

## ISSUES IN RESEARCH IN THE PSYCHOANALYTIC PROCESS

ROBERT S. WALLERSTEIN and HAROLD SAMPSON, SAN FRANCISCO

It is perhaps part of the greater vigour and compass of psychoanalytic research than is usually accorded it either by friendly critic or by sedulous proponent that this essay must begin with defining caveats that delimit its scope within the confines of what we mean by our title, and that thereby keeps the entire endeavour reasonably focused and manageable. First, it is not an overall review of the state of the field; i.e. not an effort at a comprehensive treatment of all facets of psychoanalytic research, research which within its range covers investigations, varyingly combined, into analytic problems, or built on analytic data, or using the analytic method, or guided by analytic theory. Such surveys have been essayed, though usually from the single major perspective of methodological considerations (for the latter, read difficulties or obstacles); the articles by Benjamin (1950), Escalona (1952) and Bellak (1961) have been among the more cogent and influential. And secondly, it is not a political or contextual exposition of the overall state of health of psychoanalytic research, buttressing an accounting of either its achievements or its shortcomings, as a field unto itself, or in comparison with progress across the whole spectrum of sciences. Engel (1968a) has but recently reviewed this subject from the standpoint of the obstacles that have impeded research development in psychoanalysis, both those obstacles intrinsic to the nature of analysis itself and those external to it, consequent to its particular history and organization, i.e. the play of social and political factors. And one of us has added to this dialogue in discussion of the Engel article (Wallerstein, 1968)<sup>1</sup> and in official committee reports within the American Psychoanalytic Association devoted to the struggle towards removal of some of these obstacles.

The purpose and focus of this essay, stripped of these broader considerations (context though

they be for the topic), is simply the justification (and the elaboration) of the need to formalize (i.e. to go beyond) the clinical case study method as the central research instrument and research access to the therapeutic process in psychoanalysis. It is concerned with the (in our mind, necessary) place of formal systematic research on the psychoanalytic process and the very many problems and issues thereby raised in devising and executing such research in a manner at once meaningful and responsive to the subtlety and complexity of the phenomena and at the same time scientific in the best sense of that term (loyalty to the reality principle, as here embodied in appropriate canons of scientific inference).

### WHY FORMAL PSYCHOANALYTIC RESEARCH?

#### *The traditional case study method*

This question of why more formal systematic research is required for further scientific advance is then the proper starting point of our inquiry. There is no need to document the extraordinary reach of the traditional (specifically psychoanalytic) case study method innovated by Freud. The whole corpus of psychoanalysis, the closest in existence to a general psychology, comprehending the phenomena of both normal and abnormal personality development and functioning, attests brilliantly to the explanatory power of the theory derived from the data of the consulting room. It has flourished in the hands of its founding genius and of those who have come after and has provided a truly extraordinary range of insights into the structure of the mind, the organization of mental illness, the forces at work in the treatment situation, the processes of change and the requirements of technique. By contrast, it is the sobering appraisal of Strupp, a dedicated psychotherapy researcher (whose professional origins and commitments were fashioned primarily in the

<sup>1</sup> See, in fact, the whole range of viewpoints, concurring and dissenting, represented within psychoanalysis in the printed discussions of the Engel paper (1968b).

research rather than the clinical crucible) that whatever spectacular growth formal research method and research inquiry in psychotherapy have undergone recently, that *to this point* these have exerted but very slight influence on the theory and practice of psychotherapy. He states the issue bluntly (Strupp, 1960):

clinical penetration and scientific rigor have varied inversely . . . If the advances of psychoanalysis as a therapeutic technique are compared with the experimental research contributions, there can be little argument as to which has more profoundly enriched the theory and practice of psychotherapy. To make the point more boldly, I believe that, up to the present, research contributions have had exceedingly little influence on the practical procedures of psychotherapy.

And he raises this question, he says, 'not as a therapeutic nihilist but as a researcher who feels that important answers should come from research . . . [though he is] not fully convinced that they will'. His intent is to examine the reasons for this state of affairs, 'in the belief that, somehow, we should be able to do better'. And he goes on to cogently explore the schism that for the most part divides practising therapists and investigators doing research on the therapeutic process and the issues involved in the search 'to reconcile successfully scientific rigor with the richness and subtle complexity of interpersonal dynamics'.

Valid as we recognize Strupp's argument (rather, his lament) to be—that our understandings in psychoanalysis have not been much enriched by formal research, but have been powerfully moulded by clinical case studies, starting with Freud's original cases—on balance we need be at least equally cognizant of the limitations of the case study method as a source of prospective continuing knowledge. These limitations have been clearly and variously summarized (Bellak, 1961; Benjamin, 1950; Escalona, 1952; Glover, 1952; Shakow, 1960). Shakow (1960) spoke of the intrinsic inadequacies of (psychoanalytic) data that are reported by a *participant-observer* (the analyst); as observers and reporters analysts are

handicapped sensorially, memorically, and expressively . . . Put simply, they are limited in how much they can grasp, in how much they can remember of what they do grasp, and in how much and how well they can report even the slight amount they have grasped and remembered.

Glover (1952) focused, and much more sharply, on the distorting biases fostered by the very conditions of analytic life.

Analysts of established prestige and seniority produce papers advancing a new theoretical or clinical viewpoint or discovery. If others corroborate it they tend to report that; but if others feel reason to reject it, this scientific 'negative' does not get reported. So ultimately it is canonized 'as so-and-so has shown'.

For this reason, Glover feels it almost inevitable in psychoanalysis that 'a great deal of what passes as attested theory is little more than speculation, varying widely in plausibility'. Such defect is not corrected within the training situation, which Glover feels tends rather to perpetuate error through its hot-house atmosphere and the ready ascription of dissent or question to 'resistances' that the candidate must overcome. And he concludes that the ready recourse to the example of Freud in this instance serves the field ill.

One of the remarkably rare disservices which Freud rendered his own science was when, with a tolerance born of his own scientific integrity and of his incomparable flair for reconstruction, Freud gave sanction to the individual manipulation of 'samples' . . . 'Altogether, experience shows that a reader who is willing to believe an analyst at all will give him credit for the touch of revision to which he has subjected his material' (Freud, 1912). No doubt when someone of Freud's calibre appears in our midst he will be freely accorded and will in any case freely exercise this privilege. Until that time comes the uncontrolled licence should be revoked.

Gill has commented on the limitations of the treating analyst as the sole research observer and reporter, to wit, the problems of the countertransference—to which we think we pay far more attention in our actual operations than, at least in research, we actually do. In Gill's statement (in Brenman, 1947) the research problem created by the countertransference is that of properly calibrating and recording oneself (the analyst) as the observer (in researches that rest, after all, on human assessments and not on dial readings).

The psychoanalyst's recognition of the phenomena of 'countertransference' is, in a sense, a first revolutionary step in the direction of the individual observer making correction for his own selective

blindnesses. Yet in a case report we hear nothing of the subtleties of the analyst's attempt to understand and evaluate his own role in the 'experiments' he conducts with each patient, and thus we must take on faith that individual idiosyncracies have played a minimal role in the observations and in the conclusions drawn from them. This is not to belittle the problems which would beset a man who attempted to record in his case report his own countertransference reactions and his method of dealing with them, but it must nevertheless be pointed out that in no other branch of science are we willing to refrain from inquiring into the possible sources of error inherent in a recording apparatus or in the experimental design itself.

It is hard in the face of such cogent portrayals of the inherent limitations of the traditional case study method (embedded in the traditional clinical and training practices of psychoanalysis) as a source of verifiable and verified knowledge to cavil at Benjamin's cautious summary (1950):

It may fairly be stated, I think, that psychoanalysis has clearly demonstrated some necessary conditions for the development of certain behavior patterns. The greatness of this achievement is not minimized by the recognition that in other areas necessary conditions have not been convincingly demonstrated, and that in none or almost none are sufficient conditions even approximated.

This discourse has not been one-sided, however. Not only are most analysts ready to agree with Strupp's discouraged caveat that, after all, more formal research has not been able to demonstrate to this point any real improvement over what can be learned out of the psychoanalytic process. There is the additional argument that the standard psychoanalytic situation anyway is or can be considered from a perspective that permits it to be viewed as close to a research (even quasi-experimental) method. This is not the same as the extreme argument, advanced only by some (cf. Ramzy, 1962, 1963), that every clinical analysis is not only a search, but in its essence also a research. To quote:

It may be hoped that with more scrutiny of the requirements of logic and scientific method, and of what actually happens in psychoanalytic treatment, it may turn out that, after all, the standard psychoanalytic method will be considered as the best research device for understanding the human

mind so far suggested. It may turn out that every analyst who *merely* follows the method he was taught to follow will discover that he has been doing research, just as Molière's M. Jourdain suddenly discovered that he has been speaking prose for 40 years without knowing it. Psychoanalysts, however, have dire need to learn how to read, write, and correctly spell the prose they have used since the inception of their discipline. [Italics ours.]

The more limited and more tenable espousal of this position is rather the viewpoint advanced so vigorously by Ezriel (1951, 1952) and supported as well by Bellak & Smith (1956), Hartmann (1959), Kris (1947) and Kubie (1956; and in Bronner, 1949) that the psychoanalytic situation as it naturally is can itself be considered as fulfilling *essentially* the requirements of a *quasi-experimental* research model. According to this view, the psychoanalytic situation is a relatively stabilized, recurring experimental situation in which the experimenter (the analyst) introduces independent variables (interpolations and other specifiable interventions) and can then predict and ascertain their impact on all the dependent variables within the situation in which, after all, he has the fullest conditions of access to the subjective data that enters consciousness (no matter how seemingly remote or trivial) ever devised. Using this model, of the 'controlled conditions of the analytic relationship' (1951), Ezriel feels that the analyst 'can state the necessary and sufficient conditions to produce a predictable event during that session' (1952). Kubie (1956) attempts to spell out these *controlled conditions*. He speaks of the 'formal constancy of the observational situation' powerfully established by the 'analytic incognito' which helps ensure 'that the variables which are brought into each session are brought in predominantly by the analysts'. Into this system, in which it is thus

possible to study the patient and not the analyst, to study the role of one system of unconscious ideas and feelings (i.e. the patient's) instead of two interwoven into an inextricable tangle,

the analyst then intrudes interpretations as deliberate and calculated variables. In this sense, 'An interpretation is nothing other than a working hypothesis to be tested by certain implicit or explicit predictions.'<sup>2</sup> Therefore

<sup>2</sup> But of course (also in Kubie, 1956) 'variables are many, the controls are few and the quantitative devices

still almost non-existent' so that there will be many wrong, unsustained hypotheses (interpretations).

'to an unexpected degree analysis, as an experimental design, is an excellent model'. Kris (1947) states that it is through its 'rules of procedure' that the psychoanalytic situation establishes itself as (almost) an experimental setting, an almost which Hartmann (1959) likewise advances:

This is not to claim that psychoanalysis is an experimental discipline. However, there are situations where it comes close to it. At any rate, there is sufficient evidence for the statement that our observations in the psychoanalytic situation, set in the context of psychoanalytic experience and hypotheses, make predictions possible—predictions of varying degrees of precision or reliability...

Appealing as this thesis is, that (Bellak & Smith, 1956)

the analytic situation itself resembles a scientific experiment...[in which] there is an attempt to treat the patient's behavior as the only dependent variable and to keep all independent variables at constant values (except those that are varied in a controlled way, namely, the interpretations and other actions of the analyst),

it is clear that the terms of the argument have been subtly shifted by recourse to it. Truly following it, in its implications, goes far beyond 'merely following the [case study] method he [the analyst] was taught' (the path of M. Jourdain from *Le bourgeois Gentilhomme*) and is already in fact far along the road towards more organized, i.e. more formal, research with its invoking of such words (and the concomitant implementation of the principles they subsume) as 'control, keeping factors constant, dependent and independent variables, hypotheses, and predictions'. The issue is now that such more systematized study is necessarily predicated on data more reliable, more replicable and more public than that yielded by the traditional case study alone (cf. for example, the foregoing quotations from Shakow (1960) and Glover (1952) on the *inherent* limitations of the usual consulting room data and case study method), and additionally necessarily involves more operations than just considering the analytic hour from the point of view of a quasi-experimental situation.

Because, in fact, it is in itself still far from it. For this, the criteria are more stringent. Lustman (1963) quotes Feibleman on this score, that

an experiment is controlled observation of what happens when some segment of the natural world is forced to an alternative in which it will be *as easy* to obtain a negative as well as a positive answer [*italics ours*],

and Lustman then goes on to say that

Within the clinical paradigm the closest psychoanalysis can come to experimental manipulation is to take advantage of the experiments in nature... Actually the method is an extension of the classical approach of psychoanalysis, which was the use of the neurosis to understand normal man.

Shakow (1960), in this same spirit, prefers to call the psychoanalytic hour not a semi-experimental approach, but rather a semi-naturalistic approach in which the proper controls are not experimental but observational and statistical.

In naturalistic studies the eternal hope of the investigator is that nature will in time provide the manipulations of condition and control which he would like to introduce experimentally.

Perhaps, circling back to Strupp's plaint, this goes almost as far as Loevinger's statement (leaning actually too far the other way) that

there are other ways of learning things than through scientific research. Many of us claim to have learned something from case histories, particularly those of Freud, but that does not make case studies (by themselves) into scientific research (Loevinger, 1963).

We may here foreshadow the central issues to be developed in detail later in the paper by summarizing important methodological problems of the *informal* case study method which must be considered in transforming that method into a more reliable and objective research instrument: (1) the basic observations are not ordinarily public, i.e. they are directly available only to the treating analyst, and are not available to independent, concurrent observation; (2) the ways in which the observations are usually reduced, ordered and summarized to develop or to test hypotheses are also not ordinarily public, but reflect rather the private judgements (by usually unspecified canons) of the analyst-investigator; (3) the clinical retrospective method, in which causes are inferred after the fact from the study of the consequences, involves problems of built in circularities of reasoning; and (4) there are problems of generalizing

appropriately from observations made on single or few cases.

*The dilemma of the significant v. the exact*

It should be abundantly clear from the foregoing discussion why two apparently diametrically opposed, but actually also complementary, sets of questions have arisen from the two sides of the dialogue. From that of the practising analyst, the questions have been: Why such concern with formalizing psychoanalytic research? Does that make sense? Do not all agree that it is our traditional case study method, developed by Freud, and essentially unimproved upon since, that has developed all of our really useful knowledge of the inner workings of the mind in health and in disease, of psychopathology and of its unravelling via psychotherapy? The subjectivity of our data and of our method are of its essence; why destroy its demonstrated fruitfulness in an arid quest to render it objective? From the side of the friendly non-analyst, the sympathetic clinical researcher, the questions have been: Why such lack of concern with formalizing psychoanalytic research? Does that make sense? Do not analysts realize that despite the extraordinary insights achieved by the clinical analytic case study into the workings of the mind and the way it can change, that that method is necessarily (and obviously) far less effective in verifying hypotheses (subjecting them to definitive empirical 'test'), in resolving differences between the insights (hypotheses) of different equally qualified investigators studying the same material, and thereby in moving steadily and systematically beyond what Rappaport called the 'rules of thumb' (1960) that comprise the fabric of the psychoanalytic theory of therapy (i.e. of change) towards systematized and differentiated explicit and reliable knowledge?

Among other ways, the tension in this dialectic has been expressed in Gill's aphoristic statement of the problem of formalized clinical research as that of 'the dilemma between the significant and the exact', or (in Strupp's words, 1960) between 'clinical penetration and scientific rigor' which he declares to have varied inversely in the bulk of therapeutic investigation. Loevinger (1963) aptly maintains that no one-sided resolution of this issue *should* be sought:

the function of the researcher is to look for what is objective, behavioristic, and quantifiable without losing the sense of the problem. The function of

the clinician is to preserve the depth and complexity of the problem without putting it beyond the reach of objective and quantifiable realization. As in the battle of the sexes, so in the clinical research dialogue, if either side wins, the cause is lost.

In this respect, analytic clinicians can rightly and suspiciously point to too many instances where focus on the exact (the quantifiable) has redundantly affirmed the obvious and the trivial (or has led to misleading implications because divorced of a larger qualifying context). For example, from two major and recent volumes covering the fields of psychotherapy research one can learn from one study on the 'influence of the therapist behaviors' that

the interviewers were instructed to vary their behavior along the dimensions of activity-passivity and hostility-friendliness. The interviewees were found to prefer friendly interviewers, with active-friendly interviewers best liked and passive-hostile interviewers least liked (Matarazzo *et al.*, 1966).

And from another study, that

inspection of the data suggested that persons who discontinue psychotherapy will reveal more resistance to analyzing problems than those who continue psychotherapy (Snyder, 1959).

In puristic defence of the value of these and similar 'findings', the author then states:

these facts would seem to be an elaboration of the obvious, but it is still desirable to check the obvious by research to find out whether it is really true (Snyder, 1959).

The philosopher of science, Goodman, can be quoted in effective answer to this position:

A philosophical problem is a call to provide an adequate explanation in terms of an acceptable basis. If we are ready to tolerate everything as understood there is nothing left to explain; while if we sourly refuse to take anything, even tentatively, as clear, no explanation can be given. What intrigues us as a problem, and what will clearly satisfy us as a solution, will depend upon the line we draw between what is already clear and what needs to be clarified (Goodman, 1955).

This is the point of knowing where to draw the lines so as to avoid dogma and leave room for research inquiry (with its potential for the surprise finding) and at the same time avoid the pursuit of the trivial and the compulsive need

to reprove the most obvious and best substantiated phenomena.

The serious investigator of the psychoanalytic process cannot, in our view, simply accept this dilemma at face value, and accept the necessity of choosing between significance and objectivity. He must address significant problems by as exact methods as his ingenuity and persistence can develop, recognizing that if established research methods cannot cope with the significant problems in his field, then new methods must be sought (Sargent, 1961).

*The dilemma of methods advance before substance advance*

Those who have attempted to study the psychoanalytic process and/or the outcome and effectiveness of psychoanalytic therapy in a systematic way—for example in the projects at the University of Illinois (Carmichael, 1956; Haggard *et al.*, 1965; Sternberg *et al.*, 1958), the NIMH (Bergman, 1966; Cohen & Cohen, 1961), The Menninger Foundation (Robbins & Wallerstein, 1959; Sargent, 1960, 1961; Sargent *et al.*, 1968; Wallerstein, 1964, 1966; Wallerstein *et al.*, 1956), Boston University (Knapp *et al.*, 1966) and the Downstate Medical Center and Brookdale Medical Center in New York (Gill *et al.*, 1968; Simon *et al.*, 1970)—have inevitably been more preoccupied with the development of *methods* which can be both relevant and objective (exact); which promise to validate appropriately, i.e. within the theoretical system, and yet with due safeguards against error, against circularity of reasoning and against argument by tradition and authority; and which search out the degree of precision, i.e. of mathematization, appropriate both to the nature of the data and to the investigative methods; than they have been to this point with the definitive investigation of specific *substantive* issues. We say *inevitably* because there are real and formidable scientific (as well as practical) difficulties in the way of studying the therapeutic process in psychoanalysis. These difficulties must not be allowed as arguments against formal, systematic research (the need for which is so clear in the context of the requirements for the further development of psychoanalysis as a science) but they do emphasize the compelling need for psychoanalysts with a research bent to address seriously these complex technical problems which require solution if our research is to be faithful to the complexities and nuances of mental life, and at the same time rigorous,

empirical and bold (i.e. willing to venture beyond our present understandings).

Carrying the reverse of the argument even further, exclusive reliance on the 'informal' case study method has hindered, and gives promise to more gravely hinder, the further scientific development of psychoanalysis. In its thrust to this point in history, psychoanalysis has indeed profited enormously from the natural (and the fortuitous) observations of gifted individual observers free, within a supple and comprehensive theory, to evolve ideas about how the mind hangs together, and to test these ideas by further observations of their own (and in informal consensus with their colleagues) until they arrive at an inner conviction concerning the inherent degree of plausibility (credibility) of these ideas.

Such convictions rest upon basic shared assumptions about how one acquires knowledge within psychoanalysis, assumptions that have been rendered explicit by a number of psychoanalytic theorists. Erikson (1958), in a little noticed article, spells out the nature and role of what he calls 'disciplined subjectivity' in the handling of evidence and inference. Waelder (1962) delineates this further in defending the role of 'introspection or empathy' in this regard. He decries that

the evidence that psychoanalysts are asked to supply should consist of experiments or adequate statistics, undertaken on the material of sense perceptions; no allowance is being made for the kind of reasoning that we all apply in historical matters, or for the data of introspection or empathy.

He argues that though introspection and empathy are not infallible ways to know, neither are they negligible and that this is at least one advantage our science has, *qua* science, over physics. Kris (1947) avers that interpretation works not by 'producing' recall, but by completing an incomplete memory (thereby implying that validation consists of the judgement of the goodness of fit). That is,

the situation existing previous to the interpretation, the one which 'suggested' the interpretation, must be described as one of incomplete recall (and therefore, as in some measure similar to the situation in which the memory trace was laid down).

It is Schmidl (1955) who has developed this concept of validation within the system by goodness of interpretive fit most fully, arguing

for the fit of the specific Gestalt of what is interpreted and how it fits the Gestalt of the interpretation, in which not only are inferences made from a general empirical rule to a specific case, but additionally, certain elements of the specific life experience of the patient come to be connected with each other (a homely example being the unerring fit of the two halves of the torn laundry ticket). And Bakan in his defences of introspection (1954, 1956) has provided the philosophical support of this scientific stand. He rejects the notion that we cannot know what goes on in the mind of another, since all experience is considered to be utterly and unalterably private (which he calls the 'postulate of epistemological loneliness' (1956)) in favour of the common sense view that prediction and control *can* operate to some extent in the affairs of men because 'after all, we are all pretty much alike'.

But at the same time we must agree with Kubie's cautious statement (1956) about the limits of the explanatory precision achieved within this approach to the crucial problem of validation. He stated that in analytic work we arrive at circumstantial evidence of the plausibility of an interpretation (at best), not yet of its unique necessity. If we are to move beyond these criteria of plausibility sustained by inner conviction towards the more usual scientific criterion of replicability—in the sense of repeatable, publicly demonstrable, and consensually shareable observation—then the procedures of traditional psychoanalytic case study need to be formalized, especially in their verification phases.

Psychoanalytic investigation to this date, relying almost exclusively on informal methods of case study, has underrated, and has paid insufficient attention to, just these complex problems of systematic hypothesis testing and consequent logical theory building (the 'systematized and tightly knit theory' (1960) to which Rapaport felt psychoanalysis could aspire). It has fallen short in its attention to these problems of hypothesis testing and theory building as they necessarily go beyond just satisfying the investigator and his colleagues that they are right, beyond, that is, the usual criteria of plausibility, sense of inner conviction, and absence of conflict with accepted authority. As one consequence of this limitation in its research tradition, psychoanalysis has been left without adequate scientific means for resolving disagreements. How is a scientific issue resolved

when the experts disagree in their observations or their interpretation of the data? When recourse to personal observation and personal experience leads to differing conclusions? Psychoanalysis has been notably deficient in responding effectively to this 'consensus problem' (Seitz, 1966) and has thus had to rely excessively on authority rather than on verifiable empirical evidence in determining what is 'correct' when experts differ.

Relevant here is the point that psychoanalysis is not a tightly-knit and unified theoretical structure; it is rather a congeries of theories, varying in linked, and of various degrees of comprehensiveness, logical clarity and explanatory range. The psychoanalytic theory of personality is, for example, an already well-developed and far ranging theoretical structure that enables us to array meaningfully the major phenomena of mental development and functioning in their normal and abnormal manifestations. By contrast, the psychoanalytic theory of therapy, i.e. of change, is still a fragmented series of mostly *ad hoc* and pragmatic propositions, called by Rapaport for the most part 'rules of thumb' (1960), unclearly interlocked, substantially unverified and incompletely related to the theory of personality structure. The point here is not simply that our knowledge (like all knowledge, and in even the most advanced of sciences) is incomplete, but that it has a special type of incompleteness and of informality intimately dependent upon its origins in the informal case study method that has characterized psychoanalytic research; and further, that however adequate or inadequate rules so generated may be for the art of analysis, they cannot be made to constitute an evolved, explicit scientific theory.

Our thesis then is clear, that for purposes of further development of the science, psychoanalysis must devise more formal and systematic means of clinical investigation (of investigating the phenomena of the consulting room in ways that are simultaneously clinically and scientifically relevant) to supplement our continuing accretions of insight by our informal case study methods; hence too, the priorities of methodological preoccupations amongst our clinical researchers. In the following sections of this essay we shall consider a number of these important issues in taking steps beyond informal case study towards systematic clinical research. The problems are, in many instances, straight-forward, though the solutions may not be at all

certain. The issues are not new; we have already, we think, demonstrated how repeatedly they have been raised inside and outside the psychoanalytic literature. We address them here now to convey our current perspectives (born out of current research activities and research preoccupations) on the tasks and technical problems which require consideration as research psychoanalysts and their research colleagues undertake more systematic efforts.

HOW MAY THE BASIC DATA OF PSYCHO-ANALYSIS BE MADE AVAILABLE TO SCIENTIFIC STUDY?

*The historical method of 'memory only'*

The issue here is of erecting the conditions that allow independent and concurrent observation of the phenomena of the analytic consulting room. The historical method, bequeathed to us by Freud, consists of the study of the case material as reconstructed by the analyst, perhaps after completion of the case, in more or less detail, and with indeterminate accuracy, omissions, and selective distortions (biases). Freud's advice (1912) proscribing any note-taking during the analytic hour was quite categorical (and has probably been abided by, by practising analysts, *including* those engaged in psychoanalytic research, as much as any of his technical dicta). He stated simply that note-taking in analysis, because it meant a conscious, i.e. logical, secondary-process, selection of data was inimical to the proper analytic stance of evenly hovering attention, that Freud intended as the analytic counterpart to the patient's effort at free association; such focused selection would be clinically detrimental to the analytic unfolding, in effect partially cancelling the gains in advancing the analytic work achieved through the free association effort, and also incidentally deflecting the analyst from the fullest attention to his interpreting task. Freud said:

the technique, however, is a very simple one. As we shall see, it rejects the use of any special expedient (even that of taking notes)... I cannot advise the taking of full notes, the keeping a shorthand record, etc., during analytic sessions... a detrimental selection from the material will necessarily be made as one writes the notes or shorthand, and part of one's mental activity is tied up in this way... No objection can be raised to making exceptions to

this rule in the case of dates, the text of dreams, or particular noteworthy events which can easily be detached from their context and are suitable for independent use as instances. But I am not in the habit of doing this either. As regards instances, I write them down from memory in the evening after work is over...

Freud did toy with one other (for us, very germane) possible exception to his general rule, that

Taking notes during the session with the patient might be justified by an intention of publishing a scientific study of the case. On general grounds this can scarcely be denied.

But he felt this potential gain to be not only not worth the effect on the process, but also less substantial than it might appear on first thought.

It must be borne in mind that exact reports of analytic case histories are of less value than might be expected. Strictly speaking, they only possess the *ostensible* exactness of which 'modern' psychiatry affords us some striking examples. [*Italics his*].<sup>3</sup>

In summing up, Freud, in fact, talked of the (at times) divergence of the therapeutic and research aim in analysis as follows:

One of the claims of psychoanalysis to distinction is, no doubt, that in its execution, research and treatment coincide; nevertheless, after a certain point, the technique required for the one opposes that required for the other.

One other advantage of his data reconstruction method, not mentioned by Freud at this specific point, but an issue with which he gave ample evidence of being much concerned, is that the avoidance of note-taking obviously makes for the most secure ethical (as well as necessary technical) safeguards of the patient's right to proper privacy and confidentiality.

Nevertheless the many scientific liabilities of an essentially totally private and memory-based method of data generation are obvious and have been in part already presented under Shakow's (1960) and Glover's (1952) statements on the limitations of the traditional case study method as a source of continuing knowledge in psychoanalysis. The fallibility of human memory is itself the focus of a vast body of research in general psychology. The diagnostic (one could

<sup>3</sup> See in this connection the earlier statement of Glover's (1952) quotation from Freud on a very closely related but not identical point (p. 12).



also by extrapolation, read, therapeutic) interview as a data collection instrument has been specifically investigated by Haggard and his co-workers (1960, 1968) from the point of view of the reliability of anamnestic data. Analysts will not be surprised at the astonishing variations (that is, unreliabilities) revealed by such reported systematic studies of factors determining memory report. Lustman (1963), in a comprehensive review of issues in analytic research, commented on just this point: 'I consider it peculiar that analysts who every day deal with the vagaries of memory would trust their own memories in terms of scientific data.'

Beyond these issues of fallibility, the reliance on the treating analyst's memory alone as the central (and unsupplemented) data gathering instrument in analytic research is beset by many other equally powerful limitations. These have been piecemeal adumbrated, and will here only be again listed. (1) The unchecked, or at least not systematically checked, subjectivity of the observer (see Gill's remarks on the inadequate attention to the countertransference as a potent distorter of the data perceived and reported). (2) The unknown systematic biases introduced by any process of selecting data for presentation according to unspecified canons of procedure for determining relevance (an issue linked to but by no means co-extensive with possible countertransference distortion). (3) The lack of public character so that concurrent or independent evaluation is precluded. (4) The lack of repeatability or reproducibility of a forever vanished circumstance. Clearly systematization must begin, rather than rest, at this point.

#### *Process notes as the primary research data*

At first glance, the use of process notes (a summary by the analyst-observer of his most salient observations), even assiduously written and in considerable detail right after or at least on the day of each analytic session, would not seem to promise any significant advance over the above described schema of memory only, supplemented by ad hoc note-taking, and retrospective reconstruction after the treatment. Process notes seem indeed subject in full measure to some, and in some measure to most, of the limitations already adduced for the 'memory only' method. Certainly the charges of both incompleteness and distortion seem

equally applicable. For after all, process notes, however detailed, are a highly selective sampling of the universe of events actually occurring during the analytic sessions. The sampling is biased by the expert's selective emphasis on that which he perceives as relevant, as well as by other uncontrolled factors already mentioned (limitations of perception and memory, unconscious forces making resistances) which introduce an indeterminate degree of distortion.

But despite these real and obvious limitations, process notes may have some formidable advantages for clinical research purposes, and may indeed have a critical place at this stage of research on the psychoanalytic process. Keeping and studying (working from) systematically recorded process notes as the essential research data may actually be a large step forward in the direction of formalizing and systematizing the clinical research enterprise. The notes provide a permanent and 'public' record of a systematic series of observations by a highly trained participant observer. They are a record which therefore does allow of independent and concurrent observation and study (though the record itself may, in fact *must*, distort the universe of events of which it is a sample). They may be obtained with relatively little special effort and with minimal disruption of the natural analytic situation. Daily process notes are fuller, more detailed, and more descriptive-observational than notes written for summary purposes to be periodically entered into a chart; at the same time they effect roughly a fifty-fold reduction of the material that would be obtained per session from a typescript of a tape-recorded hour. There is the vast gain then of having a (relatively) brief account (even of a long analytic treatment) which the human mind can process so that large sweeps of material, disclosing major configurations and sequences of change, can be encompassed. In specific therapeutic research projects with which we have been associated, the Psychotherapy Research Project of the Menninger Foundation (Wallerstein *et al.*, 1956) and a current one, the Therapeutic Process Study on Modification of Defenses in Psychoanalysis,<sup>4</sup> process notes have yielded (even more than) sufficient detail for the specified research purposes, which require examination of both long sweeps of analytic change and of more subtle component shifts occurring over periods of but a few sessions. Such notes are

<sup>4</sup> Referred to as PRP and TPS.

obviously unsuitable for study of minute-to-minute shifts in impulse-defence configurations, for close study of verbal interactions, or for any kind of micro-analysis; but they do allow (may indeed enhance) detection of recurrent configurations (not buried in such a morass of deadly detail). Notes for virtually a year of analysis may be read in a few hours, and the clinically experienced reader will retain a gross picture of the overall course of events, and perhaps of some obvious turning points. The material may be re-read repeatedly, and gone over together in a group, until each investigator has a highly differentiated cognitive map of the terrain. Patterns may be tentatively identified, and then checked and rechecked against observations; similarities and differences between groups of sessions become visible.

The implication should be clear in all of this that to the extent that process notes represent equally well *the phenomena of particular research interest* as do the more complete verbal account of the audio-tapes transcribed into the verbatim typescripts (a still open, empirical question); that is, to the extent that they are 'good enough' or 'equally good' for the particular research purpose at hand, process notes may indeed have research advantages over the theoretically fuller recording of the audio-tape, since they can well make not only for greater manageability of data (clearly so) but also for greater visibility and hence extractability of (centrally relevant) data. And they can have a further evident advantage in the very realm of completeness itself. For the usual assumption in the use of the (unsupplemented) tape-recording as the essential research data is 'that analysis all takes place in the space between the analyst and the patient and that it is fully represented by their verbal behavior' (Schlesinger, 1967). But it is clear that what is here necessarily missing is the unspoken activity in the analyst's mind, the overall meanings he grasped in the patient's material, and the basis on which he decided to intervene when and in the ways he did, deciding what to say, and what not to say. For in analysis, only the patient is under the injunction to try to say everything that comes to his mind. It is, however, precisely this mental activity by the analyst that is (should be) captured by properly written process notes. A proper prescription for the latter might include: what the patient was conveying; what the analyst

understood the patient's material to mean; and the basis in that understanding for the analyst's technical interventions—and his withholdings. In this area, process notes can have a 'completeness' denied to the tape-recording.

That this process can work, and in this way, is attested to by the reliance on just this data method in our entire clinical and teaching enterprise in psychoanalytic work, supervision, the case conference and the continuous case seminar. With all the acknowledged limitations in the method, people do conduct psychoanalyses, and learn under supervision how to do so, with process notes as the primary media of exchange and study. That the clinical and educational enterprise 'works', however imperfectly, is not a scientific argument but it at least makes of this an issue worthy of systematic scientific scrutiny. This issue is the particularized version of the question—how good/bad are process notes anyway?—particularized to the modifying question, for what purpose? The clinician has the impression that he *can* supervise a case from process notes, and be reasonably comfortable that he has a fairly good grasp on what is going on. There is no real body of other empirical evidence—pro or con—in regard to this issue, which is not the same issue as whether distortion occurs (on which everyone agrees). We need to know rather whether for a particular purpose process notes provide a 'good enough' account, or perhaps at times even a better one.

For example, a study group decides and demonstrates from process notes that affect expression increases, differentiates and changes qualitatively in particular ways during the first year of analysis. The real question about process notes here is not whether they picked up all evidences of affect (obviously not), nor whether they distorted the affect representations in certain instances (obviously yes), but whether the patterns of change demonstrable in the process notes would also be documented in the same ways in the actual verbal transactions between the patient and analyst as revealed in the tape-recordings of the first year's sessions. Knapp *et al.* (1966) are one of the few groups who have collected such side-by-side data (tape-recording and independent process notes) and subjected them to comparative study in order to begin to provide empirical evidence on this question of such major import for clinical practice and teaching, as well as research.<sup>5</sup> The

<sup>5</sup> Knapp *et al.* (1966), for instance, have presented a detailed side-by-side comparison of tape-recordings and

process notes from segments of two sessions and concluded that the process notes preserved a great deal of

TPS is similarly turning to such a study which will use side-by-side process notes and tape-recordings in two ways. The first is to determine whether significant changes demonstrable in or by the process notes are demonstrable in similar ways and to comparable extent in the recorded material, i.e. does the process note sampling significantly distort the universe in regard to the phenomena of interest? For example, if more affect expression is found after a particular defensive shift than before, will this finding hold in the study of the tapes or the typescripts of the relevant blocs of hours? Second, the tape-recording *can* be used to *amplify* hypotheses about changes and the sequence in which they occur by reviewing key sessions selected for relevance on the basis of the process notes.<sup>6</sup>

Nevertheless, and despite this array of real advantages, process notes continue to be a biased sampling of the universe of events which interest us. Many investigators in calling for recourse to verbatim recording have squarely challenged any reliance on the therapist's process notes. Shakow (1959) has urged that investigators

Love, cherish and respect the therapist—but for heaven's sake don't trust him . . . It is the skepticism which recognizes the problem of achieving objectivity of report in any kind of situation, and the particular limitation of the human organism as a reporting instrument in interpersonal situations. Because of this fact, I find it difficult to accept . . . the possibility of studying psychotherapy process through therapist's notes.

Our own position is clearly more tempered. Despite all that we know about distortion, and despite the striking demonstrations that distortions of magnitude and significance regularly occur,<sup>7</sup> we nonetheless dissent sharply from the widely held view that process notes are useless

the essential material, including even 'a precise pattern of associations, twice repeated'. On the basis of many such comparisons, they concluded that 'In spite of all they necessarily omit, such notes, which can be more thorough in the analytic situation than in most, give valuable information about on-going day-to-day process.' Knapp and his colleagues, however, also commented that the analyst is inevitably bound to be a poor observer about himself, and that where precise knowledge of the therapist's activity is necessary 'one must respect the inevitable limitations placed upon him as observer and find other sources of information'.

<sup>6</sup> Sargent (1961) has characterized this tactic as the 'Zoom' model of investigation. 'In Zoom finding, the entire view is first surveyed—the ordinary data of routine evaluation—but with provision for magnifying and bringing into closer focus any area which, in the first sweep, appears to be of interest.'

for serious objective work. In fact we hold the almost contrary view that for many kinds of clinical research questions they may be *as good*—if indeed not significantly better (for the reasons already adumbrated).

#### *Verbatim recordings as the primary research data*

Verbatim recording has in recent years been increasingly vigorously proposed, as the essential methodological advance if research into the psychoanalytic process is to progress beyond the inherent limitations of subjective memory-based data upon which all other approaches rest. It was introduced into clinical psychoanalytic research as early as 1933, when Earl Zinn was known to have made dictaphone recordings of psychoanalytic sessions with a patient at the Worcester State Hospital (Carmichael, 1956). Since then it has been advocated and used by a widening array of analytic investigators (Bergman, 1966; Brenman, 1947, 1948; Brody *et al.*, 1951; Bronner, 1949; Carmichael 1956; Gill *et al.*, 1968; Haggard *et al.*, 1965; Knapp *et al.*, 1956; Kubie, 1958; Roose, 1960; Shakow, 1960; Simon *et al.*, 1970; Sternberg *et al.*, 1958); certainly Freud's early expressed concern that the introduction of any such outside element would not be possible has been amply laid to rest. Freud (1916-17) had stated;

The talk of which psychoanalytic treatment consists brooks no listener; it cannot be demonstrated . . . the information required by analysis will be given by him only on condition of his having a special emotional attachment to the doctor; he would become silent as soon as he observed a single witness to whom he felt indifferent.

Patients, in fact, do not become silent under these circumstances, which have included by now not

<sup>7</sup> Kubie (1958) offers the following vignette in illustration of the subtle tricks of forgetting and false recall to which the human mind is prone. A young therapist who regularly brought the tape-recordings of his therapy hours for discussion in supervision reported on one occasion that this time he had only an incomplete record because the patient had exercised his option to ask in the middle of the session that the recording machine be turned off so that he could divulge some particularly painful material with less embarrassment. After full discussion of the possible dynamic reasons for this circumstance, the recording was turned on to be played up to the point of interruption. It turned out that, faced with the patient's evident reluctances in the hour, the *psychiatrist* had suggested that the patient would feel more comfortable if the recording were turned off. He *had no memory* of this.

only audio-taping but also the filming of analytic treatments (Bergman, 1966; Carmichael, 1956; Van Vlack, 1966).

The most obvious advantage of the recorded session is that of the greater completeness (and permanence and public character) of the data record. Gill and his co-workers (1968) contrast this with the (by now oft-repeated) limitations of the analyst's memory, necessarily distorted, since he remembers and synthesizes according to his conception of the case, since he is (unwittingly but necessarily) influenced by the nature of his therapeutic commitment, and since any such intense human relationship is inevitably clouded by *scotomata*. Kubie (in Brenman, 1947) further contrasts recording as an operational technique with the difficulties that arise in any effort at similar comprehensiveness of data recording by any other method:

Months of daily observation of a process which waxes and wanes continually by just perceptible increments and decrements gradually dulls the perceptions of even the keenest observer, paralyzes the memory through the monotony of repetition, and renders the written word literally useless as an instrument of record... Clinical theorizing, however brilliant, does not become clinical research until it is buttressed by precise, repeatable observations, accurately reported.

Elsewhere he says even more categorically that

It is an implicit denial of all that analysis has taught us when we pretend to ourselves that a student's report to his supervising analyst can provide a true and representative sample of that which has actually taken place in the analytic sessions which he is attempting to describe (Kubie, 1958).

Gill and his co-workers (1968) additionally refer to one other, to them equally compelling, advantage of tape-recording, beyond the improvement of the data record, i.e. the facilitation of the separation of the therapeutic from the research responsibility (with the possibility of thus bypassing the inevitable biases of the analyst as a contaminant of the data filter). Haggard and his co-workers (1965) in fact wonder if valid research on the therapeutic process *can* be done by the treating analyst at all, if

By objective research we mean procedures for data collection, analysis and interpretation which do not

rely on the perception, judgment, appraisal or memory of a single individual—especially if he is personally involved in the therapeutic situation or process.

Kubie (1958) again advances the argument more categorically, when he declares of the analyst as participant-observer, that

In other words, he has simultaneously to be a free reactor, a participant in a complex emotional interchange, an observer, a recorder, and an objective recounter of this whole intricate chain of events. It is wholly un-analytic to assume that this is possible.

And Freud (1912), too, can be quoted in implicit support of this position that the analyst should not attempt to do treatment and research at the same time since cases which are 'devoted from the first to scientific purposes and are treated accordingly suffer in their outcome'. With all of this we are meant to wonder in the words of Haggard *et al.* (1965) whether the treating analyst can 'do research under these conditions with any reasonable degree of accuracy, reliability and effectiveness?'

Bergman (1966), who has probably conducted the only several hundred hour-long completed analytic treatment with recording (in fact, sound filming) of every session, felt the fact of the recording itself to become only a part of the 'structure' of the treatment. Psychotherapy has many elements of structure which provide its background and stable context that we regard as 'natural', though there is nothing natural or usual about them in human discourse. For example,

In the customary office setup of psychotherapy three conditions are taken for granted as part of the background: That the patient meets the therapist in privacy; that the meeting lasts a predetermined amount of time, no more and no less; and that the patient pays the therapist for his services. There is nothing 'natural' about any of these conditions... (Bergman, 1966).

To Bergman, sound-recording is just another arbitrary convention which if used widely enough becomes taken for granted as another 'natural' element of the 'cultural mores' of psychotherapy. (Most psychoanalysts would of course dispute this stand, feeling that however much recording may be useful in the service of—at least some kinds of—analytic research, its

role in the therapeutic process cannot be dismissed as just another contextual element whose impact need be considered no further.) In summing up the argument for the necessity of recording, Shakow (in Bronner, 1949) has declared it to be essential to the proper data-gathering for the naturalistic-observational phase that psychoanalysis at times seems in a hurry to bypass, though it can no more afford to do so than can any other developing science.

Which is not to claim that recording poses no problems to analysis. The most manifest is the possible inhibiting effect (or distorting effect) upon the conduct of the analytic work. Historically, therapists have always couched their reluctances to record in terms of the presumed anxieties and sensitivities of their patients, but it seems to be an exceptionless experience that where recording has been done it is the therapist who has been the more anxious and disturbed. The patients have seemed less manifestly distressed, and have accommodated more easily and quickly. Haggard *et al.* (1965) have noted the possible contagious impact of the therapist's doubts and anxieties; they state too that therapists who are made anxious tend to exaggerate the extent to which their patients are anxious. These anxieties of therapist-analyst tend towards a particular patterning. There is the initial anxiety around putting one's professional practice and competence on the line this way. There is then some (relative) accommodation, followed often by a second wave, a recurrence of almost equally severe anxiety, over the new dread of 'seemingly never ending exposure' (Sternberg *et al.*, 1958).

These anxieties generated in the analyst have several identified sources. Concern over professional exposure is of course uppermost. In this sense, a distinction made by Bergman (1966) is apt, that the threat posed by the recording to the patient is neurotic, but to the therapist the fear of anticipated criticism from colleagues is a reality threat—hence the harder to dispel. Gill and his co-workers (1968) talk, too, of a deeper and less overt source of anxiety in the analyst. They quote Wheelis's (1956) phrase that doing analytic therapy in confidentiality 'is an unparalleled kind of intimacy within a context of impersonality' and that there are gratifications in such a situation which lead analysts to resist intrusion.

Autonomy is always relative, and the power and pervasiveness of infantile drives is such that they

must find an expression... The analyst would have to yield some of these gratifications if he opened his work for inspection. Even more, he would have to face criticism of the searching kind to which only one analyst can subject another and for which tape recording yields inexhaustible 'ammunition' (Gill *et al.*, 1968).

No wonder that Carmichael (1956) talked about how the majority of analysts whom he approached to participate as therapists in such a project expressed great interest in it but evaded a personal commitment; they did not have time or did not have a suitable patient. Carmichael said that they 'preferred to remain at a distance from it, expressing doubts about the validity with which the therapeutic process could be represented under such conditions', a viewpoint which he characterized as both a legitimate issue and a rationalization. He wondered whether it betokened narcissistic or exhibitionistic or voyeuristic or masochistic impulses to let oneself be filmed (he used all these words). Shakow (in Bronner, 1949) among others used the word 'courageous'.

Whatever the case that can be made out for the presumed necessity of recording for (at least some kinds of) psychoanalytic research, an equally salient question is then that of the impact of the fact of recording upon the analysis (again, no matter how well analyst and patient can themselves seem to be able to accommodate to the intrusion). The question here is, is the analysis vitiated in any serious way as an analysis? This is not the same as the question, will the recording have an effect (as a major parameter)?—which no one can deny. Rather, as put by Haggard *et al.* (1965):

the question is not, Would the therapy have been exactly the same if there had been no recording? but is rather, Did this particular therapy, even though recorded, possess those components—free association, transference, interpretation, and so on—which characterize and must occur in psychoanalytic therapy?

Gill and his co-workers have addressed themselves most extensively to this problem (1968). They see the (recorded) research analysis as sharing at least two major attributes with another kind of different-from-usual analysis, the training analysis; these are (1) the absence of full confidentiality, and (2) the existence of goals in addition to the therapeutic. These they acknowledge to constitute special

problems for the analysis; 'difficulties are introduced which may on occasion be enough to tip the scales against success. By analogy, however, we argue that a recorded research analysis is not in principle impossible' (any more—or less—than a training analysis). Obviously, any aspect of the reality (research) context or any technical parameter of an analysis may be used defensively. For example, the research goal may be willingly accepted and

if he accepts it he will lend himself to the research to defend himself against understanding what it means to thus 'lend himself.' And (conversely) if he rejects it he will use the research goal to justify his refusal to examine himself.

Similarly, for the usual taken for granted safeguard of utmost confidentiality, 'Whatever the desirable and rational reasons for maintaining it, it can carry hidden irrational and transference meanings as well.' Their own recorded case showed a contrary defensive use of her ready acceptance of the lack of full confidentiality: 'She has come to recognize that she would fear the *non-research* situation because it would be one of excessive intimacy with the analyst.' The proposed handling of these issues is, of course, thorough analysis, which they feel is 'not in principle impossible'. And this cuts both ways. *Any* analysis has a reality context so that 'It is not merely *deviations* from the *usual* analytic situation which must be analyzed but so must the *usual* situation'. And they conclude with a caveat:

We recognize that the demonstration that part of the usual context of an analysis may be exploited for defensive purposes does not mean that it can necessarily be dispensed with as part of an analysis.

Apropos of the inevitable distortions of the analytic process introduced by the fact of recording, Knapp *et al.* (1966) point to a mutually cancelling effect:

Actually there seems to be a continuum, proceeding from direct audio-visual observations, to utilization of sound recorded tape itself, to working with typescript, to therapist's notes, or even to the report of his contacts with a supervisor. One proceeds from less to more distortion, in terms of [understanding] what actually took place. As

regards impingement on the psychoanalytic situation itself, the spectrum of distortion is in the reverse direction.

Haggard *et al.* (1965) take this whole issue further on to empirical ground. They bring a content analytic approach to a comparative study of material from recorded and from non-recorded 'control' cases.<sup>8</sup> They found no overall difference in the amount of concern expressed, for instance, by patients over revealing intimate thoughts and feelings and having them scrutinized by another person; but with recorded subjects these concerns were mobilized earlier in the treatment course (i.e. the impingement of the reality forced the pace of the material). They summarize that there is no conclusive empirical evidence as yet, one way or the other, in regard to the existence of an undue (significant) distorting effect.

Others, however, have felt differently. Roose (1960) reported on the research analysis of an asthmatic patient in which the research required tape-recording and other intrusions as well (concomitant physiological tests conducted before and after the analytic hours). All departures from usual analytic practice were designated either as parameters (the absence of fee, the hospital setting of the treatment, etc.) or as contaminants (which included the recording), but the criteria for distinguishing between these classes were not stated. The analysis was overall unsuccessful; Roose felt that because of all the contaminants, the parameters (usually analysable) could not be analysed. He underlined his negative conclusions as follows:

Recording tends to reinforce the use of words by the patient for the function of appeal (Bühler's classification). Words become endowed with magical qualities... The recording corresponds too closely to the infantile belief in the omniscience of the parents. To give the appearance of reality to this infantile belief may diminish reality testing to the point that the reconstruction of the infantile neurosis may be impeded if not blocked (Roose, 1960).

Schlesinger (1967), in discussing Gill's paper, raised this same question of ultimate analysability of the research dimension. One can analyse neurotic embellishments but not reality.

<sup>8</sup> In the non-recorded 'control' case the material subjected to analysis consisted of the detailed notes

taken by the analyst for the (usual) purposes of analytic supervision.

'Clearly a limit is set to the psychoanalytic work that can be done when a parameter is introduced that in principle cannot be liquidated.' Surely the solution to this contention would come in making the issue over the phrase 'in principle' one to be decided empirically rather than a *a priori*!

Certainly all do agree, even the most vigorous proponents of recording, that there is an (indeterminate) impact on the analysis, more overt in the stress aroused in the analyst than in the patient, leading to undetermined distortions in the unfolding analysis, and susceptible to analytic resolution to an unknown (to be determined) degree. To aid in the handling of analytic stress under such circumstances, almost all have come explicitly to feel the need for ongoing supervisory and/or consultative help, regardless of how experienced and sophisticated—clinically and research-wise—the analyst (Bergman, 1966; Knapp, 1968; Knapp *et al.*, 1966; Shakow, 1960; Simon *et al.*, 1970). For after all, the analyst has many special needs in such a research undertaking; the need to prove something, to make a research contribution; the need to look good, to have ideal results; the needs aroused around having a 'special case' (Knapp, 1968). Simon *et al.* (1970) put this argument on a theoretical basis as well. If others than the treating analyst are party to the analytic proceedings (by definition in such research study based on recordings); then there is inevitably a judgemental and evaluative aspect to the study. In that sense the research is psychologically akin to supervision and, the authors contend, should be frankly faced on that basis. What follows as corollary to this is that the therapist must be part of the research just as the supervisee is part of the supervision.<sup>9</sup> In another place they state this more militantly:

The analyst's vulnerability to criticism makes it desirable if not essential that he have maximum opportunity to confront his potential critics (Gill *et al.*, 1968).

Our own position is fully in accord with this view not only for the analyst who tape-records a case but for every analyst whose analytic material is offered for research study in whatever form. The analyst outside the research-study group will lack adequate motivation to do his

part of the (data-producing) work; and money and promised participation in authorship are unsatisfactory rewards in such situations. The treating analyst who is initially outside the research group will want to be inside it, and if he is in it at all, he should be in it all the way. Whether in or out of the group, the analyst will be concerned with the scrutiny of his work by professional colleagues, and we share the many views quoted, that these anxieties are most likely to be mitigated by being part of a freely communicating group of friendly colleagues. The group's inhibitions about criticizing a colleague (or temptations to do so) are part of the psychological reality of this type of study, whether or not the analyst is part of the group. His participation in the group is more likely to permit mastery of these inhibitions and temptations, and achievements of relative objectivity, than the more uneasy situation of discussing the case (as it would seem) behind his back.

Enough has been written to indicate the many difficulties that beset the seemingly simple and direct solution to the data-gathering problem of psychoanalytic research via the verbatim recorded case. In addition to the major areas already delineated; the inevitable, and undetermined distortions of the process under scrutiny; the personal demands upon the participants, especially the therapist; the invasions of privacy and confidentiality, with both the scientific and ethical implications; in addition to these difficulties, there are others of a more practical kind. These, just to list them, include the sheer volume of data accumulated (Shakow, 1960); the very high costs of such an enterprise (especially if it is a matter of sound-film recording as has been proposed and carried out—for a full analytic treatment); and the many technical barriers, for instance with the sound film,

the difficulty in reconciling the physical presence and the demands of the camera—even though it is in another room and concealed—with the needs of the therapist and patient (Sternberg *et al.*, 1958).

Kubie has in fact proposed that the proper marshalling of the human and technological resources for this task can only be successfully undertaken within specially created and supported 'Neurosis Treatment Centers' (in Bronner, 1949) staffed by full-time research analysts.

<sup>9</sup> Though of course they acknowledge that it leads to other complications design-wise, those of research bias, if the analyst is a member of the research group, party to

its hypotheses. But if he is not, he will feel excluded from it and may in turn regard the research as an interference or a threat.



In all of this enterprise, of making the data of clinical psychoanalysis available for research study, the possible mutual interaction—facilitating or detrimental—of therapeutic and research goals and outcomes upon one another has been subject for much considered concern. Everyone of course agrees that given established evidence of these goals being in conflict with, or inimical to, each other, that the research ambition must give way to the therapeutic exigency. This is so even from a narrow research concern just with the value of the findings being generated. As put by Haggard *et al.* (1965):

If the purpose of the research is to study the psychotherapeutic process in its naturalistic setting, if the therapeutic goals were not met the research goals would not be met either, since they involve the approximation of the normal therapeutic situation.

The differences here are of interpretation. What some take as self-evident, that analysis is in principle irretrievably altered by the formalized research imposition, is to others a yet to be decided empirical issue, with a cogent theoretical rationale that could undergird the opposite position (cf. Gill *et al.*'s analogy to the training analysis, 1968).

There is also a less openly talked about impact of these formal organized researches based upon study of recorded material with the enormous time and energy expenditures involved, the other way, not upon the therapy, but upon the research—and upon the clinician turned researcher. This has been candidly voiced by Cohen & Cohen (1961) as follows:

For many years we have been ambivalently interested in the systematic study of the psychotherapeutic process. In this paper we wish to report some of our experiences and current ideas. As we do so, the reasons for mentioning our ambivalence may, perhaps, become clear. At any rate, we have come to believe that it does not represent only a personal neurotic problem, but rather reflects a general difficulty and common discouragement in evolving a research methodology for dealing with problems in this complex field... secured a grant to conduct a series of personality studies of successful naval officers... [Five] of us decided to carry out this study, and encountered the vicissitudes which are all too common in this field. Each of our 700 recorded interviews (the total for three subjects) was listened to by the therapist and at least one other person. There were weekly conferences between the therapist and that other person, and once a week,

the whole group gathered for a three hour session to discuss issues which seemed significant. As might be expected, differences were stirred—both emotional and intellectual. Yet each one of us at one time or another announced how much he had gained from the total experience. But the results of this tremendous expenditure of time and thought and feeling have found expression in our individual careers and not directly in scientific publications. Although a few publications grew out of this, we found ourselves submerged in the mass of data. Each positive statement we felt able to make seemed insignificant against that which was left undescribed—and we finally turned away to follow the clinicians' usual course of focussing on some isolated aspects of the therapeutic situation.

#### *Process notes v. verbatim recordings*

It is clear from all of the foregoing discussion that, other (more arguable) advantages and, disadvantages aside, the recording on film of the analytic hour has the very real and very great advantage of providing an indubitably more 'complete' data record which successive and independent researchers may hear and see, again and again, and in slow motion, so to speak, to establish the objectivity, the replicability and the validity of observations. But this completeness too needs examination, in its reality, and also in its putative advantages. For the search for ideal completeness of data is ultimately unrealizable (and—a point we wish to make—also not the issue). Even the most faithful sound movie, most minutely studied, cannot reveal whole dimensions of highly relevant data. The obvious example, already stated, is that of the feelings and thoughts of the analyst, for the most part unuttered, which comprise his various reactions to the patient's material and the processes by which he selects how and what he will respond to. This consideration has led to the need stated by Shakow (1960) to supplement the verbatim recording of the analytic hour by the additional recording of the analyst's immediate post-session elucidation of his understanding of the session including all the associations that he could give to his (unexpressed) thought processes. In the interest of the continuing quest for ever greater 'completeness' and even greater access than this, this proposal was pyramided in a round table discussion by Shakow, Brosin and Kubie (in Bronner, 1949) to include not only sound-film recording of the therapy and the post-session recording of the therapist's associations, but also the notes made by observers of the



treatment from behind a one-way vision screen, and the relevant material from the ongoing supervisions and/or psychoanalyses of the therapist and of the research-observers; i.e. a proposal both to multiply the observers of the therapeutic process and to tap, at both conscious and unconscious levels, the undivulged reactions of the therapist and the observers through concomitant access to *their* psyches. Such a top heavy method would not only be unfeasible in any practically conceivable sense but in terms of the position we are stating, beside the point. Our thesis is rather that any manner of studying a phenomenon or a process will reveal only certain orders of data, hopefully those that are most centrally relevant to the hypotheses and the theoretical framework of the investigation, and will never be 'complete' in the abstract sense.

And the very completeness, in the sense of far greater comprehensiveness, of the verbatim record imposes formidable scientific and practical problems. The scientific problems are those of salience and relevance, of the relationship between the objectives of the research study and the kinds of data germane to it. Studies of process in the sense of therapeutic interaction, for example, that of Gill, Simon and their co-workers (Gill *et al.*, 1968; Simon *et al.*, 1970) of the close responses to various kinds of interventions may (probably *do*) require the scrutiny of microsize units of interaction, the kind of precise, minute detail available only through verbatim recording. The PRP (Wallerstein *et al.*, 1956) and the TPS are, on the other hand, alike in a greater concern with repetitive and 'molar' events and processes seen in perspective over broad sweeps of time. For these purposes, the fifty- to hundred-fold reduction in data effected by process notes, if appropriately selected and condensed in terms of relevance to the issues under study, and if safeguarded to be sufficiently representative of the phenomena existing in the greater detail of the verbatim interactions but seen the more clearly because extracted and highlighted for salience—rather than buried in the minutiae—for these purposes process notes may be 'better', not just as good. Here we can gain the advantage of the 'automatic' provision for the inclusion of the analyst's overall conceptual framework and his very specific grounds of understanding. But again, to turn the other way for a moment, and again depending on purpose, the seeming obvious advantages of access to the mental

process of analyst as well as patient, need not remain unexamined, as if self-evident. That is, we also may need to know about that only for some purposes and not for others. Gill, Simon *et al.* (Gill *et al.*, 1968; Simon *et al.*, 1970) feel—we think correctly—that the relevant effects of an intervention show themselves in the material produced (immediately or ultimately) by the patient, whatever the analyst's intentions or his *ex post facto* rationalizations of his intervention. Analytic theory as a frame of understanding needs to be applied to the material produced in order to assess the effects of the intervention, but that is not the same as needing to apply the analyst's actual inner mental processes during the session to the material. The issue is simply that this is not a general question to be disposed of by general dictum, but rather a specific to the particular research question.

The practical problems of the greater completeness, i.e. comprehensiveness, of the verbatim data record are simply stated. The material is voluminous, miles of tape, with roughly 30-page typescripts of each transcribed hour, and soon becomes overwhelming. Listening to the tapes and/or reading the material, especially when it is of someone else's therapeutic work, is enormously time-consuming and rapidly becomes inordinately tedious. We become the hapless victims of a major dilemma for research method in our field that (like another field, electroencephalography) we mainly suffer from too much, rather than too little, data. This obviously can have its advantages. As Hartmann (1958) reminds us, though we treat but small numbers of cases, each case provides us with countless instances, so that 'we can say that for every case there is often a great wealth of instances in which every single hypothesis that comes into play can be tested'. At the same time, in some respects, verbatim recording maximizes the disadvantages of this advantage.

Which is all to say that tape-recording (or for that matter any other of the data-gathering methods here proposed) cannot be (or at least should not be) categorically dismissed but cannot either be (or at least should not either be) categorically defended. The major issue of the (detrimental) effect upon the analysis is an empirical one to be answered in the specifics of the investigation, and the major issue of the value to the research is a theoretical one to be determined by the relevance of the data gathering method selected to the phenomena being

elicited<sup>10</sup> Schlesinger (1967) refers to the same issue:

No universal or categorical objection should be raised against tape recording or any other research technique. The question of the effect of tape recording in analysis has to be answered in the context of a particular patient and a particular analyst and the report of any such piece of research must of necessity include an account of the analysis of the *meanings* of this unusual element of structure in the psychoanalytic situation.

In overall summary then of what we trust is by now a sufficiently focused exposition of the development of a research posture, we believe that a permanent and public record is an extremely useful (actually indispensable) measure in moving from the traditional informal to the systematic and formalized in research on the psychoanalytic process—and that further delays in instituting steps in this direction are to the detriment of continuing psychoanalytic advance. The kinds of (appropriate) steps in this direction of making the basic data of psychoanalysis more systematically available to scientific study may be several. We do defend the value of process notes, especially for studies of relatively molar processes taking place over broad sweeps of time, while recognizing all their fallibility, their biased selectivity and their potential for uncontrolled distortion. We think though that steps can (should) be taken to improve the quality of the notes used in research studies, as well as to test empirically the suitability of such notes for the intended research purpose. In the TPS, for example, the treating analyst prepares his own summary notes of each analytic hour without ever hearing the tape-recordings of the sessions. Although he is free to include in his notes any material which he feels important, he is also guided by a checklist of features of the hour which he is expected to observe and record *invariably*. These features include types of observations deemed relevant to the testing of this particular research group's explanatory concepts concerning the therapeutic process and the mechanisms of change. The instructions also facilitate direct comparisons between process notes and typescripts of tape-recordings for selected categories of material, such as 'new memories', and indeed the TPS is planning a systematic study of how well such process notes

(when compared to typescripts of the tape-recordings) pick up observations crucial to their investigation. And of course, if such demonstration be successful (for those particular purposes), the enormous research advantage would then accrue of a fifty- to hundred-fold reduction in data volume (into truly manageable proportions) without significant loss of, or alteration of, relevant information.

Similarly, we also defend the recording of analyses for certain research purposes, as the evident impossibility of studying *micro*-processes effectively without recordings, but here too, it is equally incumbent that those purposes be clear in mind beforehand and that the purposes be relevant to the kind of data (minute interactional data) being generated, because it will not really advance knowledge very far or very fast to simply have miles of tape without ideas and methods for converting that raw material into usable knowledge. At the same time we need to urge the continuing careful study of the impact of the research procedures (not only recordings) on the processes being studied. We say this not as unfriendly critics convinced that research will 'ruin' (i.e. irretrievably alter in ways that make useless for scientific research and/or render less effective as therapy) the processes under study, but rather as psychoanalytically knowledgeable investigators who recognize the powerful impact of methods of investigation on the human subjects of investigation, but who recognize too that the empirical evidence does not suggest that that impact renders analysis inherently impossible or too difficult (though it may do so in individual instances), or that the study of that impact is impossible, and lastly, who recognize that it is most consistent with psychoanalysis itself to bring the impact of the total research procedure—including recording, but not only recording—within the scope of the analytic understanding. In this way, psychoanalysis as science will have an enhanced potential for advance.

#### HOW MAY THE BASIC DATA OF PSYCHOANALYSIS BE REDUCED, ORDERED, SUMMARIZED?

Once made available, in whatever form (and in whatever quantity), another whole order of problems in psychoanalytic research revolves around the question of handling the data in

<sup>10</sup> The more 'minor' *practical* issues of actually carrying out the investigation and handling the data generated are all soluble (at least in theory) given

sufficient conviction about the worthwhileness and necessity of the task.

manageable ways, that nonetheless remain loyal to the subtlety, the complexity and the richness of the clinical phenomena, again the 'dilemma between the significant and the exact'. Here too there are thorny unresolved conceptual and methodological issues which will be delineated under the following headings: (a) Of what order are the data? (b) How do the data relate to the concepts? (c) The use of clinical judgements or inferences as 'data'? (d) The 'consensus problem', or what to do when experts disagree?

*Of what order are the data?*

At their most patent, the data of analysis are manifest behaviours. Hartmann (1959) put it thus:

The data gathered in the psychoanalytic situation with the help of the psychoanalytic method are primarily behavioral data; and the aim is clearly the exploration of human behavior. The data are mostly the patient's verbal behavior, but include other kinds of action. They include his silences, his postures (cf. F. Deutsch), and his movements in general, more specifically his expressive movements. While analysis aims at an explanation of human behaviors, those data, however, are *interpreted* in analysis in terms of mental process, of motivation, of 'meaning'; there is, then, a clear cut difference between this approach and the one usually called 'behavioristic'... [Italics ours.]

That is, the data of analysis are manifest behaviours, *as interpreted*—by (to be specified) canons, according to which such interpretations can be consistently and reliably made.

The successful solution of the problems inherent in arriving at such consistent and reliable interpretation (problems to be considered from successive vantage points through the subsequent headings of this chapter) rests initially on two necessary conditions: (1) the maintenance of meaning without significant loss or distortion, as the data are compressed and are isolated from their qualifying context; and (2) adequate definitional clarity, of the nature of the events and phenomena, and the concepts according to which their meanings are understood. On the first of these, Lustman has said of the problems of categorizing and indexing psychoanalytic material for intra-case and inter-case analysis, in the Hampstead Child-Therapy Clinic,

The success of this approach with large masses of psychoanalytic material coming from large numbers

of patients remains to be demonstrated. By that I refer to problems of compressability, fragmentation, and the loss of meaning when taken out of context... [Furthermore] yet to be solved are methods of correlating data of disparate origin—such as one has in multidisciplinary longitudinal studies, which include psychoanalysis (Lustman, 1963).

The second of these necessary conditions, that of adequate definitional clarification, likewise represents a still major unsolved issue for psychoanalysis. Though psychoanalysis has a well established theoretical structure of explanatory constructs, on at least six levels of conceptualization and generalization varying systematically in remoteness from the observational base, and in centrality and importance for the theoretical structure (Waelder, 1962), it has not been able to achieve precise definitional clarification of even its most fundamental and most pivotal concepts. This is so even in regard to the boundaries encompassed by the term psychoanalysis itself (and even when delimited only to psychoanalysis in its aspect as a therapy). For example, an official Committee on the Evaluation of Psychoanalytic Therapy (Cushing, 1952) established by the American Psychoanalytic Association, abandoned its efforts to formulate an acceptable scientific framework for the comparative evaluation of the effectiveness of the psychoanalytic therapies after five years (1947–52) of futile work because (to quote Rangell, 1954) it was

never able to pass the initial and vexatious point of trying to arrive at some modicum of agreement as to what exactly constitute psychoanalysis, psychoanalytic psychotherapy, and transitional forms.

Simple working 'definitions' can of course be formulated for ordinary clinical and heuristic purposes as in a recently published official *Glossary of Psychoanalytic Terms and Concepts* (Moore & Fine, 1967). But two examples (both from the work of the PRP) can illustrate the scientific complexities in the definitional process (and the scientific labour involved in their elucidation for research purposes) that are beyond the usual capacity of conventional, pragmatic 'working' definitions. The term and concept 'Anxiety Tolerance', crucial to the assessment of ego functioning and of major predictive import in assessing differential treatment indications and treatment prognosis, is not an official glossary term at all. In the PRP it was initially defined as follows:

We call anxiety tolerance the capacity to experience anxiety without having to act to discharge it. This can be judged in many ways . . . (Wallerstein *et al.*, 1956).

Self-evident as this definition may seem, it proved to be unduly restrictive and clearly one-sided in its actual application; therefore, not adequate (Siegal & Rosen, 1962). Through the examination of actual clinical instances arising in the research population, instances, for example, of obsessive-compulsive individuals who progressively decompensated to more regressed levels of thinking and behaving with sharply increasing manifest anxiety (more ruminative, indecisive and over-ideational) but without any weakening of the inhibition of motor discharge (i.e. without any increased tendency to discharge anxiety through action), it became evident that, *as defined*, the concept of anxiety tolerance did not (descriptively) cover this exigency. It was clear that the definition had embraced actually only the kind of behaviour arising from an alloplastic orientation, a disposition to discharge anxiety-provoking tension or a tendency to 'act out' (this latter being perhaps the most loosely used term of all in the clinical vocabulary). In order to extend the concept of anxiety tolerance to apply as well to autoplastically oriented individuals through their spectrum of degrees of decompensation, it was redefined (on the basis of the project experience) as follows: 'Anxiety tolerance is the capacity to experience signal (or 'secondary') anxiety'<sup>11</sup> (Siegal & Rosen, 1962). Simple though this reformulation may sound, it radically enlarges the definitional cover by making the concept equally applicable to all individuals, not just one-sidedly so to those who respond alloplastically to the threatened arousal of anxiety. For,

the autoplastically oriented individual for whom anxiety has become disorganizing (or primary, if you will) without an increased tendency to motor activity, has just as little capacity for the experience of signal anxiety as his alloplastically oriented, impulse-ridden counterpart (Siegal & Rosen, 1962).

(This revised definition depends, in turn, upon the concepts, i.e. *their* definitions—of two 'kinds' of anxiety, secondary (signal), and primary (traumatic).)

The other, even broader, definitional and conceptual confusion offered in illustration is over the distinction between the terms 'defence' and 'defence mechanisms'. Within PRP, this distinction was felt to be crucial, that between

*defense mechanisms*, as constructs that denote a way of functioning of the mind, invoked to explain how behaviors, affects, and ideas serve to avert or modulate unwanted impulse discharge, and *defenses* as the actual behaviors, affects, and ideas which serve defensive purposes. For example, an exaggerated sympathy can be a *defense* against an impulse to cruelty. The postulated operative mental *mechanism* by which this is explained is called reaction formation (Wallerstein, 1967).

Wallerstein (1967) then went on to develop some of the operating consequences (including for the theory of technique) of this distinction. However, it is far from a universally maintained distinction. The *Glossary of Defenses*, published by Grete Bibring and her collaborators (1961) from their investigation of the psychological processes in pregnancy, is the most comprehensive such classification yet attempted. It encompasses both the most discrete defences explicable by reference to a single defence mechanism and the more regularly recurring of the more complex defensive patterns; but it does not maintain at all the sharp conceptual distinction here advanced between defence mechanisms as constructs and defences as actual phenomena (behaviours). The need for clarification and the absence of agreement in this area of definition is presently so widespread that it almost becomes incumbent on each research group engaged in psychoanalytic research of serious scope to write its own glossary of terms, with its own idiosyncratic specifications of usage within the overall framework of psychoanalytic thinking. (As the PRP has done; unpublished Glossary of Terms.)

In summary, the data of psychoanalysis are behaviours, but only as interpreted and given meaning. These meanings must be maintained without significant deviation as the data are compressed and taken from context, and must rest on still not achieved agreed upon definitional clarification of both the phenomena and the constructs according to which they are ordered.

<sup>11</sup> This definition includes, implicitly, the ego function of discriminating the signal as such; that is not solely the tolerance for the *experience* of anxiety but the

ability, in addition, to *recognize* it as a signal. Therefore, an increase in anxiety tolerance implies improved ego functioning'. (Siegal & Rosen, 1962).

*How do the data relate to the concepts?*

This problem is multifaceted and one of peculiar difficulty for psychoanalysis because of specific problems inherent in its structure as a science. Rapaport (1960) states the general issue and its specific difficulty in application to psychoanalysis thus:

All sciences must subject observations to interpretation in order to establish their evidential significance for the theory. This is particularly conspicuous in psychoanalysis, where the concepts are by and large at a considerable distance from the observations.

This issue of the *remoteness* of the concepts from their observational base makes for especial difficulties in the task of empirical hypothesis-testing in psychoanalysis because of the additional complexities automatically introduced by the (psychoanalytically necessary) principles of *multiple determination* and of *over-determination*.<sup>12</sup> Escalona (1952) addressed these problems as follows:

The correctness or incorrectness of a hypothesis can be established only if concrete events can be shown to occur or not to occur in accordance with the hypothesis in question. Yet the very nature of psycho-analytic theory implies that altogether different kinds of overt behavior may refer to the same psychological factor. Similarly, the identical overt behavior may be the manifestation of totally different psychological states in different persons or in the same persons at different times. The principle of over-determination, to add to our troubles, asserts that one and the same behaviour at one and the same time may be the overt expression of different phases of underlying psychological processes.

Rapaport (1960), in strikingly similar language, elaborated somewhat more specifically on the need for, and the consequences of, the principle of over-determination:

Psychoanalysis' need for this principle seems to be due partly to the multiplicity of the determiners of human behavior, and partly to the theory's characteristic lack of criteria for the independence and sufficiency of causes. The determiners of behavior in this theory are so defined that they apply to all behavior and thus their empirical referents must be

present in any and all behavior. Since there is usually no single determiner which constantly assumes the dominant role in a given behavior, other determiners can hardly be neglected while a dominant determiner is explored. When favorable conditions make one determiner dominant, the investigator is tempted to conclude that he has confirmed a predicted functional relationship—as he indeed has. Regrettably, the attempt to repeat the observation or experiment in question often fails, because in the replication either the same behavior appears even though a different determiner has become dominant, or a different behavior appears even though the same determiner has remained dominant.

Though the problems here specified—of remoteness of concepts from observations, of multiple determination and of over-determination and their consequences for the issues of prediction and of hypothesis-testing, and of the lack of established correspondences between concepts and observations with the absence of accepted canons for clinical interpretation—are indeed formidable, they are added to, from the other side, by avoidable (and perhaps also, to some extent, unavoidable) *confusions between data and inferences*, and between levels of inference. For a systematic discussion of the latter difficulty (confusions between levels of inference, or of theory), see the clear—and rigorous—exposition by Waelder (1962). In relation to the former, seemingly more surprising, confusions even between data and inference (or data and concepts) Hartmann (1958) has carefully traced some of the sources. He said:

What one calls 'clinical' and 'theoretical' presentations in analysis are divergent styles of abbreviation. In this paper, I merely wish to demonstrate that due to specific features of the psychoanalytic approach the demarcation line of clinical and theoretical work is often not easily traceable... Every reading of psychoanalytic literature asks of the reader a labor of reconstruction if he wants to view it in its aspect as a scientific contribution. What was the background in terms of observables? What are the hypotheses, either presupposed by the author, or deducible from his work?... [These hypotheses] interpenetrate with fact-finding in such a way that their hypothetical character is not always clearly recognized. Highly abstract hypothetical constructs (as libidinal

<sup>12</sup> It is not implied here, or in the following, that these complications are *unique* to psychoanalytic science. Although the languages in which they are cast are different, these same orders of difficulty do exist in any

scientific field in which the theoretical foundation is complex, and in which the theoretical assumptions are at some remove from the observations.

cathexis, etc.) are then reported in a descriptive sense as data of observation.

—with consequent major confusions then in both clinical practice and theory! To describe how insidiously this process operates, Hartmann (1958) further said:

Once lawlike propositions . . . had been formulated, our knowledge of mental processes and their interactions became less tenuous and many hypotheses became less tentative in character. The analyst learned to feel more secure, more at home with them, and the corresponding concepts entered more and more the reporting of clinical material . . .

An additional danger that can arise in this connection, that of trying to resolve such ambiguities by making observational data bear the weight of too much inferential specificity, has been stated by Lustman in his catalogue of the issues in psychoanalytic research as follows:

The danger of such observational data lies in the fact that observation is of events only; *everything* else is interpretation. Thus, observational research can almost always be interpreted in many ways. Alternate theories as well as alternate hypotheses must be considered. *The greatest danger is that one may 'lend too much depth' to the observation* in order to use the data to support one hypothesis rather than another (Lustman, 1963; long italicization ours).

The implications of these considerations on the difficulties in identifying the pattern of relationship of observations to concepts—difficulties then both of proper separation, and of properly establishing linkages (the inferential process) in the face of the layered remoteness of the one from the other, and the multiply varying, multiply determined, and *even more*, multiply *over*-determined connecting links—the implications of these difficulties are that among the central tasks of psychoanalytic clinical research are those of making both descriptions and inferences, as explicit, as 'public', and as objective as possible, and of *tracing* and *specifying* as much as possible the relationships between observations and concepts, all as essential prerequisites to the establishment of reliability of inter-subjective judgements. The pitfalls in the way are many. Both descriptions and propositions advanced may fall short of being 'public' in a number of ways. Statements may be insufficiently explicit,

as is commonly the case in clinical communications between people sharing a very similar conscious and preconscious paradigm, and may thus fail to convey ambiguous information to those outside the immediate research group. The nature of the inferred links between observations and propositions may be left unspecified, so that relatively straightforward descriptions, and subtle, elaborate inferences become intermingled without distinction. And group consensus (of which, more presently) cannot be taken as adequate evidence of inter-subjective agreement between independent qualified observers. And in proper pursuit of the task of 'relating the surface to the depths', the strategic emphasis may well needs be the reverse of the more obvious direction. Since the theory specifically denies an isomorphic relationship between discrete behaviours (thoughts, actions) and intrapsychic states and since the rules for inferring underlying processes from the data of observation can only be specified in the least complex (and often in the least informative) of instances, it may not be possible to proceed very far in the discovery or the reconstruction of complex underlying states from the configurations of elementary observational building blocks. We should expect rather to work more the other way, i.e. to start with the more overarching concept, to differentiate its components, and their various possible pathways of representation, and to discover thereby their array of possible empirical referents in the data of observation.

#### *The use of clinical judgements or inferences as 'data'?*

In the whole process here described of according explanatory importance to both observation *and* inference and the fullest explication of the inferred pathway as well, the underlying state (or inferred intrapsychic organization) takes on an at least co-equal position as a relevant and manipulable base of knowledge. Sargent (1961), in an article establishing the methodological undergirding of the PRP, in fact persuasively argued the viewpoint that the *essential* data of the whole clinical research enterprise are not behaviours or verbalizations (important as these are) but rather the patient's intrapsychic organization as 'seen' clinically through these, i.e. the essential data are the clinical judgements. She stated this as:

An empirical fact has been defined as beginning with an awareness that can be communicated, hence confirmed or disconfirmed independently. Sharability and agreement, thus, become the crucial test of reality, rather than *visibility or other sensory base*. One can, in other words, perceive and compare one's own and others awareness. The essence of objectivity is, in fact, thereby achieved (Sargent, 1961; *italics ours*).

In another place she specified this simply as:

As far as the project is concerned, we start not from what the patient said, or even what the observers heard, but *what the clinical judges make of* the multiple cues available. Do they come out with the same clinical impressions and formulations? Are the predictions of one group comparable with later judgments by another? What are the areas of agreement and disagreement, and of predictive success? (Quoted in Wallerstein, 1960; *italics ours*.)

The reader is referred to Sargent's own paper (1961) for the fuller exposition of this position, the anchoring of an entire research enterprise in clinical judgements as its essential data, and the manifold implications of this position for the methodological issues of psychological science (in its own right and in its place among the array of sciences). This position has not been argued uniquely by Sargent or from the area of therapy research alone. Schafer (1967) in discussing the uses of projective test data for clinical research purposes developed the same point:

the research *unit* should be *interpretations* and not scores or theme counts. Only then may we continue to work in context, which is to say, work with clinical data clinically. Scores, content, sequences, attitudes, behavior, and style of verbalization must all feed into the interpretations. Any one of these by itself is not a reliable, specific, and hierarchically localized indicator. [*Italics ours*.]

For the practical purposes of the research enterprise, of course, additional operating complexity is thereby introduced. If the 'units' of observation are indeed to be 'interpretations', interpretations of the meanings of observational data (behaviours) in terms of organizing constructs or inferred intrapsychic states (whether inferred by research observer of the data or by treating analyst of the patient for that matter), the operational problem is not only of, 'What' is an interpretation? (distinction between inference and observation) but also even of

'When' is an interpretation? Schlesinger (1967), in his discussion of the paper on an audio-recorded research analysis by Gill and his co-workers, illustrated this 'when' problem by a rendering of a contrived, but plausible because commonplace, paragraph of the verbal flow of an analytic patient associating. In the material, the patient stated a behaviour sequence, then paused and reflected about the way the analyst would look at it, and then went on with the inferred meaning (the interpretation) with which the patient increasingly identified, adding confirmatory associative evidence. All this time the analyst remained silent. A convincing bit of analytic interpretative work had been done. By whom? The patient? For the patient was only recalling as his own the past interpretive activity of the analyst, carried out over time. Then had the interpretation been made only long before, with the present work no more than a bit of necessary but repetitive 'working through'?

Schlesinger brought this up from the viewpoint of identifying a specific difficulty in studying micro-interaction effects (like the particular consequences of specific interpretive interventions) from the minutiae of interaction bits available in verbatim records. It illustrates, however, but one aspect of a wider research difficulty, that of properly isolating and circumscribing the interpretive unit, a significant part of what makes the issue of 'consensus' in clinical judgement (to be next discussed) so stubborn.

#### *The 'consensus problem', or what to do when experts disagree?*

What Seitz (1966) has called the 'consensus problem' in psychoanalytic research, or what to do when the experts disagree, poses the curious paradox of being inherently probably the most difficult issue to contend with in clinical research, both conceptually and technically, and at the same time one that is studiously under-emphasized in the empirical and theoretical clinical research literature. This issue becomes crucial to any situation involving complex interpretative judgements (based on inference); it does not operate importantly in the instance of simple reliability tasks performed upon sensory observational data. It is such an especial problem for psychoanalysis precisely because psychoanalysis is so centrally dependent upon interpretation, and at the same time, as

Rapaport (1960) has stated, 'there is as yet no established canon [in psychoanalysis] for the interpretation of clinical observations'. Glover (1952), in his role as a polemicist on the shortcomings of psychoanalytic research, has called this fact its 'Achilles heel'. He stated

that in any given case interpretation is an essential part of the process of psycho-analytical investigation and that nevertheless there is as yet no effective control of conclusions based on interpretation, is the Achilles heel of psycho-analytical research

which is a major determinant of the unhappy situation

that so far no system exists whereby the scientific authority of research workers can be distinguished from the prestige of senior analytical practitioners and teachers.

Seitz reported one of the few clinical research projects, called the Consensus Research Project, directed squarely to the effort to surmount just this dilemma. The project was disbanded after the three years of work by Seitz and his seven senior analytic colleagues from the Chicago Institute for Psychoanalysis and the write-up called it the 'report of the failure of a research', because of the reported

inability [over that time] to make progress in developing a reliable interpretive method—i.e., a method that would yield greater consensus among a group of analysts in making independent formulations of the same case materials (Seitz, 1966).

The failure to deal effectively with this consensus issue has handicapped the effort to enhance research yield (in terms of data and knowledge accretions that are shared, public, and objective) through the logical path of research group formation. Pfeffer (1961a), in reporting from a panel discussion on the issues of psychoanalytic research, summarized first the advantages of such an effort, the use of something akin to the familiar 'continuous case seminars' for investigative instead of just didactic purposes. The advantages of group consideration and group judgement (consensus) would be several; that a larger number of alternative propositions could be evoked than is ordinarily possible with a single analyst

working alone; that the pooled data and hypotheses would be tested in terms of internal consistency against the subsequent development of the case; that conviction of plausibility and probability, having to evolve out of the deliberations of a group, would be subject to greater pressure of criticism and therefore (hopefully) greater precision of formulation. But the potential disadvantages, too, were stated; the dilution of responsibility, group suggestibility, the premature presentation of ideas and formulations, the possible inhibition of the development of new ideas, subtle pressures on the therapy, the tendency to settle scientific issues on a parliamentary basis. Several of these disadvantages stem directly from the absence of canons for interpretation leaving appeal to authority, the ultimate refuge of the unscientific, as the chief recourse in the face of disagreement. Lustman (1963) stated a not unknown extreme of this in study group formation where

the greatest risk remains that of 'consensus of opinion' research in which the *senior* member is the one who never sees the data at first hand.

The reasons for this persisting major consensus dilemma in psychoanalytic research are many. It is far more than an issue of inadequate research sophistication stemming just from an insufficient tradition of attention in clinical research to problems of method and design. Strupp and his co-workers (1966) outlined three kinds of possible sources of consensus difficulty. These were: (1) that the research *observers* are poorly calibrated, with their differing degrees and kinds of training, experience, and expertise, not to speak of possibly differing theoretical orientations; (2) that the clinical *phenomena* can be so unclear with (in extreme instances) observations little more than projected fantasy systems of the observers; and (3) that the observational *methods* are inadequate, the problems of demarcating meaningful units of observation, of specifying precisely observational dimensions or categories, etc. For the most part, these orders of difficulty specified by Strupp *et al.* do fall within the realm of the alterable (remediable) by proper attention to classical issues of research method and design. That is, they are problems of the state of the art, the *state of the theory*.

But beyond this there are issues that reside as well not only in the inherent *limitations in the power of human intellectual mastery* but also in



the *characteristics of psychoanalytic theory* as an explanatory system evolved to account for human behaviour. Pitted against the quest for consensus are such (inevitable) *human* problems as the enormous complexity of psychological data against which interpretive powers are limited and fallible; the fact that minds function differently and apply different problem-solving paradigms, for example, the tendency in some to assimilate the new to the already familiar against the opposed tendency in others to search for novelty, to look for perspectives not yet in other people's minds; the by now endlessly familiar problems of observer bias which operates in any process of selection from the complexity of clinical data, selection influenced inevitably by the observer's theoretic predilections as well as emotional blind-spots.

And, in the nature of psychological (psychoanalytic) science itself there are at least three central features that characterize the field and that, *in principle*, render consensus difficult to come by. These can be stated as: (1) the principle of multiple causation or determination, (2) the principle of over-determination (and multiple function), and (3) the probabilistic nature of psychic states. The essence of each can be stated briefly.

1. The principle of multiple causation leads clearly to the problem of the blind men and the elephant, the partial, and therefore only partially correct, views of observers with limited observational or theoretical vantage points. Strupp and his colleagues referred to this as an important contributing factor in the consensus difficulty in clinical research.

Because many events with which the therapist deals are highly complex and far from being directly observable, and because a high level of clinical inference is often required in describing the nature of the events, it is likely that an independent observer will not traverse the same roads of clinical inference travelled by the therapist. The result is that their respective descriptions may fail to agree, or because the two observers may focus on different levels of abstraction or different facets of the matrix of events, it is difficult to determine from their respective descriptions whether they agree or disagree, or to assess the extent of their agreement (Strupp *et al.*, 1966).

2. The principle of over-determination is related to but distinct from the above, though often confused with it. Here the consensus difficulty becomes further compounded, since

beyond the array of necessary and sufficient causes to psychic events called for in the concept of multiple causation, there can be an additional superfluity of *over-sufficient* causes. Once the elephant is fully and adequately described by bringing together the variety of partial perspectives into an integrated whole, there can still be additional descriptions, *each* of which (each of the metapsychological points of view, for example) can be sufficient description to account for the entire elephant (the phenomenon under description). We have already quoted in part from Rapaport's (1960) statement of the need of psychoanalysis, as theory, for this concept.

3. A much less remarked problem for the consensus search has to do with the probabilistic nature of patient-states. Chassan (1957) has especially elaborated the implications of this perspective. He said:

it is easy to argue from this point of view that the inability of coefficients of stability to become and to remain high is more a reflection of the underlying probabilistic aspects of patient-states than of any particular deficiencies in the testing procedures.

(For the word testing, we would substitute, assessing or judging.) From this it not only follows that

if a patient is first observed by investigator A and subsequently observed by investigator B, a difference in the value of a variable between the two observations need not reflect a defect in the measuring instrument or an inconsistency in the observers. The change could have been in the patient.

(A point not relevant to the consensus issue as such.) But it also follows that when two observers observe at the *same* point in time—and disagree—that there is too a probabilistic aspect to interpretation.

That is, the interpretation of the same manifest phenomenon may vary to some extent 'within' the observer because of a relevant complexity of circumstances and interactions at the instant of observation.

Here Chassan referred to the now classical studies on the major internal inconsistencies revealed when the same panel of outstanding radiologists each re-read the same chest x-rays for the presence of minimal tuberculous lesions. And of the startling extent of the failure to agree with oneself demonstrated by this study, Chassan could of course state that

the phenomenon of a tuberculous minimal lesion . . . is a state or a process which can be said to be entirely uninfluenced by the mere act of observing it on an x-ray plate, and the goal of complete agreement between all investigators is entirely unambiguous;

in contrast, psychoanalysis is a study of interpersonal relationships, participant observation is a basic phenomenon of the therapy, countertransferences operate, etc.—all reasons for increased uncertainty (variability) of interpretation.

Seitz (1966) added one further consideration into this study of the consensus dilemma, that of time in process. He wondered if perhaps the quest for consensus would become somewhat easier as one went along in the study of the psychoanalytic treatment process, as the dynamics become more clarified, the defences laid more bare, the psyche more open. After all,

the more data (arising longitudinally over time) that could be accounted for by the interpretation, and the more internal consistency that could be shown to exist among the data by applying this interpretation, the more likely it was to be considered correct (Seitz, 1966).

*A priori* there does not seem reason to derive much hope in this direction since the more one knows, the more complex and therefore perhaps less manageable the data may become.

In the face of these formidable difficulties in the way of reducing, ordering, summarizing (i.e. of making judgements about) the data of psychoanalysis, are there roads which nonetheless point in the direction of solving or at least of mitigating the restrictions imposed thereby on empirical research yield? The problem, as we have developed it, is that we cannot reasonably work from the two platforms chiefly advanced in the research literature. The first is based on the attempt to get rid of the clinician so to speak, that is, to get rid of the vagaries of clinical judgement, by the research focus on observables and measurables, on the raw observational data of manifest behaviours alone, applying then the usual statistical analytical techniques to determine their reliability—and hoping thereby to advance knowledge of how the universe of events hangs together. The limitation here (so clearly stated by clinician to researcher) is that these observables and measurables do not take on meanings (either singly or in combinations that can be

inductively built up) except by *interpretation* in the light of concepts of varying degrees of remoteness from and varying kinds of relationships to the data, all beyond the scope of the statistical manipulations. The alternative approach is based on maintaining the (skilled) clinician and attempting to work from his ordinary (which can mean highly experienced) clinical interpretative judgements. The limitation here (stated equally clearly by researcher to clinician) is the 'consensus problem' that renders these judgements not reliable enough, or rather not even sufficiently about the same things, and with a largely undetermined (and to a certain extent indeterminate) degree of difference in the 'things' being judged. This is the problem stated so cogently by Seitz; the work of his project, involving highly skilled and experienced analytic collaborators, investing large quantities of time, zeal and research sophistication in dedicated pursuit of this goal of consensus in clinical interpretative judgement, is eloquent testament of the incapability of such a quest set in these terms. To seek more from that process is to ask more of the clinician judge and of the concepts he essays to judge than they are in effect 'calibrated' to bear.

Two other strategies less tried within psychoanalysis, derivatives and extensions beyond these, seem to us more concordant with the complexities of the issue at hand. The first is that of the PRP. The basic stance here has consisted of the effort to 'refine' ordinary clinical judgement through a variety of operations, in part already alluded to in this essay—the careful definitions of the concepts as generally understood within the framework of psychoanalytic theory, and in terms as well of specific and idiosyncratic local usage; the continued redefinition and clarification of the concepts as determined by the testing of the concepts against expected judgements in the actual operations of the project; the 'training' of experienced clinician judges in the actual clinical comparisons made using these concepts (the research variables); the use of paired comparison (forced choice) judgements to reveal areas of convergence and of difference in applying the concepts (the clinical variables) against the clinical phenomena. To the extent then that variables, thus successively refined, can give rise to judgements (of heightened agreement, consensus, or of known degree of disagreement, dissensus) from which empirically testable predictions can be generated, a method

(of successive approximations) has been evolved for pushing clinical judgements of complex psychological events in the direction of more reliable, hence more measurable, statements. Put another way, such a research approach rests not on the precise measurement of behavioural observables (of varying degrees of relatedness to the underlying organizing constructs), but on complex assessments (clinical judgements about the configural meaning of behaviours) that are, however, successively refined in the ways cited, and then made to bear the burden of empirical testing via predicted consequences. It is the circling back from the observed consequences that strengthens or weakens the credibility of the inferential process built into the clinical judgements; and it is the attention to the refining process that renders these inferences more visible, hence asymptotically correctable.

An alternative approach to these same issues is exemplified in the TPS, in which rather than attempting to refine the clinical judgements relating to the organizing constructs, the effort is systematically made to more tightly link the constructs to their observable consequences. If, for example, a particular defence is said to be modified during a period of analytic work in the direction of 'greater integration within the ego', manifest phenomena indicating lesser or greater integration are specified, raters unfamiliar with the hypotheses of the study identify these phenomena in sessions arranged in a scrambled, random order, and statistical tests are applied to determine whether in fact the hypothesized change has taken place. This approach has also been applied to the study of relationships between hypothetically covarying phenomena or processes. The strategy is to use careful clinical study (i.e. clinical judgement) to hypothesize relationships and then to seek out the behavioural consequents that should be evident if the relationships are as postulated, i.e. to find the behavioural events that would correspond to the dimensions of the relationship. The final testing point is then in behaviour observation that can be subjected to the usual specifications for reliability and validity. Both these strategic approaches exemplified in the two psychoanalytic therapy research projects then take account of the problem of the remoteness of concepts from observations, as well as of the potential for confounding the two; both too are anchored at crucial points in mensurable behavioural data. The varying but very partial

degrees of success achieved to date by efforts in these directions lies in the difficulty of execution.

At the same time all these very characteristics of the theory here discussed that from the one vantage point pose such major difficulties for both clinical practice and research endeavour, from another perspective are the indications of the nature of the reality the theory is designed to encompass and the conditions that determine what it is possible and useful to investigate within it. The fact that two successively recounted versions of the same manifest dream are usually not identical but differ in particulars that systematic inquiry proves to be dynamically meaningful is in itself the very lever that helps pry open the pathway towards unravelling the associative network to the underlying latent dream thoughts. Or the fact that inappropriate and highly charged affective responses of patient to analyst—the transference—unexpectedly complicated the analytic investigation proved not only the chief resistance to the treatment, as Freud discovered to his discomfiture in the treatment of Dora, but also, of course, emerged as its central technical and conceptual vehicle, via the unfolding transference neurosis. That is, the very methodological difficulties for psychoanalytic research that stem from the discussed characteristics of the theory are also the manifestations of the richness and the reach of the theory, providing unique avenues of access to subtle and complex phenomena.

#### HOW MAY CIRCULARITY OF CLINICAL JUDGEMENT BE CIRCUMVENTED?

Closely linked to all the problems of arriving at (consensus on) clinical judgements (interpretations either to the patient by the analyst, or about the patient by the researcher—or by the analyst) is the problem of circularity contaminating the judgements so arrived at. The starting point for discussion of this problem can be taken from a statement by Rapaport (1960):

Clinical predictions are always fraught with the fact that all motivations have multiple, equivalent, alternative means and goals. Thus, such predictions usually cannot specify which of these equivalent alternatives are to be expected, and therefore, the results of experimental tests of these predictions must *first be interpreted* before their bearing on the theory can be established. [Italics ours.]

The rub rests in the italicized phrase. This Rapaport (1960) discusses directly at another point:

In the lack of a canon for clinical research, it is difficult to accept as positive evidence observations which must *first be interpreted* before it becomes clear whether or not they confirm the predictions of the theory. We must be wary lest we *smuggle in the confirmation through the interpretation*. Axiomatization and/or a canon of investigation protect other sciences from such circularity . . . [In psychoanalysis] as things stand, there is no canon whereby valid interpretation can be distinguished from speculation, though *ex post facto* the experienced clinician can distinguish them rather well.

This tendency to circularity becomes inevitably built into any science relying predominantly upon the clinical retrospective method for the gathering of its data and the confirming of its hypotheses. This becomes clear in a statement by Freud himself made to indicate his scepticism of the role of predictive efforts in relation to psychoanalytic understanding. In 'Psycho-genesis of a Case of Homosexuality in a Woman' in 1920, he said:

So long as we trace the development from its final outcome backwards, the chain of events appears continuous, and we feel we have gained insight which is completely satisfactory and even exhaustive. But if we proceed to reverse the way, if we start from the premises inferred from the analysis and try to follow these up to the final result, then we no longer get the impression of an inevitable sequence of events which could not have been otherwise determined. We notice at once that there might have been another result, and that we might have been just as well able to understand and explain the latter. The synthesis is thus not so satisfactory as the analysis; in other words, from a knowledge of the premises we could not have foretold the nature of the results.

It is very easy to account for this disturbing state of affairs. Even supposing that we have a complete knowledge of the aetiological factors that decide a given result, nevertheless what we know about them is only their quality and not their relative strength. Some of them are suppressed by others because they are too weak, and they therefore do not affect the final result. But we never know beforehand which of the determining factors will prove the weaker or the stronger. We may say at the end that those which have succeeded must have been the stronger. Hence the chain of causation can always be recognized with certainty if we follow the line of analysis, whereas to predict along the lines of synthesis is impossible.

If we truly have no way of assessing the relative strength, the balance of forces, short of the criterion behaviours that are predicted to, i.e. if we are only able to judge the antecedent state of affairs *after the fact*, by observing the outcome, then indeed we are in a circular bind. Waelder (1963), who has quite currently advanced essentially the same viewpoint as Freud on the difficulty, in principle, of making clinical predictions, relies heavily on the same argument: 'i.e., a tendency is proved to have been the stronger one by virtue of the fact that it has actually prevailed' which completes the circle to 'We cannot predict the outcome through measurement of the strength of the forces involved if we need that very result to make the measurement'. The dilemma, so stated, would be insoluble in these terms if clinical research in psychoanalysis had to be confined totally to the following of the clinical retrospective method. This was indeed the classical method of investigation of Freud in his studies of symptoms, dreams, etc., i.e. he started with the manifest dream as reported, followed the dreamer's associations to the presumed latent dream thoughts, and then could reconstruct the steps which must have been taken by the dreamwork (the reversal of the associative pathways) in transforming the latent dream thoughts, the presumed antecedents of the dream, into the manifest dream production. This very powerful model which worked so successfully for Freud in unravelling the mystery of the dream (and of his own psychic structure) became then the cornerstone of the clinical method for the study of the abnormal—and the normal—phenomena of mental life. Its limitations have only subsequently become evident as psychoanalytic science has begun the movement from the generating of hypotheses (formulations) to the effort at their more rigorous testing (when just such issues of hidden circularity arise). And it is at precisely this point that the classical experimental model, in which antecedent conditions are specified and controlled *in advance*, and the subsequent consequences then independently observed, has its locus of most potent application.

The strength of the analytic method and the powerful example of Freud's capacity to gather such a monumental harvest of insights into the functioning of the mind by the use of it has been such, however, that few psychoanalytic clinicians have taken this problem of the circularities built into (only) retrospective

analyses as seriously as the issue has scientifically warranted. Rapaport (1960) concluded his discussion of this issue with the historical statement:

In conclusion: the nature of the material Freud worked on led him to overemphasize postdiction and underemphasize prediction in building his theory... But it may be questioned whether any science in its beginnings has been free from such imbalances. The basic necessary condition for predictions and for their confirmation is present in the theory of psychoanalysis, and certain types of psychoanalytic predictions have been confirmed.

In keeping with this statement of Rapaport, the psychoanalytic discourse on prediction, and the use of prediction to cope with the problem of circularity has, however, not been one-sided. One of us has elsewhere (Wallerstein, 1964) and with collaborators (Sargent *et al.*, 1968) reviewed the psychoanalytic literature on the place of both the *principle* and the *tool* of prediction in psychoanalytic investigation (in longitudinal developmental studies, in psychosomatic studies and in therapeutic process studies), has developed a rationale and a structure for the formal use of predictions to more systematically and precisely link the data of psychoanalysis to the theory of psychoanalysis (thereby a wedge by which to test and extend the theory) and has elaborated a manual for such usage together with a fully written out case illustration including the explicit formalized predictions made in that case and the process of their empirical testing. The interested reader is referred to those very extended discussions. Here we will abstract only that aspect of the Prediction Study component of the PRP that bears on the issue of the control of circular reasoning. Suffice it that despite the difficulty in making them, clinical diagnostic and therapeutic work rests on predictions that are inherent in the clinical undertaking. As stated elsewhere (Sargent *et al.*, 1968):

Every responsible action in diagnosis and treatment involves one or more predictions derived from clinical experience or from theoretical hypotheses. When the clinical team in a case conference assigns a patient to treatment, a prediction is implied; the course recommended is expected to benefit the patient in certain ways, some of which are more, some less, spelled out in case discussion. In such deliberations, many other predictions are made, implicitly or explicitly, relating to possible contraindications, vicissitudes of the treatment course,

and/or specific outcomes to be hoped for. Furthermore, clinical predictions, followed by observations of the course and outcome of therapy, provide myriad potential experiments which could test the hypotheses by which the predictions and treatment recommendations were guided.

Recommendations and case formulations conveyed by superiors to therapists in training, and dynamic formulations in psychological test reports, also involve prediction. If a particular approach is suggested, a given response is anticipated. If a patient is placed in a diagnostic category, reactions, symptoms, and special characteristics belonging to that category are implied. Diagnosis is meaningless unless oriented toward prediction; that is, there is little use in labeling a patient as hysterical, obsessive-compulsive, or paranoid unless the label carries some connotations about what may be expected in the development of the illness and in the treatment course. The label 'paranoid', for example, implies that the patient will be rigid and resistant to change, that one will have to deal with projective defenses in therapy, and that underlying homosexual problems may emerge. In the testing and analysis of predictions thus made in clinical practice lies the hope of establishing a solid core of psychotherapeutic theory, separated from *ex post facto* dogma sometimes used to rationalize success or failure, and from *a priori* assumptions based more on belief than demonstrated by association with favorable outcome.

Thus prediction is actually a pervasive, even universal, *clinical* phenomenon, usually implicit, and as such, unremarked. The research task of the project was to make explicit (so as to set up the conditions for formal testing) a range of predictions in a sample of patients entering psychoanalytic therapy, predictions relating to

the anticipated course and outcome of the recommended therapy, the nature of the problems to arise in the therapy in terms of expected transference paradigms, major resistances, and foreseeable external events that might (favorably or unfavorably) be expected to bear on the treatment course, and prognostic estimates in regard to expected or hoped-for changes in symptoms, in impulse-defense configurations, in manifest behavior patterns, and in level and nature of achieved insights (Wallerstein, 1964).

These predictions (clinical inferences embedded in clinical context and qualification) were then recast into discrete and testable predictive statements in accord with a tripartite 'if-then-because' (conditions-consequences-assumptions) logical model adapted by Sargent to clinical analytic data (in Sargent *et al.*, 1968).

It is at this point that critical control for circularity enters by setting down *in advance* the entire predictive complex of conditions, predictions proper, and assumptions clause together with the predetermined evidence, in fact or in judgement, that will subsequently be necessary in order to sustain or refute the predicted outcome. It is by thus forcing the whole sequence of statements and of supporting reasons in advance that observation is controlled and *post hoc* reconstruction, according to which almost any outcome can be plausibly rationalized in terms of a retrospective weighing of contending forces, avoided. (Parenthetically, the more the predetermined evidence necessary to confirm or refute a predicted outcome is anchored to observables, to behaviours that lend themselves to ready objective judgement, and the less it is anchored to assessments of intrapsychic processes that rest on interpretations of the configural meaning of behaviours, the less again the danger of circularity, of sneaking in the confirmation via the interpretation, this ever-present danger to theory validation in psychoanalysis.)

In this way, despite the many conceptual and practical difficulties in implementation, the systematic use of the predictive method can overcome the problems of circularity, and permit us to conduct 'experiments in nature' (i.e. where we do not *control* the antecedent conditions, but designate their presence and hypothesize about their consequences) within the clinical research context. The same strategic use of the concept of prediction (though not necessarily so designated) marks as well other ongoing investigations of the psychoanalytic process. Gill and his co-workers planned to incorporate into their micro-study of the audio-recorded analytic process a quasi-experimental study of the impact of appropriate, as contrasted with presumed tangential, interpretative interventions, compared with respect to their differing postulated evoked consequences. At this point we can discern the very close relationship conceptually between such short-term 'predictions', say to interactions within the immediate or the next analytic hour, and the ordinary process of analytic interpretation. (And at this point, we can share in this limited sense the viewpoint quoted in our introductory discussion on the need for formalizing psychoanalytic research—Bellak & Smith, 1956; Ezriel, 1951, 1952; Hartmann, 1959; Kris, 1947; Kubie, 1956—that psychoanalysis can in some ways be considered analogous to the quasi-experimental

research model with the interpretations serving as the specifiable and manipulable independent variables within the (relatively) stable, recurring experimental situation.)

The research strategy of the TPS likewise embodies the principle of prediction. The theoretical position of the group leads to hypothesized relationships between inferred processes; and these hypothesized relationships portend (imply, or predict) that certain covariances should be manifest in the clinical material under study. To the extent that the deduced behavioural consequences to be discerned in the clinical material are specifiable *in advance* of the search for verification, prediction is again involved, and the same safeguard against circularity of reasoning has been introduced.

Which is not to say that prediction is the only way to avoid circularity or the only way to guard against confounding and error. For example, much of child analytic research is based on the direct observational method applied to the study of child development. Data from such observational and longitudinal study not only supplement but also check the retrospective and reconstructive data derived from the therapeutic process (whether of adults or of children). The congruence of the formulations derived from the data of the two (independent) observational sources can be assessed. Additionally, and from within the psychoanalytic method proper, the fate of Freud's original traumatic (seduction) theory of the psychoneuroses is a demonstration not just of the major fallibility of the retrospective method, but also of the capacity of the superior mind to discern the increasing deviations from reality that such a false formulation progressively imposes, and within the method, turn it to a successful reformulation more loyal to reality; that what was once considered fact, an experiential vicissitude, is now to be considered fantasy, a maturational unfolding of an internal drive representation. This whole process of error and of rectification took place purely within the classical psychoanalytic method, and without benefit either of 'outside confirmation' or of predictive safeguard.

#### TO WHAT EXTENT CAN ONE GENERALIZE FROM AN *N* OF 1, OR OF VERY FEW?

Probably no one would cavil with Strupp's admiring remark (from the perspective of an empirical psychotherapy researcher) that it is a

tribute to Freud's genius 'that he succeeded in making valid generalizations on the basis of exceedingly small samples' (Strupp, 1960). The study of the Schreber case is a striking example of the far-ranging generalizations, in this instance about the dynamics of paranoia and the operation of projective mechanisms (see the four formulae and their explanatory applicability to some of the major symptomatic expressions of the paranoid illness) stemming from just one case, an  $N$  of 1, and that one studied only from a written product. Yet it is a fundament of modern empirical science that generalization across cases requires a sampling of many. As Janis (1958) put it:

An obvious weakness of a single case study... is that it can provide no indication as to whether the relationship applies to all other, many other, a few other, or no other human beings. Thus, even when a causal sequence is repeatedly found in a given person, the investigator cannot be sure that his findings can be generalized to any broad class of persons because the relationship may occur only in an unspecifiable, restricted class of persons sharing a unique constellation of complex predispositional attributes. [Or] to put the matter in more technical language... a major limitation of the findings is that there are zero degrees of freedom with respect to individual differences, even though each finding may be based on hundreds of degrees of freedom with respect to the samples of the subject's behavior that enter into the correlation between the independent and dependent variables.

Margolin (1951) gave an informative example of where lack of proper attention to this basic caution led to premature closure on unchecked speculation. He obtained research access, for combined psychoanalytic and physiological study, to a young woman with a gaping stomach fistula, healthy in other respects. The several previous such research subjects had all been men and in one, subjected to intensive study, it had been reported that the affects of anger, resent-

ment and fright heightened the outpouring of gastric secretion. Those who first studied the young woman found that in her case gastric secretion seemed inhibited by these affects. This apparent sexual difference (in each case built on an  $N$  of 1) in the discerned physiological responses of the stomach during emotional stress was declared to justify the formulation that it was related to the known far greater incidence of duodenal ulcer in males!

Yet in the face of such seemingly elementary considerations, how can one account for the success of so many  $N = 1$ , or  $N =$  very few, studies in psychological science?<sup>13</sup> Support can be found both among statisticians and clinicians for the position that for many clinical research purposes more can perhaps be learned from a smaller than from a larger number of cases. From the statistical point of view Edwards & Cronbach (1952) have stated:

Information gained from an experiment mounts more or less in proportion to *factorial n* where  $n$  is the number of uncorrelated response variables. By this estimate five tests can report 120 times as much knowledge as a single test in the same investigation! ... Effort to refine measurement has the *same* beneficial effect on the power of an investigation as adding to the number of cases...

And from the clinical point of view, Gill and Brenman have stated:

The clinical researcher must compare situations in which a number of variables are varying at once, thus differing from the experimentalist who can attempt to hold all the variables but one constant. The clinical worker must find patterns and principles of relationships which must be true to account for the observed variations. The more simultaneously varying variables he must deal with, the more uniquely determined is the hypothesis he must deduce to fit the observations... Instead of saying that many variables force a multiplication of cases, we would say that they make necessary only a relatively few cases (in Brenman, 1947).

<sup>13</sup> The case to be here made, despite the cogency of Janis's warning, for the value nonetheless of  $N=1$  studies under certain conditions in psychological science, should not be confused with the different case that some of the limitations upon statistical inference imposed by the  $N=1$  model (when  $N=1$  individual) are transcended when we look at an individual case as an entire population of instances suitable for controlled statistical manipulation of the instances (not the individuals)—as in the array of individual predictions (about 50) in each case of the PRP. In that project, there is an  $N$  of 42 individuals, which in terms of the complexity of the phenomena encompassed by the study (the multiplicity of variables)

is very few indeed, but there is an  $N$  of approximately 2000 discrete predictions (each of course simpler by many quantum steps than a whole functioning person) which is a very respectably large  $N$ , suitable to many standard statistical manipulations. In a similar spirit, Luborsky (1953) has elaborated a methodology for study within the  $N=1$  situation, the method of repetitive intra-individual measurement, called P-technique. The technique is based upon repeated measurements on the same battery of tests or variables, correlating the measurements when the series is long enough to give an adequate 'population of occasions' (Luborsky, 1953).



And the literature of general psychology has, in fact, also reviewed the variety of circumstances under which the even more limiting condition of  $N = 1$  can still mark an appropriate and useful, and even the only possible, research strategy. Shapiro (1961) has described three general research strategies based on the use of the single case. These he stated as follows: (1) Experimental control, which 'implies that the conditions affecting variations in the appearance of a phenomenon in an individual case are well enough known to enable one to predict, in measurable form, the appearance of that phenomenon in situations which have never before been presented to that individual. The description of such conditions would amount, in effect, to the description of some of the laws affecting the phenomenon in that individual'; (2) Replication, 'The essence of the idea of replication is that when one has confirmed a law in a single case one then determines the degree to which the law can be found operating in other cases'; and (3) Development of Appropriate Methods, methods of measurement of complex psychological variables peculiar to the individual patient, permitting comparisons across time for the same patient, and also across patients.<sup>14</sup>

Dukes (1965), in a paper entitled  $N = 1$ , discusses the conditions that warrant employing an  $N$  of 1, under four headings: (1) 'If uniqueness is involved, a sample of one exhausts the population. At the other extreme, an  $N$  of 1 is also appropriate if complete population generality exists (or can reasonably be assumed to exist).

That is, when between-individual variability for the function under scrutiny is known to be negligible . . .'; (2) The *dissonant* nature of the findings; 'In contrast to its limited usefulness in establishing generalizations from "positive" evidence, an  $N$  of 1 when the evidence is "negative" is as useful as an  $N$  of 1,000 in rejecting an asserted or assumed universal relationship'; (3) When there is a *limited opportunity* to observe; When 'individuals in the population under study may be so sparsely distributed spatially or temporally that the psychologist can observe only one case, a report of which may be useful as a part of a cumulative record' (and 'situational complexity as well as subject sparsity may limit the opportunity to observe'); and (4) 'Problem-centered research on only one subject may, by clarifying questions, defining variables, and indicating approaches, make substantial contributions to the study of behaviour.' And in this connection Ebbinghaus's classic and still fundamental work on memory, done in 1885, on only one subject, himself, was quoted. (Italics in above quotations ours.)

In summary Dukes (1965) said: The usefulness of an  $N$  of 1 in research is viewed as extending beyond the single-case studies of clinicians and personologists. An  $N$  of 1 is seen as also appropriate when, for the function considered, inter-subject variability is low, when opportunities for observing a given class of events are limited, and when a supposed universal relationship is questioned and the obtained evidence is negative.<sup>15</sup>

<sup>14</sup> The development of individualized measurements in intensive research design is treated in detail in Chassan's recent monograph (1967).

<sup>15</sup> There is a further argument, the conclusions of which are beyond what we feel the reasoning will sustain, advanced by Svenson & Chassan (1967) on the moral and ethical advantages of the  $N=1$  study. They make two main points. The first is that an  $N=1$  study 'virtually eliminates the gap between clinical research for the general good and clinical research for the good of the individual patient who participates in it'. In the conventional extensive, or group-averaging, model of research design only general results are obtained, group effects which have no necessary applicability to any one patient. For example, in studies of pharmacotherapy in psychiatric disorder, 'All we can infer on a statistical basis is that on the average, the group on the test medication fared significantly better than the group on placebo and that, somewhere along the line, one or more patients in the test medication group improved through the pharmacological activity of the drug. We cannot say how many patients and we cannot specify which patients.' In the single case (especially if properly controlled, double blind, etc.) with the drug and placebo mixed up and repetitively used, the data acquired will indicate that 'the patient himself responded positively

to the drug'. From this they deduce their second point, the statement of ethical advance; since, under this circumstance, the research itself is more to the individual patient's best interests 'than either the group-averaging extensive double-blind study or the open non-controlled evaluation of a given new drug in a single patient'.

Of course to the extent that this is so, formal and systematized clinical psychoanalytic research partakes of these methodological and ethical rewards. No matter what the numbers (as in the  $N$  of 42 in the PRP) it is not a group-averaging phenomenon but rather a succession of single cases, built around individual prediction studies, etc. On the other hand, this entire argument of Svenson & Chassan points one-sidedly, and therefore misleadingly, only to possible benefits and not to possible injuries. It is true that under the  $N=1$  design as expounded by them, more is learned that is directly applicable to the individual from whom (and upon whom) it is learned. This does not, however, overcome the ethical dilemma in clinical research to the extent that they state it does. For the dilemma arises rather from the degree of existing ignorance of the potential effects (beneficial or hurtful) of untried clinical methods or agents. To the extent that there is such advance ignorance of the help or harm that will be (may be) incurred (inflicted) by clinical trial (and that cannot presumably be overcome by knowledge



Granting then the critical importance and even the established value of the intensive study of one or a few cases in order to discover relationships or under special circumstances, is there some point at which psychoanalytic research on the therapeutic process must become large scale in order to 'prove' the hypotheses developed on the few cases? If certain mechanisms for the 'integration of defences' can be demonstrated as components of change in the course of a completed psychoanalysis under study, just what is really involved in being able to prove that this is *generally* true in *successful* psychoanalyses. In part, the answer to this is linked to the distinction delineated by Bakan (1955) between general-type propositions and aggregate-type propositions.

General-type propositions assert something which is presumably true of each and every member of a designatable class. They are given increased support with each successive positive instance, though never 'proved' in a formal sense. With the first negative instance they are either overthrown in toto, or more likely, the class boundaries must be further circumscribed, to effect a new more limiting definition that excludes the negative event and maintains the new (more narrowed, and hence more precise) proposition. In this sense, the 'truth' is progressively approached—via a succession of single cases. The situation is of course very different with aggregate-type propositions which assert things presumably true only of the class considered as an aggregate, and where increasing exactitude and significance accrue with the increasing size and representativeness of the sample.<sup>16</sup>

Appropriate sample size thus depends on the purposes of the study as well as on the circumstances and conditions. It is still a moot question as to how much hypothesis-testing within psychoanalysis will ultimately require large *N*'s (with full awareness of the enormous difficulties of meeting the ordinary statistical assumptions upon which the manipulation of large *N*'s via

the usual statistical techniques rest) in order to approximate, in so far as possible, the kinds of certainties achievable in the 'harder' sciences. The exactness of physics, we must remember, rests on the fact that it deals with very large aggregates of (very small) particles (Waelder, 1962). Physics cannot predict for the fate of the single electron. Yet this is precisely our clinical task, to predict lawful relationships and consequences for the one person (except for certain special circumstances where we do deal clinically with very large aggregates of people, as in epidemiology or insurance medicine, and where comparable high precision is achievable). And, in coping with uncertainty, as physics also has to, we do not have the advantage of physics, expressed pithily by Chassan (1957) that there are a lot more atoms than people and that atoms are more alike than people.

Nonetheless we can and do justify our efforts on single or few cases, and the literature cited indicates that there is cogent scientific justification for this position depending on purpose and on circumstance of the study. Intensive case by case study may ultimately not be required for the testing of psychoanalytic propositions and at such time appropriate canons must be devised for the transition from hypothesis-formulating studies of single or few cases, intensively scrutinized, to hypothesis-testing studies on the appropriately larger samples. Meanwhile we must remember that our scientific interests (in contrast to our immediate clinical interests) are not anchored in the *case* but in the revealed *processes*, i.e. in lawful relationships between variables. These processes of course occur in cases, and must be observed in cases, but scientific advance requires abstracting relevant processes from individual cases. And though such abstractions might 'oversimplify' a case, they should not oversimplify the processes. Processes of course may be studied across cases with *N*'s greater than 1. But in clinical research we start with (an approximation of) *N* = 1 for at least two good reasons. The

obtainable outside the clinical trial) no research design (whether intensive intra-individual or extensive group-averaging) can really protect fully against this larger ethical dilemma that the agent employed *may* be hurtful (even devastatingly hurtful in unanticipated ways) rather than helpful.

<sup>16</sup> Bakan (1955) put this distinction as follows: 'The cogency of the distinction between the two types of propositions is revealed by the different role that is played in connection with them of the "next" case. The "next" case presents a fundamental threat to the validity of a general-type proposition. General-type

propositions are thus critically testable, since they are jeopardized by each new instance. If a general-type proposition fails to be confirmed by the observation of a member of the class to which the proposition presumably applies, then either the proposition must be rejected, or the class must be more closely delineated. The "next" case, for the aggregate-type proposition simply increases the "power" of the test, and the likelihood of the empirical proposition which is under consideration, if the study is properly conducted with respect to randomness, etc.'

first is that we want to make reasonably sure that we are not *oversimplifying* the world in abstracting certain processes, but are rather appropriately (correctly) simplifying the world by identifying salient, invariant relationships. We prefer to seek these relationships within individual cases intensively studied because this provides a needed anchorage in the complexity of clinical reality and therefore some protection against naive conceptions of how things hang together. And secondly, psychoanalytic researchers are historically just now in the process of developing research methods which formalize and systematize and render explicit the dimensions of clinical practice and clinical inference that up to now have remained informal, implicit, and intuitive (the 'art' of psychotherapy). In order to make sure that these methods do not distort the processes they are intended to investigate, the methods have to be devised on, and studied in relation to, well understood individual cases. As we become more secure about our research technology in this area, we may imagine being able to deal with more extensive designs for cross-case studies.

#### OTHER PROBLEMS OF PSYCHOANALYTIC RESEARCH

There are a number of other major issues for the developing field of clinical psychoanalytic research which considerations both of space and of central relevance to the main developmental concerns of this essay preclude detailed discussion of here. But they should not be passed over without at least some brief statement demarcating some of these issues and indicating some of the lines of connection with the themes developed here more fully. These issues will be indicated under three headings.

##### *The problem of control*

Related to the issue of sample size, just discussed, is that of sample representativeness and the linked problem of proper controls. The control issue in psychosocial research has been the specific subject of a whole monograph by the GAP Committee on Research (1959) and will not be recapitulated here. Rather a few main principles will be highlighted:

1. The special difficulties of control within the complexity of the interacting and interdependent variables operating in the clinical field, specifically the psychoanalytic situation, are too obvious and too well known to warrant recounting. Their scope and tenor is caught by

the quip attributed to Freud 'that the best control is to treat the same person twice—once with analysis and once without, and then compare results' (in Pfeffer, 1959). This contains a sad truth, for implicit in the commitment to a clinical, i.e. a naturalistic, approach to the study of the psychoanalytic process is the inability to apply the usual kinds of control manipulations, via control groups of 'normal' subjects, or matched groups of treated and non-treated cases, etc.

2. But the absence of appropriate formal control *groups* does not abrogate in any way the responsibility to operate in accord with control *principles*. This forces then consideration afresh of what to control and how to control it, by what specific control methods. In this situation the PRP introduced principles of control in four ways (Robbins & Wallerstein, 1959). One was *intra-patient* control via the individual prediction study with specification in advance of the predictions, their assumptive base, the contingencies (or conditions) to which they relate, and the evidence subsequently necessary to confirm or refute them. The second method was *inter-patient* control using profiles that derive from paired comparisons study of all the adjudged relevant variables. The profiles permit the selection of patients alike in respect of certain variables, while dissimilar in respect of others, thus controlling some variables while the variability of others is investigated. The third method was the parallel and independent assessment of variables and the making of predictions, using a different data source—in that case, projective psychological test protocols. The fourth method, called 'inadvertent controls', occurred when for reasons of geography or finances a treatment plan other than that of choice had to be worked out with the patient, or when the research group differed from the treating clinical staff in regard to treatment plan and recommendations.

3. Inherent in clinical work is the concept of control *for*, since control *of* is not always possible. As stated in the GAP Report (1959):

In *any* type of research, the emotional involvement of the human investigator is a factor for which appropriate controls may be desirable. In the psychiatric research with which we are concerned this characteristic almost invariably belongs to some aspect, at least, of the observed and may additionally require specific types of control. Furthermore, the instruments of observation and the setting have this attribute in high degree. Thus, the interaction

occurring between the various components of the research situation introduces very specific problems for control. It should be stressed that the requirement is control *for*, not control *of*, this aspect of the interaction. Emotional involvement, the interactional processes stemming from it, are an essential element of many subject matters of psychiatry, and attempts to rule out this element by controls will inevitably change the nature of the problem being studied.

4. These problems of error and control are different for clinician and researcher and this can be the source of tension and misunderstanding. Loevinger (1963) put it thus:

The commitment of the clinician is to be as right as possible about every one of his cases. The commitment of the scientist is to be as right as possible about the nature of things. The clinician must minimize his error in every case. The scientist need not worry about errors in individual cases, since his methods allow for error. What is important for him is that the errors in separate observations be independent. Not random errors but repeated errors, errors of bias, render fruitless the scientific enterprise. Just at this point one finds frequent conflict in clinical research.

5. And finally these problems of control and the human factor are not as central in other types of information gathering situations. To again quote the GAP Report, using the simple example of observing a clock to determine the time it reports:

If the common problems in psychiatric research were actually involved in reading a clock, it would be necessary to consider, not only the possibility of the time being mis-read but also these additional possibilities: (1) that the frequency with which the clock is consulted may modify the time it reports; (2) that the time the clock is expected to show may modify the time it actually reports; (3) if the observer dislikes the clock (let us say from an aesthetic viewpoint), it will report time differently than if he is fond of it; (4) if the observer sends someone else to consult the clock, it will report differently; (5) that the time indicated by other clocks adjacent to the one being consulted or the position of this clock relative to other clocks might influence the time the clock in question reports.

For a fuller discussion of the implications and consequences of these and the many other aspects of the issue of controls in psychosocial research, the reader is referred to the GAP Committee report.

### *The problems of time-span and follow-up*

The problems of control link naturally to those of observational time-span and of follow-up, which in turn set the conditions for the discovery and verification of the dimensions of stability and of durability of change. Again, rather than extended discussion, we will state some of the highlights of these issues:

1. Clinical study of the psychoanalytic situation is by its nature a naturalistic and a longitudinal study with a time span at least as long as a (often very long) psychoanalytic treatment course. It carries an entirely different time commitment from the usual psychological experiment conceived to isolate for controlled study the events of a moment in time. To the researcher it is the commitment (and the risk) of a point in time, or at least of some manageable and circumscribed portion of time, as against what may seem the commitment (and the risk) of a research lifetime over the endless reaches of a longitudinal project. (And studies of therapeutic processes over the whole span of long-term psychoanalytic treatments, and involving a succession of cases, studied only in part concomitantly, and in part sequentially, can be as truly longitudinal as are the child developmental projects which span the growing years from infancy into adulthood.)

2. Adequate research, especially into issues of stability and durability of change, involves adequate follow-up and this is not a clinical tradition in psychoanalysis. Of this, Helene Deutsch said, in an article describing the post-analytic fates of two patients who she saw again 25 and 27 years after terminating their analyses:

The analytic literature is rich in case histories, in reports of failures and successes, and in theoretical interpretations. There is, however, an evident lack of information about the post-analytic psychic state of patients whose treatment has been successfully terminated (Deutsch, 1959).

Pfeffer reported in a series of three papers (1959, 1961b, 1963) from a psychoanalytic follow-up study of a succession of completed psychoanalyses, focused not just on outcomes, but on the bases and mechanisms of the perceived changes (the outcomes), and the ways in which the 'follow-up study transference phenomena' (Pfeffer, 1963) could be used to elucidate not only the original conflicts of the neurosis, but the solutions of those conflicts as achieved in

the analysis. Deutsch feels that this 'method of individual observation' exemplified by her follow-up study (and also Pfeffer's) is obviously the more reliable of what she feels to be the only two ways in which to evaluate the therapeutic outcomes of psychoanalysis, the other being the overall statistical reports of outcomes as submitted, for example, by the Chicago and the London Psychoanalytic Institutes, the Menninger Clinic, etc. (Deutsch, 1959). Obviously, one can disagree with Deutsch's entire over-simplified dichotomization of the possible follow-up paths. Systematic clinical research as we have amply attested is more than a single or even a succession of (just) individual case studies. But neither is it simply a statistical accounting. Among others, the PRP has evolved a follow-up methodology that combines the virtues (and eliminates many of the drawbacks) of each of these approaches, singly pursued. For fuller discussion of this issue, the reader is referred to that source (Sargent, 1960).

3. In research that is longitudinal and long-term, since analysis is long and adequate follow-up study time is long, there are inevitable problems of attrition of the subjects, and since (unlike with experimental animals) the subjects have by and large the same longevity as the investigators, there are concomitant problems of attrition of the investigators. Add to this that the researchers vary over the course of time in their research commitments and in their personal commitments, and the difficulties of the task that beset the workers in this field—the scientific difficulties and the extra-scientific difficulties—become obvious.

*The extra-scientific problems (the sociology of clinical research)*

Two sets of problems will be mentioned here, interpersonal and role problems and problems of time allocation and career lines.

*Interpersonal and role problems.* These are in part the interpersonal problems of any complexly organized group, and even more so, of the multidisciplinary groups almost unavoidable in formal clinical research. And in part they are the inevitable role problems created for clinicians who participate in research tasks, and who try to keep clear about the appropriate deployment of both their research and their clinical functions. These difficulties are those, then, of the ecology of research, an area that is only occasionally discussed in the public literature of research (Bush, 1957; Campbell,

1953; Cohen & Cohen, 1961; Luszki, 1958; Mitchell & Mudd, 1957; Redlich & Brody, 1955; Simmons & Davis, 1957), but is the *constant* preoccupation of the private discussions among researchers. As one of us has said elsewhere:

These are the large issues of maintaining a more or less cohesive and harmonious working together of research colleagues of varying degrees of clinical and research motivation, sophistication, and determination, working with a population of research subjects who are at the same time clinical patients subject to the vicissitudes of varying experiences in both their treatments and their lives that may or may not allow them to remain amenable and accessible to the research requirements over the projected life span of the research design . . . These difficulties of conducting a research study into the nature of the mechanisms of change in psychotherapy are the greater in a dedicated clinical center with its clinical service ideology and its strongly positive convictions about the high value, the seriousness, and the effectiveness of the psychotherapeutic work which is its central activity. By and large clinicians believe strongly in the great usefulness of their practical clinical endeavors, and of their theoretical guiding principles. Research that explores as open questions the nature of psychotherapeutic change and the effectiveness of psychotherapists as the instruments of that change thus readily evokes anxieties and resistances in such a clinical community. This is the more so, since the research data—the clinical *judgments* about therapeutic changes and outcomes, and about the therapeutic processes by which these changes have come about—can only with difficulty be separated from the *judge* who makes them despite the utmost honesty in the striving for such objectivity. And just as the judge is always part of the judgment, so, to the therapist whose work is being studied, must the assessment of the therapy always involve an assessment of the therapist (with its inevitably heavy weighting of the unconscious predispositions and biases, the countertransference elements) . . . Such factors obviously give rise to interpersonal tensions that beset and often seriously threaten research into the basic operations of a clinical community. That most research is free of such complications has to do with the fact that most research is about phenomena that are not the object of such fierce subjective convictions. It is the difference between doing research on the things that really count as against the things that are emotionally neutral or irrelevant (Wallerstein, 1966).

Lustman (1963), though he too talks of the difficulties inherent in the multidisciplinary nature of clinical research, and the multi-

theoretical character of the overall clinical field, at least finds this counterbalanced somewhat by the advantages of 'the provocative skepticism of university heterogeneity' which offers the most fertile soil for expansion in scientific knowledge. 'In such a structure, whether through intrigued interest or irritated awareness, multidisciplinary communication and teaching [can] occur.'

*Time allocation and career lines.* As 'little science' has increasingly grown into 'big science' in the most successful areas of natural science the picture of the effective researcher has, for very substantial and evident reasons, correspondingly evolved towards that of the full-time member of a complex research team. But for the analyst-researcher, turned towards research on the psychoanalytic process, the research must with rare exceptions necessarily remain only a part-time activity; since only the researcher who secures his experiences and maintains his skills in a continuing ongoing clinical analytic work will have maximal continuing fertilization of the research inquiry by clinically originated and clinically meaningful problems, and minimal danger of the research endeavour losing its analytic rootedness and relevance. Given this seemingly necessary state of affairs, the issue is then of squaring it with the question of how much concentration can be focused on the research task, how much complex case material can be kept in mind in fulfilment of the requirements of the research design, if one is *not* able to be wholly immersed in it and free of the time pressures of fixed and unremitting daily therapeutic commitments. For the most part this is solved (optimally?) at the expense of the private life of the clinician-researcher via the 'super' working day. Perhaps in the Neurosis Treatment Centers envisaged by Kubie (in Bronner, 1949) a more appropriate balance and therefore more optimal solution can ultimately be achieved, at least for those few.

#### SUMMARY

To summarize the purposes of a discursive journey through the dilemmas posed by the many issues in research into the psychoanalytic process discussed in this essay, we have attempted to confront side-by-side, with reference to both theory and practice, two questions relevant to our central thesis: Is it

*necessary* to conduct more formalized and systematized studies of the therapeutic process in psychoanalysis? And, is it *possible* to do? We feel, on grounds that we hope are cogent and persuasive, that the answer to both questions today is an emphatic yes. And yet we also hope that we have not sought, however unwittingly, to minimize the manifold real conceptual and technical difficulties encountered by the investigator who seeks to combine clinical relevance with scientific rigour.

Our central conviction is that the informal clinical case study (which is, as we have indicated, something like an 'experiment'), in spite of its compelling power, has certain real and obvious—and very formidable—scientific limitations. The major task for research in the clinical field and into the clinical process is the formalization of this highly artistic method into a disciplined research instrument which transcends our clinically satisfactory operating criteria of inner coherence and plausibility and clinical conviction bred of experience, and approaches the scientific criteria of systematic replicability; i.e. takes us beyond our 'rules of thumb' (Rapaport, 1960) to a theoretically coherent set of canons of clinical correspondence which can then be used to build thoroughly tested accumulating data into logically sustained new knowledge in psychoanalytic science. Psychoanalysis has historically underrated these complex problems of hypothesis-testing and verification. In part this has been because it has not wished a sterile scientism to obstruct genuine exploratory and investigative zeal; but in part this has been out of a historical tradition—and a particular constellation of scientific problems which conduced to that tradition—which has placed exclusive reliance on a single method of naturalistic observation by trained participant-observers. It is our belief that it is appropriate, feasible and very necessary to supplement that tradition now in order to make further progress towards the solutions of the problems which we have here so urgently raised.

#### ACKNOWLEDGEMENTS

This study of the issues in research in the psychoanalytic process has been in relation to the Therapeutic Process Study on the Modification of Defences in Psychoanalysis of the San Francisco Psychoanalytic Institute and the Mount Zion Hospital Department of Psychiatry (supported by NIMH Grant MH-13915) and the Psychotherapy Research Project of the Menninger Foundation (supported by NIMH Grant MH-8308, and previously by the Ford Foundation and the Foundations' Fund for Research in Psychiatry). Their generous assistance is gratefully acknowledged.

## REFERENCES

- BAKAN, D. (1954). A reconsideration of the problem of introspection. *Psychol. Bull.* **51**, 105-118.
- BAKAN, D. (1955). The general and the aggregate: a methodological distinction. *Percept. mot. Skills* **5**, 211-212.
- BAKAN, D. (1956). Clinical psychology and logic. *Am. Psychol.* **11**, 655-662.
- BELLAK, L. (1961). Research in psychoanalysis. *Psychoanal. Q.* **30**, 519-548.
- BELLAK, L. & SMITH, M. (1956). An experimental exploration of the psychoanalytic process: exemplification of a method. *Psychoanal. Q.* **25**, 385-414.
- BENJAMIN, J. D. (1950). Methodological considerations in the validation and elaboration of psychoanalytical personality theory. *Am. J. Orthopsychiat.* **20**, 139-156.
- BERGMAN, P. (1966). An experiment in filmed psychotherapy. In L. A. Gottschalk & A. H. Auerbach (eds.), *Methods of Research in Psychotherapy*. New York: Appleton-Century-Crofts.
- BIBRING, G. L. et al. (1961). A study of the psychological processes in pregnancy and of the earliest mother-child relationship. Appendices: Outline of variables; Glossary of defenses. *Psychoanal. Study Child* **16**.
- BRENNAN, M. (1947). Problems in clinical research: round table, 1946. *Am. J. Orthopsychiat.* **17**, 196-230.
- BRENNAN, M. (1948). Research in psychotherapy: round table, 1947. *Am. J. Orthopsychiat.* **18**, 92-118.
- BRODY, E. B., NEWMAN, R. & REDLICH, F. C. (1951). Sound recording and the problem of evidence in psychiatry. *Science, N.Y.* **113**, 379-380.
- BRONNER, A. F. (1949). The objective evaluation of psychotherapy: round table, 1948. *Am. J. Orthopsychiat.* **19**, 463-491.
- BUSH, V. (1957). Professional collaboration. *Science, N.Y.* **125**, 49-54.
- CAMPBELL, A. (1953). Administering research organizations. *Am. Psychol.* **8**, 225-230.
- CARMICHAEL, H. T. (1956). Sound film recording of psychoanalytic therapy: a therapist's experiences and reactions. In L. A. Gottschalk & A. H. Auerbach (eds.), *Methods of Research in Psychotherapy*. New York: Appleton-Century-Crofts, 1966.
- CHASSAN, J. B. (1957). On the unreliability of reliability and some other consequences of the assumption of probabilistic patient-states. *Psychiatry* **20**, 163-171.
- CHASSAN, J. B. (1967). *Research Design in Clinical Psychology and Psychiatry*. New York: Appleton-Century-Crofts.
- COHEN, R. A. & COHEN, M. B. (1961). Research in psychotherapy: a preliminary report. *Psychiatry* **24**, suppl. to no. 2, 46-61.
- CUSHING, J. G. N. (1952). Report of the committee on the evaluation of psychoanalytic therapy. *Bull. Am. psychoanal. Ass.* **8**, 44-50.
- DEUTSCH, H. (1959). Psychoanalytic therapy in the light of follow-up. *J. Am. psychoanal. Ass.* **7**, 445-458.
- DUKES, W. F. (1965).  $N = 1$ . *Psychol. Bull.* **64**, 74-79.
- EDWARDS, A. L. & CRONBACH, L. J. (1952). Experimental design for research in psychotherapy. *J. clin. Psychol.* **8**, 51-59.
- ENGEL, G. L. (1968a). Some obstacles to the development of research in psychoanalysis. *J. Am. psychoanal. Ass.* **16**, 195-204.
- ENGEL, G. L. (1968b). Some obstacles to the development of research in psychoanalysis: discussion by D. Beres, M. Kanzer, R. S. Wallerstein and E. R. Zetzel. *J. Am. psychoanal. Ass.* **16**, 205-229.
- ERIKSON, E. H. (1958). The nature of clinical evidence. *Daedalus* **87**, 65-87.
- ESCALONA, S. (1952). Problems in psycho-analytic research. *Int. J. Psycho-Anal.* **33**, 11-21.
- EZRIEL, H. (1951). The scientific testing of psycho-analytic findings and theory. *Br. J. med. Psychol.* **24**, 30-34.
- EZRIEL, H. (1952). Notes on psycho-analytic group therapy: interpretation and research. *Psychiatry* **15**, 119-126.
- FREUD, S. (1912). Recommendations to physicians practising psycho-analysis. *S.E.* **12**.
- FREUD, S. (1916-17). Introductory lectures on psycho-analysis. *S.E.* **15-16**.
- FREUD, S. (1920). The psychogenesis of a case of homosexuality in a woman. *S.E.* **18**.
- GAP COMMITTEE ON RESEARCH (1959). Some observations on controls in psychiatric research. (Report no. 42.)
- GILL, M. M. et al. (1968). Studies in audio-recorded psychoanalysis. I. General considerations. *J. Am. psychoanal. Ass.* **16**, 230-244.
- GLOSSARY OF TERMS. Prediction Study of the Psychotherapy Research Project, Menninger Foundation, Topeka, Kansas. (Unpubl. ms.)
- GLOVER, E. (1952). Research methods in psychoanalysis. *Int. J. Psycho-Anal.* **33**, 403-409.
- GOODMAN, N. (1955). *Fact, Fiction and Forecast*. Cambridge, Mass.: Harvard Univ. Press.
- HAGGARD, E. A., BREKSTAD, A. & SKARD, Å. G. (1960). On the reliability of the anamnestic interview. *J. abnorm. soc. Psychol.* **61**, 311-318.
- HAGGARD, E. A., HIKEN, J. R. & ISAACS, K. S. (1965). Some effects of recording and filming on the psychotherapeutic process. *Psychiatry* **28**, 169-191.

- HAGGARD, E. A. *et al.* (1968). The diagnostic interview as a data collection instrument. In Feldman (ed.), *Studies in Psychotherapy and Behavioral Change*, vol. 1. Buffalo: Buffalo Studies, Univ. Press, SUNY.
- HARTMANN, H. (1958). Comments on the scientific aspects of psychoanalysis. *Psychoanal. Study Child* 13.
- HARTMANN, H. (1959). Psychoanalysis as a scientific theory. In S. Hook (ed.), *Psychoanalysis, Scientific Method and Philosophy*. New York: New York Univ. Press.
- JANIS, I. L. (1958). *Psychological Stress*. New York: Wiley.
- KNAPP, P. H. (1968). Discussion of 'Studies in audio-recorded psychoanalysis. II. The effect of recording upon the analyst' by J. Simon *et al.* (Meeting of Am. Psychoanal. Ass., Boston.)
- KNAPP, P. H., MUSHATT, C. & NEMETZ, S. J. (1966). 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.
- KRIS, E. (1947). The nature of psychoanalytic propositions and their validation. In S. Hook & M. R. Konwitz (eds.), *Freedom and Experience: Essays Presented to Horace M. Kallen*. New York: Cornell Univ. Press.
- KUBIE, L. S. (1956). The use of psychoanalysis as a research tool. *Psychiat. Res. Rep.* no. 6, 112-136.
- KUBIE, L. S. (1958). Research into the process of supervision in psychoanalysis. *Psychoanal. Q.* 27, 226-236.
- LOEVINGER, J. (1963). Conflict of commitment in clinical research. *Am. Psychol.* 18, 241-251.
- LUBORSKY, L. (1953). Intraindividual repetitive measurements (P technique) in understanding psychotherapeutic change. In O. H. Mowrer (ed.), *Psychotherapy: Theory and Research*. New York: Ronald Press.
- LUSTMAN, S. L. (1963). Some issues in contemporary psychoanalytic research. *Psychoanal. Study Child* 18.
- LUSZKI, M. B. (1958). Interdisciplinary team research: methods and problems. (Research training series, no. 3. National Training Laboratories. New York: New York Univ.)
- MARGOLIN, S. G. (1951). The behavior of the stomach during psychoanalysis: a contribution to a method of verifying psychoanalytic data. *Psychoanal. Q.* 20, 349-373.
- MATARAZZO, R. G., WIENS, A. N. & SASLOW, G. (1966). Experimentation in the teaching and learning of psychotherapy skills. In L. A. Gottschalk & A. H. Auerbach (eds.), *Methods of Research in Psychotherapy*. New York: Appleton-Century-Crofts.
- MITCHELL, H. E. & MUDD, E. H. (1957). Anxieties associated with the conduct of research in a clinical setting. *Am. J. Orthopsychiat.* 27, 310-323.
- MOORE, B. E. & FINE, B. D. (1967). *A Glossary of Psychoanalytic Terms and Concepts*. New York: Am. Psychoanal. Ass.
- PFEFFER, A. Z. (1959). A procedure for evaluating the results of psychoanalysis: a preliminary report. *J. Am. psychoanal. Ass.* 7, 418-444.
- PFEFFER, A. Z. (1961a). Panel report: Research in psychoanalysis. *J. Am. psychoanal. Ass.* 9, 698-718.
- PFEFFER, A. Z. (1961b). Follow-up study of a satisfactory analysis. *J. Am. psychoanal. Ass.* 9, 562-570.
- PFEFFER, A. Z. (1963). The meaning of the analyst after analysis: a contribution to the theory of therapeutic results. *J. Am. psychoanal. Ass.* 11, 229-244.
- RAMZY, I. (1962). Research in psychoanalysis: contribution to discussion. *Int. J. Psycho-Anal.* 43, 292-296.
- RAMZY, I. (1963). Research aspects of psychoanalysis. *Psychoanal. Q.* 32, 58-76.
- RANGELL, L. (1954). Similarities and differences between psychoanalysis and dynamic psychotherapy. *J. Am. psychoanal. Ass.* 2, 734-744.
- RAPAPORT, D. (1960). The structure of psychoanalytic theory: a systematizing attempt. *Psychol. Issues* 2, monogr. 6.
- REDLICH, F. C. & BRODY, E. B. (1955). Emotional problems of interdisciplinary research in psychiatry. *Psychiatry* 18, 233-239.
- ROBBINS, L. L. & WALLERSTEIN, R. S. (1959). The research strategy and tactics of the Psychotherapy Research Project of the Menninger Foundation and the problem of controls. In E. I. Rubinstein & M. B. Parloff (eds.), *Research in Psychotherapy*, vol. 1. Washington, D.C.: Am. Psychol. Ass.
- ROOSE, L. J. (1960). The influence of psychosomatic research on the psychoanalytic process. *J. Am. psychoanal. Ass.* 8, 317-334.
- SARGENT, H. D. (1960). Methodological problems of follow-up studies in psychotherapy research. *Am. J. Orthopsychiat.* 30, 495-506.
- SARGENT, H. D. (1961). Intrapsychic change: methodological problems in psychotherapy research. *Psychiatry* 24, 93-108.
- SARGENT, H. D. *et al.* (1968). Prediction in psychotherapy research: a method for the transformation of clinical judgments into testable hypotheses. *Psychol. Issues* 6, monogr. 21.
- SCHAFER, R. (1967). *Projective Testing and Psychoanalysis*. New York: Int. Univ. Press.
- SCHLESINGER, H. J. (1967). Discussion of 'Studies in audio-recorded psychoanalysis. I. General considerations' by M. M. Gill *et al.* (Meeting of Am. Psychoanal. Ass., New York.)

- SCHMIDL, F. (1955). The problem of scientific validation in psycho-analytic interpretation. *Int. J. Psycho-Anal.* 36, 105-113.
- SEITZ, P. F. D. (1966). The consensus problem in psychoanalytic research. In L. A. Gottschalk & A. H. Auerbach (eds.), *Methods of Research in Psychotherapy*. New York: Appleton-Century-Crofts.
- SHAKOW, D. (1959). Discussion. In E. I. Rubinstein & M. B. Parloff (eds.), *Research in Psychotherapy*, vol. 1. Washington, D.C.: Am. Psychol. Ass.
- SHAKOW, D. (1960). The recorded psychoanalytic interview as an objective approach to research in psychoanalysis. *Psychoanal. Q.* 29, 82-97.
- SHAPIRO, M. B. (1961). The single case in fundamental clinical psychological research. *Br. J. med. Psychol.* 34, 255-262.
- SIEGAL, R. S. & ROSEN, I. C. (1962). Character style and anxiety tolerance: a study in intrapsychic change. In H. H. Strupp & L. Luborsky (eds.), *Research in Psychotherapy*, vol. 2. Washington, D.C.: Am. Psychol. Ass.
- SIMMONS, O. G. & DAVIS, J. A. (1957). Interdisciplinary collaboration in mental illness research. *Am. J. Sociol.* 63, 297-303.
- SIMON, J. et al. (1970). Studies in audio-recorded psychoanalysis. II. The effect of recording upon upon the analyst. *J. Am. psychoanal. Ass.* 18, 86-101.
- SNYDER, W. U. (1959). Some investigations of relationships in psychotherapy. In E. I. Rubinstein & M. B. Parloff (eds.), *Research in Psychotherapy*, vol. 1. Washington, D.C.: Am. Psychol. Ass.
- STERNBERG, R. S., CHAPMAN, J. & SHAKOW, D. (1958). Psychotherapy research and the problem of intrusions on privacy. *Psychiatry* 21, 195-203.
- STRUPP, H. H. (1960). Some comments on the future of research in psychotherapy. *Behav. Sci.* 5, 60-70.
- STRUPP, H. H., CHASSAN, J. B. & EWING, J. A. (1966). 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.
- SVENSON, S. E. & CHASSAN, J. B. (1967). A note on ethics and patient consent in single-case design. *J. nerv. ment. Dis.* 145, 206-207.
- VAN VLACK, J. (1966). Filming psychotherapy from the viewpoint of a research cinematographer. In L. A. Gottschalk & A. H. Auerbach (eds.), *Methods of Research in Psychotherapy*. New York: Appleton-Century-Crofts.
- WAELDER, R. (1962). Psychoanalysis, scientific method and philosophy. *J. Am. psychoanal. Ass.* 10, 617-637.
- WAELDER, R. (1963). Psychic determinism and the possibility of predictions. *Psychoanal. Q.* 32, 15-42.
- WALLERSTEIN, R. S. (1960). The Psychotherapy Research Project of the Menninger Foundation: Third Report. Helen D. Sargent and the Psychotherapy Research Project. *Bull. Menninger Clin.* 24, 159-163.
- WALLERSTEIN, R. S. (1964). The role of prediction in theory building in psychoanalysis. *J. Am. psychoanal. Ass.* 12, 675-691.
- WALLERSTEIN, R. S. (1966). The Psychotherapy Research 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.
- WALLERSTEIN, R. S. (1967). Panel report: Development and metapsychology of the defence organization of the ego. *J. Am. psychoanal. Ass.* 15, 130-149.
- WALLERSTEIN, R. S. (1968). Discussion of 'Some obstacles to the development of research in psychoanalysis' by G. L. Engel. *J. Am. psychoanal. Ass.* 16, 215-222.
- WALLERSTEIN, R. S., ROBBINS, L. L., SARGENT, H. D. & LUBORSKY, L. (1956). The Psychotherapy Research Project of the Menninger Foundation: rationale, method and sample use. *Bull. Menninger Clin.* 20, 221-278.
- WHEELIS, A. (1956). The vocational hazards of psycho-analysis. *Int. J. Psycho-Anal.* 37, 171-184.

Copyright © Robert S. Wallerstein and Harold Sampson

Department of Psychiatry  
Mount Zion Hospital,  
1600 Divisadero Street,  
San Francisco, Calif. 94115