degree of Master of Arts from Temple University in 1947. From 1946 to 1948 I was employed as a vocational appraiser by Albright College and from 1948 to 1950 as an instructor at Emory University. While completing the requirements for the degree Doctor of Philosophy at the Ohio State University, I held a position as teaching assistant from 1950 to 1951 and from 1951 to 1952 as a research assistant in the Department of Psychology.