

Single-Case Methodology in Psychotherapy Process and Outcome Research

Russell B. Hilliard

The primary focus of this article is on delineating and explaining the essential characteristics of single-case research methodology applied within the domain of psychotherapy research. Single-case research is presented as a subclass of intrasubject research in which aggregation across Ss is avoided and the generality of one's findings is addressed through replication on a case-by-case basis. The basic ways in which single-case designs vary are also discussed, and 3 basic types of single-case research are differentiated: (a) single-case experiments, (b) single-case quantitative analyses, and (c) case studies. Furthermore, some of the major weaknesses in current single-case psychotherapy research are identified.

Single-case designs have suffered from association with the research design that Campbell and Stanley (1963) originally referred to as the "one-shot case study." The term was used to refer to a design that had no control group, assessed the dependent variable only once, and was not replicated. In their words, "such studies have such a total absence of control as to be of almost no scientific value. The design is introduced here as a minimum reference point" (p. 6).

Although many researchers now associate this type of design with case studies, and even with single-case designs in general, Campbell (Cook & Campbell, 1979) has moved away from such an evaluation of case study methodology, stating more recently, "Certainly the case study as normally practiced should not be demeaned by identification with the one-group posttest-only design" (p. 96), and furthermore, "While it is probable that many case studies professing or implying interpretation or explanation, or relating the case to theory, are guilty of these faults, it now seems to me clear that not all are, or need be" (Campbell, 1979, p. 57). He has gone on to write the foreword to a volume dedicated exclusively to case study methodology (Yin, 1989).

A similar shift is seen in the writings of a leading psychotherapy research methodologist. Kiesler (1981) has stated

Both Hersen and Barlow (1976) and Gelso (1979) take me to task, and rightly so, for my earlier conclusion that single case study "has little place in the confirmatory aspect of scientific activity" (Kiesler, 1971, p. 66). As it stands, this statement is inadequate, contains ambiguities, and needs correction and clarification. (p. 213)

More recently, he has gone further, claiming, "Studies seriously pursuing these [psychotherapy] change-process goals cannot

attain them by use of traditional, rigorous experimental or nomothetic designs. Instead, what seems to be most appropriate and necessary are small *N* or single-case studies" (1983, p. 13).

Kiesler's view is representative of the widespread resurgence of interest in single-case designs within psychotherapy research. However, little systematic attention has been given to describing what single-case research actually is.¹ The wide variety of terms used by both advocates and critics to refer to single-case research—*single case*, *N of 1*, *case study*, *small N*, *idiographic*, *intensive*, *discovery-oriented*, *intrasubject*, and *time-series*—have only served to confuse the underlying methodological issues involved in this type of research.

The primary focus of this article is on delineating and explaining the essential characteristics of single-case research methodology. Also, some basic ways in which single-case designs may vary will be identified, and separate terms for the different basic types of single-case research will be proposed. My goal is to create greater clarity in thinking and communicating about single-case designs and thus facilitate a more productive discussion of the potential and limitations of such research in studying the psychotherapy change process. Furthermore, some of the major weaknesses in current single-case research within psychotherapy are identified here.

The Basic Logic of Single-Case Research

I turn first to the question, What is single-case research and why would one choose such a research strategy? Single-case research is often viewed merely in terms of its sample size; that is, as *N-of-1* research. From the perspective of group research, a design with an *N* of 1 is at best highly limited, if not outright meaningless. However, viewing single-case research along the same lines as group research is highly misleading. Such a view reflects an inadequate appreciation of what single-case research actually is.

I propose that single-case research is best viewed as a subclass of *intrasubject* research in which aggregation across cases

I express deep appreciation to Hans H. Strupp, my mentor, whose unabated scientific interest in single-case methodology inspired this article. I also thank Irene Elkin, William P. Henry, Steven D. Hollon, and Andrew J. Tomarken for their insightful comments on earlier versions of this article.

Russell B. Hilliard is currently at The Cambridge Hospital/Harvard Medical School, Cambridge, Massachusetts.

Correspondence concerning this article should be addressed to Russell B. Hilliard, care of Hans H. Strupp, Department of Psychology, 301 Wilson Hall, Vanderbilt University, Nashville, Tennessee 37240.

¹ This has been done for single-case experiments within the behavioral tradition but not for single-case research as a whole.

is avoided and the generality of one's findings is addressed through replication on a case-by-case basis. Each of these basic characteristics of single-case research will be discussed in turn. The focus here is primarily on the conceptual reasons for pursuing this methodology, although I will consider pragmatic reasons in passing.

Single-Case Research as Intrasubject Research

It is necessary to clearly distinguish between two basic types of variation within psychotherapy research and other areas of psychology—namely, variation within subjects and variation across subjects. The former is commonly referred to as *intrasubject* variation and the latter as *intersubject* variation. Intrasubject variation implies that a variable is free to vary within individual subjects. Each variable can only take one value at a specific time point within an individual; thus, repeated measures of the variable(s) over time within the subject are involved. If the variable represents an ordinal, interval, or ratio measure, intrasubject variation will refer to variation over time in the amount of the respective variable and is traditionally plotted with time on the *x*-axis and quantity on the *y*-axis. If the variable represents a nominal measure, intrasubject variation will refer to variation over time in the level of the variable and is traditionally represented as a string of codes that correspond to the different levels of the nominal variable. Intersubject variation refers to differences between or across subjects and typically involves a cross-sectional perspective, in contrast to the longitudinal perspective inherent in intrasubject variation.

Here, the term *intrasubject designs* refers to research that focuses on the temporal unfolding of variables within individual subjects.² Not all intrasubject research is single-case research, but all single-case research is intrasubject research. In other words, characteristic of all single-case research is a focus on variation within subjects over time and the attempt to understand this variation as a function of other variables that vary within subjects over time or as a function of variables that vary across subjects (within a single-case paradigm, the latter requires replication across cases, as will be discussed later). Examining only one time point within single-case research would be as nonsensical as examining only one subject within group research.

Given that all single-case research focuses on intrasubject variation, the relevance of such variation to psychotherapy research is briefly considered here. Few would disagree that intrasubject variation is relevant to psychotherapy process research (the study of patient-therapist interaction within the therapy session). The term *process* implies the temporal unfolding of variables within therapeutic dyads. However, the relevance of intrasubject variation to psychotherapy outcome research is not as apparent, with its traditional focus on global outcome—that is, pre- to post-treatment change—ignoring the process resulting in such change.

Kiesler (1983) and Greenberg (1986a), among others, have argued that if we are ever to discover what aspects of a particular psychotherapeutic modality actually result in therapeutic change and how they do so, we must break down global out-

come into a series of smaller interrelated changes and attempt to discover how the therapist's interventions and patient's responses (i.e., the therapeutic process) contribute to or explain these smaller changes. In this approach, there is a shift from an exclusive focus on the therapeutic process or therapeutic outcome to a focus on the *process of change* or *process of outcome*. Greenberg (1986a) offered the following broad description of this approach:

In studying the process of change, both beginning points and end-points are taken into account, as well as the form of the function between these points. With processes of change as the focus of the investigation, the emphasis is not on studying what is going on in therapy (process research) nor only on the comparison of two measurement points before and after therapy (efficacy research) but rather on identifying, describing, explaining, and predicting the effects of the processes that bring about therapeutic change over the entire course of therapy. (p. 4)

From this perspective, a focus on variability within therapeutic dyads over time (i.e., intrasubject variability) is at the very heart of psychotherapy research.

However, psychotherapy research has tended either to simply ignore intrasubject variability (the situation in most outcome research) or to assess it indirectly through cross-sectional group correlations (the situation in most research attempting to link process and outcome). The problematic nature of assessing intrasubject variation indirectly through cross-sectional group correlations has been emphasized by numerous authors, although the field has been slow to heed such warnings.

Gottman and Markman (1978) pointed out this problem within psychotherapy research:

The kind of correlation Kiesler (1973) described [a high positive correlation between therapists' accurate empathy behavior and eventual successful outcome] has nothing to do with process notions of "reciprocity" in social interaction; Kiesler referred to the correlation *across* clients and therapists of rates of behaviors and not a correlation of a particular therapist's behavior with a particular client's behavior. The distinction is critical. . . . Kiesler's correlation does not describe contingent interaction *within* dyads but correlates base rates of behaviors *across* dyads. The process

² Although intrasubject designs always involve a time series of data in a broad sense of the term, the term *time-series designs* is not used here because it could be taken to imply that time-series analysis was the only appropriate analytic strategy in such research. The analysis of variation or covariation over time may take different forms (e.g., time-series analysis, sequential analysis, growth curve analysis), depending on the properties of one's data and one's assumptions concerning the underlying temporal processes (the reader interested in general introductions to some of the main quantitative techniques for analyzing intrasubject variability is referred to Bakeman & Gottman, 1986; Gottman, 1981; and Gottman & Roy, 1990). Furthermore, although intrasubject research involves a within-subjects analysis in a broad sense of the term, the term *within-subjects designs* was not chosen because it is customarily used to refer to a design that involves a very limited number of assessments within a given subject, not yet sufficient to constitute a time-series of data. Also, within-subjects designs involve aggregation across cases, which—as explained later—is avoided in single-case designs. The term *intrasubject* is neutral in terms of aggregation across cases.

question, "How does a particular therapist's behavior affect the client?" is a within-dyad [intrasubject] question. (p. 28)

As Gottman and Markman stated, the assessment of the impact of the therapist on the patient and vice versa requires assessing the relevant therapist and patient behaviors over time within therapeutic dyads—assessing the intrasubject variation in the relevant variables—and applying appropriate data-analytic techniques to these data. Unless one has clear conceptual, empirical, or both types of reasons for making the leap from the cross-sectional group level to the longitudinal individual level, intrasubject variability should be assessed directly whenever one's basic questions pertain to such variability.

The Problem of Aggregation Across Cases

So far the focus has been on intrasubject research in general. I have attempted to show that there are very good reasons for focusing on intrasubject variability within psychotherapy research and that when this is done, it should be done directly. After assessing intrasubject variation directly, one is faced with the question of whether to aggregate these measures of intrasubject variation across a number of subjects or to analyze the data for each subject separately, at least initially. Single-case methodology refers to the situation within intrasubject research in which the data are not aggregated across any subjects, even as few as 2; rather, they are analyzed separately on a case-by-case basis.³

To clarify why single-case research avoids aggregation across cases, I address the problematic nature of aggregating measures of intrasubject variation over even as few as 2 subjects. First, the situation in which the basic form of the relation—the pattern of intrasubject variation—varies across subjects will be considered, and then the situation in which the basic form of the relation is similar or the same across subjects. Aggregating across subjects prematurely may be highly misleading when the basic form of the relation is different across subjects. For example, one may find that there are three basic patterns across subjects in intrasubject variability. A certain class of intervention (e.g., transference interpretations) may lead to ever deeper insight over the course of therapy in one type of patient, ever greater resistance in another type of patient, and no impact on a third type of patient. Obviously, if one were to fail to segregate the three types of individuals before the data analysis, the average impact of the intervention may very well appear to be a flat curve.

What about the situation when the form of the relation is the same or similar across subjects in a group, that is, when one has "homogeneous" groups? One would be tempted to assume that, in this case, aggregation across subjects is harmless, but this is not necessarily true. Sidman (1952) was one of the first researchers to show the danger of seeking patterns in averages, even within homogeneous groups. He discussed the situation with reference to the functional relation between two continuous variables, showing that even if the form of the relation was the same across every individual, the average values of the two respective variables may be related in a way that differs funda-

mentally. For example, consider the classic negatively accelerated positive growth function. If there are individual differences in the asymptotes approached by the curves or in the rates of approach to the asymptotes, although all the individual curves are described by the same function, Sidman showed mathematically that the average curve cannot be described by the same function.

Besides distorting the form of a relation, averaging across subjects can obscure the "fine grain" that individual data can show. Mook (1982) gives the example of dark-adaptation curves, where the two-phase decrease in threshold apparent in individual dark-adaptation curves can be totally obscured by averaging over as few as 2 subjects, resulting in a gradual, continuous curve.

An example closer to psychotherapy research is provided by Henry, Schacht, and Strupp (1986), who examined interpersonal complementarity in psychotherapy using the quadrant version of Benjamin's (1974) *Structural Analysis of Social Behavior*. Because there were significant individual differences in the number of codes in each quadrant, it was necessary to calculate the base rates for each case separately and to compare conditional and unconditional probabilities on a case-by-case basis. Complementarity beyond the chance level was found in almost all cases, thus the group of subjects was homogeneous in this respect. However, if the data had been pooled across all subjects before complementarity was assessed, the finding would probably have been masked, given that the subjects were not homogeneous in terms of their underlying base rates.

The issue of homogeneity is not as simple and straightforward as it might first appear. A single-case approach would exert great caution in assuming that a group is truly homogeneous until this has been clearly shown, which would involve considerable understanding of the phenomenon of interest at the single-case level. Once a phenomenon is sufficiently well understood at the single-case level, intelligent aggregation may be possible. From this perspective, a program of research would begin with the study of single cases and then possibly move on to aggregation over groups that have been established as truly homogeneous. Three factors should be considered when exploring the issue of true homogeneity: (a) the nature of the question being asked, (b) how the pattern of intrasubject variation differs across subjects, and (c) the particular aggregational statistic being used. The question of aggregation across cases is complex and cannot be addressed thoroughly here. The attempt has been made simply to show why proponents of single-case research are suspicious of aggregation across cases, an issue that much of psychotherapy research has largely ignored.

Although the focus here has been on aggregation *across* subjects, it should be added that aggregation *within* subjects may prove to be as misleading as aggregation across subjects. For example, if one aggregates within a single subject over different

³ The terms *case* or *subject* may refer either to a single individual (the patient or therapist) or to a therapeutic dyad, depending on whether the phenomenon of interest is analyzed at the individual or dyadic level.

phases of therapy (e.g., beginning, middle, and end), and the relation between the examined variables varies across the different phases, it is very possible that the systematic pattern within each phase will be obscured. This issue is not just an aggregational issue but also an issue of the amount of data collected: To examine the relation between two variables within each phase of therapy, one must collect a sufficient amount of data from each phase to make this possible.

The terms *intensive* and *extensive* designs, introduced by Chassan (1979) to distinguish between single-case and group research, capture this difference between single-case research, which involves gathering a large amount of data on a limited number of subjects, and group research, which involves gathering a small amount of data on a large number of subjects. If a researcher is interested in gathering a large amount of data per subject, he or she will probably be limited to a small sample size given the investment of research resources necessary per subject. This is the primary pragmatic reason for engaging in single-case or small *N* research. Although one could certainly argue that a concern with both aggregation across and within subjects should be characteristic of single-case research, this type of research has not always shown sufficient sensitivity to the issues involved in the latter type of aggregation (for an interesting discussion of the issue of aggregation within subjects in single-case research, see Hill, Carter, & O'Farrell, 1983, and the ensuing comments by Howard, 1983, and Lichtenberg & Heck, 1983).

The Question of Generality Within Single-Case Research

Thus far, it has been proposed that single-case designs are intrasubject designs that avoid aggregation across cases. The question remains of how to establish the generality of one's findings within a single-case research paradigm.

There are certain situations, as Mook (1983) has argued, in which one is not interested in the external validity of the findings (e.g., when one is interested in establishing what is possible rather than what is common). In such situations, one is content with studying an isolated case. What about the more common situation in which one *is* interested in determining the generality of the findings? The basic approach to establishing the generality of one's findings within a single-case paradigm is expressed well by Thorngate (1986):

To find out what people do in general, we must first discover what each person does in particular, then determine what, if anything, these particulars have in common. . . . Nomothetic laws lie at the intersection of idiographic laws; the former can be discovered only after we find the latter. (pp. 75–76, [Thorngate's cited references omitted])

From this perspective, the generality of the findings would not be determined by means of group aggregates but by replication on a case-by-case basis.

Although single-case methodology has been identified with an exclusive idiographic focus in the minds of many, this identification is simply not warranted. Most single-case research clearly involves determining the generality across subjects of

the relations uncovered at the individual, or idiographic, level. A prime example of this would be Skinner's (1953) use of single-case research to study the principles of operant conditioning. From this perspective, it would be more accurate, although clumsier, to refer to single-case designs in which researchers are interested in the generality of their findings with a term Gottman (1973) proposed: *N-of-one-at-a-time* designs.

Within the behavioral tradition of single-case research, there is an important distinction between two types of replication across subjects (Sidman, 1960). The first, designated *direct replication*, refers to the attempt to replicate the findings in subjects that are similar in terms of the individual-differences variables that are viewed as affecting the phenomenon of interest. The second type of replication, *systematic replication*, refers to the attempt to show that the findings differ in predictable ways when one selects subjects that differ along the critical individual-difference variables.

Although single-case research always involves the study of intrasubject variability, it may also involve studying intersubject variability in patterns of intrasubject variability. However, within a single-case paradigm, the impact of any variable that does not vary within an individual can only be assessed through systematic replication. This point has not been sufficiently appreciated by certain proponents of single-case methodology. For example, Greenberg (1986b) claimed, "In single-case studies, the population to which the study refers is so well specified (initially the single case) that tests of hypotheses concerning treatment effects in relation to individual difference variables can be made directly" (p. 730). This claim begs the question of how hypotheses concerning individual differences will be tested in an individual given that there is, by definition, little or no variability in such variables within individuals.

Given the critical importance in psychotherapy research of many variables such as diagnosis or global outcome, which do not vary within individuals (i.e., within the period of time of the study), the importance of systematic replication across individuals should be apparent. The lack of both direct and systematic replication is one of the great weaknesses of most single-case research within psychotherapy research.

Strupp (1980a, 1980b, 1980c, 1980d) presented a forward-looking example of the forms that both direct replication and systematic replication can take within psychotherapy research. In a series of four case studies, he compared four sets of 2 patients, each of the 2 patients having been treated by the same therapist. All 8 of the patients were highly similar in terms of certain individual-difference variables (diagnosis, sex, age, and occupation), but within each pair was 1 patient with a good outcome and 1 with a poor outcome. The time has come for such programmatic single-case research to become the rule, rather than the exception, within the field.

Basic Categories of Single-Case Research

Single-case methodology has been presented as intrasubject research in which aggregation across cases is avoided and the generality of findings is addressed through replication on a case-by-case basis. Having attempted to identify and explain

the essential characteristics of single-case research, I would like now to point out some basic ways in which single-case designs may vary. On the basis of these differences, three basic categories of single-case research will be differentiated. Separate terms for each category will be proposed. Certain mixed categories will also be considered. The reason for differentiating between these different types of single-case designs is to highlight the diversity possible within this group of designs and to create greater precision in thinking and communicating about single-case research.

The first way in which single-case designs vary is whether one's data are quantitative or qualitative. Because these terms are not used consistently in the literature, I will clarify how they are used here. The term *quantitative* is sometimes used to designate measurement scales having ordinal, interval, or ratio properties, whereas the terms *qualitative* or *categorical* are used to designate scales that have nominal properties. In other cases, nominal scales are viewed as quantitative to the extent that one counts frequencies of different levels and applies descriptive and inferential statistics to the results. In such cases, the term *qualitative* refers to the situation in which no formal quantification occurs, even at the nominal level. Such data are expressed in prose and have the nature of a narrative. The latter distinction is followed here, given the highly sophisticated methods of quantitative analysis that can be applied to nominal data. Single-case research may involve either quantitative or qualitative data.

Another important way in which single-case designs vary is whether the independent variables are directly manipulated by the experimenter. Cook and Campbell (1979) have suggested the term *experimental* for studies in which the independent variable is directly manipulated by the researcher and the term *passive-observational* for studies in which this is not the case. Single-case research can be of either type. There is nothing inherent in single-case designs that requires either direct manipulation or passive observation.

A third aspect is whether the focus of the study is testing hypotheses that have been formulated a priori or generating hypotheses to be tested in later research. The former has come to be referred to as the *context of justification* and the latter as the *context of discovery*. The direct manipulation of an independent variable presupposes an a priori hypothesis; otherwise, one would not know how the independent variable should be manipulated. However, passive observation can be used to both generate hypotheses and test hypotheses. Although there is a general consensus that direct manipulation is necessary for the most rigorous test of a causal hypothesis, passive observation can be used to test hypotheses about concomitant variation, hypotheses that are often highly relevant to causal claims. Single-case research is not necessarily limited to the context of discovery.

On the basis of these basic ways in which single-case designs may vary, three basic categories of single-case research can be differentiated: (a) *single-case experiments*, (b) *single-case quantitative analyses*, and (c) *case studies*. Each of these categories may involve direct and systematic replication across cases or may be limited to an isolated case.

The first two categories involve quantitative data. Single-case designs based on quantitative data may involve direct manipulation or passive observation. Typically, direct manipulation is involved when one is testing hypotheses. I propose that the term *single-case experiment* be limited to single-case designs that involve quantitative data and direct manipulation of the independent variable. Single-case experiments have been conducted almost exclusively within behaviorally oriented research. This tradition has reached a high level of methodological sophistication, as reflected in volumes such as Kazdin (1982) and Barlow and Hersen (1984). The contribution to this special series by Moras, Telfer, and Barlow (1993) is an exemplar of this category of single-case research applied to a treatment that is not primarily behavioral.

Quantitative single-case designs may also involve passive observation. The term *single-case quantitative analysis* is proposed for this category. This category refers to the situation in which quantitative techniques for analyzing the temporal unfolding of variables, such as time-series analysis, sequential analysis, and growth curve analysis, are applied to single cases without the direct manipulation of any of the variables studied. Of the three basic categories of single-case research, this is the category with the least clear identity from a historical perspective, reflected by the lack of an established term to refer to this category. Either hypothesis testing or hypothesis generation may be the goal of such studies, depending on whether the researcher's hypotheses are formulated a priori. The terms *confirmatory* and *exploratory* could be used to distinguish these two types of single-case quantitative analyses. The contributions to this special series by Jones, Ghanam, Nigg, and Dyer (1993), Spence, Dahl, and Jones (1993), Silberschatz and Curtis (1993), and Horowitz et al. (1993) all represent exemplars of this category of single-case research. The variety of these four studies illustrates the breadth of this category.

The third category, *case study*, is proposed for designs involving qualitative data. Historically, case studies do not involve formal quantification.⁴ Although, in principle, qualitative single-case designs could involve either direct manipulation or passive observation, I am only aware of designs based on passive observation. For this reason, I propose that the term *case study* be used to imply both qualitative data and passive observation. This approach has been associated primarily—although not exclusively—with psychoanalysis, going back to Freud's classic case studies.⁵

⁴ It has been stated above that a focus on intrasubject variability, that is, the temporal unfolding of variables within subjects, is characteristic of all single-case research. Some individuals may have difficulty thinking of case studies in these terms given the dimension of time or sequence is not formally quantified. However, the story the case study presents is nothing other than a qualitative account of intrasubject variability.

⁵ Both advocates and critics of case study methodology have identified this approach with a hermeneutic epistemology. Whether one accepts this view or not (it has been strongly challenged by some, e.g., Edelson, 1985), the issue of quality control within qualitative research remains. Stiles (1991) has begun to address how this can be done within qualitative psychotherapy research.

In an analogous manner to single-case quantitative analyses, case studies may differ in terms of whether hypotheses have been generated before the study and an attempt is made to test them or whether the focus is on generating hypotheses. Again, the terms *confirmatory* and *exploratory* could be used to distinguish between these two situations. Although the notion of a theory-testing, or confirmatory, case study may appear foreign to some, there are numerous examples of such confirmatory case studies in certain fields; in ethnography, for example (see e.g., Campbell, 1979; Rosenblatt, 1981), researchers seeking evidence in support of prior beliefs and theories have found that they were wrong. Confirmatory case studies are almost nonexistent within psychotherapy research because the researcher's hypotheses are seldom specified with sufficient precision to permit the identification of disconfirming cases.

Soldz (1990) has introduced the term *research-informed case study* to refer to case studies that are basically qualitative in nature but in which the individual cases have been selected from a traditional between-groups design in terms of quantitative criteria, for example, on the basis of outcome measures in a comparative treatment study. Furthermore, such case studies incorporate formal quantitative data to provide greater support for the qualitative judgments they make. This qualification of the term *case study* reflects sensitivity to two important issues, although the term proposed does not clearly express them.

The first issue is whether the cases to be studied at the single-case level are chosen on some systematic basis from a group design. Strupp (e.g., 1980a, 1980b, 1980c, 1980d, 1990; Strupp, Schacht, Henry, & Binder, 1992) has provided numerous examples of single-case research in which the respective cases were chosen from the large Vanderbilt I and Vanderbilt II group design research projects. Grawe (1992) has also done this within the Bernese Comparative Treatment Study. Other examples could be given. This strategy permits one to select and place a given case in terms of where it stands in comparison to other cases. Such combinations of group and single-case research appear promising and deserve systematic exploration (see Elkin, 1991).

The other issue raised by Soldz's (1990) term is the possibility that both quantitative and qualitative data are assessed within the same single-case study. Soldz addressed the situation where a case study, in which the quantification process is not formalized, draws upon quantitative data. Also, a quantitative single-case analysis may draw upon some qualitative data. Thus, one may have quantitatively informed case studies and qualitatively informed single-case quantitative analyses. Strupp's case studies are exemplars of the former, and Grawe's (1992) single-case quantitative analyses are exemplars of the latter. The assessment of both quantitative and qualitative data within the same study has become characteristic of much of the research following a change process or change event perspective (see the different approaches presented in Rice & Greenberg, 1984). The issue of how to best relate the two types of data has received almost no systematic attention within the psychotherapy research literature (see Gaston & Marmar, 1989, for an exception), although it has received some attention within the general qualitative methodology literature (e.g., Jick, 1983).

Conclusion

To conclude, I address a disappointment one may face when first exploring single-case methodology, and make a couple of general comments about single-case psychotherapy research. Single-case experiments involve many demands that have highly limited their application to nonbehavioral treatment modalities, as Gelso (1979), McCullough (1984), and Safran, Greenberg, and Rice (1988) have pointed out. All of these demands result from the fact that single-case experiments deal with threats to internal validity by following the logic of interrupted time-series designs (see Campbell & Stanley, 1963; Glass, Willson, & Gottman, 1975). Internal validity considerations in these designs require that the treatment effect be relatively immediate, reversible, and in some cases, that it not spread to other areas of behavior (see Kazdin, 1982).

When one is dealing with the impact of changing a reinforcement contingency on a targeted problem behavior, these requirements do not pose a problem. This is captured, for example, in Leitenberg's (1973) call for ". . . demonstrations that changes in behavior are under the control of a specific therapeutic procedure, that desired behavior can be 'turned on and off, or up and down' by manipulating the therapeutic program" (p. 89). However, in nonbehavioral therapies these assumptions have proven to be highly problematic. For example, the effects of the cognitive and interpersonal treatments examined by Moras et al. (1993) are at best difficult, if not impossible, to reverse, and the effects of the treatment spread to other areas. This results in a situation in which threats to internal validity cannot be handled as thoroughly as when treatment effects are reversible and limited to one problem area.

Many researchers have come to a fairly pessimistic appraisal of the possibility of dealing with threats to internal validity at the single-case level when studying nonbehavioral therapies. This has led some to abandon single-case designs. Others, convinced of the critical role of single-case research, have felt forced to play down the importance of internal validity in single-case research, taking refuge, for example, in the context of discovery. However, as Campbell and Stanley (1963) and Kazdin (1981) have shown, internal validity is not an all-or-none affair. Often, measures can be taken to significantly increase the internal validity of a study, even if all potential threats to internal validity cannot be ruled out. Issues of internal validity remain one of the greatest challenges to single-case research, and they deserve sustained attention within the resurgence of interest in single-case psychotherapy research.⁶

Finally, as Kiesler (1983) observed, "Our available theories of psychotherapy are in a form too global or general to provide much guidance for identifying the significant change-process events occurring within therapy sessions" (p. 6). Single-case research needs theories that address the psychotherapy change process at this level of specificity. Such theories can then serve

⁶ Readers interested in exploring the issue of internal validity in single-case research are referred to the volumes by Yin (1989) and Edelson (1988) as a point of departure.

to generate specific research questions to be tested empirically. Unfortunately, much of single-case psychotherapy research appears to "tap around in the dark" without clearly formulated questions. Theory-based, question-driven, single-case research, in which disconfirmation remains a real possibility, is necessary within psychotherapy research.

References

- Bakeman, R., & Gottman, J. M. (1986). *Observing interaction: An introduction to sequential analysis*. Cambridge, England: Cambridge University Press.
- Barlow, D. H., & Hersen, M. (1984). *Single case experimental designs: Strategies for studying behavior*. New York: Pergamon Press.
- Benjamin, L. S. (1974). Structural analysis of social behavior. *Psychological Review*, 81, 392-425.
- Campbell, D. T. (1979). "Degrees of freedom" and the case study. In T. D. Cook & C. S. Reichardt (Eds.), *Qualitative and quantitative methods in evaluation research* (pp. 49-67). Beverly Hills, CA: Sage.
- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Boston: Houghton Mifflin.
- Chassan, J. B. (1979). *Research design in clinical psychology and psychiatry* (2nd ed.). New York: Wiley.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Boston: Houghton Mifflin.
- Edelson, M. (1985). The hermeneutic turn and the single case study in psychoanalysis. *Psychoanalysis and Contemporary Thought*, 8, 567-614.
- Edelson, M. (1988). *Psychoanalysis: A theory in crisis*. Chicago: University of Chicago Press.
- Elkin, I. (1991, November). *Criteria for evaluating psychotherapy process research*. Summary discussant paper presented at the National Institute of Mental Health-PARRC Workshop on Treatment Development Research, Bethesda, MD.
- Gaston, L., & Marmar, C. R. (1989). Quantitative and qualitative analyses for psychotherapy research: Integration through time-series designs. *Psychotherapy*, 26, 169-176.
- Gelso, C. J. (1979). Research in counseling: Methodological and professional issues. *The Counseling Psychologist*, 8, 7-36.
- Glass, G. V., Willson, V. L., & Gottman, J. M. (1975). *Design and analysis of time-series experiments*. Boulder: Colorado Associated University Press.
- Gottman, J. M. (1973). N-of-one and N-of-two research in psychotherapy. *Psychological Bulletin*, 80, 93-105.
- Gottman, J. M. (1981). *Time-series analysis: A comprehensive introduction for social scientists*. Cambridge, England: Cambridge University Press.
- Gottman, J. M., & Markman, H. J. (1978). Experimental designs in psychotherapy research. In S. L. Garfield & A. E. Bergin (Eds.), *Handbook of psychotherapy and behavior change* (2nd ed., pp. 23-62). New York: Wiley.
- Gottman, J. M., & Roy, J. M. (1990). *Sequential analysis: A guide for behavioral researchers*. Cambridge, England: Cambridge University Press.
- Grawe, K. (1992, June). *Figuration analysis: A new approach to a persistent dilemma of psychotherapy research*. Paper presented at the annual meeting of the Society for Psychotherapy Research, Berkeley, CA.
- Greenberg, L. S. (1986a). Change process research. *Journal of Consulting and Clinical Psychology*, 54, 4-9.
- Greenberg, L. S. (1986b). Research strategies. In L. S. Greenberg & W. M. Pinsof (Eds.), *The psychotherapeutic process: A research handbook* (pp. 707-734). New York: Guilford Press.
- Henry, W. P., Schacht, T. E., & Strupp, H. H. (1986). Structural analysis of social behavior: Application to a study of interpersonal process in differential psychotherapeutic outcome. *Journal of Consulting and Clinical Psychology*, 54, 27-31.
- Hill, C. E., Carter, J. A., & O'Farrell, M. K. (1983). A case study of the process and outcome of time-limited counseling. *Journal of Counseling Psychology*, 30, 3-18.
- Horowitz, M. J., Stinson, C., Curtis, D., Ewert, M., Redington, D., Singer, J., Bucci, W., Mergenthaler, E., Milbrath, C., & Hartley, D. (1993). Topics and signs: Defensive control of emotional expression. *Journal of Consulting and Clinical Psychology*, 61, 421-430.
- Howard, G. S. (1983). Toward methodological pluralism. *Journal of Counseling Psychology*, 30, 19-21.
- Jick, T. D. (1983). Mixing qualitative and quantitative methods: Triangulation in action. In J. Van Maanen (Ed.), *Qualitative methodology* (pp. 135-148). Beverly Hills, CA: Sage.
- Jones, E. E., Ghannam, J., Nigg, J. T., & Dyer, J. F. P. (1993). A paradigm for single-case research: The time series study of a long-term psychotherapy for depression. *Journal of Consulting and Clinical Psychology*, 61, 381-394.
- Kazdin, A. E. (1981). Drawing valid inferences from case studies. *Journal of Consulting and Clinical Psychology*, 49, 183-192.
- Kazdin, A. E. (1982). *Single-case research designs: Methods for clinical and applied settings*. New York: Oxford University Press.
- Kiesler, D. J. (1981). Empirical clinical psychology: Myth or reality? *Journal of Consulting and Clinical Psychology*, 49, 212-215.
- Kiesler, D. J. (1983, October). *The paradigm shift in psychotherapy process research*. Summary discussant paper presented at the National Institute of Mental Health Workshop on Psychotherapy Process Research, Bethesda, MD.
- Leitenberg, H. (1973). The use of single-case methodology in psychotherapy research. *Journal of Abnormal Psychology*, 82, 87-101.
- Lichtenberg, J. W., & Heck, E. J. (1983). Use of sequential analysis in counseling process research: A reply to Hill, Carter, and O'Farrell and to Howard. *Journal of Counseling Psychology*, 30, 615-618.
- McCullough, J. P. (1984). The need for new single-case design structure in applied cognitive psychology. *Psychotherapy*, 21, 389-399.
- Mook, D. G. (1982). *Psychological research: Strategy and tactics*. New York: Harper & Row.
- Mook, D. G. (1983). In defense of external invalidity. *American Psychologist*, 38, 379-387.
- Moras, K., Telfer, L. A., & Barlow, D. H. (1993). Efficacy and specific effects data on new treatments: A case study strategy with mixed anxiety-depression. *Journal of Consulting and Clinical Psychology*, 61, 412-420.
- Rice, L. N., & Greenberg, L. S. (Eds.). (1984). *Patterns of change: Intensive analysis of psychotherapy process*. New York: Guilford Press.
- Rosenblatt, P. C. (1981). Ethnographic case studies. In M. B. Brewer & B. E. Collins (Eds.), *Scientific inquiry and the social sciences: A volume in honor of Donald T. Campbell* (pp. 194-225). San Francisco: Jossey-Bass.
- Safran, J. D., Greenberg, L. S., & Rice, L. N. (1988). Integrating psychotherapy research and practice: Modeling the change process. *Psychotherapy*, 25, 1-17.
- Sidman, M. (1952). A note on the functional relations obtained from group data. *Psychological Bulletin*, 49, 263-269.

- Sidman, M. (1960). *Tactics of scientific research: Evaluating experimental data in psychology*. New York: Basic Books.
- Silberschatz, G., & Curtis, J. T. (1993). Measuring the therapist's impact on the patient's therapeutic progress. *Journal of Consulting and Clinical Psychology*, 61, 403-411.
- Skinner, B. F. (1953). *Science and human behavior*. New York: Macmillan.
- Soldz, S. (1990). The therapeutic interaction. In R. A. Wells & V. J. Giannetti (Eds.), *Handbook of the brief psychotherapies* (pp. 27-53). New York: Plenum Press.
- Spence, D. P., Dahl, H., & Jones, E. E. (1993). Impact of interpretation on associative freedom. *Journal of Consulting and Clinical Psychology*, 61, 395-402.
- Stiles, W. B. (1991). *Quality control in qualitative research*. Paper presented at the annual meeting of the Society for Psychotherapy Research, Lyon, France.
- Strupp, H. H. (1980a). Success and failure in time-limited psychotherapy: A systematic comparison of two cases (Comparison 1). *Archives of General Psychiatry*, 37, 595-603.
- Strupp, H. H. (1980b). Success and failure in time-limited psychotherapy: A systematic comparison of two cases (Comparison 2). *Archives of General Psychiatry*, 37, 708-716.
- Strupp, H. H. (1980c). Success and failure in time-limited psychotherapy: With special reference to the performance of a lay counselor (Comparison 3). *Archives of General Psychiatry*, 37, 831-841.
- Strupp, H. H. (1980d). Success and failure in time-limited psychotherapy: Further evidence (Comparison 4). *Archives of General Psychiatry*, 37, 947-954.
- Strupp, H. H. (1990). The case of Helen R. *Psychotherapy*, 27, 644-656.
- Strupp, H. H., Schacht, T. E., Henry, W. P., & Binder, J. L. (1992). Jack M.: A case of premature termination. *Psychotherapy*, 29, 191-205.
- Thorngate, W. (1986). The production, detection, and explanation of behavioral patterns. In J. Valsiner (Eds.), *The individual subject and scientific psychology* (pp. 71-93). New York: Plenum Press.
- Yin, R. K. (1989). *Case study research: Design and methods* (2nd ed.). Newbury Park, CA: Sage.

Received August 5, 1992

Revision received October 15, 1992

Accepted October 16, 1992 ■