Report by Dr. Donald Watterson

November 3, 1952

As revised by Dr. Lester Luborsky

November 10, 1952

RESEARCH PROPOSAL FOR EVALUATION OF PSYCHOTHERAPY

Table of Contents

I. Introduction

During the last fifty years different forms of psychotherapy have gradually come into widespread use in the treatment of mental illness. As yet however it has never been proven rigorously that psychotherapy is even of benefit in the various forms of mental or emotional illness (although the writers are, of course, all personally convinced of its value) let alone that a particular method of psychotherapy is more beneficial than some other method for this or that variety of mental illness (although here the writers have personal differences). Such demonstration of the value of psychotherapy, or the comparable value of different methods of therapy, is lacking for a number of reasons. First there exists no valid instrument by which behavioral and psychological change in the patient during the course of treatment can be measured.* In the absence of such procedures the state of the patient at the termination of treatment is expressed (if expressed at all) in such terms as "recovered," "improved," "unimproved," "regressed," and these terms are usually left undefined or ill-defined so that one therapist is unable to compare his results with those of another. At that, there have been no studies in which independent observers apply these terms to the same patients. Second, there is as yet no adequate or generally accepted calssification of psychiatric illnesses; a nosological label such as schizophrenia, for example, would be applied to a given patient by one group of workers but not necessarily by other groups. Third, the various methods of psychological treatment are very poorly described and defined in the literature. Even colleagues who profess non-

* Many devices are in the process of development; none yet properly validated for this task, e.g., Hunt's Movement Scale; Ferguson's Hospital Behavior Check List, et

inally to employ the same method of treatment may not be able to agree on a definition of that method; recently for example the members of the American Psychoanalytic Association were unable to arrive at an agreed definition of psychoanalytic treatment even though the suggested definitions were formulated in the most general terms. Fourth, the therapist remains an unknown quantity, though every worker would agree that his personality (to the extent to which it impinges on the treatment process) is a most important factor in the treatment. Fifth, it is very difficult to get adequate control groups. And, lastly, a difficulty of another order, the quantity and complexity of the necessary data for research of the proper scope have scared off researchers.

Serious research has been attempted by few. The group of workers under Rogers in Chicago has concentrated fruitfully, for example, on the delineation and quantification of changes which occur in the patient during the course of "client-centered" treatment, and has also attempted to describe its method of treatment concretely. A more eclectic psychiatric group of therapists at Washington University is cooperating in a project run by Watson and Mensh.

It is still a fact, however, that virtually the whole field of psychotherapy remains nebulous (Feinsinger, Hunt). We are almost as far as ever from knowing how one method compares with another (or with other modes of treatment, such as milieu therapy, the physical treatments, the approaches of the general practitioner, the pastoral counselor, etc.) in treating this or that condition. And this state of ignorance reflects the complexity of the subject.

II. Evolution of the Project at The Menninger Clinic

Almost throughout the history of the Clinic some of its psychotherapists have tried to understand the process and evaluate the results of our treatment efforts. Studies were done (to mention a few, by Knight, 1942 (outcome of 100 cases)

Brenman and Gill, 1944-48 (Hypnotherapy, etc.), Wexler, 1950 (Treatment of Schizophrenia), Bergman, 1945- (Studies of directive vs. client-centered therapy, etc.). There was no attempt in these excellent papers to do a systematic job of evaluating treatment results.

The conditions of work at the Clinic in recent years are especially suitable for research of this nature. The various specialities at the Menninger Clinic work together in "teams," so that a given patient is always considered (and usually seen in examination) by at least two or three people. The examinational data (and subsequent data concerning the patient during his treatment) are recorded with reasonable thoroughness (extremely thoroughly compared with psychiatric practice in general). A large amount of time is devoted to psychotherapy (over 2,000 hours a month), and every therapist has the opportunity (and often the obligation) to discuss his psychotherapeutic work in detail with one or more colleagues. Thus there exists in the ordinary practice of the Clinic at least the basis for intercommunication concerning psychotherapy.

During the past two years a group (of psychiatrists, psychologists, psychoanalysts) has, at the suggestion of Dr. Paul Bergman, formed itself into a team for the general purpose of making an EVALUATION OF PSYCHOTHERAPY. It has met weekly for two hours, and in addition has carried out certain small preliminary research projects. Most of its time has been taken up in attempting to agree on aims and methods, refine concepts, develop hypotheses, and to devise an instrument of measurement of change during treatment.

III. Aims

We have come along far enough to stake out the areas of our own tastes:

1) To evaluate the outcome of psychotherapy (especially psychoanalytically oriented expressive psychotherapy and psychoanalysis). (Adequate studies of the

results of these therapies have not been reported in the literature heretofore).

2) To formulate and test fruitful hypotheses on relationships between factors in the therapist, treatment procedure, and change occurring in the patient, during and at the end of treatment. (Especially in the area of the nature of the therapists' personality very little has been reported in the literature.)

The more immediate specific plans for beginning the research as given in the descriptions of subprojects are:

1) To construct a Health Rating Scale (Project A).

2) To make a research-useful division of methods of treatment (Part of Projec

3) To examine some hypotheses about the relevance of certain personality characteristics of the therapist to the process and outcome of treatment (Project D, see Appendix).

IV. Methods

We tried to go at the job systematically by:

First, considering all the hypotheses in each of the three areas in which we were personally interested. (Appendix B) We then narrowed these down to those which are most readily testable.

Second, devising data collection procedures to provide valid data bearing on the hypotheses.

Third, parcelling out among the committee members and assistant the responsibility for clear-cut sections of the data as their special "subprojects." The subcommittees will periodically report the progress of their work to the full committee.

A major investment in data collection is required for these aims. We re-examined the data on the patient, treatment, and therapist in comparison with the demands of the projected subprojects and came out with the detailed data collection plans given below. Much is already available as part of the customary practice of the Clinic. Supplementation is necessary mainly in the areas covered by the subprojec

A, B, D, and C (Retests of patients at the completion of treatment.) We hope to be able to take advantage of a procedure for follow-up of patients in the years after completing treatment which is being considered as part of D.A.P. practice.

Schema of Data Collection
(About patient and method of treatment)

Source \ Time	Apply for rx	Begin rx	During rx	End rx	After 1 - 3 years
Section Eval.	Abstract H.R.* by members. Pred's				
Therapist		H.R. Pred's	H.R. @ 6 mo. intervls Process notes Sample 3 recorded hrs if possible	H.R. Disch. note	H.R.
Psychol. Tester	Tests & report. H.R.			Tests & report. H.R.	
Therapy Superv.		H.R.		H.R.	
Patient	Self-descrip or cue S.	Self-descrip or cue S.	Self-descrip or cue S.	Self-descrip or cue S.	Self-descrip or cue S.
Psychoth. Comm & other indep. observers		H.R.		Int. wth T. -H.R. Int. wth Sprv-H.R. H.R.	H.R.

* H.R. = Health Rating

The "control group" problem is still in the discussion stage. None of the control groups that we could think of could be considered as more than comparison groups, for example, a) Patients who apply for treatment and then through some act of fate can't get it. b) A "self-control" group of people who apply for

treatment and then must wait a long time before they get it. I suppose it would be necessary to test these people when they start therapy, as well as when they applied. c) A group of patients at Winter hospital who do not have psychotherapy for a long period of time. d) A group of patients who have some neurotic illness and spontaneous remit. (We can get this group from insurance statistics quoted by Eysenk in his recent book.)

Brief Description of Subprojects underway:

Project A: Development of a scale for rating mental health.

The scale and preliminary results are given in Appendix A. We intend to use the instrument to quantify judgments of the mental health status of patients at various points before, during, and after treatment. The scale consists of 100 points, where 100 represents ideal health and 0 total disintegration. Many of the intermediate points are defined by general criteria and exemplified by 52 graded case descriptions. In making a rating the judge must balance off our seven criteria of health, consider the approximate position of his case in the sample case series and then arrive at a single global rating.

The whole staff is being trained in the use of the scale. We have carried out several reliability studies comparing ratings made by therapists vs. therapists' supervisors, case presenters vs. professional audiences, etc. and find that we have been successful in making a 100 point scale which gives us interrater reliability within 10 points for most raters. In view of the difficulties the committee members experienced in agreeing on and communicating our criteria for "health," it is encouraging that we were able to get so much agreement. "Health" appears to be a concept (like "ego strength") which therapists feel they can usually describe with considerable certainty for a particular patient, and as experience with our scale piles up we demonstrate they are usually "right" (right in terms of being able to obtain fair consensus with closely working co-workers).

The directions for future work are:

1) Construct rating scales for the separate criteria. These can be rated with "separate" and "non-separate" sets as Dr. Murphy suggested. Then do a correlativ study. This will tell us something about the composition of the global rating. If we use the separate criteria for separate ratings we will occasionally learn the basis for divergent global ratings by different raters.

2) Examine cases that are given a health rating in each decile to see in what characteristics they are distinguishable. Examine cases which show most and leas change during the course of treatment.

3) Examine cases in which we predict change most and least successfully.

Thus a beginning has been made in filling one of the lacunae mentioned as causing the dearth of information on the value of psychotherapy, namely the lack of any single instrument by which change can be measured. We believe that the Health Rating Scale (after further refinement) will be applicable in many settings where psychiatric treatment is carried out.

Project B: Interviews with therapists at the termination of patients' treatment. (Pilot Study)

Purposes:

1) To describe the changes in the patient from the beginning to the end of treatment.

2) To describe the types of treatment given.

3) To study the effects of treatment by clinical impressions and tests, and comparing these two sources.

The method is still being developed and we have not worked out a way of summarising the findings. We use a semi-standardised form with a list of questions and the material is written up in a free fashion. Two members of the Committee do

each of the interviews, after reading some of the therapist's psychotherapy process notes. So far ten cases have been completed, on all but one of which there have been retests.

Project C: Retests of patients completing psychotherapy and psychoanalysis. (Pilot Study)

Purposes:

1) To describe the change in the patient from the beginning to the end of treatment.

2) To study the effects of psychoanalysis as revealed in clinical impressions and psychological tests, comparing the two sources.

Most patients have taken a battery of psychological tests as part of their initial evaluation. If the patient is willing, the same battery of tests is readministered.

The Topeka Psychoanalytic Society allotted to us $300 for this purpose, which the Committee agreed to take responsibility for administering. The testing is done by Dr. Holsman and Dr. Holt. About 10 cases can be retested with the money available. It is preferable to do the testing in the last 6 months of the treatment or within 6 months after the treatment. The tester writes his report independently of information from the therapist and in two parts (a) the patient as seen now (without referring to the original tests), (b) the patient's changes from the first testing. The tester and the therapist also make independent ratings of the patient on our health rating scale.

Other inroads into the rich body of data have been planned but not work begun.

Project D: A study of the personality of psychotherapists. (Appendix ____)

Choice of subjects for the total program:

The subjects of the studies are the patients and staff of the Clinic. Two criteria of selection of subjects are presently being used - first the willingness of the various patients and staff-members to take part, second the availability of the mass of detailed information which is required (for example in certain cases lengthy process notes of the treatment are kept hour by hour, in others not). It also seems likely that, for some time to come, the group will confine itself largely to the group of psychotherapies which may be designated as "expressive." It is hoped that data can be accumulated at the rate of one to two patients completing treatment per week until a total of 100 cases has been achieved.

Research staff, time perspective, and some estimated costs:

The members of the Psychotherapy Research Committee will serve as a planning body, as consultants to participants in projects, and as members of projects. The committee members are Doctors Aronson, Bergman, Fabian, Holt, Kaiser, Luborsky, Murphy, Rubenstein, and Watterson. Dr. Watterson has, as of this date, succeeded Dr. Rubenstein as chairman. Dr. Luborsky continues as secretary.

Some estimates:

1) Meetings of the Psychotherapy Research Committee — No Cost*
*Paid out of D.A.P. funds

2) To oversee the data collection procedures, a half-time to full time person is necessary, either a competent experienced secretary or research assistant — Estimated Cost: $3-4,000 per yr (for approximately 2 years)

3) To develop "Health Rating Scale," a full time research assistant or thesis student with help of a member of committee, Dr. Holt or Dr. Luborsky. — Estimated Cost: $3-4,000 per yr (for approximately 1 year)

4) To interview therapists on completion of patients' treatments. (Under the supervision of Dr. Watterson) Rotation of two members of committee. For about 20 patients. No Cost*

*Paid out of D.A.P. funds.

5) Retesting of patients after treatment. Doctors Holt and Holzman. For about 10 patients Estimated Cost: $300

($300 allotted by Topeka Psychoanalytic Society. Covers cost of 10 patient retests but does not cover cost of making report of the findings.)

APPENDIX

A. Copy of the Health Rating Scale
and two summaries of Results of Reliability Studies.

B. Enumeration of the various working hypotheses that have been of interest to the group.

C. Formulations of Subprojects
which take in blocks of hypotheses or single hypotheses.

D. A study of the personality of psychotherapists.