To the Graduate Council:

I am submitting herewith a dissertation written by Louis Sebastian Berger entitled "The Logic of Observation in Psychotherapy Research." I recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Psychology.

Howard R. Pollio, Major Professor

We have read this dissertation
and recommend its acceptance:

Accepted for the Council:

Vice Chancellor
Graduate Studies and Research

THE LOGIC OF OBSERVATION IN PSYCHOTHERAPY RESEARCH

A Dissertation

Presented for the

Doctor of Philosophy

Degree

The University of Tennessee

Louis Sebastian Berger

August 1974

ACKNOWLEDGMENTS

It is a great pleasure to acknowledge my debt to my committee, Professors James Anker, Lida Barrett, Stephen Handel, and Howard Pollio (Chairman). Each gave generous and valuable help throughout all stages of the dissertation.

My very particular and special thanks go to Dr. Pollio. He not only encouraged an unconventional effort from the very beginning, provided major guidance and support as chairman, and made technical contributions in his many areas of interest and expertise, but even more important, he unfailingly was a very human and accessible friend. To a large degree it was this relationship that turned the traditionally painful process of dissertation writing into an almost pleasurable experience.

1158615

ABSTRACT

An analytical study of methodology in psychotherapy research was performed. Although the study selected clinical research for analysis it addressed basic, quite general methodological issues. Thus, the outcomes of the study--analytical tools, systems of observation, and approaches to mathematization of data--apply not only to psychotherapy research, but are broadly relevant to experimental psychology in general.

The study began with a comprehensive review of psychotherapy research methodology, emphasizing traditional approaches, problems, and recommended solutions. As a result, it was concluded that two fundamental processes in which research is formally grounded--observation and mathematization of acquired data--had so far failed to receive adequate analytical treatment congruent with their importance.

Consequently these two principal processes were analyzed through formal models. Observation was examined in terms of a state process model; 14 observational systems prominent in psychotherapy research were analyzed. A parallel analysis of mathematical operations was based on relational systems concepts familiar from measurement theory.

The dual analysis led (1) to a design philosophy for observational systems and (2) to the recommendation to return for the present to a logically primitive mathematical system, namely, the relation scale. The design philosophy suggested an ideal of complete, pragmatically conceived observational systems: A family of multimodal, multilevel, modular systems was envisioned.

A prototype system was constructed to illustrate the recommended approach to experimental observation. It consisted of loosely-knit diverse subsystems drawn mostly from the literature and of newly-proposed subsystems which would allow idiographic representation of events. This system appeared to be highly flexible, and easily capable of improvement and modification.

Future research areas were identified and discussed. These areas included completeness, redundancy, reliability, as well as the use of relation scales to mathematize data. The topic of other, alternative observational systems in the proposed family was also considered.

TABLE OF CONTENTS

CHAPTER		PAGE
I.	RESEARCH IN DYADIC PSYCHOTHERAPY	1
	Introduction	1
	Traditional Conceptualizations	5
	Traditional Difficulties and Traditionally Recommended Solutions	19
	A Different Critique	36
II.	OBSERVATION AS A FORMAL PROCESS	44
	Ingredients of Observation	44
	A Proposed Model	48
	Analysis of Some Major Systems of Observation	50
	Critique	89
	The Search for a Comprehensive System	96
III.	THE MATHEMATICS OF OBSERVATION	100
	Standard Approaches	100
	Mathematics and the Proposed "Complete System" of Observation	108
IV.	DEVELOPMENT OF A SYSTEM	115
	A First Approach	116
V.	SUMMARY AND FUTURE DIRECTIONS	145
	Future Directions	146
REFERENCES		154
VITA		165

LIST OF TABLES

TABLE		PAGE
1.	Major Reviews of Research in Psychotherapy	4
2.	Organization of Core Reviews	7
3.	Difficulties and Solutions in Experimental Design Issues . .	22
4.	Difficulties and Solutions in Measurement Issues	23
5.	Difficulties and Solutions in Variables Issues	24
6.	Difficulties and Solutions in Data Analysis Issues	25
7.	Major Systems of Observation	52
8.	Outline of the Assembled System	142

LIST OF FIGURES

FIGURE		PAGE
1.	Strupp's System	55
2.	Jaffe's System	58
3.	Butler's System	61
4.	Auld and Dollard's System	63
5.	Gottschalk's System	64
6.	Leary's System	65
7.	Lennard's System	67
8.	Mahl's System	68
9.	Matarazzo's System	70
10.	Rogers' System	71
11.	Scheflen's System	72
12.	Deutsch's System	75
13.	The Menninger Project's System	77
14.	Pittenger, Hockett, and Danehy's System	82
15.	Model of a Typical Content Analysis Process	85
16.	Principal Levels of the Proposed System of Observation	122
17.	Major Subdivisions of Patient or Therapist Behavior Levels	123
18.	The Patient's System--Verbal Levels	128
19.	The Patient's System--Paralinguistic Levels	130
20.	The Patient's System--Body Movements and Gestures	133
21.	The Patient's System	135

FIGURE PAGE

22. The Therapist's System 138

23. Interactional System 141

CHAPTER I

RESEARCH IN DYADIC PSYCHOTHERAPY

Introduction

Ask a practicing psychotherapist, a researcher active in the area, or a methodologically oriented scientist (these may or may not be the same individual) about his impressions of research in psychotherapy, and you are likely to get a less than sanguine response. Criticisms might include comments about the low utility of the yield (from the practitioner), fragmentation of research effort (from the researcher), or methodological inadequacy (from the formalist). It seems that research in psychotherapy is, and has been for a long time, in an unsatisfactory state. Yet, as evidenced by the size and growth rate of the literature, research is burgeoning; the area is felt to be highly important because those involved in the field recognize the profession's responsibility to separate dogma and myth from fact, and to offer the public they serve "more adequate, efficient and effective method of psychotherapeutic intervention [Frank, 1961, p. 92]."

If it is true, as is generally accepted, that the research discipline has been plagued from its beginnings by deep, persistent, intractable problems that have withstood the assaults of a considerable number of very able, dedicated, imaginative researchers, what could one expect to do to improve this state of affairs? The proposed approach will emerge from the discussion of this chapter, but to give an idea of

its direction, the argument will be that an approach calling for a different methodology is in order.

Pointing to methodology as a solution to fundamental problems is a recommendation that is viewed with suspicion in many quarters, and properly so. This distrust is quite charmingly expressed in a letter from Bergin to his coauthor which he closes with "I always enjoy talking to them (that is, methodologists, mathematicians and philosophers of science), but privately I always have my guard up to keep from being 'snowed' [Bergin and Strupp, 1972, p. 180]." It may help reassure the reader to say that the methodological avenues to be explored rather than being esoteric, arcane, generally inaccessible or incomprehensible, will on the contrary focus on basic processes, although perhaps from a different perspective.

The plan of the chapter is to first discuss the way in which the field has been traditionally conceived, and then to describe traditionally recommended approaches to well-known, traditional difficulties. Throughout the literature there have been occasional references which have implications for unorthodox research approaches. The last section of this chapter will present those approaches that appear to be precursors to this study as well as a sketch of the approach to be pursued.

Not the least of the problems involved in undertaking a methodological study in this area--of doing research on research in psychotherapy--is the enormous bulk of published material. The problem of mastering the material in reasonable time is illustrated by Bergin and Strupp's (1972) remarks. These highly regarded, experienced

researchers report that in the early stages of their extensive survey they realized that there was still an enormous amount of new, unfamiliar research material. Thus, even long-time, intimate familiarity with psychotherapy research does not enable the would-be methodological analyst to be cognizant of every one of its important aspects. As Malan (1973) notes, it does take an effort of the magnitude, in terms of resources and time, of Bergin and Strupp's four-year study to prepare a complete review. Since this study is methodological, not substantive, however, it is not as essential to be aware of all research results; it should be sufficient to be familiar with major, representative literature reviews which cover the important methodological issues. The present study therefore has relied heavily on major reviews and other key writings which, as one becomes familiar with the literature, are easily identifiable almost thrusting themselves on the reader.

The thirteen major reviews which have been used extensively in the present work are listed in Table 1. It can be seen that these reviews span a period from 1959 to 1973, thus providing a broad historical perspective for a study of the important issues and characteristics of research is psychotherapy. They cover a wide range of approaches to reviewing this area, including such diverse works as proceedings of three important conferences (Rubinstein & Parloff, 1954; Shlien, 1968; Strupp & Luborsky, 1962), a general state-of-the-art review in the Annual Reviews of Psychology (Howard & Orlinsky, 1972), a comprehensive handbook covering virtually every major topic (Bergin & Garfield, 1971), a detailed review of outcome studies (Malan, 1973), as well as other

Table 1

Major Reviews of Research in Psychotherapy

Authors	Abbreviations
Bergin and Garfield, 1971	BG
Bergin and Strupp, 1972	BS
Fiske, Luborsky, Parloff, Hunt, Orne, Reiser, and Tuma, 1970	F
Gottschalk and Auerbach, 1966	GA
Golstein and Dean, 1966	GD
Howard and Orlinsky, 1972	HO
Malan, 1973	MAL
Meltzoff and Kornreich, 1970	MK
Millon and Diesenhaus, 1972	MD
Rubinstein and Parloff, 1959	APA I
Shlien, 1968	APA III
Strupp and Bergin, 1969 (bibliography)	SB
Strupp and Luborsky, 1962	APA II

Note.--These abbreviations will be used in Tables 2-6 and in Figures 1-14.

kinds of summary documents. The rich variety of approaches provides a very complete, multifaceted view of the general research area. Of course, in the present study the core literature of Table 1 was supplemented by a large number of additional references.

As a consequence of the size of the literature, a long list of references could be given for almost any point made in a study such as this one. For example, remarks concerning the identification and selection of important research variables can be found in more than 50 instances in five of the core volumes. Quite apparently, there is much overlap; there are also a great many relatively minor points made about different topics. The overlap should assure the researcher that redundancy will guard against omission of major factors on issues while the plethora of minor points forces one to resign himself to less than perfect completeness. Where references seem required, one or two major representative sources have been selected; minor material has usually been included without explicit citations.

Traditional Conceptualizations

Traditionally, research in psychotherapy has been set within a framework of certain underlying assumptions. These concern (1) the nature of psychotherapy as it is conceived by researchers--that is, its characteristics as an object of experimental study; (2) the broad, fundamental research questions posed--the basic, general concerns from which the more specific questions addressed by particular research studies derive; and (3) the appropriate, accepted ways of carrying our

research--the standard steps and common practices which constitute traditional research. In this section, each of these will be considered.

Psychotherapy as the Target of Research

It is important to understand how a discipline conceptualizes the phenomena which it believes constitute the subject matter of its research area. Such understanding may help reveal tacit assumptions which could circumscribe research. One approach to learning how researchers in psychotherapy view the object of their research--how psychotherapy as a research object is constituted--is to identify the relevant terms appearing prominently in the literature.

Table 2 lists major descriptors used in thirteen core publications. Terms which are used to head chapters or other major sections in the core reviews presented in Table 1, page 4, have been assembled and grouped into classes. Table 2 lists four major classes of descriptive terms with subcategories identified within a class. Specific descriptors actually used are also listed. The table shows how often descriptors from a given subcategory appeared as a chapter or section heading in each review. Summary tabulations for totals within each category and subcategory appear in the last column. In addition, for each review, totals for a category are also shown.

The four major categories are (1) experimental method and design; (2) variables, (3) techniques of therapy, and (4) outcome/process. A fifth category ("mixed") of descriptors combining terms of the above four classes is also listed, as is a sixth ("miscellaneous"), minor category for which only summary data are given.

Table 2

Organization of Core Reviews

CATEGORY, Subcategory, descriptors	APA I '59	APA II '62	GA '66	GD '66	APA III '68	SB '69 (bib.)	F '70	MK '70	BG '71	BS '72	HO '72	MD '72	MAL '73	Total
EXPERIMENTAL METHOD, DESIGN														
General (6): design; statistics; control; quantitative research; specifying the problem; general studies.	1			1		1	2	1	1	1		2	1	11
Data (4): collection; analysis, analysis and interpretation; measurement; collection and measurement.			3	1								2		6
Techniques and Methods (10): content analysis; methods; multifactorial; psychophysiological; analogue studies; systems analysis; verbal processes and techniques; methods of assessment of change; measuring personality change; theory, method, experimentation.	1	1	1	1		3		1	3		1			12
Total:	2	1	4	3		4	2	2		1	1	4	1	29
VARIABLES														
Referent (7): client; patient treatment; therapist-patient relationship; patient-therapist interaction; expectancies, outside features.	1				1	1	4	2	1	1	1			12
Modal (4): input; temporal; dependent; psychopharmacological					1		1	1						3
General (1): general.			.											
Total:	1		1		2	1	5	3	1	1	1			16

Table 2 (continued)

CATEGORY, Subcategory, descriptors	APA I '59	APA II '62	GA '66	GD '66	APA III '68	SB '69 (bib.)	F '70	MK '70	BG '71	BS '72	HO '72	MD '72	MAL '73	Total
TECHNIQUES OF THERAPY														
Schools (8): behavior therapy; counter conditioning; operant approaches; modeling; social psychology approaches; learning; analysis of client-centered, psychoanalytical, related eclectic; nonprofessional.					1	2		7						10
General: (6) comparative, techniques; therapist skills; prognosis, duration, prediction; clinical innovation; treatment.						2	1		2	1				6
Total:					1	4	1		9	1				16
OUTCOME/PROCESS														
(no categories) (8): outcome criteria; output; effect of therapy; Eyesenck's position; outcome studies, review and evaluation; outcome evaluation; process; outcome.														
Total:			2	2		2		3	1	1	2		2	15
MIXED:														
(Variable, Methodology) (2): variables, measures, and procedures; definition, measurement, analysis of significant variables.		1								1				2

Table 2 (continued)

CATEGORY, Subcategory, descriptors	APA I '59	APA II '62	GA '66	GD '66	APA III '68	SB '69 (bib.)	F '70	MK '70	BG '71	BS '72	HO '72	MD '72	MAL '73	Total
MIXED (continued)														
(Variable, Process/Outcome) (5): therapist as a process variable; patient as outcome variable; therapist as outcome variable; therapy methods as an outcome variable; patient contribution to treatment process.		1						4						5
(Variables, Technique of Therapy) (1): therapist, therapist's behavior, personality and technique.						1								
Total:		2				1		4		1				8
MISCELLANEOUS														
Historical, Philosophical, Programmatic, Teaching of Psychotherapy (14)														
Total:			2	2			2	1	2	2	1		2	14
Total, All Categories:	3	3	9	7	3	12	10	13	7	7	5	4	5	98

Table 2 will be considered later in this chapter; for the moment it is used mainly to point out the frequent appearance of descriptors in the "variables" category suggests that the concept of variables is an important one. Indeed, when major reviews such as Meltzoff and Kornreich (1970) or Gottschalk and Auerbach (1966) are examined, one finds that variables do occupy a highly prominent place. They are numerous and varied; variables are used to measure such diverse phenomena as anxiety, therapist warmth, or eyeblink rate and are ubiquitous in research discussions. Many methods for classifying variables have been proposed or used in the literature (Parloff & Rubinstein, 1959, p. 288; Robbins & Wallerstein, 1959, p. 30; Shakow, 1959, p. 111; Marsden, 1971, p. 346; Levinson, 1962, p. 15; Strupp, 1962, p. 273; Auerbach & Luborsky, 1968, p. 157; Lang, 1968, p. 92; Meltzoff & Kornreich, 1970, p. 24; Millon & Diesenhaus, 1972, pp. 67-148). These methods usually fall into one of two types of categorizations. One type of taxonomy classifies variables according to the object, person, or situation the variable refers or "belongs" to; variables classified by means of such schemes may be called "referent" variables. An example of this approach to taxonomy is the frequently used scheme which classifies variables into therapist, patient, interaction (between patient and therapist), psychotherapeutic technique, and situational variables. The second type of taxonomy may be called "modal," as it classifies variables according to modality or format--their physical or psychological properties or attributes. Some examples of categorizations which use the modal approach are (1) biophysical, intrapsychic, phenomenological,

behavioral, and documentary (Millon & Deisenhaus, 1972) or (2) verbal, contentual, localization quality, kinesic, and covert physiological (Shakow, 1959, p. 111) taxonomies.

These two kinds of taxonomies for variables can be applied equally to general experimental psychology. There, variables are also frequently classified in terms of their referent--for example, as "belonging to" the environment (stimuli; situational variables) or to an organism (organismic; subject; experimenter; response variables). Alternatively, experimental variables can be classified according to an attributional, modal dimension--for example, physical characteristic or modality (color, pitch, loudness, movement, etc.), informational value, and so on.

These parallels between conceptualization of variables in psychotherapy research and general experimental psychology are suggestive. (Other similarities will be mentioned below.) One may infer that researchers traditionally view psychotherapy basically as an experimental situation just like any other in experimental psychology. We readily find confirmation in the literature. An extended analysis is given by Sargent (1961); she demonstrates that methodological behaviorism, the basic position of traditional experimental psychology, is the dominant position in psychotherapy research as well. The matter is succinctly summarized by Bordin (1959, p. 238): "We treat the therapist as a person, his office setting, and the tasks he sets the patient, as a stimulus field towards which the patient must react." (See also Butler, Rice & Wagstaff, 1962, p. 189; Lennard, 1962, p. 218.)

The parallels between psychotherapy research and general experimental psychology are quite straightforward. In the general experimental (associationistic) model the variables define a subject, his organismic characteristics, the properties of impinging stimuli, and the characteristics of the subject's behavioral responses. Analogously, in therapy the therapist, psychotherapeutic intervention, and environmental situation are considered stimulus variables for the subject/patient, who in turn has attributes described by organismic variables, and behavior defined in terms of response variables.

In its simplest form, then, the general experimental model assumes that the therapist/situation/technique does something to the patient which elicits or reinforces a response. This model can easily be expanded by including the therapist as subject as well. As Beck notes, "The patient talks to his therapist and the therapist to the patient [1966, p. 551]." The obvious extension is to interpret the patient's responses as stimuli for the therapist. In this way the original model is elaborated into a chaining model, sometimes also referred to as an action-reaction, or coder-decoder model (Scheflen, 1966, p. 268). It is like a somewhat peculiar, alternating, double version of the basic linear, associationistic model (see also Pittenger, Hockett & Danehy, 1960, p. 224). The identification of psychotherapy as one more standard situation in experimental psychology is thus reiterated.

Basic Research Questions

The basic, broad questions which are the concerns of a scientific research area may be overt and recognized, or not. In either case, they

are the matrix in which specific questions asked in a given research program are grounded. Thus, they evidently are important because they channel traditional research. The different perspectives and interests of therapists, methodologists, and theoreticians are reflected in differences in their conceptualization of what the fundamental questions are that constitute the proper domain for research in psychotherapy.

The practitioner wishes to have research yield information that is directly useful in his practice. His concerns are pragmatic and he seeks guidance in practical matters such as evaluation of patients for optimum selection, accurate diagnosis and prognosis, and appropriate selection of treatment approach. One major concern thus is to identify appropriate, effective therapeutic techniques by which he can best help a specific patient to achieve specific goals (Bergin & Strupp, 1972, p. 8). A second general area of great interest to many clinicians is the evaluation of outcome. Ever since Eysenck's unsettling (1952) review challenging the effectiveness of psychotherapy, therapists have looked to research for support of their empirically based conviction that therapy is effective.

The theoretician is concerned with mutual relationships between theory and experiment and thus wants to know how theory accounts for phenomena and what implications phenomena have for theory-building. Faced with experimental data, one of his tasks is theoretical explication; he attempts to account for the data within an existing theoretical framework, or, if this attempt fails, he seeks to develop a suitable,

adequate theory. Alternatively, given a theory and its experimental implication, the theoretician asks experimental questions which will enable him to test his constructions, and which can provide support or perhaps modifying information for his theoretical position. It should be noted that frequently the theoretician is also therapist and therefore a rigid distinction between theoretician and therapist cannot always be easily maintained.

The methodologist's responsibility is to insure that research is performed effectively, efficiently, and in general according to the accepted criteria of good scientific practice. Therefore, the questions he asks reflect his role as an expert who on the one hand is charged with supplying useful, supportable, well-grounded answers to therapists' and theoreticians' questions, and who on the other hand is motivated to pursue relevant methodological questions in their own right, questions whose answers would provide useful techniques or tools for subsequent research. Given his perspectives, a methodologist's questions will focus on the major methodological areas--experimental design, measurement problems and techniques, and the mathematical treatment of data. These areas are considered below in the discussion of traditional research practices.

Research Procedures

Having considered the situation to be researched and the basic questions to be asked, the traditional conceptions of how research is to be accomplished are explored next. What is the general approach, and what are the typical steps to be followed? As a consequence of

adhering to the implicit model of psychotherapy described above, research in psychotherapy is traditionally conceived as being no different in principle from research in general experimental psychology (Saslow & Matarazzo, 1959, p. 157; Parloff & Rubinstein, 1959, p. 286). While it is conceded that the therapy situation may be considerably more complex than most laboratory investigations such as animal experiments--there may be more variables, they may be more difficult to measure, control may be more difficult to achieve--it is generally assumed that the standard methods and approaches of experimental psychology are also valid here (Edwards & Cronbach, 1952). Basically, this approach assumes that the phenomena are ultimately quantifiable and measurable; that they will be adequately understood when all the variance is accounted for in terms of independent variables and their interaction; that at our stage of imperfect understanding it is necessary and appropriate to study pieces or simplified aspects of the total process, relying on additivity (or, at most, on inclusion of nonlinear interactions) to arrive at increasingly better solutions by piecing together partial solutions, solutions to selected aspects of the total phenomena.

In short, the research paradigm in psychotherapy is that of general experimental psychology, which in turn is the paradigm classically followed by the natural sciences. Thus, the traditional pre- and post-research steps are the usual ones found in experimental psychology. Since these steps are widely recognized, they will be presented here in brief outline only.

Preresearch phase. The preresearch phase usually begins with translation of the general questions described above to a researchable form. This translation is a complex, only partially formalized process. Some of the factors that influence and guide it are the current level of knowledge or understanding, the researchers' theoretical orientation, the availability of methodological tools, and the characteristics of the available subjects and therapists. The translation from basic research questions to researchable questions usually displays two features: a narrowing of scope, and reduction in the level of abstraction (Shlien 1962, 142; Robbins & Wallterstein, 1959, pp. 29-31). That is, a limited aspect of the general basic questions is selected, and operationalized. Ideally, after translation from the broad, global basic research questions to the specific questions selected for a particular research program the researcher should arrive at a formulation of the question that is meaningful, important, and researchable. A commonly accepted classification of the areas which are the "appropriate concern for the research in psychotherapy" are "Form of Therapy and Technique, The Therapist, The Patient, and Role of Theory [Parloff & Rubinstein, 1959, p. 288]."

An experimental design must be chosen or developed, appropriate to the nature, scope, and operational characteristics of the research question. Good design will help ensure that the proposed research leads to logically defensible results. Usually this step relies on the "superscience known as methods and design [Sargent, 1961, p. 96]" which is well-known and extensively described in the general literature (see, for example, Edwards & Cronbach, 1952; Campbell & Stanley, 1966).

Design issues are widely discussed in the psychotherapy research literature as well. Prime references are Sargent (1961), Kiesler (1971), Millon and Diesenhaus (1972), and Parloff and Rubinstein (1959, pp. 281-288).

Another important task is the selection of variables. Their choice reflects the area selected for study and is usually made on the basis of "taste and clinical judgment" of the researcher (Parloff & Rubinstein, 1959, p. 288), availability of suitable and accepted techniques of measurement, past work in the area, and guidance derived from a given theoretical orientation. Some of the decisions which need to be made concern the number of variables, their operational definition, and the use of multiple criteria or measures. As has been mentioned above, much of the core literature listed in Table 1, page 4, focuses on variables (and their measurement, to be discussed below); they are extensively discussed in Bergin and Garfield (1971, Ch. III & Chs VII-XI), Gottschalk and Auerbach (1966, Parts II & IV), and Rubinstein and Parloff (1962, Part III).

The identification of appropriate tools, development of new measurement procedures required by a particular research question, and the achieving of objectivity, acceptable reliability and validity are aspects of measurement procedures that psychotherapy has in common with general psychological research. All the usual modalities of measurement--the use of raters and judges, psychological tests, clinicians' notes, self-reports and interviews, and behavioristic measures including biophysical methods--have been used at one time or another (Millon &

Diesenhaus, Ch. 4). Selection is guided by much the same considerations that determine the selection of variables--convenience, availability, researcher preference, and appropriateness.

Yet another set of preresearch decisions deals with selection of appropriate clients and therapists for a research program. It is evident that the specific formulation of the research question, as well as the ready availability of a research population, will greatly influence the researcher's decision in this area. Meltzoff and Kornreich (1970, p. 26) discuss this research step, particularly the question of sampling, while descriptions of actual procedures followed in selecting and sampling of populations are given by Frank (Rubinstein & Parloff, 1959, p. 12), and Robbins and Wallerstein (1959, p. 40).

Although the above steps were listed sequentially, in practice it usually is necessary to make interdependent decisions. As in other areas of experimental psychology, the formulation of research question, the choice of experimental design or model, selection of variables, specification and sampling research populations, choice of treatment, and selection of measurement methodology all have mutual implications and thus form a complex interdependent network of choices. For example, choice of variables reflects available measurement techniques; specification of population and treatment implies a selection of a system of descriptor variables; the nature of the research question has direct bearing on design model and variables choices, and so on. An appreciation for the complexities which the preresearch tasks and decisions can assume in a carefully conceived project may be obtained from a series

of publications which report the large-scale "outcome-process" study carried out for the past 16 years at the Menninger Clinic (Robbins & Wallerstein, 1959; Luborsky, 1962; Sargent, 1961; Wallerstein, 1968; Wallerstein & Robbins, 1956).

Postresearch phase. The postresearch work revolves mainly about analysis, interpretation and future utilization of research data. These steps have received relatively little attention and emphasis in the methodological literature. The analysis of data requires the ordering of data and study of relationships, including mathematical and other manipulations and operations on the data. The interpretation of results consists of three steps: a study of the internal consistency of the experimental data, analysis of the relationship between a study's findings to those of related studies, and a judgment concerning the degree of support obtained for the research hypothesis or hypotheses (Millon & Diesenhaus, 1972). In addition, it is usual at the conclusion of an experimental project to explore the future implications of the results, perhaps indicating the direction of needed studies or making general programmatic suggestions.

Traditional Difficulties and Traditionally Recommended Solutions

Psychotherapy research would be straightforward enough if the steps and tasks listed in the preceding section could be implemented without serious difficulties. Unfortunately, efforts to carry out the traditional steps have encountered major problems ever since the period of intense

research activity began in the early 1950's. Consequently, a significant part of the research literature concerns the identification and analysis of difficulties as well as suggestions for their resolution.

Several extensive recent reviews can be found in the core literature listed in Table 1, page 4. One source of comprehensive information is a series of publications which have been prepared from several projects supported by the National Institute of Mental Health. Important documents deriving from these projects are a task force report on outcome studies by Fiske et al. (1970), and publications (Bergin & Strupp, 1972; Strupp & Bergin, 1969) reporting Bergin and Strupp's extensive investigations into the possibility of large-scale coordinated research. Detailed summary information is also provided in selected review chapters of Bergin and Garfield's Handbook (1971); of particular methodological importance are Kiesler's chapter on experimental designs, Heller's on experimental analogue studies, Bergin's on the evaluation of therapeutic outcomes, Garfield's on client variables, Truax and Mitchell's on therapists' skills in relation to process and outcome, and Marsden's chapter on content analysis studies. Other useful summaries can be found in Meltzoff and Kornreich (1971, Chs. 1 & 2) and Millon and Diesenhaus (1972).

The discussion of the great many traditional difficulties and associated suggested remedies that have been identified and analyzed in the literature can be conveniently organized to follow the previous discussion of the various pre- and postresearch phase steps--formulation of research questions, measurement, variables, and the phase of analysis,

interpretation, and extrapolation. To avoid cluttering the present discussion with much disconnected detail, the narrative will consider problem areas very generally, commenting primarily on their broad-ranging implications. Specific problems and proposed solutions as well as literature references are presented in greater detail in Tables 3-6 which supplement this discussion. Each table concerns a major methodological area (experimental design; measurement; variables; analysis, interpretation, and extrapolation). The left half of each table lists and references difficulties, while the right half of the page lists and references the corresponding solutions. In each table, further subcategories of the problem area are identified. For example, in Table 3, Experimental design, the two subcategories are Selection of design, and Control. Subcategory headings are given in capitals in the left column.

Preresearch Phase

The definition of research problems is considered first. The difficulties associated with translating broad, global research questions into focused, researchable, operationalized statements are twofold. First, some research questions cause problems because they constrain the answers to a form which is not useful for the researcher (or perhaps the theoretician). The second class of problems arises because the nature of the question is such that it makes the research difficult to carry out.

Examples of the first type of problems are those raised by research questions which lead to trivial studies, or which may make it difficult

Table 3

Difficulties and Solutions in Experimental Design Issues

Problem	References	Solution	References
SELECTION OF DESIGN			
How to choose between experimental/correlational (or exp'tl/naturalistic)?	F., p. 728; Parloff & Rubinstein, APA I, p. 281	(1) match design to research question (2) investigator's values & assumptions (3) naturalistic for generating, experimental for testing hypotheses	Kiesler, BG, p. 69 Parloff & Rubinstein, APA I, p. 281 Butler et al., APA II, p. 181
conventional designs are deficient because they ignore data of individual pts.	Chassan & Bellak, GA, p. 478;	Use "intensive" designs	Chassan & Bellak, GA,p. 478 BS, pp. 69-70; MK, p. 26
How should sampling be treated?	MK, p. 26	Standard methods of experimental design	MK, p. 26
CONTROL			
vagueness of description, and heterogeneity of population	Robbins & Wallerstein, APA I, p. 40; Frank, APA I, p. 13; Chassan & Bellak, GA, p. 484; F, p. 732	(1) gather more complete population (2) use more complex designs (3) single-case, own control studies	F, p. 732 APA I, pp. 40-42; GA, p. 484; APA I, p.40
small N, even in large hospitals, implies that controls have questionable validity	Chassan & Bellak, GA, p. 484	Use "intensive" design	same
No-treatment control groups are difficult to realize	BS, p. 61; MK, p. 21; F, p. 729	(1) within-pt. controls (2) study help-receiving events from outside therapy (3) compare only different therapies, or different levels of same therapy (4) use different design	F, p. 729 BS, p. 61 F, p. 729 BS, p. 69, 441; F, p. 729
How to assess influence and effects of nontherapy factors during therapy era?	F, p. 731; MK, p. 24	(1) randomize subjects (2) gather data on extra-therapy sources (3) use sophisticated designs	F, p. 731 F, p. 731 MK, p. 21, 24
How to identify task confounding (Underwood)?	Kiesler, BG, p. 61	manipulate therapist behavior	same

Table 4

Difficulties and Solutions in Measurement Issues

Problem	References	Solution	References
METHOD			
How to measure private data?	Gottschalk et al., GA, p. 95; Bergin, 1970, pp. 61-69; Shlien, 1970, pp. 95-128; Sargent, 1961, p. 101	(1) patient self-report; inference from psychoanal. theory	GA, p. 95
		(2) legitimize use of clinician as recorder	Sargent, 1961, p. 105
		(3) develop a "sophisticated technology for studying private experience"	BS 1970, p. 25
		(4) be satisfied to obtain agreement between like-minded observers	Sargent 1961, p. 103
		(5) ignore inner experience	Scofield, GA, p. 537; Strupp, GA, p. 367
How to avoid affecting data by measurement operations?	F, p. 731; MK, p. 14	Use unobtrusive measures	F, p. 731; Shakow, APAI, p. 268; Snyder, APAI, p. 270; Parloff & Rubinstein, APAP, p. 284
(inference problems--see "variables")			
OPERATIONALIZATION			
How to operationalize meaning?	Scheflen, GA, p. 282; Auerbach & Luborsky, APA III, p. 155	Define increases in meaning as decreases in ambiguity (information theory - like approach)	Scheflen, GA, p. 282
(Criteria problems--see "Interpretation")			
VALIDITY			
Agreement between judges may mean little	Scheflen, GA, p. 283	study material in context, do not excise data	same
context problem: how to furnish sufficient information to judges, raters?	Harway, Dittman, Raush, Bordin, & Rigler, 1955, p. 248	study effects of reduced information	same
generally, how to validate judgments?	Sargent, 1961, p. 109	reexamine and redefine psychometricians' standard criteria	same
RELIABILITY			
poor rater agreement	Strupp et al., GA, p. 362	better specification of categories, rate only observable behavior, better training of judges	same
INFERENCE			
How inferential should variables be?	Auerbach & Luborsky, APA III, p. 155; Robbins & Wallerstein, APA I, p. 30; Sargent, 1961	(1) Since rater agreement and low level of inference are highly correlated, remain close to observables	Strupp et al., GA, p. 371
		(2) use flexible approach, mix variables; rely on clinical judgment for highly inferential variables	Robbins & Wallerstein, APA I, p. 32; Sargent, 1961

Table 5

Difficulties and Solutions in Variables Issues

Problem	References	Solution	References
SELECTION			
What guides their selection?	Auld & Dollard, GA, p. 86; Kiesler, BG, p. 53; Marsden, BG, p. 387; Parloff & Rubinstein, APA I, p. 288; BS, p. 432	(1) theory, intuition	Auld & Dollard, GA, p. 86 Bordin, APA I, p. 236
		(2) use culturally defined communication units	Scheflen, GA, p. 277
		(3) Stay close to observables	Strupp, Chessan & Ewing, GA, p. 372
		(4) naturalistic observation; partly idiographic approaches	Knapp, Mushatt, & Nemetz, GA, p. 421
		(5) use methodical, integrated approach: relate clinical practice, field trials, theory, planned case experiments, approaches and data from related sciences ("cyclic model")	BS, p. 433
		(6) choose those with interactional implications	Bordin, APA I, p. 237
		(7) Choose those that allow critical tests of theories	Bordin, APA I, p.237
		(8) use flexible, "zoom model" approach	Sargent, 1961, p. 101
How many variables does one need?	Harway et al., 1955, p. 247; Chance, GA, p. 127; Ekman & Friesen, APA III, p. 179; Parloff & Rubinstein, APA I, p. 282	(1) use all modalities	Lorr, APA I, p. 262; Shakow, APA I, p. 111; McV. Hunt, 1952
		(2) simplify (gives 3 general rules)	Bordin, 1965; BS, p. 431
		(3) Study effects of partial information on ratings (ex., transcripts vs. audiotapes)	Harway et al., 1955, p. 24 p. 248
		(4) study part-processes	Bergin, BG, p. 254
		(5) include internal (within-patient) dialogue	Beck, GA, p. 557
		(6) study "mental operations" of therapists	BS, p. 71
UNIT SIZE			
How to choose?	Harway et al., 1955, p. 248; Marsden, BG, p. 387; Ekman & Friesen, APA III, p. 192; Auerback & Luborsky, APA III, p. 155; Robbins, APA I, p. 120; Sargent, 1961, p. 109	(1) study & compare effects of extreme range of sizes	Harway et al., p. 248
		(2) use "natural" units	Ekman & Friesen, APA III, p. 193; Leary, APA I, 76.
		(3) distinguish between scoring, contextual, and summarizing units	Marsden, 1965, p. 300

Table 6

Difficulties and Solutions in Data Analysis Issues

Problem	References	Solution	References
ANALYSIS OF DATA			
Organization:			
How should clinical research data be organized?	Chance, GA, p. 129	(1) via personality theory (2) sequential analysis (3) discover the way in which units at one level combine to form units at next level (hierarchical "Context analysis")	Chance, p. 129 Snyder, APA I, p. 262 Scheflen, GA, p.279
tools:			
is frequency an important measure?	Sargent, 1961, p. 101	single events may be of crucial importance; attend to perspective.	same
should one quantify at all?	Sargent, p. 106	reconcile idiographic needs of clinicians with nomothetic goals of experimentalist; Menninger uses combined verbal/quant. approach; look for logical relationships, patterns.	same
INTERPRETATION OF DATA			
conflicts:			
conflicting interpretations from different viewpoints	Seitz, GA, p. 209; F, p. 731; Lorr, APA II, p. 175; HO, pp. 641, 657; Betz, APA II, p. 53; Kiesler, BG, p. 54; Butler, APA II, p. 181; Bordin, 1965, p. 502; Leary, APA I, p. 93; Parloff & Rubinstein, APA I, p. 282; Block, APA II, p. 167	(1) remain close to observable, noninferential data (2) construct multidimensional scales (3) take into account "modifying transformations" made by observers with different viewpoints (4) define objectivity as agreement between observers (5) use multiple viewpoints flexibly and appropriately	Strupp et al., GA, p. 371 Strupp et al., GA, p. 371 Betz, APA II, p. 53 Butler et al., APA II, p. 181 Sargent, 1961, p. 101
ambiguity of data	Isaac, APA I, p. 225	examine context of data to guide interpretation	same

Table 6 (continued)

Problem	References	Solution	References
spurious interpretations of data are always possible, but what do they mean (Bavelas)?	Parloff & Rubinstein, APA I, p. 286	no solution offered	
different criteria (particularly in outcome studies) lead to different conclusions from same data	Isaacs & Haggard, GA, p. 228; Cartwright, GA, p. 518; Bergin, BG, p. 257; Bordin, APA I, p. 242; Shlien, APA II, p. 143	(1) individualized goals and measures (2) Better definition of research questions will lead to less ambiguous criteria.	F, pp. 729, 735; Shlien, APA II, p. 143
lack of agreement on criteria	BS, pp. 7, 54; Parloff & Rubinstein, APA I, p. 278	(1) use multiple measures (2) individualized criteria for each patient (3) problem is unsolved	BS, p. 54 BS, p. 54 Bergin, BG, p. 257
noncomparability: can't relate results of various researchers--no standardized systems	Strupp et al., GA, p. 362; Snyder, APA I, p. 262	(1) coordinate research (2) standardize measurements (3) use multiple measures (4) obtain more complete description of treatment, setting, patient, measurement operations	Hunt, 1952, p. 245; BS, p. 1 BS, p. 102; Snyder, p. 262; F, pp. 728, 736 F, pp. 728, 736 F, pp. 728, 736
EXTRAPOLATION Generally, results are not useful for clinical practice	BS, p. 6; F, p. 728	(1) use better research techniques, design & control (2) perform more naturalistic studies, including effectiveness of "good" therapists (3) ask "right questions"	BS, p. 7 BS, p. 69 BS, p. 437
triviality: questions are trivial, chosen for convenience rather than importance--avoid complex or meaningful problems	Rubinstein & Parloff, APA I, p. 280; BS, p. 437; Sargent, 1961, p. 95; Shakow, APA I, p. 264; MD, p. 11; Harway et al. 1955, p. 247	(1) use problem-method-centered approach (2) use "cyclic," methodical approach (see variables selection, solution 5, Table 5, page 24) (3) acquire more complete data	Sargent, p. 99 BS, p. 433 Shakow, APA I, p. 111

Table 6 (continued)

Problem	References	Solution	References
generalizability: can't extrapolate data because populations are poorly defined	Chassan & Bellak, GA, p. 478	use intensive design	same
research is fragmented, different theories lead to noncomparable questions, questions are too narrow	BS, p. 7; Sargent, 1961, p. 94; Hunt, 1952, p. 242; MAL, p. 720	seek large-scale collaborative studies	Hunt, p. 245; BS, p. 1
experimental analogue studies can't be related to therapeutic practice (in part, the problem of applying statistical results to individual cases)	BS, p. 38	Need new concepts, better experimental taxonomy, more use of "intensive design" (Chassan) studies (Colby)	same
studies in "experimentally uncontrolled therapy" are less useful than controlled experiments (contradicts preceding point?)	BS, p. 38	use controlled experimental studies based on learning theories	same
ambiguity: outcome/process distinction is confused, unproductive, carelessly used	Sargent, 1961; Kiesler, BG, pp. 71, 45, 46; HO, p. 645	(1) careful semantic analysis of the distinction, differences (2) use "in-therapy" and "extra-therapy" studies dichotomy instead	Sargent, p.98 Kiesler, BG, p. 46
outcome studies are ambiguous, inconclusive	Parloff & Rubinstein, APA I, p. 277, 298; HO, p. 645; F, p. 735; MK, p. 13; BS, p. 8; Sargent, 1961, p. 99	(1) live with ambiguity of multiple measures (2) do not tackle broad questions, restrict research goals to narrower, specific ones (3) use better, more diversified research designs, criterion measures	HO, pp. 646, 650 G, p. 735 BS, p. 8

to compare research results with those found in other published studies, or which are asked in ways leading to statistical results difficult to apply to individual patients. These problems of utilization of results arise principally in the postresearch phase, and therefore will be discussed in that context below.

The second class of problems associated with formulation of the research question are those which impede carrying out the research. They derive from those aspects of the research question which raise difficulties in design, control, measurement, or in definition and choice of suitable variables. For example, questions which attempt to deal broadly with outcome issues are difficult to operationalize because a variety of releavant but contradictory criterion variables exist. Or, as another example, questions which use highly inferential concepts in their formulation cause problems in measurement, variable selection, etc. Because problems of this type relate principally to the various preresearch areas, they will be discussed under the applicable categories of design, measurement, and variables.

Experimental design problems will be considered next. The major problems associated with experimental design are the selection or development of acceptable designs as well as issues of experimental control. Researchers traditionally have had a choice between two major approaches to design--experimental and naturalistic--with many subcategories to choose from in each approach (Millon & Diesenhaus, 1972). The various pros and cons of each type have caused much debate and no clear-cut solutions have emerged. Control problems arise from various

difficulties associated with the nature of patient populations, lack of control over events outside the therapeutic hour and difficulties encountered in attempting to identify and adequately describe the important therapy variables.

The conventional solutions offered for these problems point to the accepted procedures of good research practice in general as described, for example, by Campbell and Stanley (1966) or Edwards and Cronbach (1952), writers whose methodological precepts appear to be widely respected by the research community. Particularly, Campbell and Stanley's work serves as a standard: so, for example, it forms the basis of an extended normative discussion in Fiske et al. (1970); it also provides the criteria on which a critical analysis of psychotherapy research studies is based (Kelly, Smits, Leventhal, & Rhodes, 1970). The strong implication drawn from the literature is that adherence to good design principles will solve most of the design questions. Here and there we find other approaches suggested: Bergin and Strupp (1972) recommend more experimental analogue studies, studies of "experiments of nature," naturalistic observations of outstanding therapists, and, in agreement with Chassan and Bellak (1966, p. 478), suggest a greater emphasis on the study of individual cases (Chassan's "intensive" method). These suggestions, however, apparently have failed to influence research practice significantly.

The measurement problems discussed in the literature and in Table 4, page 23, fall into four classes: techniques, operationalization, and the two conventional and well-known problem areas of the psychometrician--validity and reliability. From Table 4, page 23, it appears that the

long-standing, traditional measurement problems are still with us. A typical example is the problem of acquiring data concerning private experience; it has a long history in psychotherapy research (as well as in psychology in general), yet today we still have only Bergin and Strupp's (1972) plea to develop a sophisticated technology for studying private experience.

The major questions regarding variables concern selection, decisions, questions of appropriate choice of unit, and the inferential status of variables. To label these problems, "variables" problems is to some extent arbitrary. Some of the questions at least could equally well have been considered measurement problems, which illustrates the previously mentioned interdependence of research factors, a topic that will be mentioned again later. Details of major issues concerning variables are given in Table 5, page 24.

The kinds and number of variables chosen still constitute one of the least formalized research problem areas. The impression obtained from the literature is that this step is as much black art as science. Table 5, page 24, lists 14 general solutions that have been advocated over the past 20 years to this question. The status of the solutions to the other two variables problem areas, selection of appropriate unit size for a given variable and inference questions, are equally unsatisfactory unit size decisions arise especially in content analysis studies, while inferential problems increase as one moves from the behavioristic end of the theoretical spectrum to the more dynamically oriented studies. Thus, as was the case in the "measurements" problem area, these issues are long-standing and unsolved.

Postresearch Phase

The postresearch phase deals with providing a rationale for the research yield. There are two different perspectives from which the results can be viewed. One may be called the "internal" perspective--it views a project without reference to outside factors. That is, one considers whether the data can be organized to make sense, whether research hypotheses are supported by results, etc. The second, "external" perspective considers the results in relationship to factors stretching beyond the research project. It leads one to questions concerning the relationships between one's own and other researchers' results, questions of application of results to practice, and the like. The postresearch phase problems can be discussed conveniently in terms of analysis, interpretation, and extrapolation of results.

The analysis of data is not usually emphasized as a problem area in the literature probably because there are common-practice, routine, procedures whose application is generally accepted. The usual analysis operations on the data are simple counting and summing as well as the well-known statistical manipulations of correlation, factor analysis, and statistical tests. On rare occasions, more elaborate procedures such as matrix methods, stochastic techniques, or time series analysis are used. The meaningfulness of standard analytic procedures may not be convincingly demonstrated but since it is seldom identified as a major problem in the literature and since mathematical issues will be the subject of Chapter III, this problem area will not be pursued further here. The first portion of Table 6, page 25, describes issues in the area of analysis.

Two kinds of difficulties may be encountered during the interpretation of research data. One set of problems reflects conflicting implications of the data; Table 6, page 25, describes "viewpoints" and "criteria" problems; these arise because in various ways, data can lead to multiple, incompatible conclusions. Results that allow alternative or ambiguous interpretations clearly create a problem. An extreme example is given by Block (1962, p. 167), who reports that the results of the same experiment were interpreted in diametrically opposite ways by two different researchers. No truly satisfactory solution has been offered, as can be seen from Table 6, page 25.

The second set of postresearch problems reflects the isolation of various researchers from one another. Many researchers proceed in idiosyncratic ways with no apparent contact with other work in their particular area. Strupp has identified this as a problem area for some time (Malan, 1973, p. 720). As evidence for this fragmentation the following are cited: the proliferation of rating scales and other measures, each developed independently for one-time use by diverse researchers; the difficulties of translating research results obtained by one study into a form which would allow comparison with other results; apparent lack of awareness of previous related work; and repetition of research practices which have been demonstrated to be less than satisfactory. Traditional recommendations include greater coordination of separate research programs, use of standard techniques, and reporting of more comprehensive data to facilitate comparison.

The last problem areas to be considered in this section are those concerning extrapolation and application of research results, particularly in regard to clinical practice in general.

A major criticism is that research questions and conclusions are trivial--safe, researchable, but meaningless. For a variety of reasons--convenience, tractability, or naiveté--the nature of many research questions is such that their answers are unfortunately of little importance to other potential users, especially clinicians. It seems that the level of difficulty in doing research in a given area varies directly with the importance of the question asked--this may not be an unavoidable characteristic of research in psychotherapy, but it does appear to be true of such research to date. The problem of triviality is to a great extent a reflection of many of the difficulties already discussed, those associated with choice of variables, description of patient populations, use of group data, measurement problems, and inadequacies of theoretical bases. To cope with these difficulties, it has been suggested that researchers use the recommended alternative research approaches already mentioned in conjunction with problems of experimental design; traditionally, it is hoped that greater emphasis on experimental analogue studies, naturalistic observation of excellent therapists, studies of "experiments of nature," and the use of intensive designs would lead to more meaningful results.

A second group of difficulties concerns the problems of generalization. If a research question is very narrow results may not be legitimately extrapolatable to other situations, for example, to

other populations of patients or other therapists. The issue is a difficult one because while questions with highly restricted scopes are not often useful, it has been found that researching extremely broad questions is also not desirable (Fiske et al., 1970). The solution offered in some publications is coordinated research. According to this approach, various researchers would be encouraged to choose problems which are related to one another and whose solution would allow a gradual accumulation of a broadly based, useful, and generalizing body of knowledge. This was Strupp and Bergin's initial approach, although after exhaustive investigation they concluded that undertaking a large-scale program of coordinated research was undesirable at that time (1972). A major solution proposed in that document is a broad-scale research model which will join researchers to one another and to efforts in related scientific disciplines. The proposed "cyclic model" would have questions selected not in isolation, but as part of a methodical, planned research cycle relating clinical practice, planned case studies, and new experimental techniques. In addition, the whole research system would also exchange data, techniques, and related information with other branches of psychology (experimental, social), as well as with sociology, anthropology, and ultimately the biological medical sciences. Strupp and Bergin (1972, p. 433) suggest that were researchers to follow this cyclic model, research questions would be cumulatively produced which would be more relevant to the clinician, more tractable experimentally, and sufficiently broadly based to allow connection with other research, thus producing a growing, cumulative fund of practically useful experimental results.

A third set of extrapolation problems derive from logical flaws in research questions leading to ambiguous, contradictory, or vague (and thus not useful) conclusions. To some extent, this problem has already been considered in the context of data leading to, or allowing, conflicting interpretations. Viewed here in terms of difficulty of application of results, it seems that the studies that have suffered most because of logically ill-conceived questions are those attempting to evaluate therapeutic outcomes. An important example is given by research questions which attempt to assess "improvement" of patients in a simple fashion (Howard & Orlinsky, 1972). Because of these logically flawed questions, outcome studies have been the subject of endless controversy. The history of the problem is well summarized in Malan's recent (1973) review. Apparently, there are still no truly satisfactory, generally accepted methodological solutions.

What can be said to summarize the traditional problems and their recommended solutions? One can infer from the data of Tables 3-6, pages 22-25 that, by and large, workers in the discipline still continue to assume that traditional conceptions previously described as underlying research are basically valid. The main approaches to, and the major ingredients of, traditional research continue to be experimental design, dependent and independent variables, measurement, and analysis. Their continued prominence in the literature implies that solutions to persistent research questions continue to be sought along traditional lines: The implicit approach seems to be to continue doing the same things, but to do them better--to ask better questions, develop improved designs,

make better selection of variables, seek greater refinement in measurement techniques, and so on. The message is, improve, but don't change the basic paradigm. It seems that we continue to heed the advice given by a researcher quoted by Sargent (1961), to whit: don't change scientific ground rules to conform to problems.

A Different Critique

Neglected Methodological Areas

From the preceding discussion, it appears that in spite of a considerable amount of research activity, there are still major problems and difficulties associated with research in psychotherapy. Symptomatic is the persistence of certain problem questions--for example, those dealing with the privacy of data, the outcome/process confusion, general questions of design, definition of variables, measurement methodology, and, in general, the persistence of broad, general difficulties, particularly in the failure of research to be meaningful for the practitioner (Bergin & Strupp, 1972; Fiske et al., 1970, 727).

Considering the great amount of time and effort that has gone into working on these problems, why haven't the proposed solutions been more productive? One possible answer is that researchers simply have failed to follow the recommended solutions. Though this may be true in part, it is difficult to believe that had truly adequate solutions been available they would not have been implemented by now with noticeably beneficial impact. Another possibility is that the technical and methodological problems are just too difficult in this area. Yet, major

identified technical problems have apparently been attacked successfully: We now have logically adequate designs, reliable scales and other measures, set techniques, and, for analysis, sophisticated statistical tools. Still another frequently cited source of difficulties is the inadequacy of existing theories of psychotherapy. The lack of progress in research is thought to reflect the lack of basic understanding of psychopathology, characteristics of treatment, and of its effectiveness (Fiske et al., 1970).

Whatever the reasons a review of traditional problems and traditionally offered solutions must conclude that the basic view of how research in psychotherapy should be conducted has not changed significantly over the past 20 or so years. The traditional research model still persists.

Since this approach is and for some time has been in so much difficulty, perhaps it is not unreasonable to consider alternatives. The next sections will point to some clues concerning the nature of an alternative approach. Such clues seem to emerge from two quite different perspectives: First, from considering how research observations are represented or conceptualized, and second, from considering how material is organized in the major review literature. Implications of these considerations for a new approach are presented in the last section of this chapter.

Three themes appearing in the literature have been selected for consideration here, although the reasons for selecting these topics may not become fully apparent until later chapters. It is interesting

to note that these are distinctly minor themes; if they are mentioned at all, they are usually only mentioned briefly in passing, as unfortunate curiosities. They are almost never explored in depth. The three topics are the viewpoints issue, redundancy, and additivity.

The viewpoints issue raises the question, why do different observers looking at the same event (phenomena, data) see different things? Two kinds of explanations are suggested in the literature. The first is that the divergent views come from partial observation, that they are due to observers selecting "different parts of a total picture [Seitz, 1966]"--the well-knwon "fable of the elephant" explanation. The second suggestion is that alternative explanations are equally satisfactory, and can be accounted for, at least in part, by the principle of overdetermination--that events are multiply determined (Seitz, 1966). In conjunction with this latter explanation it is sometimes suggested that the data and research results obtained by using different systems ought to, and probably could be, translated from one system of observation to another (Bordin, 1959) and that different schools and different viewpoints could be compared by rearranging data (Chance, 1966).

The second class of hints concerns the duplication of information. For example, Strupp et al. (1966) suggest that variables selected in an experiment may turn out to be related to one another. In a way, this reflects the well-known factor analytic philosophy of the search for a parsimonious representation. In another way it justifies implicitly or explicitly the use of partial data for analysis--for example, the common assumption that verbal material alone is sufficient, that, say, kinesic

or paralinquistic variables would provide only redundant or duplicate information.

The last set of hints concerns the problem of piecemeal research. Although the majority of conventional approaches recommend researching partial situations, occasionally the opposite view is advocated. Scheflen's view is representative of this less orthodox approach: for the clinician, the "world of bits and pieces doesn't represent the richness of his experience [1966, p. 264]." Similarly, Bakan is quoted as considering reductionism as "a disease of our time [Bergin & Strupp, 1972, p. 381]"; or, "small events are not merely strung out like beads on a string, all on a par, to constitute a larger event [Pittenger, et al., 1960, p. 226]."

What formal, methodological topics do these three areas concern, and, taken together, what are their implications? The viewpoints issue can be rephrased. In asking about contradictory or redundant observations, it raises questions that can be characterized formally as questions of completeness and independence. Using a direct analogy to systems descriptions in physical science (see, for example, Goldstein, 1950), we infer that the issues can be restated as follows: What is a minimum, nonredundant framework (such as a coordinate system) which will completely specify the phenomena under investigation? If one uses too many descriptive dimensions, one will have duplication or dependence among the descriptors. Too few descriptors imply that the phenomena cannot be completely described (by definition, that is what "too few" means), and that therefore different incomplete systems of description

may describe different parts or aspects of a phenomenon, leading to different and perhaps incompatible conclusions about what "really" is going on. Similarly, the question of duplication of information leads one to consider the problem of necessary and sufficient data; again, it is the issue of completeness and independence.

Finally, problems of additivity point to questions about the characteristics of the class of phenomena that one is dealing with. It is widely recognized that some systems can be adequately represented as linear--the precise definition is not needed here, but roughly, a linear system is one whose analysis can be approached piecemeal: Its behavior is studied under a variety of simple conditions, and its behavior under a complex composite of these conditions can then be predicted or understood in terms of simple responses. Linear systems are the bread and butter systems of physical science, and are very well understood. All other nonlinear systems are only very incompletely understood--there is no systematic overview of the behavior of nonlinear systems, and in science their study is avoided if possible; they are approximated by linear systems whenever that is feasible.

Together, these thoughts focus on the very basic question of representation--what are the general system properties and how are they appropriately specified? How are the phenomena conceptualized appropriately and represented (described, specified) adequately? Considered in terms of their formal implications, the viewpoints, redundancy, and additivity questions converge and point to a single, fundamental area.

In addition to these three topics, there is still the question of the organizational structure of reviews. Trying to organize the subject

matter of psychotherapy research can be a very frustrating experience. There is a fundamentally elusive quality to the material as though one is caught in an endless loop of disorganization. From Table 2, page 7, already briefly considered, it would appear that there are at least four major and separate categories of subject matter: design and experimental method, variables, therapeutic technique, and outcome/process topics. Further examination of the internal content of chapters, however, reveals that chapters or sections headed by a label from one of the major topical categories almost invariably concern themselves equally with problems in the other three categories. As an example, one may consider chapters or sections headed by labels from the general area of "experimental design." Kiesler's chapter on design (1971) contains an extensive analyses of variables--patient, therapist, and inter-actional--as well as major sections concerning therapeutic technique; a section concerning data analysis (Gottschalk & Auerbach, 1966) while ostensibly a section on experimental method and design contains lengthy discussions of variables problems, personality theory, and psychotherapy techniques. Thus, on closer examination, sections presumably dealing with design areas are seen to be equally concerned with the other classes of problems listed in Table 2, page 7. Furthermore, the design issue is typical. For instance, one could equally well have chosen sections presumably dealing with variables problems; one would then find that they contain major discussions concerning design, therapeutic technique, and outcome/process issues as well.

In short, there is no real organizational structure under the apparent structure furnished and implied by chapter headings. The material is arranged in a circular, amorphous fashion, strongly reminiscent of Kuhn's description of the natural histories of the 17th century. He describes these Baconian "histories" as composed of an indiscriminate collection of "facts that will later prove revealing (e.g., heating by mixture) with others (e.g., the warmth of dungheaps) that will for some time remain too complex to be integrated with theory at all [1970, p. 16]." These writings, Kuhn states, are difficult to call scientific. Rather, they produce a morass. The parallels to psychotherapy research seem inescapable. Its traditional conceptualizations apparently do not permit creation of a logical system for organizing the area.

An Alternative Approach

The preceding section may appear to concern two very different kinds of issues: the first dealing with problems of formal representation, and the second with the confused organization of the research subject matter organization. Viewed together, these considerations suggest that current, traditional conceptualizations may be unable to provide a coherent, logical structure capable of guiding and systematizing research. Apparently, the subject matter can neither be rationally, logically organized nor adequately represented.

We have seen that in psychotherapy the traditional approaches to research follow the analytic approach of general experimental psychology. This approach structures research practice in far-reaching ways by its

methodological and representational characteristics and assumptions. The preceding discussion has suggested that in this way the associationistic model may contribute to the unsatisfactory status of the research field. Perhaps a different model, another approach to representation of what takes place in dyadic psychotherapy, might help resolve traditional difficulties, and in so doing, reconceptualize the events in therapy and thereby move research into more productive areas. As a result of these considerations, it seems that a return to the foundation on which research rests is advisable and necessary. Observation, implying representing and organizing events in a particular way, seems to be just such a foundation. Accordingly, the focus of the present study will be on the representation of events in psychotherapy. The processes of observation and representation will be systematically and formally studied and analyzed. This approach is intended to provide a unifying logical framework.

CHAPTER II

OBSERVATION AS A FORMAL PROCESS

Ingredients of Observation

There are highly complex, controversial philosophical issues underlying the processes by which data are initially lifted from the temporal stream. These processes reflect an observer's assumptions, values, choices, biases and perspectives which in turn significantly define and circumscribe the basic nature of the acquired data. The concern of the present work is with a limited aspect of this process: the logical, formal features of observation. A consideration of relevant philosophical issues is beyond the scope of this work; applicable analyses may be found in Scheffler (1967), Turner (1967), Weimer (1973), and Yankelovich and Barrett (1971).

Viewed as a logical process, observation begins with operations which impose a structure (or recognize one, depending on one's philosophical persuasion) on the temporal flow of events. "Raw data" are captured from the unstructured, primordial material through a system of classification, that is, a set of categories; the set may be formally and explicitly defined, or implicit, perhaps even unrecognized. To analyze the formal aspects of the observational process means to examine the structure of the category system used to classify data and to analyze subsequent operations performed on these data. In the present work the analyses will be performed through application of formal models. First,

however, an informal review of several typical categorization systems will be presented; several familiar data acquisition methods will be considered in terms of the classification systems initially used to code data.

Audio tape recordings are frequently used in psychotherapy research as the first step in data acquisition. Selection of this data acquisition mode implies a number of research assumptions and decisions. The present analysis, however, will focus on the classification scheme inherent in the process. An audio tape recording viewed as a formal procedure for encoding initial data classifies these data by means of an alphabet--a set of signal amplitudes. This "alphabet" contains a limited although very large set of members ("letters"): information theory has demonstrated that noise and bandwidth limitations in any system restrict the number of available signals. Thus, after being recorded the original signal is represented in terms of a limited alphabet. The recorded data may be conceptualized as a sequence of data elements drawn from the alphabet or classification set.

This analysis illustrates certain general features: the rich interview phenomena are coded into a temporal sequence by means of a limited set of descriptors. Other systems typically used to acquire initial data during an interview can be similarly conceptualized. For example, videotaping is logically equivalent to audiotaping; although the videotape "alphabet" is usually larger, because of technological characteristics of the system, the same logical considerations apply. Even a highly restricted system such as Jaffe's (1968) displays analogous

logical features. That system, described further below, uses an alphabet set having only four members; the entire interview is reduced to and represented by a data sequence composed from these four "letters." Even most methods of acquiring kinesic data can be analyzed using the above concepts; the analytical approach is very general.

The preceding examples illustrated methods used to acquire data on-line, as an interview takes place. There are other important classes of observational data such as clinicians' notes, psychological tests, case studies, etc. In these observations, initial data are generated outside therapy. It is not obvious that these processes can be conceptualized in terms of the formalism just presented. It will be shown that these methods of observation can be analyzed within the same logical framework; the discussion, however, will be postponed until an analytical tool, a formal model, has been developed.

If, as it seems, the analytical procedure outlined above can be used to conceptualize and represent within-interview (process) observations as well as observations based on clinicians' notes and other extra-therapy events, it should be possible to use that approach to analyze the logic of data acquisition used in most, if not all, classes of psychotherapy research.

The initial acquisition of data is usually followed by other operational steps which further reduce the material. Typically, unless the intial acquisition is already in the form of written text, the next, second order reduction produces a written transcript from the first reduction (say, an audiotape). It may seem artificial to separate such

common procedures into several logical steps. If such procedures are taken for granted, however, one tends to lose sight of some of their features; for example, some observational approaches such as Pittenger's (1960) transcribe additional, paralinguistic data from audio tapes, reminding one that the usual transcription from tapes brings with it a second, and often nontrivial, loss of information.

In any case, further operations on data commonly include still other, higher order reductions. Generally, these (1) further segment data into other units and (2) code units in terms of another classification set, typically one which is more constrained and formal than the preceding category system. This process may be repeated at several successive levels.

Marsden's method of classifying units (1965) will facilitate the analysis which follows. He categorizes units into scoring, contextual, and summarizing units. The scoring unit is the item, such as a word, sentence, or gesture, that is assigned to a category. The contextual unit is the surrounding material, for example a paragraph, which provides the context for defining or identifying the scoring unit. The summarizing unit defines the interval over which cumulative measures are obtained--for instance, the temporal unit used to sum or count the frequency of scoring units.

To summarize: to this point the discussion has sketched the logical steps which reduce and segment phenomena in terms of a category system. That system may be a content analysis scheme, a set of clinically derived categories, communication categories, or any other formal or informal alphabet. It should be remembered that the logical aspects of observation

are being discussed. These may not be apparent from a casual observation of research practices. For example, a content category system used in a project may not have been defined formally in advance by a researcher; he may have developed his system during the course of his experimental studies. Nevertheless, from the point of view of a logical analysis his observational methods can subsequently be analyzed in terms of the framework just presented.

Generally, one of the final logical operations on data is mathematization. This process can be conveniently analyzed in terms of the above concepts as well. A discussion of mathematization, however, will be postponed until the earlier sterps of observation have been analyzed more fully. Mathematical issues will be addressed in Chapter III.

A Proposed Model

If one wishes to examine the logical aspects of an observational process, why should a model be developed for analysis? First, the availability of a symbolic representation usually is helpful in manipulating and analyzing concepts. It has been frequently pointed out that the solution of problems in algebra or arithmetic, for example, is very clumsy, if not impossible, if attempted by using words only, and that their solution is greatly facilitated by using an appropriate symbolism (Reichenbach, quoted in Leary & Gill, 1959, p. 60). Indeed, even the choice of a particular symbolism can make a great deal of difference, as anyone who has attempted long division using roman numerals will have discovered.

Second, a formal model can help reveal the logical features of observation. For example, a formal representation may identify previously hidden issues, perhaps issues that might become important when one attempts to relate different observational systems to one another; a model may serve to clarify observational muddles such as those stemming from confused usage of the process/outcome dichotomy; or, it may explicate the process of mathematizing data.

Third, a formal model of observation may be useful for comparing alternative systems of observation. It may help to assess the relative loss of information inherent in the use of any one particular observational system.

Fourth, a model should provide a general, unifying analytic framework for conceptualizing diverse research approaches. Observation produces the initial data on which the subsequent research operations are grounded; it is fundamental. A model which would clarify that process should therefore illuminate traditionally troublesome research areas such as design, measurement, or variables issues.

Having decided to model observation, the next question is, how does one approach the task? Unfortunately, the psychotherapy research literature fails to provide a precedent, except perhaps for Kiesler's (1971) "grid" model. This is not really a model of observation, but rather one which represents relationships between patient and therapist subgroups and the progress of treatment.

Since the model is to represent a process which unfolds in time, it seems appropriate to consider state models. These are very general,

flexible, theoretically neutral models which are widely used in the physical sciences, engineering, economics, and mathematical psychology (Coombs, Dawes, & Tversky, 1970) to represent the behavior of time-varying systems. Basically, such models contain a suitable descriptive scheme or method which can appropriately and adequately characterize the state--that is, the condition, circumstance, situation--of a system (a concept used in a very general sense) at some instant of time, or over some temporal interval. Successive changes in the state representation are intended to mirror significant aspects of temporal changes in the system. The representations used in science usually are scalars, vectors, tensors, or other familiar mathematical objects. It will be argued later that these choices, though natural from a mathematical point of view, may be inappropriate and unduly restrictive in psychology. These restrictions, however, are not inevitable byproducts of using state representations; the state of a system could, perhaps, be represented by more flexible "objects." This issue will be reconsidered briefly toward the end of the present chapter, and explored at length in Chapter IV.

Analysis of Some Major Systems of Observation

So far the formal model of observation has been described only in general terms; a more specific description will be introduced by analyzing several systems of observation used in clinical research. The general model, though straightforward in principle, is somewhat complex; it therefore seems best to begin with simpler specific examples, and

allow the model's features to emerge as generalizations from particular cases.

The observational systems which will be modeled in this section have been selected because of their prominence and recognized importance in the psychotherapy research literature. They are the products of very well-known clinical and research endeavors and have been developed with care and thought, usually over a considerable period of time. Of the 14 systems considered, seven are the subject of a separate book (Butler, Dollard and Auld, Lennard and Bernstein, Leary, Matarazzo and Weins, Scheflen, and Pittenger et al.), while the other seven have been described extensively in journal publications. These systems seem to fall into five major categories, with one nonconformist system (Pittenger's) escaping easy classification. This latter system, while perhaps not qualifying truly as an observational system in the usual sense, has so much relevance to the present study that it could not be omitted. The systems, grouped by categories (to be discussed below) are listed in Table 7.

The first class of systems is characterized by an assignment of verbal material (usually reduced to a standard transcript) to a set of psychologically descriptive, meaningful categories. The six systems in this group are Strupp, Butler, Dollard and Auld, Lennard and Bernstein, Gottschalk, and Leary and Gill. These systems basically take relatively small verbal units--sentences, propositions, clauses--and assign them to a fixed set of (usually mutually exclusive) content categories.

The second class may be called acoustic/mechanical systems. They focus mainly on the acoustic energy aspects of verbal communications--who

Table 7

Major Systems of Observation

I. CONTENT ANALYSIS (LINGUISTIC)

Strupp (for therapist)
Butler et al.
Dollard & Auld
Lennard & Bernstein
Gottschalk et al.
Leary & Gill

II. INSTRUMENTED SYSTEMS

Jaffe
Matarazzo

III. KINESICS

Mahl

IV. MIXED (KINESIC, VERBAL, PSYCHOLOGICAL)

Scheflen
Deutsch

V. "GLOBAL" EVALUATION (ASSESSES OVERALL PATIENT STATUS)

Rogers
Menninger

VI. MISCELLANEOUS

Pittenger et al. (paralinguistic)

speaks, when, how often, how long, singly or simultaneously, what a speaker's pattern of silence is, etc. The two systems chosen to exemplify this approach are Jaffe and Matarazzo.

Systems in the third group are those which consider primarily kinesic variables, identifying units of body gesture, motion, or position, and assigning them to a category system of descriptors. The example used is Mahl's system.

The fourth group is a mixed group whose common feature is the simultaneous observation of a variety of modalities (verbal, kinesic) of behavior. Illustrations are provided by Scheflen's and Deutsch's systems. Scheflen's system, although it has not been applied to the observation of dyadic psychotherapy, is relevant to the analysis of observational processes, and in addition has achieved major prominence. It was included for these reasons.

The fifth type of system describes global, overall characteristics of clients. In these systems, a client is characterized in terms of a stable state description which varies little at any one stage of a client's treatment career. Two dissimilar approaches of this type are illustrated by Rogers' and the Menninger Foundation's systems.

The final example of an observational system, already referred to earlier, is Pittenger et al. This interesting analysis, based on linguistic and paralinguistic transcriptions, uses a combined psychodynamic and linguistic analysis approach to make sense of the opening five minutes of a first interview.

Each of these classes of systems are examined next using a state representation analytic model; the taxonomy of observational systems will be reconsidered following the analysis.

Many observational procedures or plans can be classified as content analysis systems (see, for example, Marsden, 1971; Meltzoff & Kornreich, 1970). Accordingly the first models of observation represent this class of observations.

A typical and logically straightforward example is Strupp's system for analyzing therapist responses. The logical stages are shown in Figure 1. In this system, the researcher chooses to focus on the spoken word--the auditory mode--and begins by reducing the complex events of an interview--sounds, sights, smells, etc.--to an audiotape (data sequence A, Level 1). As has been stated earlier, an audiotape is a data sequence using a notational system or alphabet of very high dimensionality; that is, it has many degrees of freedom, limited only by bandwidth and noise of the recording system. From this first, relatively unconstrained record a second, reduced subset of data is derived (data sequence B, Level 2). In the present example, the second operation changes the audiotape to a transcript. The second reduction uses a more limited alphabet. The next, third logical operation segments the typescript into units--therapist and patient statements (data sequence X, Level 3). Each dyad member's statement (scoring unit) is identified by its surrounding material (contextual units), the other dyad member's statements. (The report fails to account for simultaneous speech.) It may be noted that these last two logical steps--reduction

General Comments

A system for classifying therapist responses; said to be theoretically neutral.

Reductions and Operations

Standard, by raters from typescripts.

Units

scoring: all therapist material between patient's statements.
contextual: patient's responses which frame therapist's statements.

Category Set

General features: five dimensions, with mutually exclusive subcategories within each dimension.

Dimensions:

(1) Type of therapist activity--8 subcategories: Facilitating communication; exploratory operation; clarification; structuring; interpreting; direct guidance; activity not clearly relevant to therapy; unclassifiable (there are 16 subcategories).
(2) Dynamic focus--frame of reference; 2 categories: goes along with patient focus; introduces different focus (5 subcategories).
(3) Dimensions of intensity--3 subcategories:
degree of inference--1 to 5;
degree of initiative--1 to 4;
therapeutic climate (warmth)--five levels, -2 to +2.

References

Strupp, 1957, 1962.

Figure 1
Strupp's System

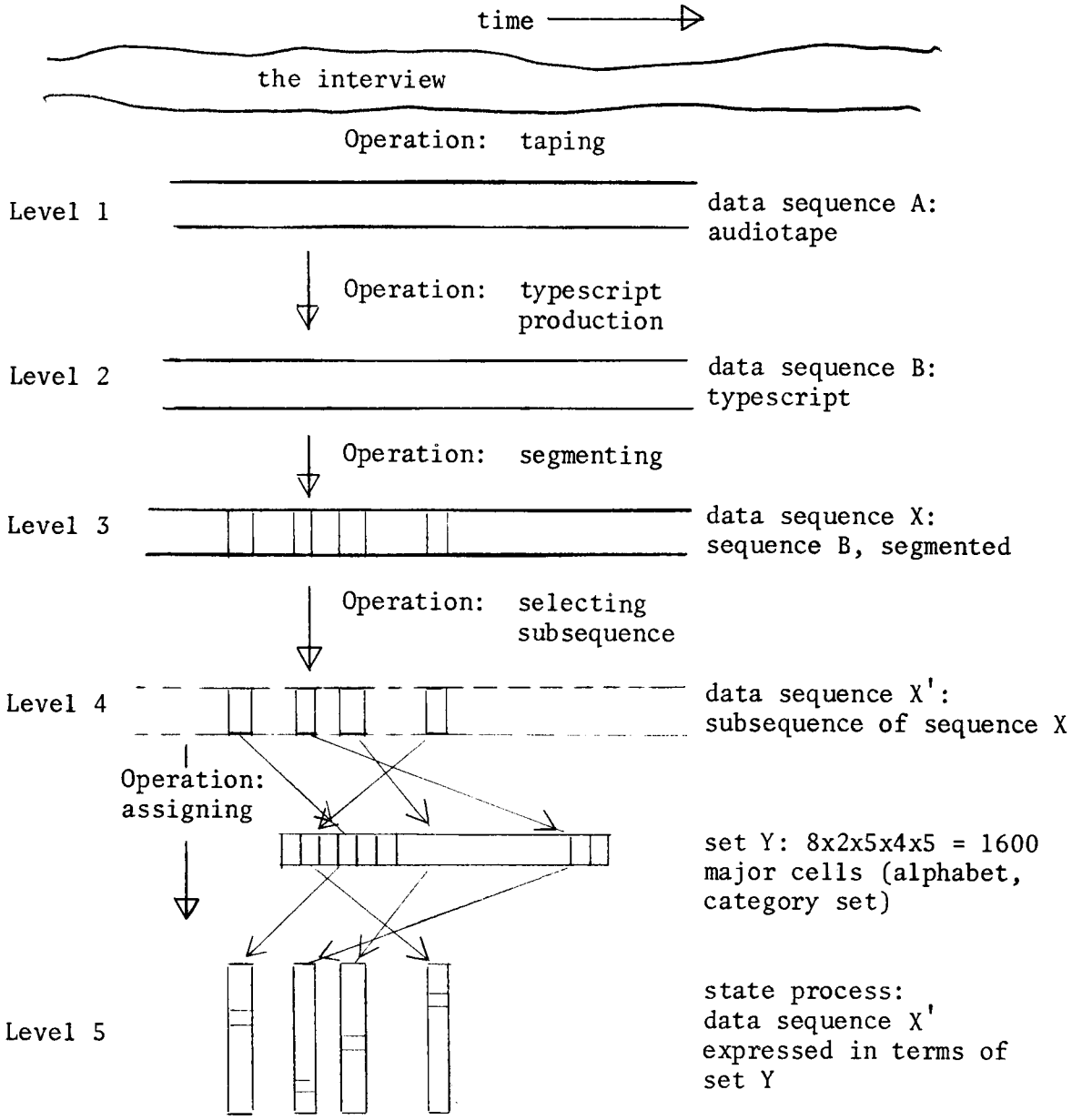


Figure 1 (continued)

to written text and segmentation--are in practice likely to be performed simultaneously by a typist. Subsequently in a fourth operation a partial sequence (X', Level 4) is selected from sequence X. This subsequence consists of the therapist's statements. Following this operation each unit element (therapist statement) of subsequence X' is assigned by a rater or judge to one element of the set Y, the content category labels. At this level, set Y takes the role of an alphabet or state representation set. (In Strupp's system the categories of set Y are defined along five dimensions listed in Figure 1. Y may alternatively be conceptualized in terms of a single dimension, in this example a dimension containing 1600 mutually exclusive alternatives.) This representational alphabet or set is used in the final description of an interview (Level 5); the events have been reduced to a data sequence or succession of states in which each state, an element of set Y, represents a therapist statement. Thus the interview data can be presented as a 1600xN array, where N is the number of therapist statements occuring in the time interval of interest.

A second relatively simple example is Jaffe's system (Figure 2). There are some differences between this and Strupp's system. Here the "observations" are performed automatically by an instrument. Once again, however, the researcher focuses on the acoustic aspect of the interview, only this time he is not concerned with verbal content, only interested in the presence or absence of speech and the identity of the speaker(s). His instrument therefore has been designed (1) to segment the temporal flow into 300 m sec intervals and (2) to identify, for each

General Comments

An electroacoustic system for recording the presence or absence of speech energy.

Operations and Reductions

The technical system simultaneously segments and classifies the acoustic stream (presumably using some predetermined threshold value). The stream is continuously segmented into 300 millisecond intervals.

Units

Scoring: 300 millisecond intervals.

Category Set

General features: four possible states defined by each participant's acoustic output.

Dimensions:

1. Both persons (A and B) are silent--state "0"
2. A's monologue--state "1"
3. B's monologue--state "2"
4. simultaneous speech--state "3"

References

Jaffe, 1968.

Figure 2
Jaffe's System

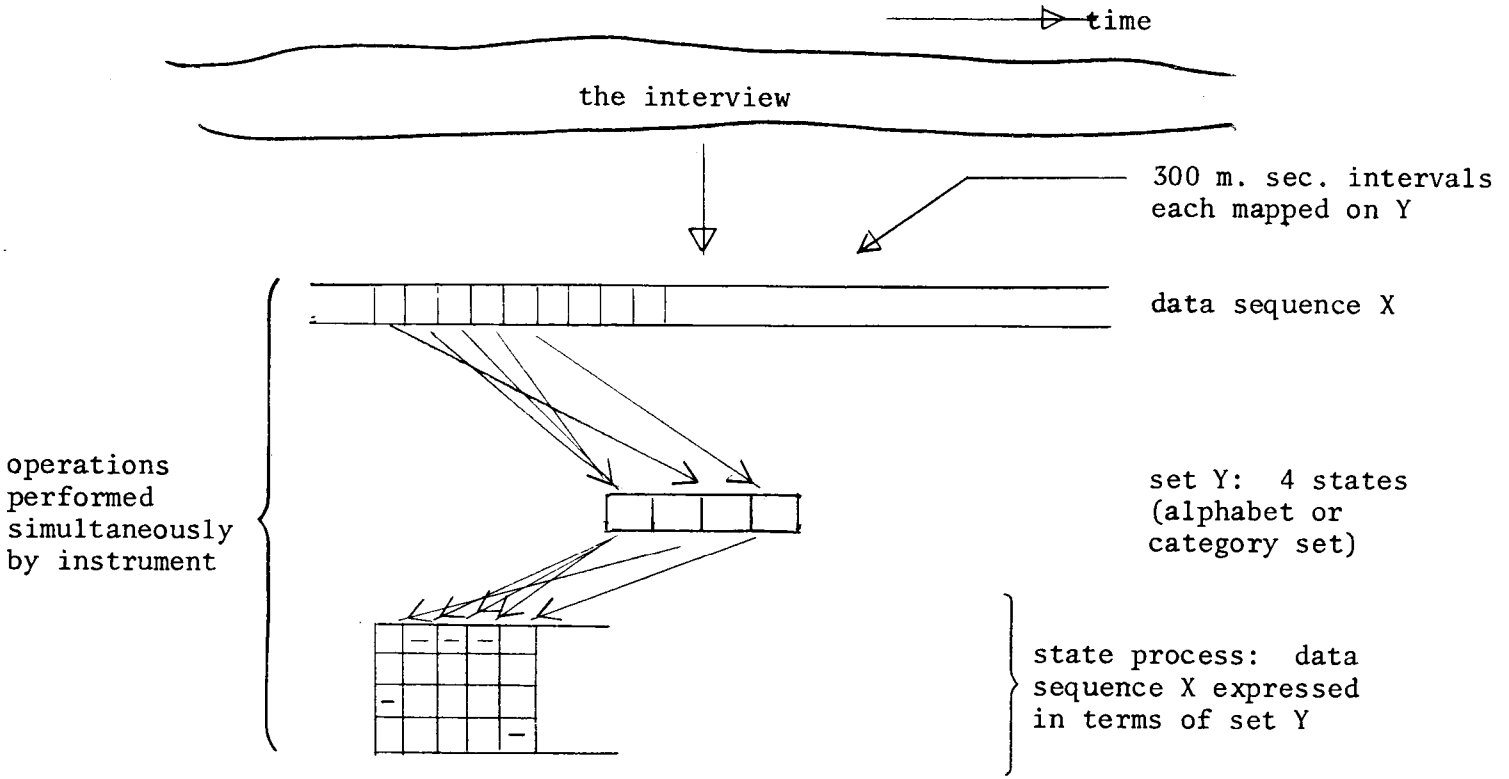


Figure 2 (continued)

interval, which of the four states (set Y, see Figure 2) obtains. These two operations are performed simultaneously. The state representation set can be conceptualized as four column vectors, each having three "zeros" and one "one"; each vector specifies one of the four states which can occur during a 300 m sec interval. Jaffe's data can be described as a temporal succession of these state vectors; the description parallels in all essentials the description in Figure 1.

From this analysis, we can see some logical differences between Jaffe's and Strupp's systems. In Jaffe's system, there is only a single formal step between the interview and the representation. In addition, the data sequence is rigidly segmented according to a fixed unit defined by a physical parameter (time).

The next example is Butler's system (Figure 3). The similarities to the preceding two systems should be evident. The significant logical difference is that now there are two different category sets (Y_1 and Y_2), one for client statements and the other for therapist statements. The state representations (sets Y_1 and Y_2) consist of members which can be depicted by column vectors; one 24x1 vector represents each type of client state, and one 280x1 vector, each type of therapist state. The interview data are represented as an alternating sequence of client and therapist states (Level 3).

Of the 11 remaining systems listed in Table 7, page 52, 8 (Figures 4-11) have logical features very similar to those displayed by the systems just described (Figures 1-3). Therefore, the operations and characteristics of the 8 systems have not been diagrammed; their

General Comments

A system for classifying therapist and client statements based on a theory of stimulus hunger as a basic drive; procedures and variables chosen from a client centered position, using a combined theoretical and empirical approach.

Reductions and Operations

Standard, by raters from audiotapes.

Units

Scoring: "natural unit of response," a complete response by therapist or client.

Category Set

General features: two separate sets, one for client, the other, for therapist responses. Three major dimensions in each set with mutually exclusive choices in each dimension. There are 3x4x2 = 24 mutually exclusive cells in the client's category set, and 5x8x7 = 280 mutually exclusive cells in the therapist's set.

Dimensions:

I. client--(A) level of expression; client attempts (1) analysis of action, (2) of feeling, (3) expresses a feeling or inner experience directly.
(B) Voice quality and manner of speaking--(mixture of amount, degree of control, direction--inward or outward--of energy in voice)--(1) emotional, (2) focused, (3) externalizing, (4) limited.
(C) quality of participation--(1) participating, (2) observing.

II. therapist--(A) Freshness of words and combination (roughly, degree of poetic language, novel metaphor); 5 operationally defined categories.
(B) Voice quality--how live, forceful is it? 8 operationally defined categories.
(C) Functional level of response--how, and in what direction, the statement evokes focus, experiencing, empathy; 7 classes.

References

Butler et al., 1962; Butler et al., 1963.

Figure 3
Butler's System

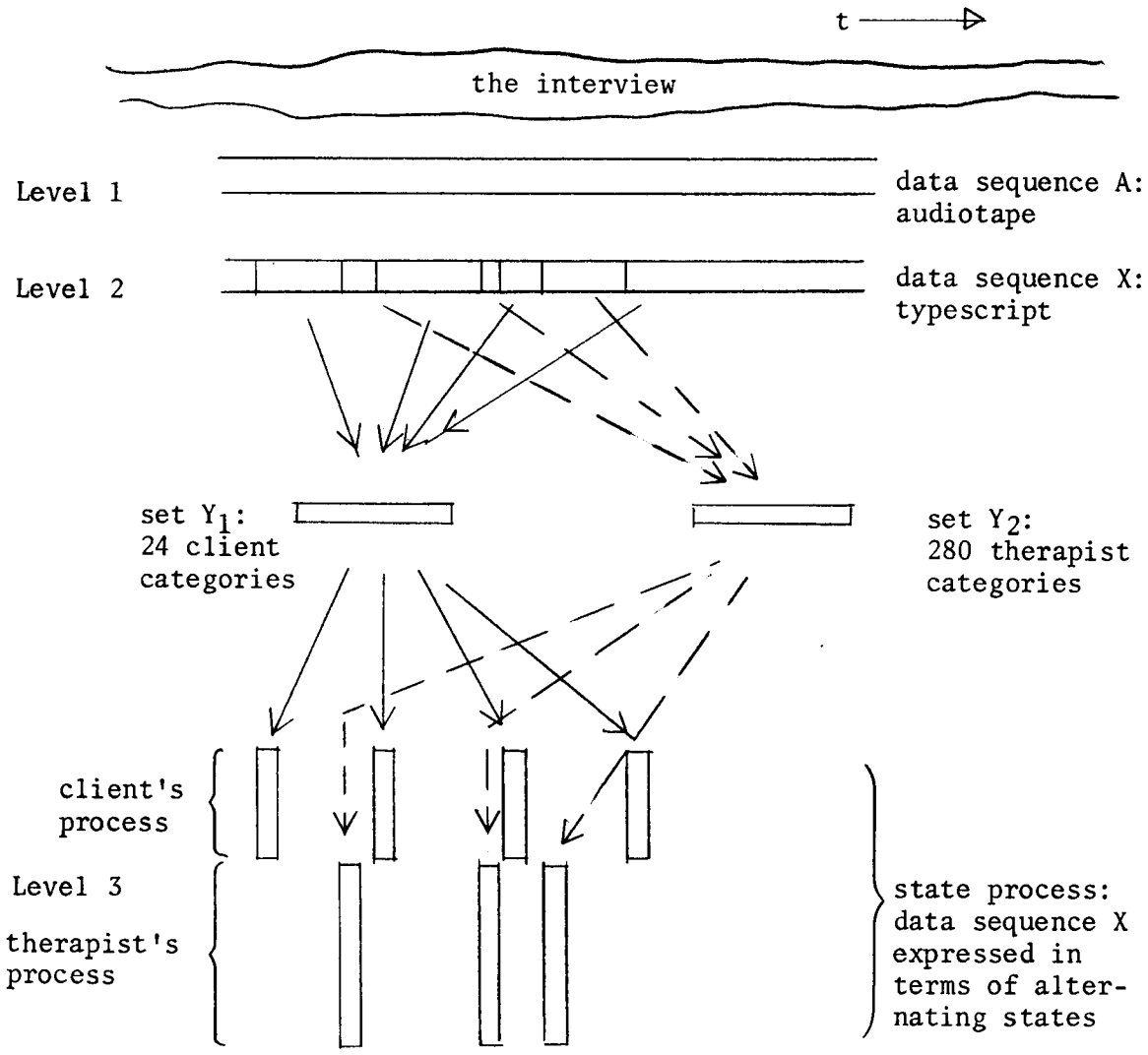


Figure 3 (continued)

General Comments

According to Marsden (1965, p. 312) this is the most thoroughgoing example of the pragmatic (see Marsden) model. It has been developed from naturalistic observation of material from the perspectives of psychoanalytic theory, behavioristic learning theory, and cultural anthropology. It deals with measurement of motivational variables.

Reductions and Operations

Direct classification by raters from audiotapes

Units

Scoring: the sentence

Category Set

General features:

Twenty mutually exclusive (but see below) patient's categories; each may be cs or ucs, making 40 basic choices. For therapist, 6 mutually exclusive categories. States that the pt. categories are mutually exclusive, but examples show that a category assignment may be modified by a subsidiary assignment to a second category. (Logically the distinction is unimportant.)

Dimensions:

Patient--(1) anxiety, (2) confirmation of an interpretation, (3) dependence, (4) dream, (5) hostility, (6) self-criticism, (7) aim-inhibited love, (8) laughter (may be anxious or relaxed--two categories?), (9) striving for social mobility, (10) negation, but not resistant, (11) obsessional, (12) report of psychosomatic symptom, (13) rewarding experience, (14) reasoning, insight, (15) resistance, (16) sex, (17) sigh, (18) weeping, (19) a simple statement which cannot be scored otherwise, (20) misc. (unscorable).

Therapist--(1) interpretation, (2) motivating statement, (3) rewarding statement, (4) signal of continued attention, (5) therapeutic error, (6) unscorable.

References

Auld, 1968; Auld & Dollard, 1968; Marsden, 1965, 1971, p. 378.

Figure 4
Auld and Dollard's System

General Comments

A system for "quantification of affects." The frame of reference is intrapsychic and psychoanalytic, using direct report and inference. The approach also explicitly assumes that a typescript is an adequate source for measuring affect, and that paralanguage or other modalities only clarify affect, being essentially redundant to written textual material. Some evidence supporting this approach is cited. Unconscious affects are not scored by this scheme.

Reductions and Operations

Standard from tapes and typescripts, using raters.

Units

scoring: clause
contextual: sentence

Category Set

General features: four separate dimensions of affect, separately applied. That is, each clause of therapist or patient can be scored for any or all dimensions, or not scored at all. Each dimension ("scale") scores each affect qualitatively into subcategories, and quantitatively as a weighted intensity.

Dimensions:

1. Anxiety; subcategories are death, mutilation, separation, guilt, shame, and diffuse anxiety. A clause is weighted from 1 to 4.

2. Hostility outward; subcategories are killing or fighting others, robbing or abandoning others, criticizing others, depriving or disappointing others. A clause is weighted from 1 to 3.

3. Hostility inward; subcategories are killing self, wanting to die, discouragement. A clause is weighted from 1 to 4.

4. Ambivalently directed hostility; subcategories are others killing, others injuring, others criticizing, and others abandoning, self. A clause is weighted from 1 to 3.

References

Gottschalk et al., 1966, p. 93.

Figure 5
Gottschalk's System

General Comments

Based on a psychoanalytic model, yet claimed to be applicable to any type of therapy. The system is intended to be able to classify every statement consisting of a subject and characteristic. The statements to which the system has been applied either come directly from interviews, transcripts, or from subsequent clinical observations, i.e., a written clinical commentary on therapy material. Each statement is mapped on the matrix described below.

Reductions and Operations

When the system is applied directly to interview material, the reduction is standard; when it is applied to clinical observations, an extra step is introduced: a clinician writes a free, unconstrained narrative which is then treated by the system in the same way as it treats material obtained directly from the interview.

Units

Scoring: "statement," not well defined operationally, consisting of a "subject" and "characteristic."
Contextual: complete sentence (or surrounding sentences).

Category Set

General features: an 11x5 matrix, having 11 major categories (characteristics of statement") along one dimension, and 5 major categories ("subjects of statement") along the other. The 11 categories are grouped into 3 types: Categories 1-5, the clinical categories; categories 6-8, the special clinical categories; and categories 9-11, nonclinical. Each of these 11 categories as well as the first one of the 5 "subjects of statements" categories is further subdivided. A third major dimension of classification is "special modifiers"; their application is obscure in the text.

Dimensions:

I. "Characteristics of statements," 11 major categories; Clinical--(1) insight, (2) admission into speech, (3) admission into awareness, (4) admission into expression, discharge in action, and (5) static condition at some structural level. Special clinical--(6) psychological symptoms, (7) psychotherapeutic, (8) vocal-kinesic. Nonclinical--(9) somatic, (10) social, and (11) inanimate.

Figure 6
Leary's System

II. "Subject of statement," 5 major categories; (1) people or animals (patient in therapy; therapist; patient outside of therapy; people known personally to the speaker; people not personally known to the speaker; and animals), (2) psychological subjects, or abstractions, (3) somatic subjects, (4) sociological subjects, and (5) inanimate subjects.

III. "Modifiers," 3 groups; (1) mixed--grammatical, evaluative, states of consciousness; (2) dynamic; and (3) connectives--logical, causal, change (improvement), or attribute of the subject.

References

Leary & Gill, 1959.

Figure 6 (continued)

General Comments

A content analysis system derived from learning theory, social psychology, and information processing viewpoints, used to score therapist and patient propositions.

Reductions and Operations

Standard, via audio tapes, transcripts, using raters.

Units

scoring: proposition--"verbal expression of a simple idea, usually of subject/predicate form."
contextual: either the sentence, or a boundary of the other speaker's utterance.

Category Set

General features: 5 dimensions, with mutually exclusive subcategories in each dimension; the number of subcategories at the 5 levels is 8, 3, 3, 3, and 4, providing 864 mutually exclusive cells.

Dimensions:

1. Informational specificity; 8 subcategories: passive encouragement, active encouragement, limits discussion to specific old propositions, introduces specific new proposition, interpretation, limits to specific answer, and excludes discussion.

2. Grammatical form; 3 subcategories: declarative, imperative, interrogatory.

3. Affective content; 3 subcategories: refers to feelings, does not refer to feelings, indeterminate.

4. Interactive processes; 3 subcategories (simplified from Bales): descriptive, evaluative, prescriptive.

5. Role system reference; 4 subcategories: primary system (treatment), secondary (patient or therapist in other than their primary roles), tertiary (family, other social systems), and self.

References

Lennard & Bernstein, 1960; Lennard, 1962.

Figure 7
Lennard's System

General Comments

A system for categorizing body movements and gestures in dyadic psychotherapy.

Reductions and Operations

First reduction (to sequence A, Figure 1, page 55) is to a verbal narrative recorded by observers who cannot hear, but only watch, an interview. It is not clear how this free clinical narrative is coded to the 15 categories described below.

Units

scoring:	the maintenance or sustaining of a gesture or movement classified by the category set.
contextual:	not explicitly mentioned, but inferentially it would seem to be the contiguous scoring units.

Category Set

General features: "natural" gestures or movements, classified into four major groups: communicative gestures, autistic gestures, hand-arm positions, and general postural changes.

Dimensions:

Communicative gestures

Pointing (out)
Making fist
Shrugs shoulders
Shakes head "no"
Shakes head "yes"
Turns palms out-up

Autistic gestures

Rubs, touches nose
Scratches self
Finger-mouth
Rubs self
Picking, smoothing, cleaning, arranging
Pats hair

Figure 8
Mahl's System

Hand-arm positions

Folds arms across waist
Hands held or clasped together

General postural changes

References

Mahl, 1968; Marsden, 1971, p. 375.

Figure 8 (continued)

General Comments

A content analysis system based on Chapple's Interaction Chronograph. It was begun in 1954, and used extensively.

Reductions and Operations

Performed on line, by observers operating a two-channel instrument for recording the two participants' "acts"--talking, nodding, gesturing, and in general interacting and communicating. Keys are operated by observers to record beginning and end of each person's episode of "acting."

Units

No formal temporal units.

Category Set

General features: formally identical to Jaffe's four-element set Y (see Figure 2, page 58). Since only binary (acting/nonacting) data are acquired for each participant, derivative dimensions must be defined in terms of the temporal characteristics of the binary data.

Dimensions: 14 derived dimensions, grouped into patient, interviewer, and two simple dimensions as follows--

1. Patient--units, action, silence, tempo, activity, adjustment, initiative, quickness, dominance, synchronization.

2. Interviewer--adjustment, units.

3. Number of interruptions.

4. Length of interview.

References

Matarrazzo & Wiens, 1972; Saslow & Matarazzo, 1959; Matarazzo, Weins, Matarazzo, & Saslow, 1968; Marsden, 1971, p. 351.

Figure 9
Matarazzo's System

General Comments

A system for evaluating the overall status of a client. The process scale is composed of seven strands, each measuring some aspect of a fluidity/fixity continuum. The strands merge as one moves from fixity; in general, the strands are expected to covary.

Reductions and Operations

Standard to either tape or typescript.

Units

scoring: 4 minutes, randomly selected from an interview.

Category Set

General features: seven strands representing seven dimensions which are thought to be dependent. The first three represent the main scales. Each dimension is directly rated by judges on a 7-point scale.

Dimensions:

1 - client experiencing
2 - personal constructs (manner in which experience is construed)
3 - problem expression--relationships to problems
4 - manner of relating to others
5 - communication of self
6 - manner of experience
7 - degree of incongruence

References

Marsden, 1971, p. 348; Rogers, 1959, 1961; Hart & Tomlinson, 1962.

Figure 10
Rogers' System

General Comments

This is a system applied to small group processes (family therapy), drawing on General Systems Theory notions of hierarchical structure. A communicational frame of reference is used, heavily influenced by Birdwhistell's kinesic methodology. It is assumed that communicational messages are concatenations of definite "structural units" in particular fixed arrangements. Motives, defenses, and other references to intrapsychic data are apparently ignored. The framework is described as complementary, but not alternative, to personality theories. The system is multileveled and multimodal.

Reductions and Operations

Standard first reduction to videotape. Raters identify events at several levels according to a very incompletely developed, fragmentary category set, and prepare "multimodal" transcripts; at any level mixed modes (kinesics, vocal data, psychological descriptors) may be used.

Units

Scoring: various "structural units" at various levels, phenomenally defined at present. In increasing order in the hierarchical structure (to larger units), a "point" is a small unit, "positions" are next (these may last from seconds to 20 minutes or more), and "periods" are constellations of positions. These 3 contiguous levels are envisioned as an intermediate sector of the to be finished system.

Contextual: neighboring scoring units at the same level.

Category Set

General features: Category sets at the various levels (see "scoring units" above) are in early stages of development. Cited examples of units at various levels are:

Points-- "taking out a cigarette and bringing it and a pack of matches to the patient's mouth"; "putting cigarette in mouth"; "discarding the match."

Positions--"cigarette lighting"; "smoking"; "postural shift"; "looking up"; "offers interpretation."

Periods-- "refraining from conversation"; "interpreting."

References

Scheflen, 1966, 1973.

Figure 11
Scheflen's System

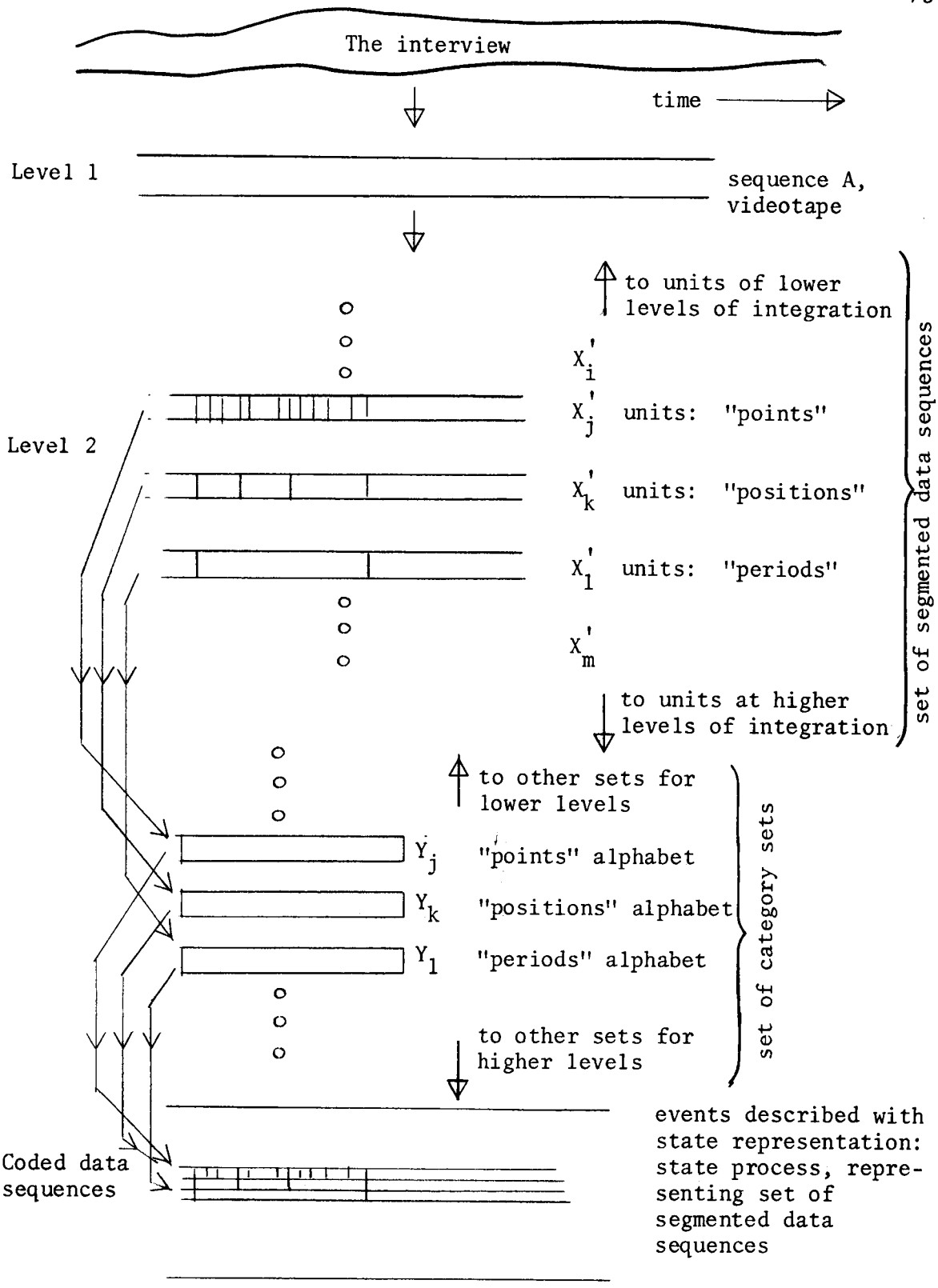


Figure 11 (continued)

features are described in the text of these figures. One exception is Scheflen's system (Figure 11). A diagram of that system has been prepared, since the methodology is more elaborate than usual. As can be seen in the figure, the first reduction (Level 1, data sequence A) is restructured with several different category sets (Y_i, Y_j, etc.); sequence A is changed into corresponding data sequences (X_i', X_j', etc.). The relationships between these various representations are described in the text of Figure 11.

The last three systems to be considered are Deutsch's (Figure 12), Menninger's (Figure 13), and Pittenger et al.'s (Figure 14) system. These differ from the 11 preceding systems in that they do not use a formal state representation set; the 3 systems describe events in free narrative. Thus, there is no formal, unambiguously defined data sequence. These narrative data could be conceptualized in terms of a variety of alphabets; for example, the set of English letters, or words, phrases, sentences, or any other suitable set of structures could be used as an alphabet in a formal analysis of observation. For free narratives the representation set is to a great degree in the eye of the beholder, the methodological analyst; he can define the representation set on the basis of convenience, appropriateness to the analysis, tradition, etc.

Figures 1-11 have presented eleven systems of observation which have broad logical characteristics in common: By a series of information-reducing changes, the initial record acquired from the complete situation (interview) is ultimately represented as a temporal succession

General Comments

A free narrative of kinesic observations, based on psychoanalytic theory.

Reductions and Operations

Standard reduction to sound film and typescript. Two independent observers recorded their free descriptions of motions and postural changes, one observer describing each interview participant. A clinician then freely interpreted these observational narratives in conjunction with the interview transcription. No further reductions were performed. The system has not been used in formal experimental research.

References

Deutsch, 1966; 9 other references dating from 1949-1959 are given in GA, p. 169.

Figure 12
Deutsch's System

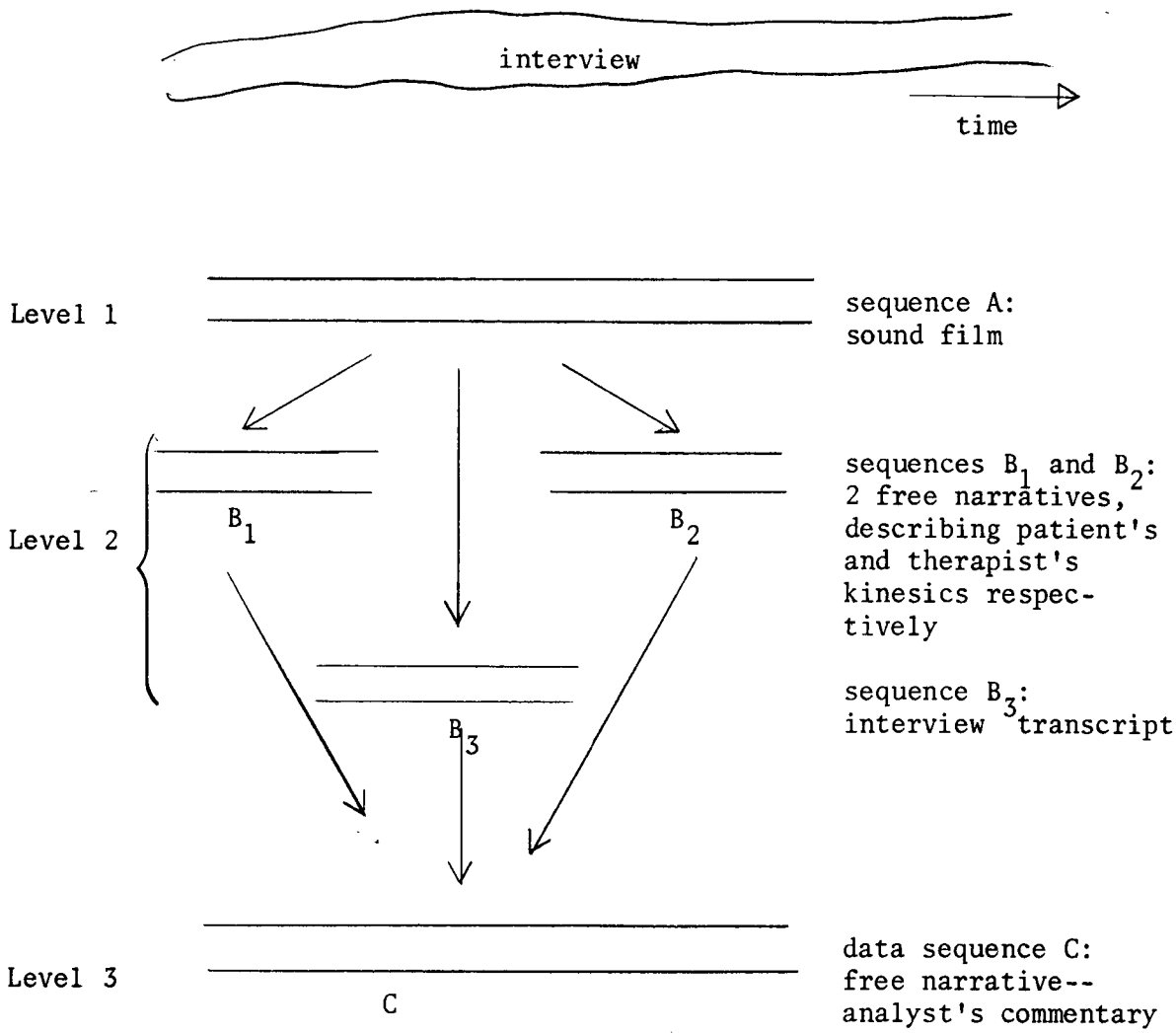


Figure 12 (continued)

General Comments

The Psychotherapy Project of the Menninger Foundation, begun in the mid-50's, is distinguished by its reliance on narrative clinical data as a basis for developing research data. Most of the restructured data remain in narrative form (see "reductions and operations" below). The body of clinical narrative was the basis of several derivative studies--the Clinical Prediction, Differential Treatment, Therapeutic Process, and "Special," Studies.

Free clinical data ("Form A") were interpreted by research clinicians at three observation points (initial, termination, and follow-up). At the initial observation, Form A consisted of all material obtained from Menninger's usual evaluative procedure--psychiatric case study, psychological testing, and summaries of social work contact with significant family members. Appropriate further clinical data were added to Form A at termination, and at follow-up (approximately 2 years after termination).

Reductions and Operations

From Form A, researchers evaluated 28 patient (Form B), 30 treatment (Form T), and 7 situational (Form S) variables. These evaluations were presented in narrative form, although some variables suitable for quantification (such as intensity of anxiety) were rated numerically also.

Units

No formal units were used.

Category Set

General features: A list of patient, treatment, and situational variables.

Dimensions:

Form B: Assessment of Relevant Patient Variables

I. Sex and Age
II. Anxiety and Symptoms
    1. Anxiety
    2. Symptoms
    3. Somatization
    4. Depression and guilt feelings
        a. Depression
        b. Conscious guilt
        c. Unconscious guilt
    5. Alloplasticity

Figure 13
The Menninger Project's System

III. Nature of Conflicts
  1. Core neurotic problem
  2. Current life problem
IV. Ego Factors (and Defenses)
  1. Self-concept
  2. Patterning of defenses
  3. Anxiety tolerance
  4. Insight
  5. Externalization
  6. Ego strength
V. Capacities Factors
  1. Intelligence
  2. Psychological-mindedness
  3. Constitutionally endowed aspects of ego strength
  4. Capacity for sublimations
VI. Motivational Factors
  1. Honesty
  2. Fee
  3. Extent of desired change
  4. Secondary gain
VII. Relationship Factors
  1. Quality of interpersonal relationships
  2. Transference paradigms
VIII. Reality Factors
  1. Presence of "neurotic life circumstances"
  2. Adequacy of finances to the treatment requirements
  3. Attitudes of significant relatives
  4. Physical health

Form T: Assessment of Relevant Treatment Variables

I. Formal Elements of the Treatment
  A. Treatment Modality
    1. Variations determined by the nature of the patient's illness
    2. Variations determined by reality events
    3. Variations determined by the therapist
    4. Variations determined by the supervisor
  B. Treatment Instructions, General and Special
  C. Special External Circumstances that Affect the Treatment Course
  D. The Basic Techniques of the Treatment
    1. Suggestion
    2. Abreaction
    3. Manipulation
    4. Clarification
    5. Interpretation

Figure 13 (continued)

E. Types of Subject Matter
   1. The patient's past life and memories
   2. Reconstructions and memories of the conflicts of infancy and childhood
   3. The patient's current life situation
   4. The events of the treatment hour
   5. Dream and fantasy material
   6. Style and form of the patient's expression
   7. Accompanying affect of the patient's expression

F. The Goals of the Treatment
   1. The goals of the patient
   2. The goals of the therapist
   3. Changes in goals

II. The Process of Treatment
   A. Life of the Patient During Treatment
   B. Significant New History
   C. Course of Treatment
      1. Major themes
      2. Major transference paradigms
      3. Changes in symptoms
      4. Changes in impulse-defense configurations
      5. Changes in manifest behavior patterns
      6. Structural changes in the ego
      7. The acquisition of insight, and its relation to changes in attitudes towards the self, others, things, and the illness
      8. Style and form of communication
   D. Termination of Treatment

III. Variables in the Therapist and the Climate of the Patient-Therapist Interaction
   A. Qualities of the Therapist
      1. Skill and experience
      2. Need for supervision and how he uses it
      3. Personal style and attitudes
      4. Countertransferences
      5. Age and sex
   B. Climate of the Therapy

Form S: Assessment of Relevant Situational Variables

I. Situational Factors
   A. Background
      1. Cultural
      2. Religion, ethics and values
      3. Education

Figure 13 (continued)

B. Interpersonal Relationships
   1. Family relationships (parental)
   2. Family relationships (own family)
   3. Friends, coworkers, colleagues, etc.
C. Marital-Sexual
D. Living Situation
   1. Home arrangements
   2. Financial
   3. Responsibilities
E. Occupation
   1. Stability
   2. Occupational demands
   3. Compensation
F. Community and Leisure Time
   1. Recreation and avocation
   2. Group participation
   3. Civic and cultural activities
G. Physical
   1. Appearance and physique
   2. Handicaps
   3. Somatic illness

II. Each of these situational factors is evaluated in respect to the following psychosituational dimensions:
   A. Degree of relevance in the individual life
   B. Amount of stress
   C. Degree of support
   D. Degree in which the situation is a crucial conflict area
   E. Extent of opportunity for self-realization growth, autonomy
   F. Congruence of the situation to needs, interests, capacities
   G. Degree of situational mutability

References

Robbins & Wallerstein, 1959; Luborsky, 1962; Wallerstein, 1966, 1968; Sargent, 1961; Wallerstein and Robbins, 1956.

Figure 13 (continued)

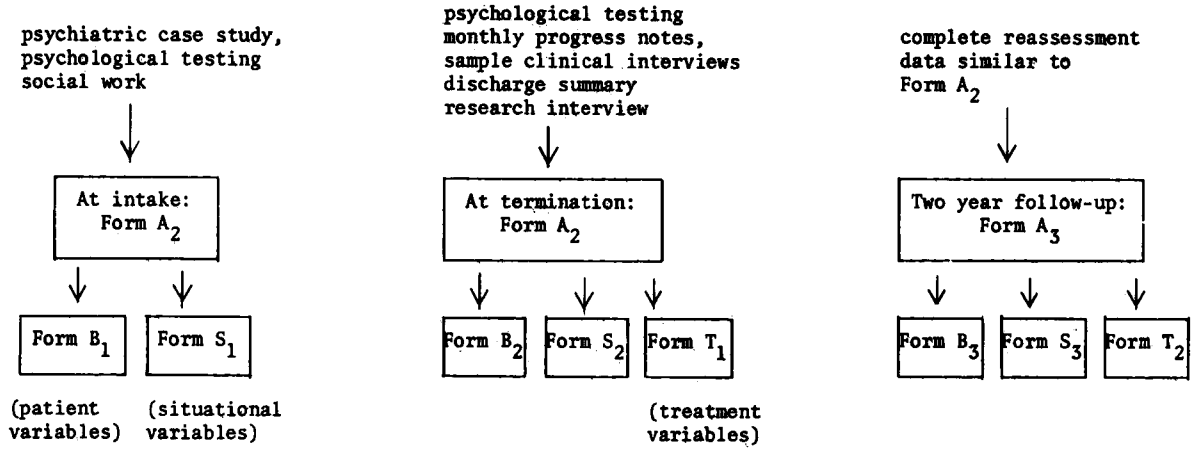


Figure 13 (continued)

General Comments

This system uses linguistic and paralinguistic data as a basis for an intensive analytical interpretation of interview material. The analysis ranges over the total interview segment analyzed, relating various portions to each other; thus, even though the main focus is on individual statements (alternating patient and therapist utterances), inferential use is made of remote, noncontiguous material. The study emphasizes the use of paralinguistic cues to discover material concealed in the overt text; these cues can call attention to rehearsal; covert messages; gambits; use of "opaque intonation" as message; preoccupation, etc.

Reductions and Operations

Standard reduction to tape recordings. A standard transcript is prepared from the interview, together with a multilevel paralinguistic and linguistic transcription.

Units

No formal units are used, other than the standard alphabet and linguistic codes.

Category Set

General features: the transcript is standard English text. The transcription uses special symbols for paralinguistic (musical and noise sounds) elements and standard phonemic symbols to transcribe linguistic data.

References

Pittenger et al., 1960.

Figure 14
Pittenger, Hockett, and Danehy's System

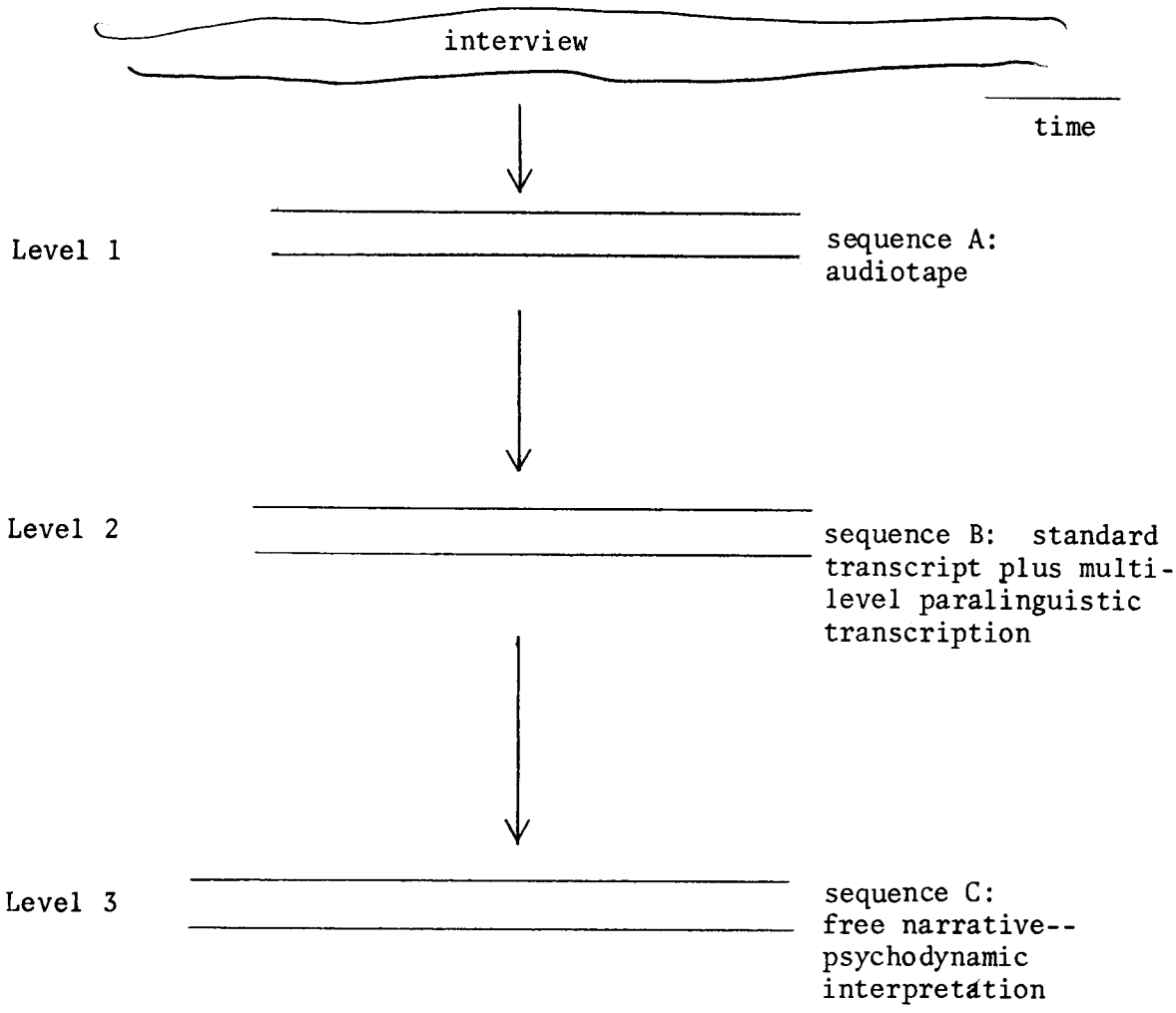


Figure 14 (continued)

of states selected from a representational category system (variously referred to in the present work as the state representation set, category set, category scheme of observation, or classification set). This set often displays typical features: Researchers generally use some multidimensional content analysis classification with each dimension consisting of a finite, relatively small, set of mutually exclusive subcategories which exhaust that dimension. The assignment of a particular scoring unit to one or more dimensions (typically, to all dimensions) can be represented as assignment to one of a mutually exclusive, exhaustive subsets obtained by the Cartesian product of all the dimensions subcategories. For example, in Butler's system (Figure 3) the client category system uses 3 major dimensions having 3, 4, and 2 subcategories, respectively. Every client statement is assigned to a subcategory in each dimension, and one can therefore represent this process logically as one which assigns each client statement to <u>one</u> of 3 X 4 X 2 = 24 subsets in the cartesian product of these 3 subsets.

The common features of observational systems using formal category sets, that is, state representation sets or alphabets, can now be presented in terms of a general model. That model is shown in Figure 15. Specific, typical operations (such as audiotaping, Level A) have been used to illustrate the general process; these specifics in no way limit the generality of the figure. It summarizes the salient characteristics of the systems shown in Figures 1, 2, 3, and 11, pages 55, 58, 61, and 72, respectively, using the schematic representations familiar from these earlier figures.

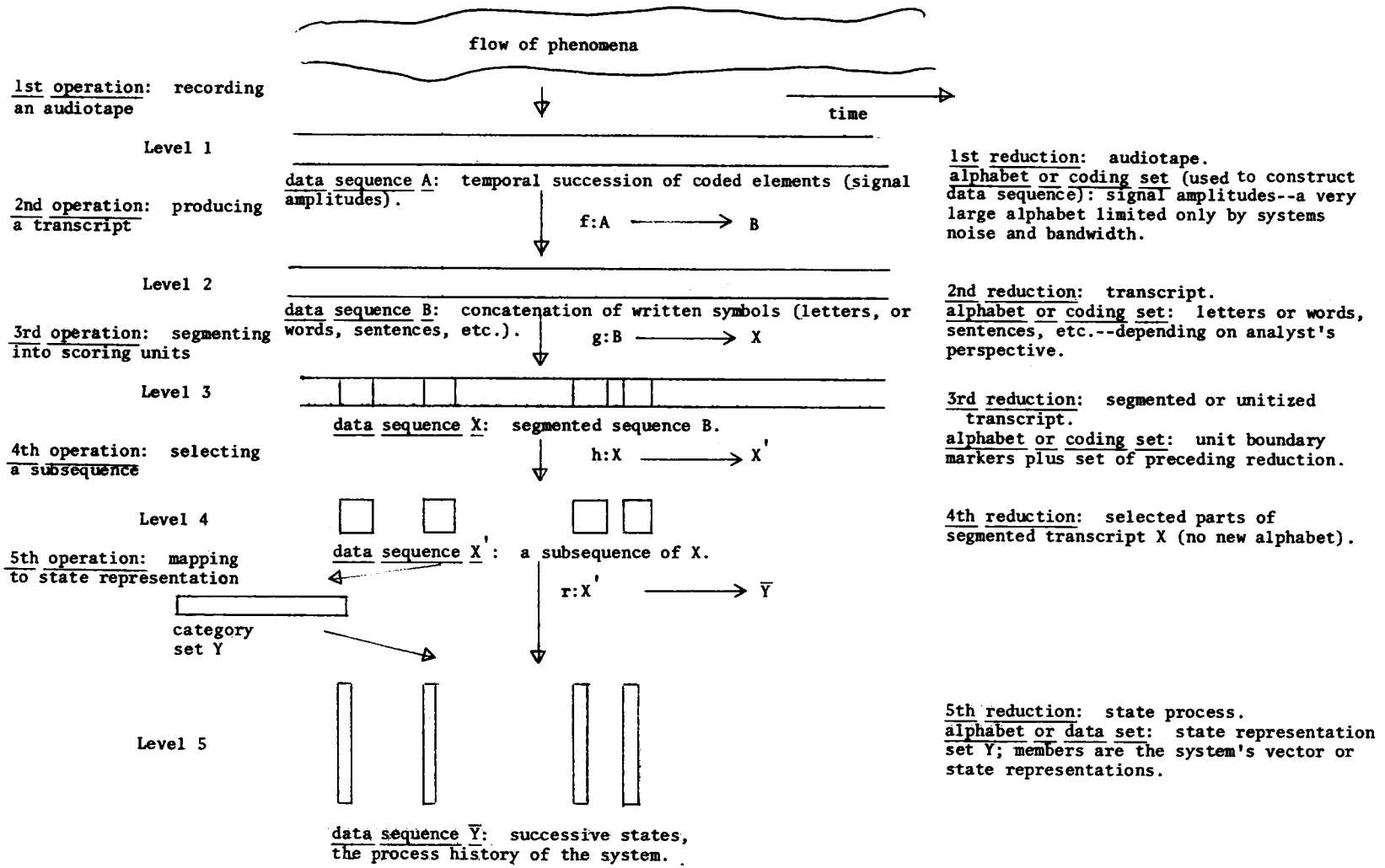


Figure 15
Model of a Typical Content Analysis Process

A separate presentation of the common features of the last three systems (Figures 12-14, pages 75-83) would add little to the discussion. These systems are logically simpler than the others; they do not code data by means of formal alphabets, and therefore mainly use operations and reductions such as the first two shown in Figure 15. Narrative data, however, can be processed formally--implicit examples were contained in Mahl's (Figure 8, page 68) and Leary's (Figure 6, page 65) systems. It will be recalled that in Mahl's method of observation an observer using only visual cues records free, narrative comments; similarly, Leary's system although frequently used to code interview transcriptions can also be used to "observe," i.e., analyze, a clinician's free narrative analysis of interview material.

An important thing to note in these two examples is that a free narrative can be viewed as a research datum and processed formally just as though it were a data sequence such as sequences A or B of Figure 15; thus, a narrative can be considered as logically equivalent to interview data such as a typescript prepared from an audiotape.

In Mahl's and Leary's systems free narratives were reduced to formal data sequences by processing these narratives in terms of well-defined content analysis categories. Some observational systems, however, do not process data beyond a narrative level. For example, in the systems illustrated in Figures 12-14, pages 75-83, the data remain in narrative form after processing is completed. Free narratives apparently do not lend themselves well to research manipulations (other than formal operations similar to those just discussed). However

informative, interesting, or clinically useful such narratives might be, the formal research potential of an observation appears to end when an observational system produces free data--in the absence of formal processing traditional research methodology has difficulties in dealing with these data. This conclusion is supported by the lack of published formal research based on Pittenger's, Deutsch's, or other similar narrative observational methods. Hence, of course, the prevalence of formal systems such as content analysis or psychological testing.

Several general comments can be made about free clinical or other narratives as methods of observation. First, although one does not gain in a formal sense, one may gain in a semantic sense by recasting a narrative from Level 2 to Level 3 in Figure 12, page 75, for example): This step frequently does represent the application of expert opinion to data, and one's understanding of the original material should therefore be enhanced. Second, two different narratives may be differently constrained although both may be relatively free. For example, the data in a typical interview transcript are likely to be loosely constrained by general overt and covert ground rules (such as those that govern the linguistic behavior in clinical interviews) while an expert commentary may be constrained differently--perhaps it is addressed specifically to fellow professionals. Third, although narratives frequently concern, or are based on, other free data such as typescripts, narratives may also be commentaries on formal, highly constrained data sequences. For example, a clinician could perhaps enlarge on the implications of a specific succession of four element vectors such as those

generated in Jaffe's data (Figure 2, page 58). This hypothetical example illustrates that in the class of systems modeled in Figure 15, page 85, an operation could reverse the direction of data reduction; a highly constrained low-dimensional data sequence could be expanded into a higher-dimensional, considerably less constrained narrative sequence by means of an insightful observation. Thus, a somewhat more general model could display two directions of data manipulations--reductions and expansion. Fourth and last, it may be noted that usually operations of these types (free narrative comments or interpretations) cannot be readily formally characterized: An observer's free narrative typically arises not only from his scientific, clinical training, but also from intuitive, unreflected thought.. Thus, the step that changes one verbal stream to another represents a highly complex, poorly understood logical operation which resists definition at present.

A discussion of the taxonomy of observation concludes this section. The analysis of various observational techniques has led to identification of a dichotomous classification: (1) systems which formalize observations by using a state representation set, and (2) those whose final data sequences remain at a narrative level. This dichotomy appears to apply very generally to research data in psychology. For example, psychological testing can be categorized in this manner: Structured tests tend to generate "observational" data sequences coded with a limited, formal alphabet (for example, multiple choice tests), while projective tests tend to generate free "observational" narrative data (for example, The Thematic Apperception Test). From the standpoint

of the analytic model, test data sequences are identical to the classes of sequences illustrated in Figures 1, 2, 3, 11, and 15, pages 55, 58, 61, 72, and 85, respectively. Thus, it would seem that most if not all classes of clinical research data can be classified according to these logical categories.

Traditional taxonomies could be examined from this logical perspective. A prominent, traditional classification of clinical research is into process and outcome studies. As noted in Chapter I, this dichotomy has led to confusion and been seriously criticized, yet it is still very much used today (Sargent, 1961; Howard & Orlinsky, 1972; Malan, 1973). Much of this confusion may result from the fact that both process as well as outcome studies can use (1) free or formal observational systems (2) within or outside therapy sessions. That is, in terms of the analytic framework of the present work, outcome and process studies could both be based on any one of four classes of research implied by the two independent dichotomous dimensions (free/formal, and within/outside therapy) mentioned above. A taxonomy of research based on these logical distinctions would seem to be more serviceable than the traditional, ambiguous process/outcome dichotomy.

Critique

The general features displayed by the observational procedures analyzed in the preceding section show certain shortcomings which are characteristic of observational systems used in psychotherapy research. This section will examine five issues; three of these are primarily

formal--they relate to logical aspects of observational processes--while the other two are concerned with more empirical questions. The first three topics are (1) formal features of the state representations, (2) completeness and redundancy, and (3) translatability among systems. The two remaining topics are (1) representation of interactional phenomena and (2) pragmatics of observational systems.

Formal Characteristics

The preceding discussion has shown that the major systems listed in Table 7, page 52, could be divided into two broad classes: formal systems, using state representations to describe interview processes (as well as some psychological test data), and "free" systems which process observed data only to a level of narrative commentary. The latter systems do not display any particular formal features and will not be examined further here. Formal systems, however, do exhibit certain common broad characteristics.

With only one apparent exception (Scheflen) the formal systems use a state representation set whose elements can be conceptualized as column vectors. The interview process is then represented by a succession of these vectors--the interview is seen as marching along in time, interval by interval, one column vector representing each temporal segment. These systems thus emphasize a vertical view of phenomena. Scheflen's system is multileveled; his hierarchical structure, it will be remembered, has different units at different levels. At any one level the process is symbolized by a scheme which can be interpreted as a column vector formalism. In order to relate levels to one another,

various units at the level are combined according to a specific presciption to define a single unit at the next, higher level. The important feature to note from the standpoint of logical analysis is that a group of units at one level is identified and rigidly defined as constituting a unit at the next level. Therefore, although it may not be readily apparent, in Scheflen's system any level of data can in principle be reconstructed mechanically from data at the lowest level by following the defined procedures for combining units. Ultimately, then, this system also represents the interview as column vector movement, although in somewhat disguised form.

Constraining the formal observational features to a column vector format appears to do violence to some of the characteristics of multi-level interview phenomena. Different aspects of an interview may move in different tempos and units; levels are not necessarily in temporal coherence or in register with one another. For example, Mahl (1968) has shown that certain classes of gestures and movements can be precursors to later verbal behavior; verbal expression can be delayed by moments, or by months. Furthermore with multilevel systems there seems to be no a priori justification for Scheflen's assumption that data at one level invariably combine to form next higher data units according to a fixed formula or pattern.

The first point of the present criticism, then, is that the appropriateness of the vertically conceived column vector formalism as an adequate representation of interview phenomena is open to question.

Completeness, Redundancy

The traditional assumption that piecemeal research is a proper approach to research in psychotherapy has already been mentioned; it is commonly assumed that partial (and sometimes very partial) data from various research endeavors can be assembled to give a coherent picture of clinical phenomena (Bordin, 1965). On the other hand, researchers frequently assume that multilevel or multimodal approaches to data acquisition will only yield data which are largely redundant (Bordin, 1968, p. 418). Partial data are therefore viewed as able to represent adequately the important features of the phenomena under study. Both these assumptions support the view that only partial observations are necessary.

Results of some researchers suggest that this approach is open to question. Mahl's previously mentioned work in kinesics shows that movements and gestures provide considerably more than redundant information already available from interview transcripts. Similarly, Pittenger et al. (1960) use linguistic and paralinguistic data unavailable from the usual transcript to significant clinical advantage. By using all these data produced in a given utterance, by ranging over the data of large portions of adjoining material, and by attending as well to psychodynamic and other semantic implications of the narrative a great deal of clinically useful (although experimentally intractable) information can be culled from their five-minute interview segment. Data from the usual transcript would not allow this kind of analysis.

The second criticism, therefore, is that observations are thinned out--that is, logically reduced--far too casually. Questions of what

constitutes an adequate, yet nonredundant, representation need to be carefully reexamined.

Translatability

The literature mentions the desirability of translating research data from one project for comparison with another, but it seems that the formal prerequisite requirements remain unexplored. Given two systems of observation, what, if any, properties must these systems exhibit to permit translation? One obvious characteristic is dimensionality. If two representational systems of state vectors have different dimensionalities it is not possible to achieve full, direct bidirectional translation of basis vectors. Data from the first, higher-dimensional system must be collapsed to allow assignments to the second system, while assignments in the reverse direction, from the second, lower dimensional system to the first could use only a partial set of the first system unless additional subcategories are created in the second system by combinations of the original vectors. A similar analysis holds for two systems using differently segmented data. In addition, there is the very difficult problem of creating a dictionary for data translation. If two projects are based on widely different theoretical positions in psychotherapy, how can research data from two such projects be directly related to one another? This problem area in addition to facing the difficulties deriving from diverse theoretical orientations must also solve the formal questions raised by the mapping problems alluded to above. (This topic will be considered further in Chapter V.)

The third criticism, then, is that observational systems traditionally are developed without giving adequate thought to questions of data translatability.

Interactional Phenomena

A number of experimental methodologies identified by their users as interactional methods are reported in the literature (Marsden, 1971, pp. 367-370). Indeed, a chapter in Meltzoff and Kornreich (1970, Ch. 16) is entitled "patient-therapist interactions." Some of the better known so-called interactional systems are Chapple's Interaction Chronograph (Saslow & Matarazzo, 1959), Bales' system (Meltzoff & Kornreich, 1970, p. 439), and Isaac and Haggard's methodology (Meltzoff & Kornreich, 1970, p. 446). On closer examination, however, these and most other reported efforts to capture interactions experimentally reveal themselves as familiar chaining models discussed in Chapter I. For example, Bales' categories define an individual's behavior. They classify his social-emotional behavior (positive--agrees, shows solidarity; negative--disagrees, antagonistic), and task-area behavior (gives or asks for suggestion, orientation, opinion). These behaviors clearly fit the traditional "subject response" label. Some investigators (Gottschalk, Winget, Gleser, & Springer, 1966, p. 105) have simply pooled dyad data and consider the result an interactional measure, an approach without great intuitive appeal.

On the whole, it would seem that the set of conceptual tools available to researchers in psychotherapy does not include an observational category system designed to genuinely describe interactional phenomena. (Possibly this is a reflection of the scarcity of such terms in the English language.) The problems of talking about such phenomena have also been mentioned by Pittenger et al. (1960, p. 222).

The fourth criticism, then, is that observational systems have failed to provide significant, meaningful methodologies that are well able to reflect interactional phenomena.

Pragmatics of Observation

To introduce this topic, let us consider the far-ranging analysis by Panzetta (1974) of the problems inherent in developing a psychiatric nosology. Although he focuses on these problems, his article has particular relevance to the present section because he analyzes problems of nomenclature, thus of categorization, and therefore to an important degree problems arising in analyzing formal observation.

Panzetta presents the notion of "arbitrary focus," explaining that a system of categorization begins with inherently arbitrary ways of carving up the world--this is a reality, not a deficiency of such systems. In that case, there are no "true," theoretically privileged classification systems including systems of observation. On what basis, then, would one choose between alternatives? Panzetta proposes several classes of criteria, but the one most relevant to this analysis is the

pragmatic aspect of the classification system--what function is to be served by a particular classification scheme?

In his discussion Panzetta mentions interventional, descriptive, and administrative functions required of psychiatric nosologies; he suggests that different systems must be created to perform different services--he advocates what he calls "nosologic pluralism." These ideas also are applicable to psychotherapy research where little attention seems to have been paid to pragmatic issues. Rarely if ever does a researcher in that field ask, Why (in a pragmatic sense) are we observing the phenomena of psychotherapy? Are we observing in order to serve the clinician, the theoretician, the methodologist? It seems to be tacitly assumed that "the" scientific approach to experimental observation can serve all pragmatic needs equally well. May not this assumption be partly responsible for the failure of psychotherapy research to be useful?

The fifth criticism is that psychotherapy research tends to try to develop general systems of observation, ignoring the implications that pragmatic aspects hold for a system to be employed in a given research effort.

The Search for a Comprehensive System

The approach to be developed in the present work follows in part directly from the five points of criticism made in the preceding section. Thus a "complete" system of representation would not be tied to a rigid vector formalism; it would be complete yet try to avoid

redundancy; it would probably be composed of a family of pragmatically specified systems; ideally, members of the family would at least to some degree be transformable from one to another; and a set of descriptors would include truly interactional representations rather than only variants of a chaining, input-output specification. An ideal family of descriptors is beyond reach at present; neither an adequate theoretical framework nor an appropriate data base is available. What can be done, however, is to suggest some general features of such a system, and then partially sketch in details of some of its aspects. These tasks will be addressed in Chapters IV and V after a detour through some mathematical investigations in Chapter III. This chapter concludes with an impressionistic picture of a general structure which might be exhibited by a "complete" system.

It has been argued that the rigid form of traditional representations is inadequate; it locks the description of psychotherapeutic phenomena into an inflexible model and thus fails to take into account properly the natural characteristics of these phenomena. Traditional systems are also highly incomplete; granted that descriptive systems in psychotherapy most probably will always fall short of the ideal of completeness, there are some major inadequacies that could be at least partially remedied by a more inclusive approach (Perhaps one should speak about "nearly complete" systems, rather than "complete" ones.) Considering now a single member of a representational family of "nearly complete systems," what are the broad features that such an observational system

would display? It would certainly have multiple levels. Three likely sectors would be patient, therapist, and interactional representational levels. In turn, the patient and therapist sectors each would be multilevel, containing verbal and nonverbal levels; we can not specify as yet structural details of an interactional sector. In each level there would be a multiplicity of units, flexibly defined on the basis of semantic and pragmatic considerations.

The emerging picture of the proposed descriptive system has features which have been alluded to in the literature. Thus, two researchers who have struggled for some time with representational problems describe the phenomena in psychotherapy thus:

> We are dealing with ordered patterns of change during change, which exhibit rhythmic and varying patterns in the temporal sequencing of such changes. Metaphorically, there are waves within waves within waves, with complex yet determinable relationships between the peaks and the troughs of the levels of waves, which serve to express organized processes with continually changing relationships [Condon & Ogston, 1967, p. 224].

The events of psychotherapy also have been compared to drama:

> It is interesting to note that the construction of a simulation of an initial psychotherapy interview has many features in common with the writing of a play or a scenario. There is a beginning, a middle and an end; there are two characters; there is the tension or conflict which arises when strangers meet, and there is a resolution of this conflict [Bellman & Smith, 1973, p. 149].

Yet another reference to a complex structure can be found in Pittenger et al., who note that

> In our transcription we try to represent, mainly through arbitrary graphic symbols, all the linguistic and paralinguistic behavior of the speaker. The transcription can be thought of as more or less like a musical score, in which

> things that happen simultaneously are represented by symbols vertically aligned with each other, while something that happens later is represented by a symbol to the right of that representing something that happens earlier [Pittenger et al., 1960, p. 186].

The last quote may provide an appropriate reference model for an envisioned system of representation. The parallels between the proposed class of systems and a musical score, although metaphorical, are compelling. We are in a highly symbolic, nondiscursive world, where the very small--an interval, a thematic fragment--may be as significant as a total work; where a fragment illuminates and is illuminated by larger units; where references to preceding and subsequent material predominate, yet the elements are ordered temporally; where the process displays concurrent voices, each having its own sense and boundaries yet logically related to every other voice; where the primary structural feature is variation, thematic development and the growth and unfolding of material fragments; and finally, where silences are crucial and cannot be defined as the mere absence of acoustic energy deriving their meaning contextually from the total musical whole. In this Gestalt of events, all is meaningful, yet meaning is beyond the analysis of any atomic component element.

It is not going to be an easy research task to develop a family of representational systems that will adequately symbolize and capture these characteristics of the psychotherapy interview. A musical score at least provides us with a simple directional signpost.

CHAPTER III

THE MATHEMATICS OF OBSERVATION

Standard Approaches

The preceding discussions have carried the analysis of observational processes to the level where phenomena are reduced to data sequences expressed in terms of state representation sets ("alphabets"). At that level, the data are still nominal. If there is a next step, it usually involves mathematization--the ordering and analyzing of data by means of a formal axiomatic system. Many of the mathematical techniques employed in psychotherapy research have been used in the fourteen systems analyzed in Chapter II; accordingly, in this section mathematical techniques will be illustrated through examples from these systems, supplementing the presentation with other illustrations as required.

It is proposed to classify mathematical operations into primary (or direct) and secondary (or derived) operations. The first class of operations are those that change nominal data sequences from strings represented in terms of a state representation set or alphabet to numerical form. Secondary operations are those that further process these data once they are in numerical form.

Primary Operations

Of the fourteen systems of Chapter II, three do not employ primary operations: Scheflen, Deutsch, and Pittenger. The remaining systems

use methods which can be classified into five categories of direct operations: counting, pattern analysis, transitional probability, pair comparison, and rating scales. Each of these will be discussed in turn.

Counting. The ingredients of this operation are (1) specification or definition of a subset of the state representation set and (2) definition of a summarizing unit (commonly, a temporal interval); counting then consists of tallying the occurrences of the subset within the summarizing unit. This operation is of course a very frequently encountered primary operation. It is used in six of the fourteen systems: Strupp, Mahl, Gottschalk, Lennard, Leary, and Matarazzo (although this last system also uses another primary operation to be discussed later).

Strupp's summarizing unit is the complete interview. His system, it will be recalled, represents phenomena along five dimensions (Figure 1, page 55). In any one specific interview, counting is performed in each dimension, frequencies being counted for each subcategory of that dimension. Mahl (Figure 8, page 68) counts frequency data for his classification system of 15 kinesic categories. The summarizing unit is not clear from the data, but probably also is a complete interview (Marsden, 1971, p. 375). For each kinesic category Mahl calculates the average temporal interval between the occurrence of a gesture (the inverse of frequency), and then qualitatively compares these averages for men and women. Gottschalk, using a summarizing unit of five minutes, counts subcategory frequencies for each of his four emotion scales (Figure 5, page 64). Lennard and Matarazzo (Figures 7 and 9, pages 67

and 70, respectively) also count subcategory frequencies. Their counting procedures are essentially the same as those used by Strupp, Mahl, or Gottschalk. Leary summarizes interview data over ten-minute intervals. He orders the totals according to a more complex scheme. Summaries for certain of his data subcategories (emotional interactions) are displayed in one arrangement--in octant sectors of the Kaiser interpersonal circle (Leary & Gill, 1959, p. 79); totals of other subcategories (intellectual interactions) are displayed in a table divided into positive, i.e., nondefensive, and negative sections.

Pattern analysis. This class of mathematical techniques directly applied to nominal, state process data sequences seeks to identify or evaluate the patterns contained in these data. In effect, the process basically amounts to defining a new set of state descriptors--a new alphabet is derived by identifying specific patterns or sequences of states at one level with specific states at the subsequent level. Used in this fashion pattern analysis involves both a redefinition of states as well as counting; that is, frequencies of the new states (particular sequences of old states) are tallied. (Pattern analysis can also be a secondary operation, as will be considered below when Butler's system is discussed.)

An example of this approach is furnished by Matarazzo; it will be recalled that in his system (Figure 9, page 70) observers record only whether in any instant one, the other, both, or neither dyad member is "acting." The interview state at any instant therefore can be formally specified by one of four elements of a state representation or alphabet

set: State 1, both inactive; State 2, dyad member A (patient) is active; State 3, dyad member B (therapist) is active; and State 4, both members are active. Various patterns thought to be significant are defined a priori by the researcher in terms of a specific succession or sequence of states. The interviewer is trained to initiate certain patterns in preselected portions of an interview. For example, one pattern is: State 1 (silence) followed by three seconds of State 2 (patient is acting); therapist then speaks, thus changing the system into State 4 (both participants acting); if this state is maintained for about five seconds by the patient, therapist ceases acting, and the system returns to State 2. This particular pattern is called "talking the interviewer down," and is counted and used as an index of "patient's dominance." Similarly, other patterns are defined and used as indexes of other interactional characteristics.

Transitional probability. Several observational systems analyzed their data directly in terms of transitional probabilities, and a mathematically sophisticated example is provided by Jaffe (1968). His system, like Matarazzo's, uses a state representation set with only four states. These allow 16 first order transitional probabilities to be defined. Various types of stochastic models are used to obtain theoretical values for these probabilities. The models were developed according to different kinds of logical assumptions, as indicated by some model names: "independent decision," "imperfect politeness," "single source." Strings of data were generated by a computer according to these different models. The generated data as well as experimental data obtained from a 60-minute interview yielded frequency distributions for durations of a given state.

Auld and Dollard (Figure 4, page 63)) also have investigated first order transitional probabilities, exploring various hypotheses about the transitional patterning of states (Auld & Dollard, 1966). For example, they hypothesized that a given state tends to persist; the hypothesis was supported experimentally. Other hypotheses concern the patterns of silences, of resistant behavior, etc.

Pair comparisons. The analysis of the Menninger project (Figure 13, page 77) showed that the observational processes carried data only to the stage of a free descriptive narrative. In several related studies at Menninger's, straightforward mathematical/formal procedures were used to formalize these data. One study (Wallerstein, 1966, p. 505) involved a pair comparison method in which clinicians evaluated groups of patients at three historic junctures in their treatment: intake, discharge, and follow-up. Comparisons were based on cumulative clinical data. Pairs of patients were compared on each patient variable suitable for ordinal treatment. For each dimension and for each patient pair, clinicians rated one member as dominant (in the logical sense). This primary operation on relatively unconstrained clinical data thus assigned a numerical value--the number of judges rating A dominant to B along a given dimension--to each pair directly from free narrative data, without intermediate formal steps.

Rating scales. Yet another method of mathematizing data is direct scaling. Judges or raters are trained to assign a numerical value to nominal data such as clinical material--this direct assignment is made

subjectively. Two well-known examples are Rogers' process scale (Figure 10, page 71) (Rogers, 1959, 1961) and the Health-Sickness Rating Scale developed in the Menninger project (Wallerstein, 1966, p. 503). In Rogers' system, each judge assigns a scale value from 1 to 7 to randomly selected four-minute segments of an interview. In Menninger's study, a therapist evaluates global clinical data (interviews, tests, clinician's case notes) along seven dimensions, scaling each on a 100 point scale. Scores have also been obtained on occasion via an overall rating on a single 100 point scale.

Secondary Operations

The five direct operations just discussed represent the different ways in which numbers become associated with basically nominal data. Very frequently these initial numerical steps are followed by other numerical operations. Many of these mathematical techniques are quite familiar to the research community, and therefore require neither introduction nor explanation. These standard techniques of experimental psychology include statistical procedures--tests of significance, correlation, regression, factor and cluster analysis. They also include simple arithmetical processing of the data yielded by primary mathematical operations (examples are computations of percentiles or of simple summary indexes) as well as psychophysical scaling techniques such as those used to process pair comparison data.

An example of a secondary operation is furnished by Butler's quite complex approach; the mathematical properties of his process matrix (Figure 3, page 61) are used to obtain a measure of patterning, or in

his terminology, of the internal productivity of an observational category system. Butler constructs a composite matrix from responses of groups of subjects (patients or therapists) and uses the rank of that matrix to assess the internal productivity of a given observational scheme--the lower the rank (compared to the highest rank theoretically possible), the less productive the observational system (Butler et al., 1963, pp. 12-19, 76-85). Butler believes that regularities of the matrix which account for its reduced rank reflect a deficiency or inadequacy of the observational system; that is, identical or highly similar patterns between patients or between interview samples indicate that the chosen observational categories do not discriminate adequately.

An important assumption is concealed within the above secondary operation. Butler's original data matrix reflects a nominal classification scheme, that is, a category set or alphabet. He has chosen to represent that set in terms of zeros and ones, although any other analogous notation could serve equally well. His secondary operation, however, immediately treats his nominal symbols (zeros and ones) as real numbers in a matrix; this step and its logical assumptions or justification are not discussed by Butler. This issue will be reconsidered in Chapter V.

Two other, less familiar classes of secondary operations are time series analysis, and control theory techniques. Although neither has been used in conjunction with the 14 observational systems, both are beginning to appear in the psychotherapy literature. They are well-known

from their successes in science and technology (Bellman & Kalaba, 1965; Box & Jenkins, 1970), and they are probably on the verge of becoming popular and fashionable in psychology (Bellman & Smith, 1973; Gottman & Leiblum, 1974).

Time series analysis can be used to analyze any random scalar function of time (or, for that matter, of any other independent variable). Various versions exist; models of varying complexity have been developed to study different empirical classes of time series data. The technique uses a parametric statistical model to assess apparent changes in the behavior of random series. The technique is predicated on the assumption that the status of a system can be assessed validly from an available scalar measure (the dependent variable). Changes in the temporal characteristics of that variable then presumably reflect changes taking place within the system. To illustrate: Consider a patient (the system) and some criterial dependent measure of his health status (intensity of a physical symptom, for example). Using time series analysis fluctuations in the severity of his symptom are evaluated before, during, and after clinical intervention; parametric changes revealed by the analysis furnish evidence for or against effectiveness of treatment.

Several attempts have been made to apply control theory to problems of psychotherapy (Bellman & Smith, 1973; Carroll & Houghton, 1970; Houghton, 1968) although this approach has not yet been used to process actual psychotherapy research data. Typically, current studies have sought to identify promising mathematical models. For example, Bellman, a major figure in modern control theory, has proposed so-called

influence-response models of stochastic multistage decision processes ("influence" seems to be synonymous with "stimulus"). These models represent a system state (state of patient or therapist) as a column vector; the theory allows one to investigate how relationships between the state and other components of the model--influence, decision-making policies, and response--affect the overall behavior of the system, usually in terms of some performance criterion. Results of the analysis are used to develop optimum decision policies (for example, "best therapeutic responses") or to simulate psychotherapy interviews.

Mathematics and the Proposed "Complete System" of Observation

The preceding section has shown that in psychotherapy research the primary mathematization of data is typically achieved (1) by tallying--either frequencies of representational categories (simple counting) or frequencies of selected combinations of subcategories (pattern analysis, transitional probability), or (2) by acquiring numbers directly via judges' ratings (pair comparisons, rating scales). These observations concerning mathematical practices can be profitably viewed from the perspective of measurement theory. This theory provides a powerful general approach to analyzing and conceptually unifying diverse measurement procedures--any procedures used in science to generate numbers from observations. A full exposition of the somewhat complex theoretical framework is not possible here, but complete discussions are readily available in the literature (Coombs et al., 1970, Chapter 2; Suppes & Zinnes, 1963, especially pp. 4-22). The question addressed by this theory, then, is how to measure a set of real world objects or phenomena.

According to measurement theory there are three basic ways in which data are mathematized: fundamental scales, derived measurements, and pseudopointers. The first type of measurements, a fundamental scale, is constructed by associating a set of objects or phenomena and their empirical relations (these will be discussed later) with a numerical set and its relations. A derived measurement can be obtained for a set of objects that have already been measured by fundamental scales, that is, there already exist assignments from the objects to a set of numbers; the derived scale then defines additional functional relationships between these existing functions and thus indirectly defines new functions from the set of objects to a set of numbers. Pseudo-pointers are devices or schemes which assign objects to be measured to a set of numbers without a supporting logical framework of the kind provided by either fundamental or derived measurements.

An example of a fundamental scale is the measurement of mass (Coombs et al., 1970, p. 12)--here the empirical relations between objects are based on empirical dominance ("is heavier than") and on concatenation ("the combined mass of two objects"); the corresponding numerical relations are logical dominance ("greater than") and addition. An example of derived measurement is density (Suppes & Zinnes, 1963, p. 18)--density is defined via a function on a set of objects which have already been measured with respect to mass and volume; the new function assigns a density to each member of the set of objects by specifying division of mass by volume. Pseudopointers are the measurements most commonly employed in psychology: numbers are associated with

objects or phenomena without reference to an underlying mathematical logic; this describes the procedure by which judges score events or by which objective tests "produce" numbers, and it also describes counting operations.

Returning now to the five classes of primary measurements employed in psychotherapy research, one notes that four of these are pseudopointers, as they are based on ratings and analogous procedures, or on counting. Only one of the five direct measurement procedures previously discussed is not a pseudopointer measurement: Pair comparison methods do not produce fundamental measurements, as explained in detail by Coombs et al. (1970, p. 19), because they formally identify empirical relations such as "heavier than" with axiomatic, logical relations. This last class of measurements, however, is infrequently used in psychotherapy research.

As has been noted, pseudopointers lack a mathematical foundation. They thus represent a logically weak class of primary measurements. (That is not to say they may not be useful; that question, however, is an empirical, and not a formal/logical concern.) Pseudopointers are often used; they are convenient, and the alternatives (fundamental and derived measurements), although well-known in psychophysics, have received little attention in psychotherapy research. It appears that the very basis of mathematizing data--the first step from nominal to formally mathematical representation--has been virtually ignored in psychotherapy research. Apparently in that field the focus of mathematically minded researchers has been on what have here been called

secondary operations: Researchers usually use sophisticated mathematical tools and techniques which can be applied to already quantified or otherwise mathematically formalized data. As discussed above, two major advanced mathematical techniques are being used as secondary operations--time series analysis and modern control theory. As secondary operations they bypass initial problems of quantification: they assume experimental phenomena can be, and have been, appropriately represented by scalars, vectors, or tensors. For this reason such analyses typically ignore the steps preceding the application of more interesting and sophisticated mathematical techniques. An excellent illustration is provided in an excerpt from a book review (Menge, 1974), although the book concerns another branch of psychology (pattern recognition):

> Andrews makes almost no reference to the origin of his objects. He assumes that they are transduced into a vector of finite dimension R whose components are real numbers. Most of the book is concerned with techniques based on the assumption that the transduced objects (which are points in R dimensional space) form clusters which constitute the different classes [p. 244].

If the preceding line of argument is valid, it follows that primary operations merit greater attention. Specifically, the two alternatives to pseudopointer measurements should be investigated. Since one of these, derived scales, however, is based on fundamental scales, one is led back directly to a consideration of this latter category of measurement. As explained in Suppes and Zinnes (1963, pp. 4-17), fundamental measurements involve two relational systems, an empirical and a numerical relational system. Each system consists of a set and relations on that

set. The meaning of relations, briefly, is as follows. First, a cartesian product of n sets A, B, . . ., P, Q (some or all may be the same set) is the set of all ordered n-tuples of elements of these sets. Then, a relation between the n sets is a subset of the cartesian product; it may be defined by enumeration or axiomatically.

In empirical relational systems the set consists of real world research "objects" or phenomena and the relations arise empirically (such as, "is heavier than," "is as valuable as," etc.), while in numerical relational systems usually the set consists of numbers and the relations are axiomatically defined. A fundamental scale is composed of an empirical and a numerical relational system as well as of an assignment from the empirical set to the numerical set.

Different axioms in numerical systems produce different scales, and it is the task of the scientist to select a system which fits and appropriately reflects the empirical relations obtaining between real world events. What axiomatic choices might be appropriate for measuring psychotherapy research data? According to the discussions and conclusions of Chapter II there are difficulties: one does not at present know what constitutes an adequate empirical (state representation) set of "objects"--research phenomena in psychotherapy--and one therefore certainly does not know what relations exist on such a set. It follows that imposing any but the most primitive mathematical structure on data must be premature.

The simplest step from nominal data to a mathematical system is to what Coombs, Raiffa, and Thrall (1954) call a relation scale--a scale in

which the numerical relations are essentially unconstrained by axioms. This scale does not seem to have been used in psychology, probably because its lack of structure provides so little support and guidance for formal mathematical operations. Nevertheless, it does seem to be appropriate to the current status of psychotherapy research and it is the scale recommended here. All it presupposes is that empirical objects can be assigned to mathematical objects and that some relations hold on the mathematical set.

What are the consequences of this recommendation? Chapter II has raised questions about the configuration of data elements: the issue of appropriate state representations of research data in psychotherapy is still open--the traditional vector formalism may be inadequate. Furthermore the discussion of the present chapter suggests that the logical structure of empirical relational systems in psychotherapy research is not yet apparent: one also still seeks appropriate logical relations on research data. It therefore seems reasonable to suggest that these two investigations--the search for representation as well as for logical structure--should go hand in hand. As appropriate systems of observation are developed, the emerging empirical structure should lead to a mathematical formalism properly grounded in empirical relations exhibited by data. There are no known routine mechanical methods by which these possibly quite novel types of relations could automatically be discovered in data, just as there are no automatic ways of identifying appropriate state representations in psychotherapy.

At the present time there is no substitute for a researcher's imaginative, inventive, and laborious exploration of his data. It is unfortunate that the recommended mathematical formalism needs to be so general and rudimentary at this time; perhaps this approach and its implications will be further clarified through the discussion of proposed observational systems to be presented in the next chapter, as well as through the discussion of directions of future research presented in Chapter V.

CHAPTER IV

DEVELOPMENT OF A SYSTEM

The analysis of Chapter II has shown that observation viewed as a formal process involves classification of events into a chosen set of categories variously known as an alphabet, representation set, or category set. As a classification process, observation shares the general logical features of any process of nomenclature: (1) differentiation--how something is apprehended; (2) standardization--how it is measured; and (3) pragmatics--why one wishes to classify it (Panzetta, 1974). In terms of observational systems used in psychotherapy research, differentiation implies an orientation or position that defines events constituting the process of therapy; standardization denotes the formal aspects of an observational system, that is, representation and mathematization; and pragmatics refers to the domain and rationale of application--how, and in what ways, a system would be useful. Although these three areas of concern should be considered during the formal development of a system of observation, psychotherapy research seldom uses such a formal approach; methodological issues of observation are rarely addressed explicitly. This omission was discussed in a session of the American Psychological Association's Second Conference on Research in Psychotherapy. During the session specifically devoted to "definition, measurement, and analysis of significant variables in psychotherapy," the following comments were made:

Our area of concern in this session is threefold:

1. Which variables in psychotherapy are significant?
2. How can they be defined rigorously?
3. How can they be measured adequately?

> To start with, let us set up the proper hierarchy of values. Needless to say, the most important of the three questions is the first--the selection of significant variables. Yet I am afraid that too often psychologists tackle the latter two problems first--particularly the problem of the adequacy of measurement. They concern themselves to such an extent with measurement that they show insufficient regard for the importance or nonimportance of the variable measured [Shakow, 1962, p. 237].

> It is often the aim in process studies to examine in a microscopic way what is going on in therapy with the aim of deriving variables for further study. Actually, we rarely do this. . . . One reason we don't do it, I think, is that it is very hard to get it published. It just does not fit the model of the usual journal article. You are supposed to end up with conclusions about the way in which a variable operates rather than simply saying on the basis of a lot of qualitative observations, "I propose that this is an important variable; it is as far as I have gotten so far in my research, but it is still worth publishing on the basis of findings to date [Levinson, 1962, pp. 297-298].

The present attempt to develop a rationally-based system of observation is necessarily hampered by the paucity of previous relevant research. The system will be incomplete at best; some suggestions for future improvements and modifications will be explored following the presentation of the proposed system.

A First Approach

Pragmatic Considerations

A pragmatic approach to observational systems rejects the ideal of a universal and theoretically neutral system of observation, thereby supporting the position that observation is always implicitly or

explicitly biased by an observer's approach (Panzetta, 1974; Weimer, 1973; and Yankelovich & Barrett, 1971). Different orientations to research thus imply different structurings of phenomena; it, therefore, seems reasonable to assume that different types of research may be served best by different systems of observation. These could be matched to the research in terms of perspective, terminology, operations, and the like. For example, research performed by different therapists using different methods of psychotherapy or by theoreticians of various different schools of thought might give rise to correspondingly different families of observation systems. A system would be tailored for a particular research orientation, reflecting an approach to differentiating phenomena. In general, from this point of view one would expect that a clinician's system of observation would categorize events in clinically meaningful terms, at least at the initial levels of data manipulations, while theoreticians might categorize in more experimental terms reflecting their interest in physiological, biochemical, or micromomentary behavior measurements. In the present work a single system will be constructed as an example of the present approach. The principles of construction used should generally apply to all systems development; construction of other systems using the recommended procedures could be carried out in an entirely analogous manner.

Representation

The task of designing a system of representation is a formidable one: in spite of years of research in psychotherapy there is "no generally accepted or thoroughly investigated system" (Luborsky & Strupp,

1962, p. 309) of observation--many researchers have developed their own, usually short-lived systems. These are commonly quite limited in scope, as illustrated by the 14 representative systems analyzed in Chapter II. Typically they use a relatively small number of dimensions and a single size of scoring unit. It has been realized for some time, however, that it would be desirable to use systems of observation capable of representing psychotherapy events more completely. A major impediment has been the intimidating volume of data in the absence of an effective methodology (Bordin, 1959, p. 236; Ekman & Friesen, 1968, p. 179; Lennard, 1962, p. 237; Luborsky & Strupp, 1962, p. 315; Shakow, 1959, p. 110). It has been proposed in the present work that there is an alternative approach to the traditional answer, circumventing the problem via standard techniques of simplification: Chapter II concluded with a set of very general criteria for a proposed approach. These design considerations must be made more specific at this point.

To begin with, it has repeatedly been stated in the literature that modalities (visual and auditory, verbal and nonverbal) all ought to be represented in a complete system. The main reason for advocating acquisition of multimodal data is that different modes are often nonredundant. Bordin (1968, p. 418), Deutsch (1966, p. 167), Duncan, Rice and Butler (1968, p. 566), Ekman and Friesen (1968, pp. 180-192), Kiesler (Hart & Tomlinson, 1966, p. 248), and Shapiro (1968, p. 237) are among those who strongly believe that different modalities emphasize, contradict or otherwise modify one another, generally providing independent information.

Second, the observational record should monitor data from each participant continuously (Luborsky & Strupp, 1962, p. 319; Mahl, 1968,

p. 301; Parloff, 1962, p. 257; Parloff & Rubinstein, 1959, p. 282). In observational systems based on traditional chaining models the observer's attention usually has followed the verbally active member of the dyad and thus missed any nonverbal behavior that may have been exhibited by the other, silent member. Continuous observation would avoid this loss of information.

Third, if possible, the system should make appropriate use of specifically interactional terms to define, describe, and classify interactional phenomena. At this time, however, appropriate interactional concepts and descriptors are virtually nonexistent; the proposed system therefore can meet this design requirement only very incompletely.

The fourth design area concerns unit size. The need to represent phenomena as completely as possible suggests that observation involves different size units. This approach would enable a researcher to analyze a given modality from several different perspectives. A range of unit sizes implies multiple level representations of a given modality. Each level would be segmented according to differently specified units, and could use entirely different approaches to classifying the modality. For example, the verbal sector could be multiply analyzed at different levels by means of systems designed to categorize clauses, sentences, paragraphs, or still larger and/or smaller units.

The question of using idiosyncratic rather than general representational sets is another design issue. Traditionally systems of observation are intended to be general purpose systems in the sense

that they are to be used without modification to classify and interview material. Most content analysis systems (such as those cited in Chapter II) share this design feature. Researchers have pointed out, however, that idiosyncratic behavior in therapy may be very important (Deutsch, 1966, p. 168; Mahl, 1968, p. 329; Shakow, 1959, p. 111; Shlien, 1962, p. 153). It therefore seems advisable to include a capacity for representing idiosyncratic data in the proposed system.

The following material will be a first attempt to construct an observational system according to the design philosophy developed above. The system is to serve as an illustrative example of the advocated approach. It will inevitably display shortcomings of a first try based on very incomplete theoretical and empirical foundations. Nevertheless, it should provide a starting point for future refinements and improvements. Its general features can be summarized as follows. The total system will be composed of major sectors, each subdivided into subsectors or sublevels; each sublevel will represent phenomena from the perspective of a particular unit size or scale. For each sublevel a descriptive subsystem of categorization will be chosen from systems already available in the literature. If at some levels no suitable existing systems can be found, simple, provisional systems will be suggested. These intuitive and untested provisional subsystems may serve as temporary stopgaps until better alternatives become available. For the present, it will be assumed that videotapes (or sound films) and verbal transcripts of interviews are available.

The proposed system has three prime descriptive foci: patient, therapist, and interactional behaviors. The next eight figures will

depict various aspects of the system. The three major divisions of the representation are shown in Figure 16; in analogy to a musical score, the principal divisions or sectors are arranged vertically as three major "orchestral sections" in temporal registration. Next, two of the sections, the representations of patient and therapist behaviors, are each divided into verbal and nonverbal levels; the nonverbal level is further separated into a paralinguistic and a kinesic (body movement and gestures) subsector. Figure 17 shows these major subdivisions of the patient's or therapist's sector.

Concerning interactional representations it will be recalled that preceding discussions pointed out that virtually no genuinely interactional conceptual framework is currently available. This problem area is discussed later in this section; a diagram of the interactional sector's constituents will be presented then.

Up to this point it has been possible to describe the system in very general terms. One must now begin to identify specific category descriptor sets. The system to be developed is to be a therapist's observational system. Since it has been suggested that pragmatic considerations imply a need for matching a system to the theoretical position of a therapist, it is necessary to select a particular clinical orientation as a basis for developing a specific system. As has been previously noted, a selected orientation will influence a system's microstructure, but the overall characteristics of any system designed according to the proposed approach will remain substantially the same. The demonstration system will be based on a psychodynamic theoretical

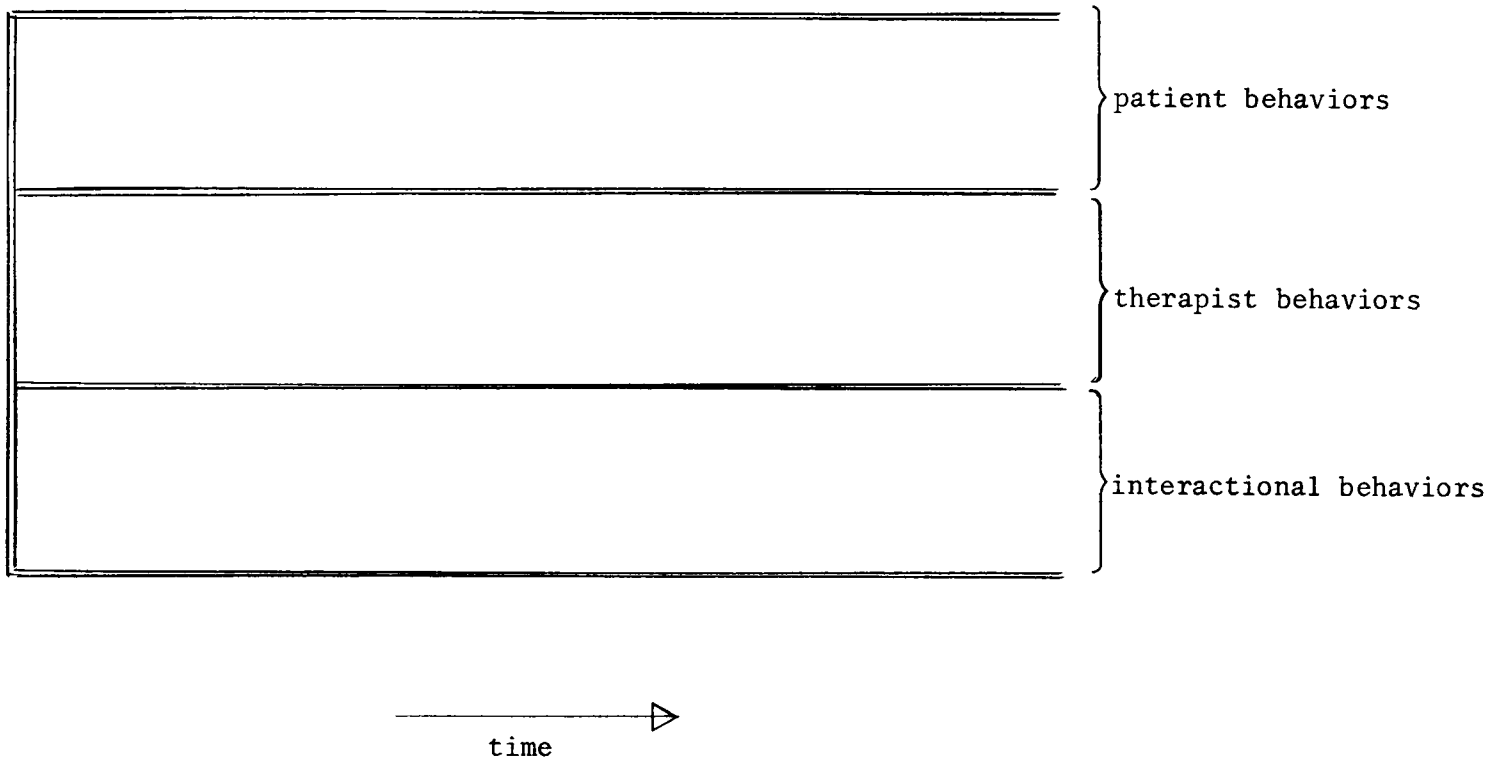


Figure 16
Principal Levels of the Proposed System of Observation

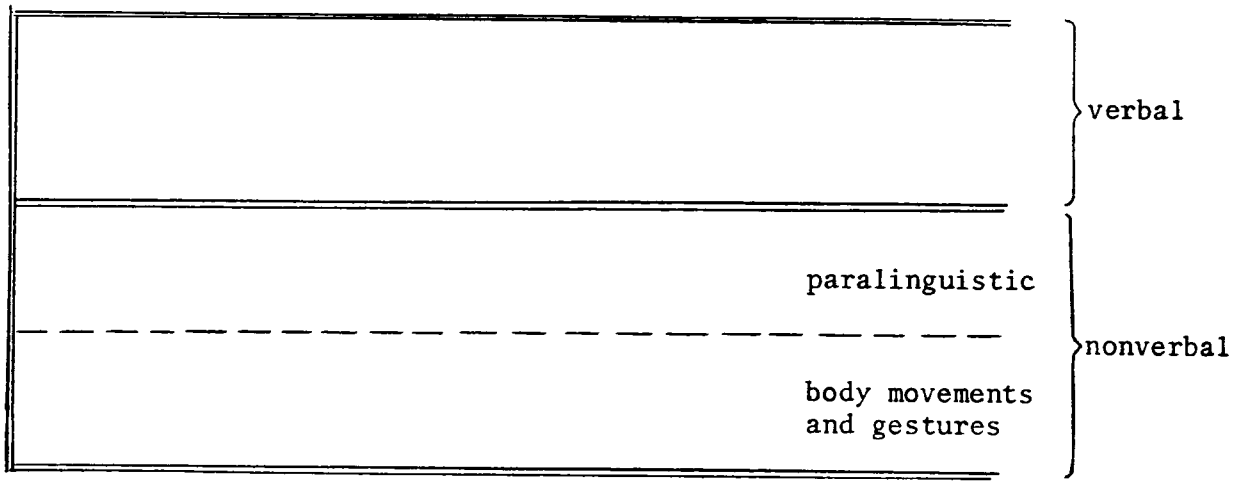


Figure 17
Major Subdivisions of Patient or Therapist Behavior Levels

orientation; modifications required by other orientations will be discussed later.

The patient's system. This major subsystem contains verbal, paralinguistic, and kinesic sectors as shown in Figure 17. The first modality to be considered is verbal behavior. The most frequently used traditional approach to analyzing that behavior is through a content analysis scheme using a fixed size scoring unit; typically, the unit is a relatively short utterance--word, phrase, or sentence. This approach implies first that each utterance can be uniquely and unambiguously assigned to one category and second that "content" can be adequately defined in terms of a single, fixed, preselected scoring unit. The first assumption fails to account for the multiple levels of meaning that a unit may have. The usual remedy is to augment the dimensionality of the analysis system. The present system will follow this approach, selecting high dimensional classification systems to be used at a given level; in addition it will analyze material at several levels in the hope that this approach will help capture additioanl dimensions of meaning. The second characteristic of traditional systems, using single sized scoring units, may result in a very incomplete analysis; small unit systems may miss larger thematic trends, while large unit systems may fail to capture highly significant details. The suggested multilevel approach should help with this problem as well. The use of several systems each considering a different scale of phenomena can provide fine, intermediate, and coarse grained analyses. Thus, both considerations of traditional system assumptions support the choice of a multilevel, diverse unit system to represent verbal behavior.

There are numerous available content analysis systems (Marsden, 1966, Chapter 8) to choose from. The most prominent systems applied to short utterances (having scoring units on the order of a sentence length) appear to be Butler's, Auld and Dollard's, Leary's, and Lennard's (Figures 3, 4, 6, and 7, pages 61, 63, 65, and 67, respectively). Leary's is the most comprehensive but it has been avoided by researchers because of its highly complex and unwieldy design (Marsden, 1966, p. 355; Millon & Diesenhaus, 1972, p. 100). Lennard's and Butler's systems are both based on relatively restrictive assumptions; Auld and Dollard's system is very highly regarded (Marsden, 1965, 1971; also see general comment, Figure 4, page 63) and for this reason appears to be a suitable choice for categorizing patient's verbal behaviors at the level of the sentence. One minor modification is required; the system includes three paralinguistic dimensions (sighing, weeping, and laughter) which should be excluded from purely verbal sectors of representation. The first level of the proposed verbal representation, then, will use the remaining 17 mutually exclusive categories of Auld and Dollard's system to categorize the patient's sentences at the first level.

The next larger verbal unit can be chosen in many different ways. One "natural" division could be a topical theme. It is provisionally assumed that an interview is capable of being segmented into thematic episodes (Pittenger et al., 1960, p. 135b, and Sheflen, 1973). A useful category scheme for classifying thematic content might be the "subject of statements," a subset of Leary's classification set. The five categories in his dimensions are (1) persons and animals (further

subcategorized as shown in Figure 6, page 65), and (2) psychological, (3) somatic, (4) sociological, and (5) inanimate subjects or topics. Because of its thematic focus it is suggested that this category set be used to classify data at the second or thematic level of a patient's verbal behavior.

It is proposed to add a second subsector at this thematic level for the following reasons. Deutsch (1966, p. 168), Mahl (1968, p. 329), Shakow (1959, p. 111), and Shlien (1962, p. 153) believe that in general a very important part of a person's behavior is idiosyncratic, not readily understandable or scoreable in terms of fixed, predetermined, general descriptors. To allow for representation of idiosyncratic material, it is proposed to use mixed scoring systems throughout at appropriate levels of analysis, an approach already used in the Menninger project. Thus, part of a representation would use a traditional, fixed category set, while a flexible, individually tailored second system may be added to capture idiosyncratic data. At the second level, it seems appropriate to use a mixed system to represent thematic verbal behavior at that level, supplementing Leary's fixed category systems, since a patient's themes arelikely to display important idiosyncratic features. To develop a representation set for this idiosyncratic component a researcher would need to identify a patient's individual themes, thus defining a small set of categories specifically for each patient. Using this mixed system, each thematic episode would be categorized along two different sets of dimensions: Leary's five fixed subcategories and the idiosyncratically developed representational set. It

will be noted that since the idiosyncratic components can be selected and categorized by individual researchers, the type of representational system proposed in the present work can provide an additional dimension of flexibility.

A third level can be used to categorize the dominant theme of the session considered as a single scoring unit. The category system proposed for the second level could be used also to characterize thematic behavior at the third level. The verbal sector of the proposed patient's coding system is shown schematically in Figure 18. This figure exhibits the three subdivisions corresponding to the three kinds of scoring units--sentences, thematic episodes, and single dominant theme of an interview--together with the other sublevels just discussed.

The other major sector of the patient's "score" is the nonverbal section, divided into paralinguistic and kinesic subsections. Paralinguistic coding will be considered first. Briefly, this term refers to the nonverbal aspects of speech, identifying musical components such as spectral register, loudness, tempo, and quality, as well as to nonmusical, noisy aspects of speech. The usefulness of paralinguistic information has been demonstrated in the literature (Boomer, 1963; Duncan et al., 1968; Goldman-Eisler, 1968; Pittenger et al., 1960), and a "well articulated taxonomy" (Duncan et al., 1968, 566) has been developed. One would anticipate that it would be difficult for most researchers in psychotherapy to use this taxonomy, as it requires a linguist's skills and training. A simplified system has been developed which apparently can be learned by nonspecialists, although not without

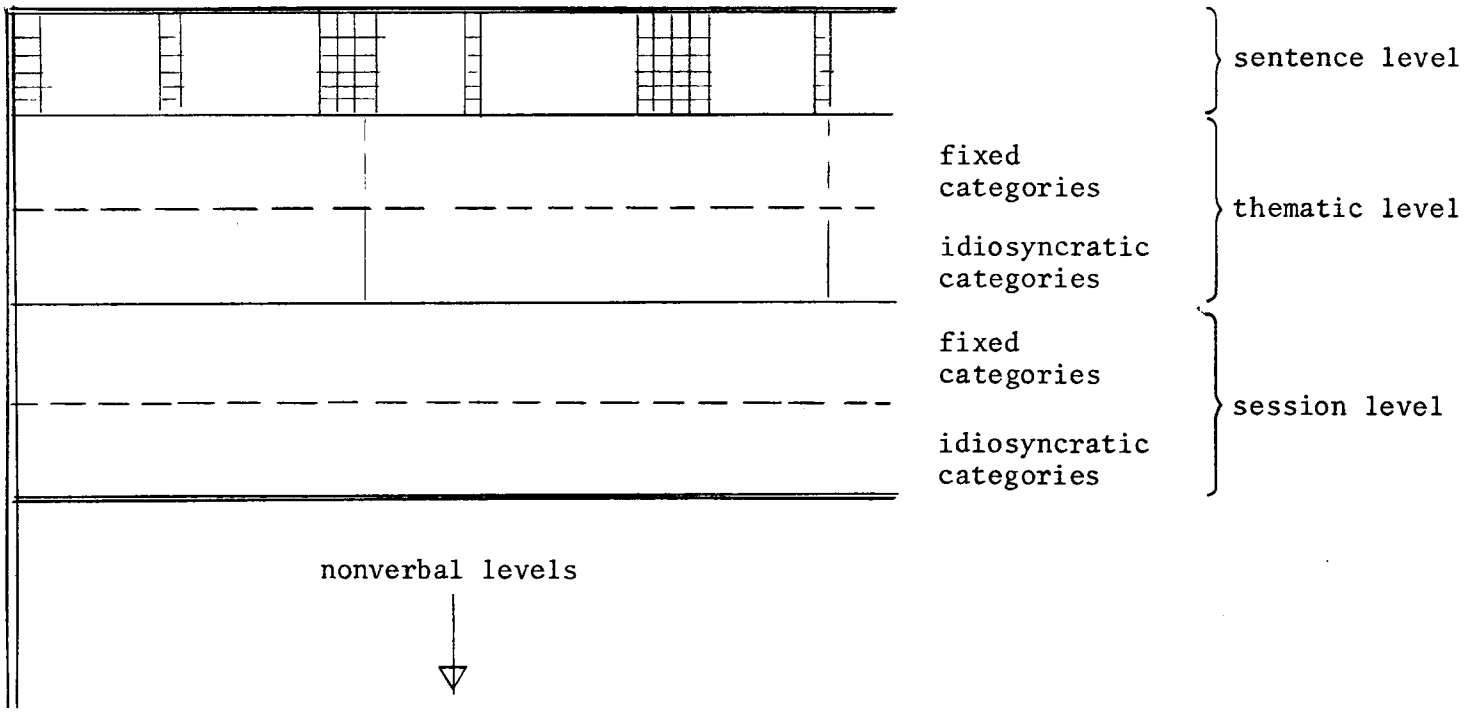


Figure 18
The Patient's System--Verbal Levels

some effort; this system, however, is not particularly applicable to coding patient's paralinguistic behaviors since it was developed specifically for classifying therapist's paralanguage (Duncan et al., 1968). The system will be considered further during the discussion of the therapist's coding system. It thus appears that no simple taxonomy is available for describing patient paralanguage. Until such a system becomes available one could perhaps use a rather crude category set containing the three previously mentioned patient paralinguistic category developed by Auld and Dollard (laughter, sighing, and weeping) and supplement these by three categories of relatively easily scored speech nonfluencies (filled, unfilled pauses; stuttering or repeating) used by Duncan et al. (1968). Thus, a stopgap system of six dimensions is proposed. Should a linguist be available, the much more elaborate system described in Pittenger et al. (1960) could of course be used instead. The proposed crude provisional paralinguistic system is probably applied most appropriately at a fairly coarse level of analysis, that is, at a scoring unit level of complete utterance--a patient's statement bounded by the therapist's responses.

Larger paralinguistic units analogous to the thematic level in the verbal domain probably could be constructed but as they are not considered in the psychotherapy literature they will not be developed further here. The paralinguistic segment of the patient's representational system thus must be left at a very primitive, single-level, incomplete stage of development at this time; this system is diagrammed in Figure 19. Contributions by knowledgeable linguists would be

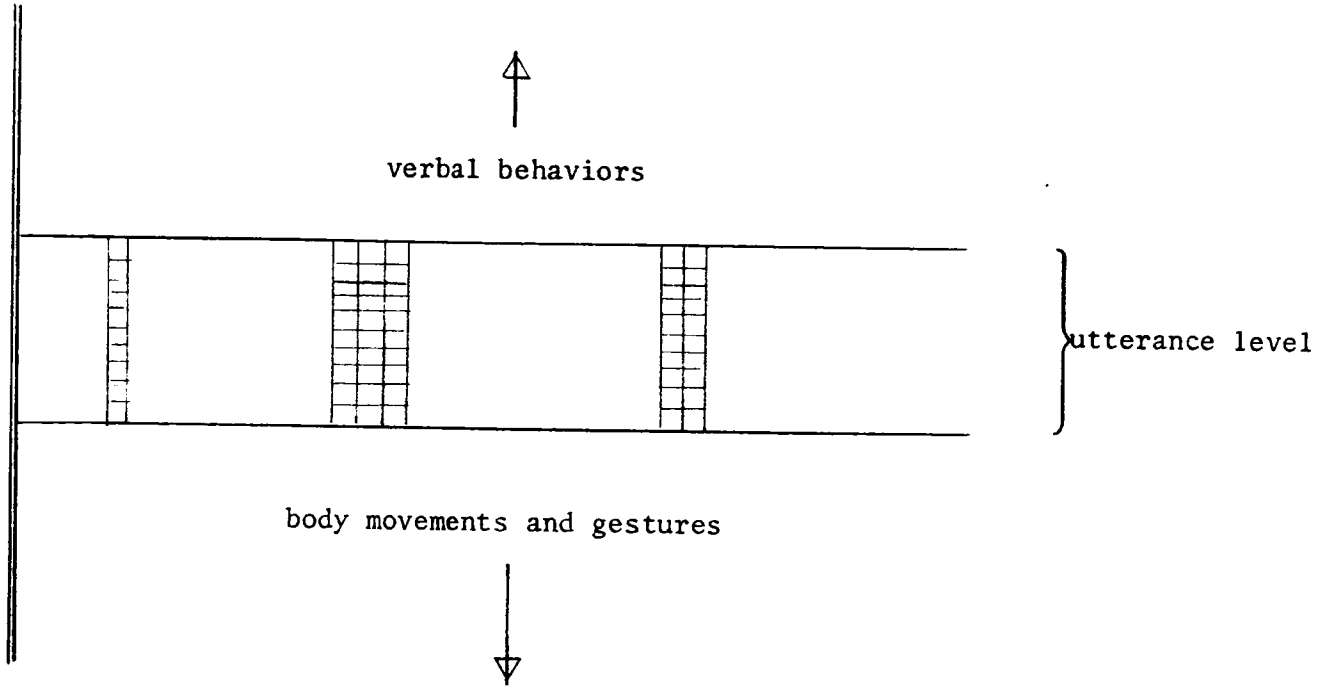


Figure 19
The Patient's System--Paralinguistic Levels

required to improve this subsystem, refining it while retaining characteristics allowing its broad use.

The remaining sector of the system for observing a patient's nonverbal behavior is the section representing body gestures and movement. According to Mahl, the kinesic modality may amplify, anticipate, contradict, or generalize concomitant verbal behavior (Ekman & Friesen, 1968, p. 185). He has, however, also pointed out (Mahl, 1968, p. 298) that the absence of a standard notational system and of convenient, inexpensive, unobtrusive data acquisition methods have held back work in kinesics; in comparison, verbal analysis has flourished, aided by availability of a standard notation and a highly satisfactory technology. For the researcher, the situation in kinesics is analogous to the problem of acquiring paralinguistic data. Most therapists are unable (and/or unwilling) to use complex notational systems such as those developed by Birdwhistell, Ekman, or Condon and Ogston, unless the services of a specialist in kinesics are available. Therefore once more a simplified stopgap system is needed. One apparently useful, practical, but rough system is Mahl's (Figure 8, page 68). In that system the three major categories are (1) general postural changes, (2) communicational gestures, and (3) autistic gestures (Mahl, 1968, p. 301). These are further subdivided as listed in Figure 8, yielding a total of 15 categories. The scoring unit for these categories is naturalistically defined by the observed boundaries framing individual kinesic events.

Mahl's system will be used to categorize kinesic data; in addition, however, it is proposed to use a level representing idiosyncratic kinesic behaviors, a representation that would be analogous to the previously

discussed record of idiosyncratic thematic verbal behavior. Mahl and other psychotherapy researchers interested in body movements and gestures note that kinesic behavior is only partially conventionalized, often including highly idiosyncratic components. It therefore seems advisable for the sake of completeness to include a provision for representing idiosyncratic kinesic data. This approach parallels that presented in an earlier section on verbal behavior. Thus, Mahl's 15 fixed categories would be supplemented by a second category system developed and defined by a researcher for each patient.

At this level of representation, the units of kinesic behavior are of relatively brief duration, perhaps lasting from tens of seconds to several minutes. Larger episodes could be envisioned, but once more the literature fails to provide a ready-made system for observation. Consequently, one simple ad hoc suggestion is to describe larger kinesic episodes in terms of the degree of general, overall physical level of activity. Perhaps a five-point scale defining a spectrum from immobility to agitated movement would be a suitable first approach. The unit would be naturalistically defined by intervals in which activity remained at the same level. Once again, a researcher would be free to develop his own alternative system.

Figure 20 shows the suggested kinesic notational segments. On the first level relatively brief events are represented by Mahl's system supplemented by an idiosyncratic category system, while longer units defined by a patient's remaining generally at one level of overall activity are represented in a second subsector.

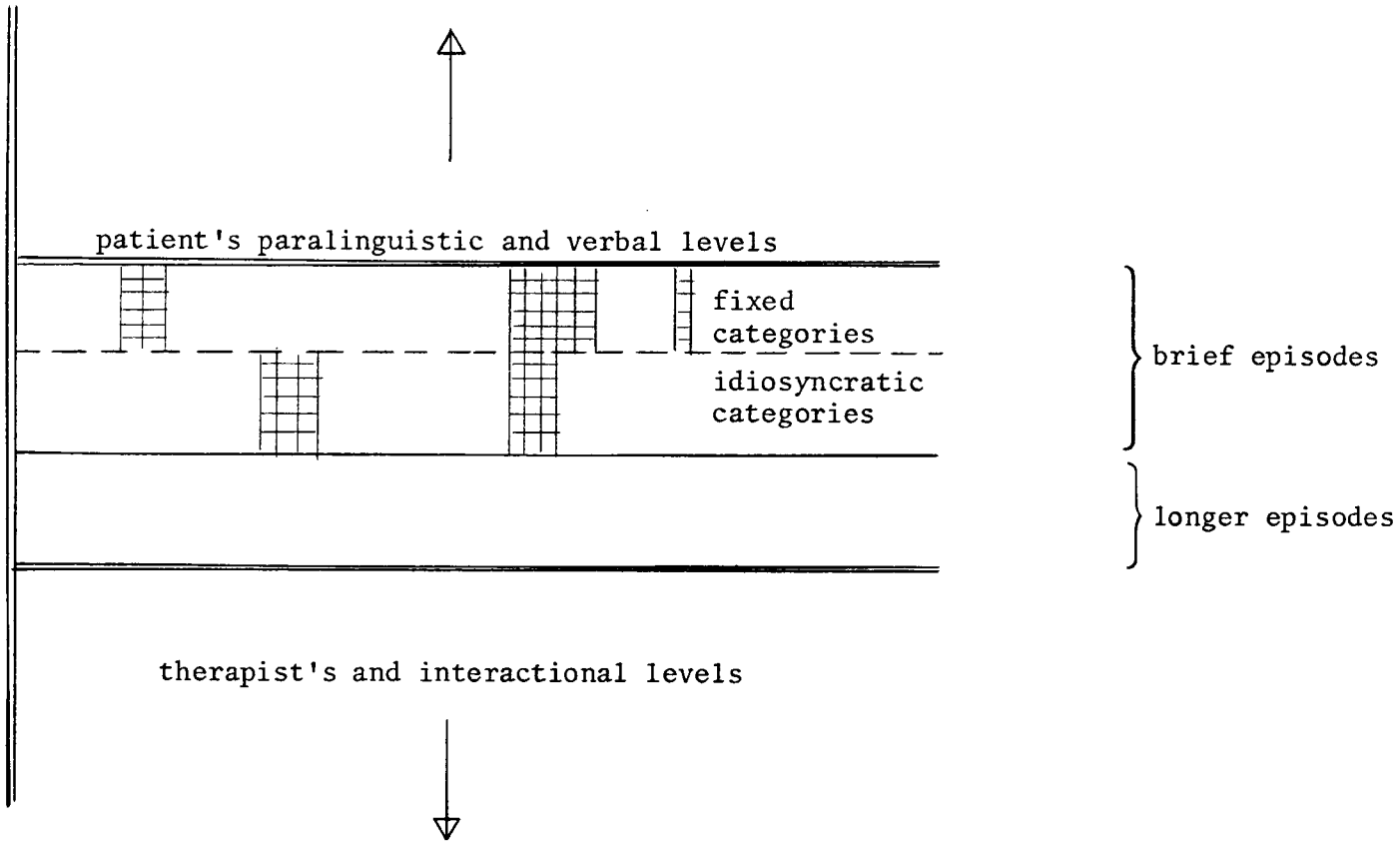


Figure 20
The Patient's System--Body Movements and Gestures

The three major levels--verbal, paralinguistic, and kinesic--shown in Figures 18, 19, and 20, pages 128, 130, and 133, respectively, complete the patient's system of representation. His total system is shown in Figure 21, which simply assembles the preceding three figures into a final composite record.

The therapist's system. In principle the development of this system parallels the construction of the patient's system; several relatively minor differences, however, need be suggested. With respect to verbal data, Strupp's system (Figure 1, page 55) has been developed specifically as a content analysis system for classifying therapists' verbalizations. Its scoring unit is a complete therapist response, the utterance framed by patient verbalizations; it thus is a logical choice for serving as a counterpart to Auld and Dollard's scheme for categorizing patient sequences. Strupp's system is suggested as the category set to be used at the first level of verbal analysis.

It will be remembered that a subset of Leary's categories was suggested for use as a taxonomy for the patient's longer verbal unit, the thematic episode. For the therapist's larger verbal units, however, it is proposed to use Strupp's rather than Leary's system; the former system, focusing on major clinical dimensions of therapists' verbal behavior, seems preferable to the latter system which emphasizes topical subject matter. Strupp's system will therefore be used to classify the dominant therapist verbal behavior at the second level also. Concerning units at this level, it will be assumed that the naturalistically observed boundaries which defined the patient's thematic episode at his

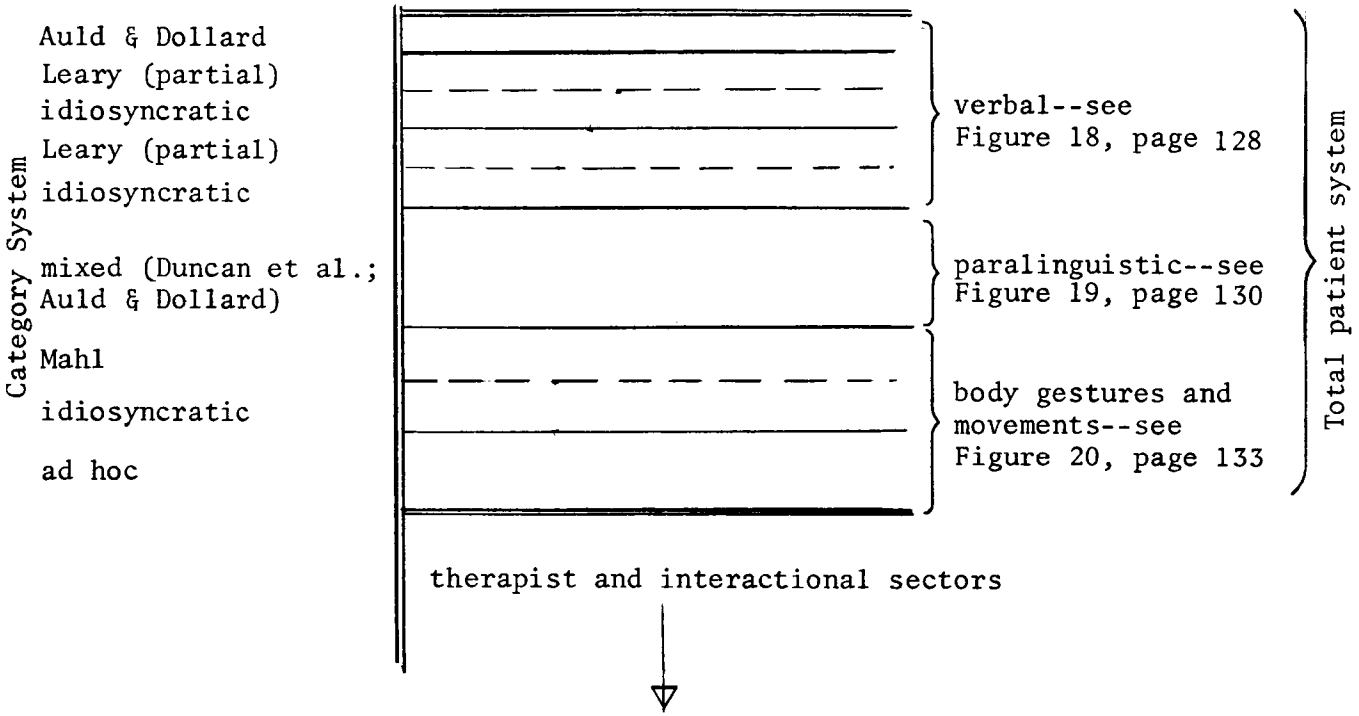


Figure 21
The Patient's System

second level of verbal behavior (see Figure 18, page 128) will also serve to define the therapist's unit at the second level.

In comparison to the patient's verbalizations, therapist's speech is likely to show less idiosyncratic content. Therapist utterances are likely to be heavily constrained by his role, orientation, training, etc., and therefore, no provision has been suggested for classifying such verbal behavior. A researcher may of course choose to add such a sublevel following the outlines presented in the patient's subsystem.

Leary's subsystem has been used at the first and second levels of therapist's verbal behavior, and it can be used also to characterize that behavior at a third level, the complete session. The three levels suggested for classifying a therapist's verbal behavior correspond to the three levels developed for the patient which were shown in Figure 18.

Turning to the next paralinguistic segment one finds an already available system for coding the therapist's paralinguistic behaviors (Duncan et al., 1968). The linguist's specialized, elaborate paralinguistic techniques used successfully by Pittenger et al. (1960) were considerably simplified by Duncan et al. for this more limited application. Nevertheless, it still does require some effort to learn the system. Their category set does provide a compromise between the requirement of a full paralinguistic transcription and omission of this class of data, and is therefore suggested for use in the present system. It is quite elaborate and not amenable to simple description, although it is not difficult to understand; researchers who plan to use it should refer to the cited literature.

Very briefly, the system's scoring unit is the phonemic clause, an utterance defined as having "one and only one primary stress and one and only one terminal juncture" (Duncan et al., 1968, p. 567). In Duncan's work, a clause was characterized (1) by intensity and pitch at the main stress, and (2) by four paralinguistic characteristics (intensity, pitch, openness of vocal cords, and pause nonfluencies) displayed by the remaining portion of the clause. On the basis of this scheme, seven clause categories were created. In addition to the scoring unit, a summarizing unit consisting of a complete therapist's response was defined. The paralinguistic scheme was found to be meaningful and provided a reasonable index for distinguishing good from poor therapy hours.

Kinesic categories are considered next. One may assume as a first approximation that in general the characteristics of therapist kinesic behaviors are similar to those of patients'. It is therefore suggested that the third major sector of the therapist's system, the kinesic sector, be coded with the same representation set used to classify a patient's kinesic behaviors (Figure 20, page 133).

The verbal, paralinguistic, and kinesic sectors of the therapist's system look very much like the corresponding patient sectors, and for this reason no separate figures were prepared for these three subsectors. The complete therapist's system is shown in Figure 22 which is directly analogous to Figure 21, the complete patient's system.

Interactional systems. A truly interactional system would described aspects of phenomena in a manner transcending the composite picture

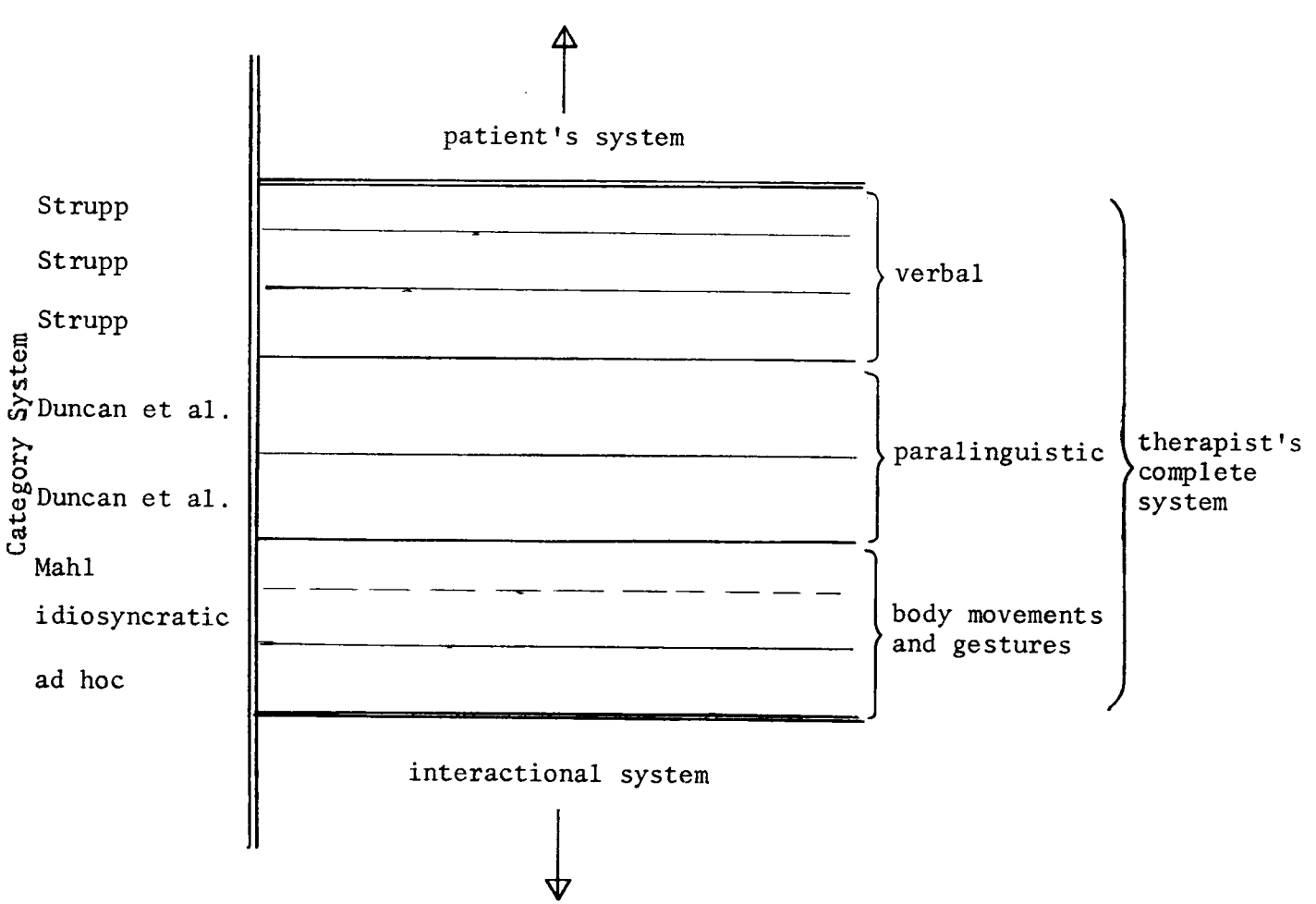


Figure 21
The Therapist's System

obtained from adding each participant's contribution. For example, in a dance the observed phenomena are more than a simple sum of each dancer's individual movements (Langer, 1957, Chap. 1). These considerations apply equally to the behavior of therapist and patient considered as a single system. It has been stated previously, however, that the research discipline fails to provide suitable tools for conceptualizing interactional phenomena. Pittenger and his coworkers have come to the same conclusion:

> . . . but there are few ready-made terms that seem appropriate for describing the state of a two-person system as a whole (perhaps "in rapport" and a few others). Efforts to deal with such a system therefore tend to degenerate into an oscillating series of assertions about the participants . . . [Pittenger et al., 1960, pp. 223-224].

To develop a representation of psychotherapy interactions is clearly a major conceptual task quite beyond the scope of the present study. What is sought, however, is a set of concepts which identify and define interactional phenomena in a two-person system framework. Rapport, the example given in the preceding quotation, is illuminating; it is a relational phenomenon which cannot be explained or described by synthetically adding together contributions from each member of a group. It is irreducibly a phenomenon displayed by a social system as a whole. Since fully developed interpersonal systems of observation are unavailable and since development of original subsystems is beyond the scope of the present work only a very tentative suggestion is made here for yet another interim, provisional set of categories. Until a suitable interactional system becomes available, it is proposed to construct a crude system from portions of Chapple's (Matarazzo et al., 1959, p. 127) and

Bales' (Lennard, 1962, p. 222) "interactional" notions. In these two systems, the terms closest to the desired concepts are: (1) Dominance-Submission, (2) Interruption (Chapple); (3) Descriptive--asking or giving information, (4) Evaluative--asking or giving appraisals or value judgments, and (5) Prescriptive--asking or giving suggestions (Bales). In addition, as suggested by Pittenger et al., another suitable interactional term would be "rapport." The tentatively proposed category system then would be composed of these six descriptors. The suggested temporal unit for this level of analysis is once again naturalistically observed interval defined as the interval over which behavior remains in a single observational category. The simple proposed provisional system is shown diagrammatically in Figure 23. The six interactional concepts are primarily verbal, and obviously a great deal of interaction must also occur in nonverbal modalities. It is hoped that researchers will develop this sector more fully, creating appropriate formal interactional systems of observation and representation.

This completes the overview of the proposed system. The three major components shown in Figure 21, page 135, Figure 22, page 138, and Figure 23 are assembled into the full system as shown in Figure 16, page 122. A separate diagram of this full assembled structure is not provided. Instead, Table 8 lists the sublevels and the selected category sets and units.

Modification of the system. The observational system sketched in the preceding section is a modular system, and thus is flexible and readily modifiable. In principle, changes could be made in three ways:

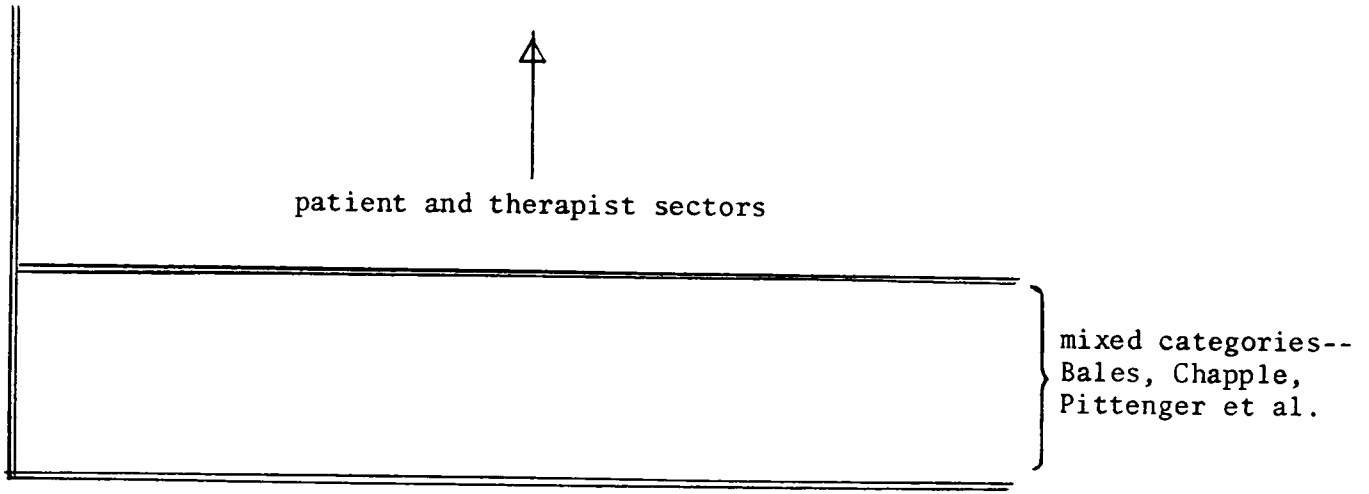


Figure 21
Interactional System

Table 8

Outline of the Assembled System

Level number	Focus	Scoring unit	Originator and reference
1.	Patient		
1.1	verbal		
1.11		utterance	Auld & Dollard; Fig. 4, page 63
1.12		thematic episode	
1.121			Leary; Fig. 6, page 65
1.122			idiosyncratic
1.13		interview	
1.131			Mahl, Fig. 8, page 68
1.132			idiosyncratic
1.2	paralinguistic	utterance	Duncan et al., 1968; Auld & Dollard (mixed system), Fig. 4
1.3	kinesic		
1.31		single gesture, movement, or position	
1.311			Mahl, Fig. 8, page 68
1.312			idiosyncratic
1.32		naturalistic (constant level)	ad hoc
2.	Therapist		
2.1	verbal		
2.11		utterance	Strupp; Fig. 1, page 55
2.12		thematic episode	Strupp; Fig. 1, page 55
2.13		interview	Strupp; Fig. 1, page 55

Table 8 (continued)

Level number	Focus	Scoring unit	Originator and reference
2.2	paralinguistic		
2.21		phonemic clause	Duncan et al., 1968
2.22		response	Duncan et al., 1968
2.3	kinesic		
2.31		single gesture, movement, or position	
2.311			Mahl, Fig. 8, page 68
2.312			idiosyncratic
2.32		naturalistic (constant level)	ad hoc.
3.	interactional		
3.1	primarily verbal (mixed at this time)		Bales; Chapple; Pittenger et al. (Lennard, 1962; References in Fig. 9, page 70; Pittenger et al., 1960, respectively)

(1) in the selection of the representational systems applied at any one level, (2) in the way in which a sector had been subdivided into levels, and (3) in the definition of the units at any level. Thus an investigator's orientation or individual preferences may lead him to apply different analysis systems, i.e., representational sets or alphabets at one or more levels; he could also add to or subtract from the sector levels outlined in the present work; furthermore, at most levels he could define longer or shorter units according to his needs. In the class of observational systems proposed in the present work, each level is defined as an independent, autonomous subsystem of observation; in general, its units need not correspond to units of other subsystems at other levels, and its category set need have no particular relation to a set used at another level.

The system should be modified in accordance with experience. If it should become apparent that an important class of phenomena had been omitted in the proposed system, the omission could be rectified by simply inserting a suitable subsystem. Similarly, should certain subsystems prove to be redundant these superfluous subsystems could be removed. Furthermore, if naturalistically defined scoring units should prove too elusive for reliable segmenting and coding, tighter operational definitions could be developed along lines familiar from traditional approaches to content analysis and to classification of behavior in general.

CHAPTER V

SUMMARY AND FUTURE DIRECTIONS

The human psyche is a "fearfully devious organ" (Shakow, 1959, p. 109), and human behavior is "incredibly complex" (Pittenger et al., 1960, p. 211). The present work has advanced the thesis that complex phenomena not only justify but also demand appropriately complex represenations. The study began with an examination of traditional psychotherapy research--its underlying methodological assumptions and current practices. Such an investigation led to the conclusion that research has tended to slight the basic processes in which it is formally grounded: observation and operations on the data acquired. Consequently, observation and mathematization were analyzed in detail in Chapters II and III respectively. Chapter II modeled observation as a state process, presenting 14 systems of observation prominent in psychotherapy. Chapter III modeled the formal processes used to mathematize research data with the formal model developed by Suppes and Zinnes (1963) used to perform such analysis.

A return to fundamental issues was advocated on the basis of the above analyses. This in turn implied the need to construct new types of observational systems, systems that were modular, multilevel and multimodal. Further mathematical analysis suggested that a logically simple mathematical system, the relation scale, ought to be explored more extensively for use in psychotherapy research.

As a consequence of this approach a prototype observational system was constructed in Chapter IV in which verbal, paralinguistic, and kinesic modalities could be coded for patient, therapist and interactional domains. These modalities were each represented at several levels of analysis. Wherever possible already available systems were used supplemented as necessary by ad hoc systems constructed and proposed as interim solutions. Various immediately available modifications were discussed illustrating that the system could be easily and flexibly adapted to the individual needs of a researcher.

The proposed approach to observation is certainly complex and makes heavy demands on researchers. This complexity, however, is unavoidable; simplification, a traditional research tool in psychology and a technique successful in the physical sciences has not been as productive in psychotherapy research. It is hoped that the burdensome complexity of the system will be partially offset by the advantages offered by its flexibility. The major hoped for gain, however, is in its productivity: complexity is only justified if and when the proposed approach yields meaningful results unobtainable by other, simpler methods. The system thus represents a hypothesis to be tested and the various directions in which it can be tested are explored in the next and final section.

Future Directions

Four broad areas of research suggest themselves when one considers the characteristics displayed by the class of unconventional observational systems explored in the present work. The first three areas

concern assessment of performance and interpretation of data of a specific system, while the last area concerns construction of other systems in this observational family, and their interrelationships.

The First System

Completeness and redundancy. A "nearly complete" observational system of the type constructed in Chapter IV could be evaluated with respect to completeness and redundancy.

Redundancy is a systems characteristic that can be deduced from a study of data. The proposed system presents data in a complex, multi-channel array. On the one hand, state representations at various levels may be unrelated. That is, various levels may have their own characteristic temporal patterns and features. On the other hand, it may be found experimentally that different levels do display close relationships. A consistent display of a tighter coupling of levels would imply redundancy; in that case, one or more of these closely associated levels could be removed from the system--data at one level would be predictable from other channels. One area of study then would be a search for redundancy.

Completeness refers to a system's ability to represent phenomena adequately. Thus, should a researcher using a system find that he is unable to code important research phenomena, it would indicate that a system is incomplete. Were these phenomena similar (in terms of concepts, characteristics, units, etc.) to phenomena already coded at some sublevel of the system, lack of completeness could be remedied by simply

expanding the state representation set at that sublevel to include the desired data. Should, however, the omitted phenomena prove to constitute a major new class of data it would become necessary to construct a new representation set. That set could then be inserted into the appropriate subsector of the array.

Reliability. The reliability of most already existing subsystems has been assessed by its originators. It must be remembered, however, that such reliability studies have evaluated an individual system only in conjunction with the particular scoring unit specifically defined for that study. Since most of the systems in their proposed use would categorize data at several different levels in terms of different size scoring units the original reliability studies would need to be extended. It would be desirable to verify the reliability of these established systems as they are applied to classifying new sizes of scoring units.

Another class of reliability studies should be considered. The several new, ad hoc alphabets (levels 1.32 and 2.32, Table 8, page 142) have not as yet been used. Should a researcher decide to use these untested interim representational sets, he may wish to perform traditional reliability studies.

Still another type of reliability study also need be considered. The reliability of the total, highly complex data array may in some sense be different from the reliability assessed by averaging or otherwise simply incorporating the reliabilities of individual subsystems taken one at a time. A research effort could be devoted to developing

holistic concepts of reliability more appropriate for assessing a complex whole; such a concept would have to take the nonlinear aspects of interacting subsystems into account.

Operations on data. Chapters II and III considered two different types of operations: (1) operations which changed (reduced) data from one nominal level to another (the type of operations shown in the figures in Chapter II and (2) mathematical operations which change empirical to mathematical data, or mathematical data from one form to another. Typically, the purpose of the first class of processes is to consolidate data; a researcher capitalizes on formal, logical features exhibited by one array to restructure data into other, more compact, efficient, or meaningful arrays. Further discussions of operations of this type must wait until the proposed system has been used in actual studies; the features of empirical data should guide the development of this first kind of operation.

Two types of mathematical procedures were distinguished in Chapter III: those that change nominal, categorized data to mathematical data (direct, or primary), and those that process mathematical data (secondary). It was suggested that at first researchers should focus on direct operations; these should be viewed at this stage as a search for meaningful relations, for patterns, among empirical data. At present, the only available way in which to study complex patterns seems to be through a laborious, subjective, intuitive study of logical relations. From this perspective the burdensome complexity of the proposed system may be an advantage since the rich empirical data set

provided by the proposed system may offer the researcher a wealth of options for constructing cartesian products and for exploring relations.

The data acquired by the observational system constructed in Chapter IV are basically nominal data; the alphabets used at each level (Table 8, page 142) are conceived as classification systems. Therefore, at any one level the specific notation or symbolism chosen to represent that classification scheme is optional. For example, each "letter" or member of a state representation set could be symbolized by an English letter or a number. Should numbers be used it must be remembered, however, that they are used as indexes. Thus, numbers would be only descriptive, not quantitative; a researcher must not be misled and ascribe inappropriate mathematical significance to them--for example, it would be a mistake to assume that these indexes have arithmetical properties. Thus, at this stage of systems construction numbers, letters, or any other representational symbols would be used only to display the patterns assumed by observed, categorized data.

The record obtained by such systems, then, would be a multilevel array; each level would exhibit its own segmental characteristics, and data at that level would be nominal--perhaps coded by a column of ones and zeros. A first approach to a study of relations would very likely involve a search for meaningful patterns along two loci: across data levels and along the temporal dimension. One sophisticated technological approach to analysis of such patterns is very tentatively suggested: Optical analog computers may provide a useful analytical tool. These computers can handle very large quantities of data flexibly, quickly,

and relatively inexpensively. Currently their application to analysis of complex arrays of data is being explored. For example, studies in the analysis of multichannel electroencephalogram data are being performed. Research data obtained with the proposed system could easily be prepared into a binary array coded onto a transparent film; at each sublevel of the observational system, a state vector could be represented by a column of binary symbols. It may be worthwhile to explore such analytical technology in future research programs.

Other Systems of Observation

Construction of new systems. It has been previously mentioned that different pragmatic considerations should correspondingly lead to different observational systems. Any individual system must be viewed as a member of a family of systems. The most signficant differences between these members probably would be the particular classification sets (alphabets) chosen by a researcher. For example, a therapist's verbal behavior would not be appropriately characterized by Strupp's categories for use in a behavioralistically oriented research program; a category system based on more suitable concepts (perhaps one using Skinner's terminology to describe verbal behavior) would be more appropriate.

In principle the construction of other observational systems should be straightforward. A basic rationale such as that displayed in Figures 16 and 17, pages 122 and 123, is defined; that is, some broad framework is selected for the basic categorization of phenomena. (In the present work the basic schemata selected were (1) patient-therapist-interactional and (2) verbal-nonverbal categories.) Thereafter,

appropriate category sets and concomitant units are either selected from available systems or created according to a researcher's unique needs and preferences. Thus, the approach to new systems closely parallels the approach demonstrated in Chapter IV.

Comparisons and transformations between systems. After one or more additional members of this system's family had been constructed, one could consider questions of relationships between such systems. For example, assuming that several of these systems had been used to describe an interview, what could be said about similarities or differences between data? By analogy to mathematical systems, one might anticipate that bilaterial transformations (transformations in either direction) should be possible between "complete" systems. Traditionally, transformations to different descriptive bases involve dimensionality, which in turn is related to issues of completeness and redundancy. These two latter topics thus arise also in conjunction with collections of systems, as well as in the previously explored context of a single system.

It is one thing to assert that a transformation may exist, and another, to discover it--particularly if it is between two complex, multilevel systems. Should the two systems have the same number of levels (a situation roughly analogous in conventional systems to an equal number of degrees of freedom or like dimensionality) one might optimistically look for a simple transformation; one might hope that the systems could be related by a set of separate functions, one function for each level. In this case each function would be composed of a category set used at one level of the first system, a corresponding

set from the other, second system, and an assignment from one set to the other. (Even in this simple case, complications would arise if the two category sets contained a different number of elements; assignments could then no longer be made symmetrically, one to one, in either direction.)

Systems could differ significantly in the number of levels used, and in their approach to temporal segmentation. Transformations between such highly dissimilar systems would be even more difficult to conceptualize.

The question of transformation is thus seen to be a complex one touching on problems of completeness, redundancy, temporal segmentation, and dimensionality of systems and their component subsystems. Perhaps this difficult area will be clarified as data are acquired through use of several, alternative systems of observation.

In conclusion, it appears that the type of observational system conceived here would require a great deal of difficult study, much of it in uncharted territory, before its value can be assessed. It is easier and thus tempting to continue using traditional approaches to observations; it also may be that simple, yet highly productive observational systems may be discovered in future psychotherapy research. Past experience, however, does not encourage that hope; hence the complex, difficult alternative offered here. It is hoped that regardless of the ultimate practical value of the specific system presented here, the analyses, methodological approaches, and general concepts developed in this study, as well as any insights gained from exploratory research generated thereby, will make a contribution to the advancement of research in psychotherapy.

REFERENCES

REFERENCES

Auerbach, A. H., and Luborsky, L. Accuracy of judgements of psychotherapy and the nature of a "good hour." In J. M. Shlien (Ed.), Research in psychotherapy. Vol. 3. Washington, D.C.: American Psychological Association, 1968, pp. 155-168.

Auld, F. Vicissitudes of communication in psychotherapy. In J. M. Shlien (Ed.), Research in psychotherapy. Vol. 3. Washington, D.C.: American Psychological Association, 1968, pp. 169-178.

Auld, F., Jr., & Dollard, J. Measurement of motivational variables in psychotherapy. In L. A. Gottschalk & A. H. Auerbach (Eds.), Methods of research in psychotherapy. New York: Appleton-Century-Crofts, 1966, pp. 85-92.

Beck, S. J. The Rorschach test, communication, and psychotherapy. In L. A. Gottschalk & A. H. Auerbach (Eds.), Methods of research in psychotherapy. New York: Appleton-Century-Crofts, 1966, pp. 551-572.

Bellman, R., & Kalaba, R. Dynamic programming and modern control theory. New York: Academic Press, 1965.

Bellman, R., & Smith, C. P. Simulation in human systems: Decision-making in psychotherapy. New York: Wiley, 1973.

Bergin, A. E. Psychology as a science of inner experience. In J. T. Hart & T. M. Tomlinson (Eds.), New directions in client-centered therapy. Boston: Houghton-Mifflin, 1970, pp. 61-69.

Bergin, A. E. The evaluation of therapeutic outcomes. In A. E. Bergin & S. L. Garfield (Eds.), Handbook of psychotherapy and behavior change. New York: Wiley, 1971.

Bergin, A. E., & Garfield, S. L. Handbook of psychotherapy and behavior change. New York: Wiley, 1971.

Bergin, A. E., & Strupp, H. H. Changing frontiers in the science of psychotherapy. Chicago: Aldine/Atherton, 1972.

Betz, B. J. Experiences in research in psychotherapy with schizophrenic patients. In H. H. Strupp & L. Luborsky (Eds.), Research in psychotherapy. Vol. 2. Washington, D.C.: American Psychological Association, 1962, pp. 41-60.

Block, J. Group discussion: Research problems relating to measuring personality change in psychotherapy. In H. H. Strupp & L. Luborsky (Eds.), Research in psychotherapy. Vol. 2. Washington, D.C.: American Psychological Association, 1962, pp. 164-176, 271-272.

Boomer, D. S. Speech disturbance and body movement in interviews. Journal of Nervous and Mental Disease, 1963, 136, 263-266.

Bordin, E. S. Inside the therapeutic hour. In E. A. Rubinstein & M. B. Parloff (Eds.), Research in psychotherapy. Vol. 1. Washington, D.C.: American Psychological Association, 1959.

Bordin, E. S. Simplification as a strategy for research in psychotherapy. Journal of Consulting Psychology, 1965, 29, 493-503.

Box, G. E. P., & Jenkins, G. M. Time series analysis: Forecasting and control. San Francisco: Holden-Day, 1970.

Butler, J. M., Rice, L. N., & Wagstaff, A. K. On the naturalistic definition of variables: An analogue of clinical analysis. In H. H. Strupp & L. Luborsky (Eds.), Research in psychotherapy. Vol. 2. Washington, D.C.: American Psychological Association, 1962, pp. 178-205.

Butler, J. M., Rice, L. N., & Wagstaff, A. K. Quantitative naturalistic research. Englewood Cliffs, N.J.: Prentice-Hall, 1963.

Campbell, D. T., & Stanley, J. C. Experimental and quasiexperimental designs for research. Chicago: Rand McNally, 1966.

Carroll, E. T., & Houghton, G. Mathematical reflections on the analytic process. Psychoanalytic Quarterly, 1970, 39, 103-118.

Cartwright, R. D. A comparison of the response to psychoanalytic and client-centered psychotherapy. In L. A. Gottschalk & A. H. Auerbach (Eds.) Methods of research in psychotherapy. New York: Appleton-Century-Crofts, 1966, pp. 517-529.

Chance, E. Content analysis of verbalizations about interpersonal experience. In L. A. Gottschalk & A. H. Auerbach (Eds.), Methods of research in psychotherapy. New York: Appleton-Century-Crofts, 1966, pp. 127-144.

Chassan, J. B., & Bellak, L. An introduction to intensive design in the evaluation of drug efficacy during psychotherapy. In L. A. Gottschalk & A. H. Auerbach (Eds.), Methods of research in psychotherapy. New York: Appleton-Century-Crofts, 1966, pp. 478-499.

Condon, W. S., & Ogston, W. D. Sound film analysis of normal and pathological behavior patterns. Journal of Nervous and Mental Disease, 1966, 143, 338-347.

Condon, W. S., & Ogston, W. D. A segmentation of behavior. Journal of Psychiatric Research, 1967, 5, 221-235.

Coombs, C. H., Dawes, R. M., & Tversky, A. Mathematical psychology. Englewood Cliffs, N.J.: Prentice-Hall, 1970.

Coombs, C. H., Raiffa, H., & Thrall, R. M. Some views on mathematical models and measurement theory. Psychological Review, 1954, 61, 132-144.

Deutsch, F. Some principles of correlating verbal and nonverbal communication. In L. A. Gottschalk & A. H. Auerbach (Eds.), Methods of research in psychotherapy. New York: Appleton-Century-Crofts, 1966, pp. 166-169.

Duncan, S., Jr., Rice, L. N., & Butler, J. M. Therapist's paralanguage in peak and poor psychotherapy hours. Journal of Abnormal Psychology, 1968, 73, 556-570.

Edwards, A. L., & Cronbach, L. J. Experimental design for research in psychotherapy. Journal of Clinical Psychology, 1952, 8, 51-59.

Ekman, P., & Friesen, W. V. Nonverbal behavior in psychotherapy research. In L. A. Gottschalk & A. H. Auerbach (Eds.), Methods of research in psychotherapy. New York: Appleton-Century-Crofts, 1966, pp. 179-216.

Eysenck, H. J. The effects of psychotherapy, an evaluation. Journal of Consulting Psychology, 1952, 15, 319-324.

Fiske, D. W., Luborsky, L., Parloff, M. B., Hunt, H. F., Orne, M. T., Reiser, M. F., & Tuma, A. H. Planning of research on effectiveness of psychotherapy. Archives of General Psychiatry, 1970, 22, 22-32.

Frank, G. H. On the history of the objective investigation of the process of psychotherapy. Journal of Psychology, 1961, 51, 89-95.

Frank, J. D. Problems of controls in psychotherapy as exemplified by the psychotherapy research project of the Phipps Psychiatric Clinic. In E. A. Rubinstein & M. B. Parloff (Eds.), Reserach in psychotherapy. Vol. 1. Washington, D.C.: American Psychological Association, 1959, pp. 10-26.

Goldman-Eisler, F. Psycholinguistics. New York: Academic Press, 1968.

Goldstein, A. P., & Dean, S. J. (Eds.). The investigation of psychotherapy: Commentaries and readings. New York: Wiley, 1966.

Goldstein, H. Classical mechanics. Reading, Mass.: Addison-Wesley, 1950.

Gottman, J. M., & Leiblum, S. R. How to do psychotherapy and how to evaluate it. New York: Holt, Rinehart and Winston, 1974.

Gottschalk, L. A., & Auerbach, A. H. (Eds.). Methods of research in psychotherapy. New York: Appleton-Century-Crofts, 1966.

Gottschalk, L. A., Winget, C. M., Gleser, G. C., & Springer, K. J. The measurement of emotional changes during a psychiatric interview: A working model toward quantifying the psychoanalytic concept of affect. In L. A. Gottschalk & A. H. Auerbach (Eds.), Methods of research in psychotherapy. New York: Appleton-Century-Crofts, 1966, pp. 93-126.

Harway, N. I., Dittman, A. T., Raush, H. L., Bordin, E. S., & Rigler, D. The measurement of depth of interpretation. Journal of Consulting Psychology, 1955, 19, 247-253.

Houghton, G. A systems-mathematical interpretation of psychoanalytic theory. Bulletin of Mathematical Biophysics, 1968, 30, 61-85.

Howard, K. I., & Orlinsky, D. E. Psychotherapeutic processes. Annual Reviews of Psychology, 1972, 23, 615-668.

Hunt, J. McV. Toward an integrated program of research on psychotherapy. Journal of Consulting Psychology, 1952, 16, 237-246.

Isaacs, K. S. Group discussion: Methods for assessment of change. In E. A. Rubinstein & M. B. Parloff (Eds.), Research in psychotherapy. Vol. 1. Washington, D.C.: American Psychological Association, 1959, pp. 226-239.

Jaffe, J. Computer assessment of dyadic interaction from chronographic data. In J. M. Shlien (Ed.), Research in psychotherapy. Vol. 3. Washington, D.C.: American Psychological Association, 1968.

Kelly, J., Smits, S. J., Leventhal, R., & Rhodes, R. Critique of the designs of process and outcome research. Journal of Counseling Psychology, 1970, 17, 337-341.

Kiesler, D. J. Experimental designs in psychotherapy research. In A. E. Bergin & S. L. Garfield (Eds.), Handbook of psychotherapy and behavior change. New York: Wiley, 1971, pp. 36-74.

Knapp, P. H., Mushatt, C., & Nemetz, S. J. Collection and utilization of data in a psychoanalytic psychosomatic study. In L. A. Gottschalk & A. H. Auerbach (Eds.), Methods of research in psychotherapy. New York: Appleton-Century-Crofts, 1966, pp. 401-421.

Kuhn, T. S. The structure of scientific revolutions. (2nd ed.) Chicago: University of Chicago Press, 1970.

Lang, P. J. Fear reduction and fear behavior: Problems in treating a construct. In J. M. Shlien (Ed.), Research in psychotherapy. Vol. 3. Washington, D.C.: American Psychological Association, 1968, pp. 90-102.

Langer, S. K. Problems of art. New York: Sribner's, 1957.

Leary, T., & Gill, M. The dimensions and a measure of the process of psychotherapy: a system for the analysis of the content of clinical evaluations and patient-therapist verbalizations. In E. A. Rubinstein & M. B. Parloff (Eds.), Research in psychotherapy. Vol. 1. Washington, D.C.: American Psychological Association, 1959, pp. 62-95.

Lennard, H. G., & Bernstein, A. The anatomy of psychotherapy. New York: Columbia University Press, 1960.

Lennard, H. L. Some aspects of the therapeutic system. In H. H. Strupp & L. Luborsky (Eds.), Research in psychotherapy. Vol. 2. Washington, D.C.: American Psychological Association, 1962, pp. 218-237.

Levinson, D. J. The psychotherapist's contribution to the patient's treatment career. In H. H. Strupp & L. Luborsky (Eds.), Research in Psychotherapy. Vol. 2. Washington, D.C.: American Psychological Association, 1962, pp. 13-24, 292-298.

Lorr, M. Discussion: Therapist-patient relationship. In E. A. Rubinstein & M. B. Parloff (Eds.), Research in psychotherapy. Vol. 1. Washington, D.C.: American Psychological Association, 1959, pp. 260-263.

Luborsky, L., & Strupp, H. H. Research problems in psychotherapy: A three-year follow-up. In H. H. Strupp & L. Luborsky (Eds.), Research in psychotherapy. Vol. 2. Washington, D.C.: American Psychological Association, 1962, pp. 308-330.

Mahl, G. F. Gestures and body movements in interviews. In J. M. Schlien (Ed.), Research in psychotherapy. Vol. 3. Washington, D.C.: American Psychological Association, 1968, pp. 295-346.

Malan, D. H. The outcome problem in psychotherapy research--a historical review. Archives of General Psychiatry, 1973, 29, 719-728.

Marsden, G. Content-analysis studies of therapeutic interviews: 1954 to 1964. Psychological Bulletin, 1965, 63, 298-321.

Marsden, G. Content analysis studies of psychotherapy: 1954 through 1968. In A. E. Bergin & S. L. Garfield (Eds.), Handbook of psychotherapy and behavior change. New York: Wiley, 1971, pp. 345-407.

Matarazzo, J. D., & Wiens, A. N. The interview. Chicago: Aldine, 1972.

Matarazzo, J. D., Wiens, A. N., Matarazzo, R. G., & Saslow, G. Speech and silence behavior in clinical psychotherapy and its laboratory correlates. In J. M. Schlien (Ed.), Research in psychotherapy. Vol. 3. Washington, D.C.: American Psychological Association, 1968, pp. 347-394.

Meltzoff, J., & Kornreich, M. Research in psychotherapy. New York: Atherton, 1970.

Menge, B. A. Review of H. C. Andrew, Introduction to mathematical techniques in pattern recognition. American Scientist, 1974, 62, 245-246.

Millon, T., & Diesenhaus, H. I. Research methods in psychopathology. New York: Wiley, 1972.

Pande, S. K., & Gart, J. J. A method to quantify reciprocal influence between patient and therapist in psychotherapy. In L. A. Gottschalk & A. H. Auerbach (Eds.), Methods of research in psychotherapy. New York: Appleton-Century-Crofts, 1966.

Panzetta, A. Toward a scientific psychiatric nosology. Archives of General Psychiatry, 1974, 30, 154-161.

Parloff, M. B. Summary report: Therapist's contribution. In H. H. Strupp & L. Luborsky (Eds.), Research in psychotherapy. Vol. 2. Washington, D.C.: American Psychological Association, 1962, pp. 256-258.

Parloff, M. B., & Rubinstein, E. A. Research problems in psychotherapy. In E. A. Rubinstein & M. B. Parloff (Eds.), Research in psychotherapy. Vol. 2. Washington, D.C.: American Psychological Association, 1962, pp. 276-293.

Pittenger, R. E., Hockett, C. F., & Danehy, J. J. The first five minutes--A sample of microscopic interview analysis. Ithaca, N.Y.: Martineau, 1960.

Robbins, L. L. Group discussion: Methods for assessment of change. In E. A. Rubinstein & M. B. Parloff (Eds.), Research in psychotherapy. Vol. 1. Washington, D.C.: American Psychological Association, 1959, pp. 116-124.

Robbins, L. L., & Wallerstein, R. S. The research strategy and tactics of the psychotherapy project of The Menninger Foundation and the problem of controls. In E. A. Rubinstein & M. B. Parloff (Eds.), Research in psychotherapy. Vol. 1. Washington, D.C.: American Psychological Association, 1959, pp. 27-48.

Rogers, C. R. A tentative scale for the measurement of process in psychotherapy. In E. A. Rubinstein & M. B. Parloff (Eds.), Research in psychotherapy. Vol. 1. Washington, D.C.: American Psychological Association, 1959, pp. 96-107.

Rogers, C. R. The process equation in psychotherapy. American Journal of Psychotherapy, 1961, 15, 27-45.

Rubinstein, E. A., & Parloff, M. B. (Eds.). Research in psychotherapy. Vol. 1. Washington, D.C.: American Psychological Association, 1959.

Sargent, H. D. Intrapsychic change: Methodological problems in psychotherapy research. Psychiatry, 1961, 24, 93-108.

Saslow, G., & Matarazzo, J. D. A technique for studying changes in interview behavior. In E. A. Rubinstein & M. B. Parloff (Eds.), Research in psychotherapy. Vol. 1. Washington, D.C.: American Psychological Association, 1959, pp. 125-159.

Scheffler, I. Science and subjectivity. New York: Bobbs-Merrill, 1967.

Scheflen, A. E. Natural history method in psychotherapy: Communicational research. In L. A. Gottschalk & A. H. Auerbach (Eds.), Methods of research in psychotherapy. New York: Appleton-Century-Crofts, 1966, pp. 263-288

Scheflen, A. E. Communicational structure--Analysis of a psychotherapy transaction. Bloomington: Indiana University Press, 1973.

Scofield, W. The Structured Personality Inventory in measurement of effects of psychotherapy. In L. A. Gottschalk & A. H. Auerbach (Eds.), Methods of research in psychotherapy. New York: Appleton-Century-Crofts, 1966, pp. 536-550.

Seitz, P. F. D. The consensus problem in psychoanalytic research. In L. A. Gottschalk & A. H. Auerbach (Eds.), Methods of research in psychotherapy. New York: Appleton-Century-Crofts, 1966, pp. 209-225.

Shakow, D. Discussion: Methods for assessment of change. In E. A. Rubinstein & M. B. Parloff (Eds.), Research in psychotherapy. Vol. 1. Washington, D.C.: American Psychological Association, 1959, pp. 264-275.

Shakow, D. Discussant: Research problems relating to the definition, measurement, and analysis of significant variables in psychotherapy. In H. H. Strupp & L. Luborsky (Eds.), Research in psychotherapy. Vol. 2. Washington, D.CL: American Psychological Association, 1962, pp. 164-177.

Shapiro, J. G. Relationship between visual and auditory cues of therapeutic effectiveness. Journal of Clinical Psychology, 1968, 24, 236-239.

Shlien, J. M. Toward what level of abstraction in criteria? In H. H. Strupp & L. Luborsky (Eds.), Research in psychotherapy. Vol. 2. Washington, D.C.: American Psychological Association, 1962, pp. 142-154.

Shlien, J. M. (Ed.). Research in psychotherapy. Vol. 3. Washington, D.C.: American Psychological Association, 1968.

Shlien, J. M. Phenomenology and personality. In J. T. Hart & T. M. Tomlinson (Eds.), New directions in client-centered therapy. Boston: Houghton-Mifflin, 1970, pp. 61-69.

Snyder, W. U. Group discussion: Therapist-patient relationships. In E. A. Rubinstein & M. B. Parloff (Eds.), Research in psychotherapy. Vol. 1. Washington, D.C.: American Psychological Association, 1959, pp. 264-275.

Strupp, H. H. A multidimensional comparison of therapist activity in analytic and client-centered therapy. Journal of Consulting Psychology, 1957, 2, 301-308.

Strupp, H. H. Group discussion: Therapist-patient relationship. In E. A. Rubinstein & M. B. Parloff (Eds.), Research in psychotherapy. Vol. 1. Washington, D.C.: American Psychological Association, 1959, pp. 264-275.

Strupp, H. H. Group discussion: Measuring personality change. In H. H. Strupp & L. Luborsky (Eds.), Research in psychotherapy. Vol. 2. Washington, D.C.: American Psychological Association, 1962, pp. 273-283.

Strupp, H. H. Patient-doctor relationships: Psychotherapist in the therapeutic process. In A. J. Bachrach (Ed.), Experimental foundations of clinical psychology. New York: Basic Books, 1962, pp. 576-615.

Strupp, H. H., & A. E. Bergin. Research in individual psychotherapy--A bibliography. U.S. Department of Health, Education & Welfare, 1969.

Strupp, H. H., Chassan, J. B., & Ewing, J. A. Toward the longitudinal study of the psychotherapeutic process. In L. A. Gottschalk & A. H. Auerbach (Eds.), Methods of research in psychotherapy. New York: Appleton-Century-Crofts, 1966, pp. 361-400.

Strupp, H. H., & Luborsky, L. (Eds.). Research in psychotherapy. Vol. 2. Washington, D.C.: American Psychological Association, 1962.

Suppes, P., & Zinnes, J. L. Basic measurement theory. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), Handbook of Mathematical Psychology. Vol. 1. New York: Wiley, 1963, pp. 1-76.

Tomlinson, T. M., & Hart, J. T., Jr. A validation study of the process scale. Journal of Consulting Psychology, 1962, 26, 74-78.

Turner, M. B. Philosophy and the science of behavior. New York: Appleton-Century-Crofts, 1967.

Wallerstein, R. S. The psychotherapy project of the Menninger Foundation: An overview at the midway point. In L. A. Gottschalk & A. H. Auerbach (Eds.), Methods of research in psychotherapy. New York: Appleton-Century-Crofts, 1966, pp. 501-516.

Wallerstein, R. S. The Psychotherapy Research Project of the Menninger Foundation: A semifinal view. In J. M. Shlien (Ed.), Research in psychotherapy. Vol. 3. Washington, D.C.: American Psychological Association, 1968, pp. 585-605.

Wallerstein, R. S., & Robbins, L. L. The Psychotherapy Project of the Menninger Foundation. Vol. IV. Concepts. Bulletin of the Menninger Clinic, 1956, 20, 239-262.

Weimer, W. B. Psycholinguistics and Plato's paradoxes of the Meno. American Psychologist, 1973, 28, 15-24.

Yankelovich, D., & Barrett, W. Ego and instinct. New York: Vintage, 1971.

VITA

Louis Sebastian Berger was born in Prague, Czechoslovakia, in 1927. He emigrated to the United States in 1940 and became a naturalized citizen in 1945. Following his graduation in 1945 from Newtown High School in Elmhurst, New York, he attended Princeton University, receiving a Bachelor of Science degree in Electrical Engineering in 1948. He also holds a Master's degree in Music from the University of Texas at Austin (1953) and a Master of Science degree in Physics from Trinity University in San Antonio (1971).

After receiving his engineering degree, Mr. Berger studied music at Eastman, Juilliard, and Mannes. He acquired professional experience as assistant principal cellist of the Kansas City Philharmonic and as principal cellist of the Austin Symphony. On completing his graduate studies in music he joined the Boston Symphony Orchestra under Charles Munch and was a member for ten years until his resignation in 1963.

During his tenure in Boston Mr. Berger resumed technical work on a part-time basis, joining AVCO's Research and Development Division, Wilmington, in 1956. There he worked in applied physics, leaving in 1964 to join the staff of Southwest Research Institute in San Antonio. When he left the Institute in 1972 to return to graduate school he was a senior research scientist in the Behavioral Sciences Section of the Department of Biomedical Engineering.