Reprint (1-4)2

Single-case Research Designs for the Science and Practice of Neurotherapy

by Neville M. Blampied, M.Sc., FNZPsS, Arreed Barabasz, Ph.D., ABPP,
and Marianne Barabasz, Ed.D.

The dominant research tradition in psychology and psychiatry requires that numbers of subjects be randomly allocated to form treatment groups. Treatment effects typically are assessed by testing hypotheses about group mean differences. This paradigm seriously inhibits the implementation of the scientist-practitioner model embraced by practitioners of neurotherapy, stifles innovation and precludes the scientific investigation of the exceptional or novel case. Single-case research designs make it possible to draw scientifically valid conclusions from the investigation and treatment of individual cases. The key elements of these designs are outlined and particular designs of potential utility to neurotherapy are discussed.

"The individual is of paramount importance in the clinical science of human behavior change. Until recently, however, this science lacked an adequate methodology for studying behavior change in individuals" (Hersen & Barlow, 1976, p. 1).

Several distinctive features characterize the dominant research paradigm in contemporary psychology and psychiatry. These include the recruitment of as large a number of participants as is practicable, the random allocation of these participants to treatments, the aggregation of individual data in group descriptors, and the use of inferential statistics to detect the signal of experimental effects from the noise of error variance. So dominant has this tradition become that it is often identified with "science" and seen as the only way in which substantive contributions to scientific progress can be made (Barlow, Hayes, & Nelson, 1984; Johnston & Pennypacker, 1980). However, neurotherapy is equally heir to an older, alternative scientific tradition. In physiology, neurophysiology experimental medicine (e.g., Claude Bernard, see Thompson, 1984) and early research in psychology (e.g., Pavlov, Thorndike; see Boakes, 1984) numerous examples of scientific research based on the intensive study of the individual subject are evident. In contemporary psychology, this tradition (Baer, Wolf, & Risley, 1968; Sidman, 1960; Skinner, 1956, 1957) remains alive within the experimental and applied analysis of behavior. This paper explains how research and practice in neurotherapy can benefit from adopting research practices incorporating the contemporary expression of this earlier tradition.

The Merits of Single-Case Research

Despite its research practice dominance, the tradition of between-groups procedures and statistical hypothesis testing has been subject to long and trenchant criticism (e.g., Berkson, 1942; Cohen, 1990; Dar, Serlin & Omer, 1994; Meehl, 1978; Rozeboom, 1960). There has been growing concern about institutional practices (e.g., journal reviewing) which help to entrench and perpetuate this research style within psychology (e.g., Dar et al.,1994; Sedlmeier & Gigerenzer, 1989). Rather than review this critical perspective, this article will discuss factors that recommend single-case research to scientists and practitioners of neurotherapy.

1. The scientist/practitioner of neurotherapy, like the behavioral scientist in general, is concerned with individual human behavior. It is this focus which sets our endeavor apart from the social sciences concerned with phenomena which are products of aggregate and/or historic human action (Johnston & Pennypacker, 1980).The first scientific challenge for neurotherapy is to account for aspects of individual human functioning in terms of the interaction of that individual's central neural processes and their past and concurrent environmental experience. The second challenge is to show that specific therapeutic actions change both neural processes and behaviors, and thereby solve or ameliorate the individual's disorder. Demonstrations that, at some level of probability, some average property of a group of human beings is related to some other averaged variable provide only indirect evidence useful to the explanation and remediation of individual human performance.

2. Reliance on data obtained by averaging the scores of large groups brings with it serious risks of making both descriptive and inferential errors. This is especially true when the phenomena under investigation involve systematic changes of the sort observed when new skills are being learned (e.g., enhancing beta while inhibiting theta). As any experienced neurotherapist has observed, individual trajectories of change may differ substantially over time. Some individuals may change rapidly at first, slowly later; others may show the reverse pattern or some other pattern. Some changes may be so abrupt as to be step-like rather than gradual. A few may not change at all.

Averaging such varied data together may give the spurious appearance of orderly and regular change, quite unrepresentative of any actual individual's experience. Since neurotherapists are (a) typically concerned with learned changes and, (b) often work with individuals during major developmental phases (e.g., childhood, adolescence) these risks inherent in reliance on group average data are especially pertinent. At the very least, those who report data in the form of group averages must take responsibility for showing that the group average pattern is representative of that of at least some individuals in the sample. Even better is the practice of reporting individual data in sufficient detail that others can form judgments about the representativeness and typicality of the group data. This practice has the additional advantage of permitting cases with specific attributes to be matched or compared with clients and participants in other studies and treatments by other investigators (Barabasz & Barabasz, 1992; Barabasz, Barabasz, & Blampied, 1996).

3. Because of the infinite variety of human attributes and human experiences, clinicians are likely to encounter individuals who present personal histories, current circumstances, and particular difficulties which do not precisely fit standard diagnostic criteria, and who are, at least in the experience of that clinician, exceptional. To insist that scientific research requires the use of large-N designs precludes valid science from being done in the investigation and treatment of such

individuals, given the impracticality and unlikelihood of the clinician ever gathering a group of similar individuals together more or less at the same time for research purposes. In contrast, single-case research designs permit scientifically valid inferences to be drawn despite the uniqueness of cases or circumstances.

Related to this is a concern with the potential misuse of diagnostic systems such as DSM-IV (American Psychiatric Association, 1994). Diagnostic classifications serve useful functions, but they are no substitute for individual functional and contextual analyses. No two individuals with the same diagnosis will be exactly alike in all pertinent aspects of either their (neuro)biology or their past or present life experience and social contexts. Scientist-practitioners of neurotherapy need to deal with the uniqueness and singularity of every client. Single-case research designs make it possible to do so while continuing to meet rigorous standards of scientific practice.

4. As noted above, insistence on the use of large-N designs as a prerequisite for doing valid science is likely to make any attempts by clinical practitioners to undertake scientific investigations impractical. This renders the scientist-practitioner ideal of clinical training and performance defeated at its outset. As Barlow et al. (1984) have eloquently shown, only by the adoption of single-case research methodologies can the scientist-practitioner ideal be given practical substance (see also Hayes, 1981).

The implications of this go further than it may at first appear. What we are pointing out is not that adoption of single-case research principles permit the scientist-practitioner occasionally to do valid science (although that is true) but that adoption of these principles means that all interventions conducted under their aegises are scientifically valid investigations from which defensible causal inferences may be made. Each case thus becomes not a "case study" but a full-fledged experiment. This is the key methodological step that is prerequisite for the scientist-practitioner model to be fully implemented. Of course, other issues must be addressed, including the use of valid measures, scientifically defensible rationales for choice of intervention protocols, and the abandonment or modification of treatments that are shown to be ineffective or harmful.

5. Another feature of single-case research procedures that increase their "goodness-of-fit" to the realities of clinical science is the fact that they are flexible and can be used creatively. "Single case research should be a dynamic, interactive exercise in which the design is always tentative, always ready to change if significant questions arise in the process" (Hayes, 1981, p. 196). Although we shall describe some specific single-case research designs below, it is important to remember that they may be used creatively to meet special or unusual circumstances and permit an attitude of "investigative play" (Hayes, 1981).

6. A further feature of single-case research designs which should enhance their appeal to those developing innovative research or therapy procedures, such as neurotherapy, is that innovations may be evaluated rigorously, without the need for large numbers of individuals to be exposed to experimental or unproven procedures. The notion of a "pilot study" is rendered largely redundant by single-case research designs. Each case is an experiment and all experiments can be conducted with equal scientific rigor. This feature enhances the ethics of innovation without compromising scientific standards.

7. Finally, but not unimportantly, the use of single-case research designs permits researchers, and perhaps more importantly, practitioners, to be accountable. Every single case can be supported with scientifically valid data to justify claims of efficacy of treatment if their work is challenged by critics, third-party payers, clients, family members, ethical or peer review bodies, etc. Given the exigencies and realities of contemporary clinical practice, this is a substantial advantage.

Nomenclature

A bewildering variety of terms have been used to refer to single-case research designs (see Hayes, 1981, for a review). These terms have included N-equals-one designs, own-control designs, small-N designs, intrasubject replication designs, time-series designs, single-subject designs, and single-case designs. Contemporary practice favors the use of either the term time-series designs or single-case designs. The term "time-series" draws attention to the consistent use of time-sequential data in these research designs, but it invites confusion with mathematical time-series data analysis.

The descriptor "single-case" is preferred over the alternative "single-subject" because it is now recognized that the "subject" to which the research procedures may apply need not exclusively be individual human beings, but might be single entities (cases), such as a couple, a family, a school class, a hospital ward, a factory or mine, or even a community (Valsiner, 1986). In the context of neurotherapy, the data will normally be from a single person, but the use of the term 6csingle-case" is still preferred because it avoids referring to research participants or patients as "subjects."

Core Features of Single-Case Research Designs

All research designs must accomplish two fundamental goals. They must permit reliable changes in the dependent variable to be detected despite variance, and they must permit valid inferences to be drawn as to the cause of any changes observed. They must achieve these goals while permitting the unambiguous identification of the nature and application of the independent variable. Single case research designs have distinctive strategies to achieve each of these objectives. These include the following:

1. The use of time-series (repeated) measures data collection. Single-case research designs rely upon gaining many repeated measures of the dependent variable from a single individual (or a few individuals) over some period of time. This feature is also characteristic of neurotherapy protocols. Compare this with between-groups designs where there are many subjects but few measures per subject and often little spread of time.

Data gathered under the same conditions are grouped together to constitute the data from a specific phase of the experiment. One phase, generally but not exclusively the first one measured, is termed the baseline" and may be followed (or preceded) by other treatment or baseline phases. Treatment phases are those in which an independent variable (treatment) is present.

2. Data are presented graphically for visual analysis. Rather than relying on computational

statistics to test hypotheses about group mean differences, single-case research procedures feature the display of the data in standard ways which facilitate viewers making accurate judgments about' differences in data paths or trajectories and the relationship of these differences to changes in the independent variable(s) (Parsonson & Baer, 1978). Cooper, Heron, and Heward (1987) should be consulted for a comprehensive and authoritative exposition of how such graphs are drawn and interpreted. Visual analyses of data do not, of course, preclude statistical analyses where they can be validly applied, but are prerequisite to any subsequent computational procedures, either descriptive or inferential.

3. The core module of analysis in single-case research is the graphic presentation of a baseline phase and treatment phase pair. This reliance on pair-wise comparisons is a feature shared with traditional research, where the irreducible minimum is the control group-treatment group comparison. All single-case research designs are built from combinations of this core module.

The first purpose of the graphic presentation of the data paths in the baseline and the treatment phase is to permit the detection of any reliable difference in the dependent variable between phases. This requires that the viewer attend to three aspects of the data: variability, level, and trend (Cooper et al., 1987). Variability must be assessed both within and between phases. It may be useful to add to the graphic display descriptive information about the phase mean, standard deviation, median, range, etc.

By the way it attends to variability, single-case research practice exhibits a very different philosophy of science than that employed by conventional between-groups designs. In these conventional research procedures, variability, other than that due to the treatment, is treated as due to "error" and is accounted for statistically- hence the "analysis of variance." Variance (other than that due to treatment) is viewed negatively, and is dealt with primarily by increasing sample size, and thereby statistical power (Cohen,1969).

In contrast, the philosophy of single-case research designs is to view variance as a phenomenon to be explained, or at least ' experimentally controlled (Skinner, 1956, 1957). This means that if the baseline data path exhibits high variability, the investigator's first task is to explore the reasons for the variability These reasons are then tested by adjusting the data-gathering procedures and circumstances in attempts to reduce the variability (see Cooper et al., 1987, and Hayes, 1981, for discussions of what typically might be done to reduce variability).

Given that variability is sufficiently low, trend and level may then also be assessed. Trend refers to any systematic increase, decrease, or cyclic pattern evident in the data path. Accurate estimates of trend require a minimum of three data points, but more are desirable. Software for the analysis of multichannel neurometric assessments (e.g., Neurosearch 24, Lexicor Medical Technology, 1992) permits data to be instantly checked for trend using the trend analysis feature.

Ideally, baseline data exhibits no systematic trend (i.e., are stable around a trend parallel to the X-axis). However, the presence of systematic trend in baseline is not fatal to the investigation so long as the baseline trend is in a direction opposite to that expected to be shown when the independent (treatment) variable is introduced. For instance, if a problem is steadily worsening over baseline observations, then a successful treatment effect will still be clearly evident. On the

other hand, if the baseline trend is for systematic improvement, it may be difficult to claim that the treatment made any substantive difference to the trajectory of the data (Hersen & Barlow, 1976).

Level refers to the location of the data path relative to the known or anticipated maximum and minimum values of the dependent measure. There may be little point in proceeding with further investigation if baseline data show that a therapeutic effect is unlikely to be evident following treatment because the data in baseline are almost as high (ceiling effect) or as low (floor effect) as they can go. If that is the case, then either some other more sensitive measures must be taken or the nature of the problem under investigation needs to be reviewed.

Ideally, the baseline phase is not terminated until the investigator is satisfied with the observed level, variability, and stability of the baseline data. How long baseline data should be collected is determined by a number of factors, some external to the measurement process. There may be ethical and practical limits to the length of time that treatment may be withheld. Equally, there is no point taking repeated measures of what are known to be highly stable properties of the person (e.g., theta/beta ratios during reading, IQ, personality measures, psychopathology scores). Furthermore, some additional baseline data may be available retrospectively, e.g., the family of a child diagnosed with ADHD may well be able to accurately report how many drug-free days their child has had in the past week or month, while school counselors or referring clinicians may have records of scholastic performance, IQ, neurometric assessments, etc. All of these data may compensate for relatively short investigative baselines.

Careful attention to the quality of baseline data is important and repays effort, because it is the foundation upon which all further analysis and inference is built. Evaluating level and trend, reducing variability, and getting as long a baseline data path as possible all enhance the likelihood that when the independent (treatment) variable is introduced, any changes accompanying exposure to treatment will be clearly shown as conspicuous changes in the data path during the treatment phase. The visual contrast of the baseline with the treatment phase is how the first goal, that of detecting reliable change, is accomplished.

Making Valid Causal Inferences

Given that a change in the dependent variable is clearly evident between baseline and treatment phases, the goal of confirming the independent variable as the cause of the changes can only be accomplished if valid inferences can be drawn, Note that thus far we have not, inferentially, gone beyond the scope of any competent traditional case study (see Barabasz, Barabasz, & Blampied, 1996). The hypothesis that it was the introduction of the treatment that caused the subsequent change in performance may be challenged on many grounds (e.g., non-specific therapeutic effects, placebo effects, concurrent pharmacological treatments, developmental processes, other coincident changes in life circumstances, etc.), since all we have is a demonstration of coincidence in time, not causality.

For single-case research designs to move from demonstrating coincidence to causal inference and achieve the second goal of experimental design, replication is the critical step (Sidman, 1960). "Replication means repeating the previously observed change with further manipulations of the

reproducing the previously observed behavior change reduces the probability that a variable other than the independent variable was responsible for the twice observed behavior change. Second, replication demonstrates the reliability of the behavior change; it can be made to happen again" (Cooper et al., 1987, p. 159).

Replication may take two forms in different single-case research designs: within-person replication and between-person replication. These forms of replication may also be combined in one investigation. For any claim to valid causal inference, one complete replication is required, but the more replications available, the stronger the validity of the causal inference made.

Before considering how combinations of replication strategies yield the common single-case research designs, two other common, although not essential features of these procedures need to be noted. These are concern for the generalizability and durability of effects. Effective treatments for clinically significant problems need to endure over time, and transfer, or generalize beyond the training setting to the other contexts and circumstances in which patients and their families experience the problem (Baer et al., 1968, 1987).

To evaluate the occurrence of these properties of effective interventions, single-case research designs (a) assess the durability of treatment effects by taking follow-up measures over as long a time period as possible, and (b) gather collateral data which should detect the transfer of therapeutic impact to the situations and circumstances where such impact is required. Not to be overlooked in the gathering of such collateral data is the possibility that there may be negative as well as positive effects of treatment. These will not be detected unless explicitly looked for.

Standard Single-Case Research Designs

1. Designs based on within-person replication

(a) Reversal (Withdrawal) Designs. Given that the minimum module for any single-case research design is a baseline treatment pair, the simplest replication involves a repetition of these two components. First, the treatment is withdrawn (hence the terms withdrawal or reversal) to institute a second baseline. When sufficient data have been collected to reliably detect any subsequent change in the dependent variable, the treatment is reinstated, and any concomitant changes observed. If performance changes are replicated, i.e., they track the instatement and removal of the treatment, the hypothesis that the treatment is causally responsible for producing the effect is supported and the argument that the changes are coincidental is weakened.

In clinical work, reversals may be planned or unplanned. Unplanned reversals may occur where for one reason or another, treatment is suspended. So long as the patient's subsequent circumstances resemble those of the first baseline condition, and so long as one or more measures of performance can be taken during or at the end of the period of suspension, then the requirements of a reversal design may have been met.

Whether planned or unplanned, reversals may not suffice to demonstrate a treatment effect. Some

behavior may rapidly become irreversible, in which case, performance will not return to baseline levels when treatment is withdrawn. Furthermore, there may be ethical problems or client resistance to withdrawing treatment. In general, in planning a reversal design, brief reversals, early in training are to be preferred to long reversals after long periods of treatment.

Another possible application of the reversal design is to substitute a second, possibly less effective treatment, for the withdrawn condition. If baseline is designated A; treatment 1, B; and treatment 2, C; this may be represented as an A B C B design. In this design it is the impact of treatment 1 which is replicated, with comparisons made to both baseline and to the less effective treatment 2.

A further variant of this design (the treatment-combination reversal) uses combinations of treatment alternating with a single treatment in the replications. This was used in the case study of Juan (pseudonym) reported by A. Barabasz and M. Barabasz (1995, 1996). This study employed two treatments for this boy diagnosed with ADHD: standard neurotherapy and standard neurotherapy supplemented by Instantaneous Neuronal Activation Procedures (INAP, see Barabasz & Barabasz, 1995, 1996). After baseline assessment, Juan received 12 sessions of combined neurotherapy plus INAP For three sessions INAP was then withdrawn, and then from sessions 16 to 32 the combined treatment was reinstated. This may be described as A B+C C B+C (where B=INAP and C = neurotherapy). Follow-up assessments occurred immediately after treatment and at 6 and 12 months. In this case, rate of progress during the withdrawal of INAP was distinctively lower than during the previous or subsequent sessions, a judgment confirmed by a second observer blind to treatment conditions. Figure 1 uses synthetic data to illustrate a possible reversal design modeled on the Barabasz and Barabasz procedure.

Figure 1

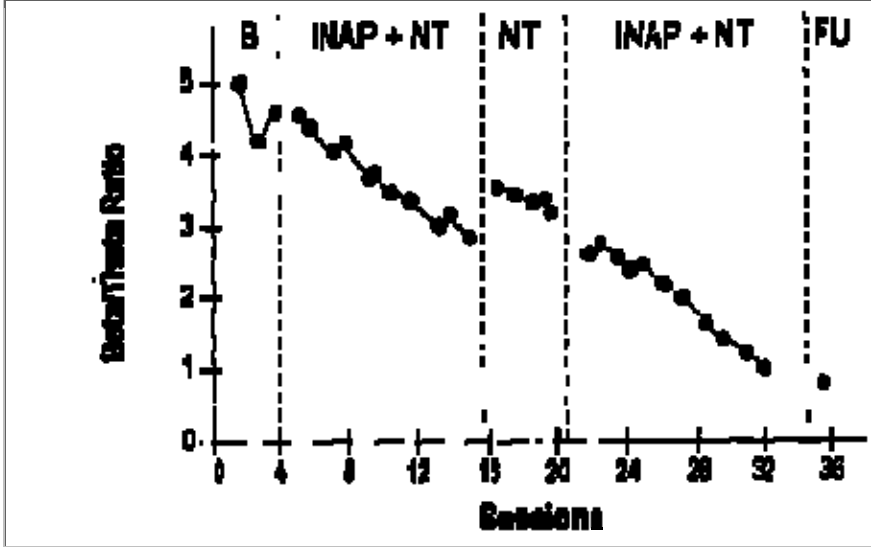

Fig. 1. Synthetic data illustrating a Reversal Design. Following Baseline (B) a combined treatment (INAP + Neurotherapy) is applied. In Phase 3, INAP is withdrawn, leaving only Neurotherapy (N). The combined treatment is re-applied in Phase4. Phase 5 is Follow-up (FU). After Barabasz & Barabasz,

1995, 1996

(b) Changing- criterion designs. Changing-criterion designs are especially appropriate for those investigations where performance is expected to change in an incremental or graded way as therapeutic criteria for change are set and then modified. Neurofeedback procedures often have this property (e.g., Packard & Ham, 1996).

In the changing-criterion design, following an initial baseline, the investigator begins treatment by setting a goal level of attainment that is slightly improved over baseline levels. Once this level has been attained, a second, more stringent criterion is set, and so forth. Criterion changes should be large enough for an effect to be detectable but not so large that the patient cannot achieve them (Cooper et al., 1987). To enhance the demonstration of experimental control, it is desirable to change the duration of each criterion period, and to include some reversals of criterion as well as enhancements of criterion. Figure 2 illustrates such a procedure, again using synthetic data.

Figure 2

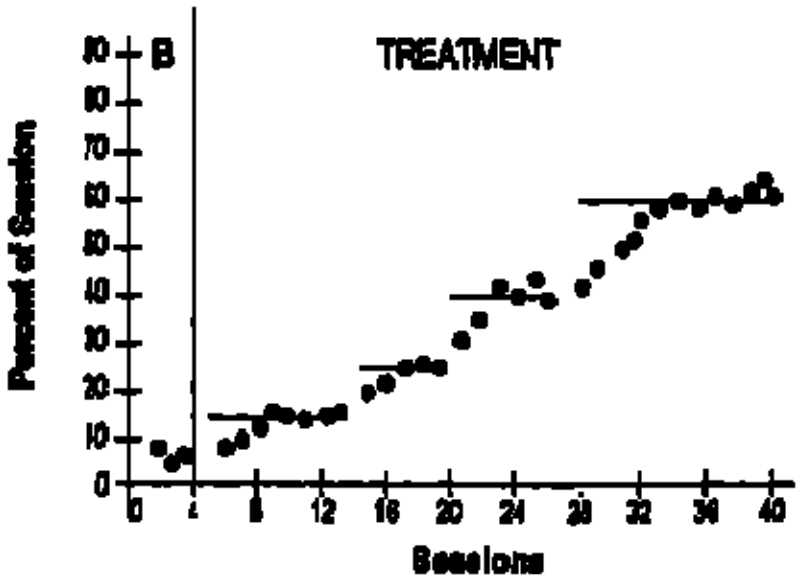

Fig. 2. Synthetic data illustrating a Changing Criterion Design. Y-axis reports the percentage of each session in which output in a particular EEG bandwidth meets or exceeds a particular criterion. Following Baseline (B), criteria are respectively 15%, 25%, 40% and 60%. These values are indicated by horizontal lines.

(c) Alternating-treatments designs. One weakness of treatment combination procedures such as those described above is that it is not possible to control for order of exposure effects, and so it might be argued that a particular treatment was effective only because it had been preceded by another treatment. Alternating treatments designs are especially appropriate for comparing two or more treatments in such a way that order effects largely may be ruled out.

Suppose that an investigator has two treatments available, but does not know which is likely to be

the more effective in a particular case. Following a baseline phase, she/he institutes treatment in phase 2 in the following way Each session is divided into two equal periods. At the beginning of each session a coin is tossed to decide which treatment is employed first, and each treatment is associated with a distinctive discriminative stimulus (SD) which signals to the participant which treatment is operating. At the appropriate time in the session, treatment and the associated SD is switched to the alternative treatment (hence the term alternating treatments), sometimes with a brief time out in between. If there is likely to be carry-over of effects from one treatment to the other it may be necessary to separate the alternative treatments by a longer time, perhaps alternating treatments between sessions rather than within sessions.

This alternation of treatments is continued until the plotted data (see the middle panel of Figure 3 for an illustration) show clear performance differences between the two treatments. At that point (phase 3), the less effective treatment is dropped and the same treatment is applied in both components of the session but performance in the separate components is still recorded. This permits a phase three replication of the treatment effect shown by the more effective treatment in phase two (see Figure 3), since performance in the lagging component should catch up with that in the other component. Alternating-treatments designs cannot, of course, be used where there is transfer of training from one component to the other and they require good control of performance by the SD associated with each treatment.

Figure 3

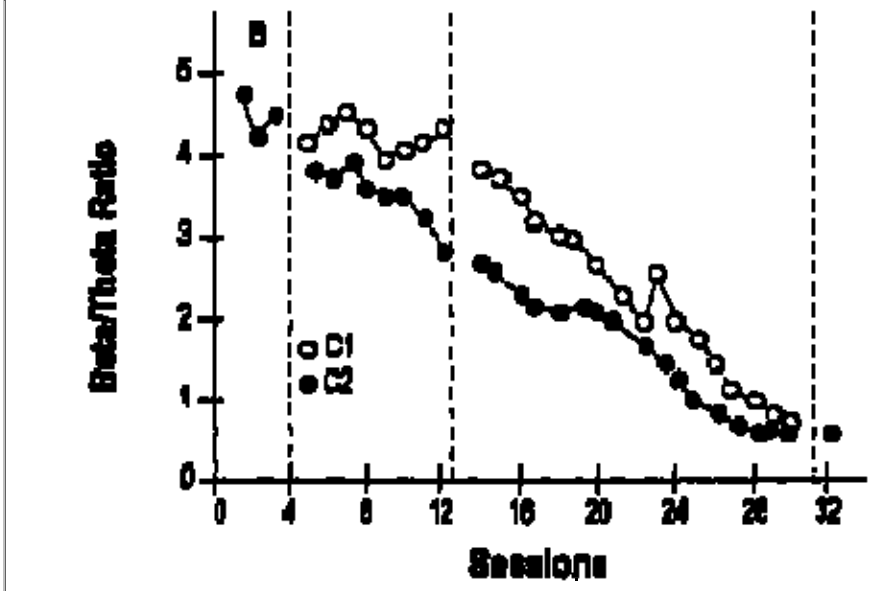

Fig. 3. Synthetic data illustrating a Alternating Treatments Design. Following Baseline (B) phase, each session is divided into two components. Treatment A (closed circles) in component 2. In Phase 3, the more successful treatment (Treatment A) is applied in both components of each session. Phase 4 is a Follow-up.

(d) Multiple-baseline-across-behaviors design. Alternating treatment and treatment-combination

performance. Multiple-baseline-across-behaviors (performance) designs, in contrast, replicate the impact of the same treatment across several different behaviors. A major limitation on the use of these designs is when there is transfer or generalization of treatment effects across different response systems.

Whenever a neurotherapy investigator has an interest in two or more response systems, a multiple-baseline-across -behaviors design is possible, such as an investigation in which both SMR and skin conductance were to be monitored and modified by biofeedback (e.g., Quirk, 1995, see Figure 4 for an example), or where several different frequency bands of EEG are successively targeted for modification (e.g., Brown, 1995; Byers, 1995).

Note, in this context, that relative or proportional measures such as theta/beta ratios, or percent of signal (e.g. beta) in particular bandwidths, are often unsuitable for use in multiple -baseline analyses, because the requirement that the responses be independent is not met. However, use of absolute power measures at different sites or in different bandwidths may be suitable.

In using this design the investigator must begin by taking baseline measures of all response systems of interest. When sufficient measures have been taken for the initial baselines to be judged stable, treatment is introduced to one response system, while the others remain untreated but continue to be monitored. When comparison of the first baseline-treatment pair reveals a treatment effect, treatment may be introduced to the second target response, while continuing to monitor any other baselines. Following the same principles, treatment is successively introduced to all target responses.

Two response systems are the minimum needed, since that permits a full replication over two baseline-treatment pairs, but more replications over additional response systems strengthen our confidence that the treatment is responsible for the observed changes, given, of course, that responses change when and only when treatment is introduced. Note that no withdrawal of treatment is necessarily involved in this design. However, following the initial treatment phase for each response it is possible to add additional phases, such as withdrawal or treatment combinations. These may then successively be replicated over the other response systems.

A useful modification of this design is the multiple-probe procedure (Horner & Baer, 1978). This substitutes an occasional probe of additional target responses for constant baseline monitoring, which may be intrusive or expensive. Continuous baseline monitoring is then instituted shortly before the treatment is to be introduced to the target response.

Figure 4

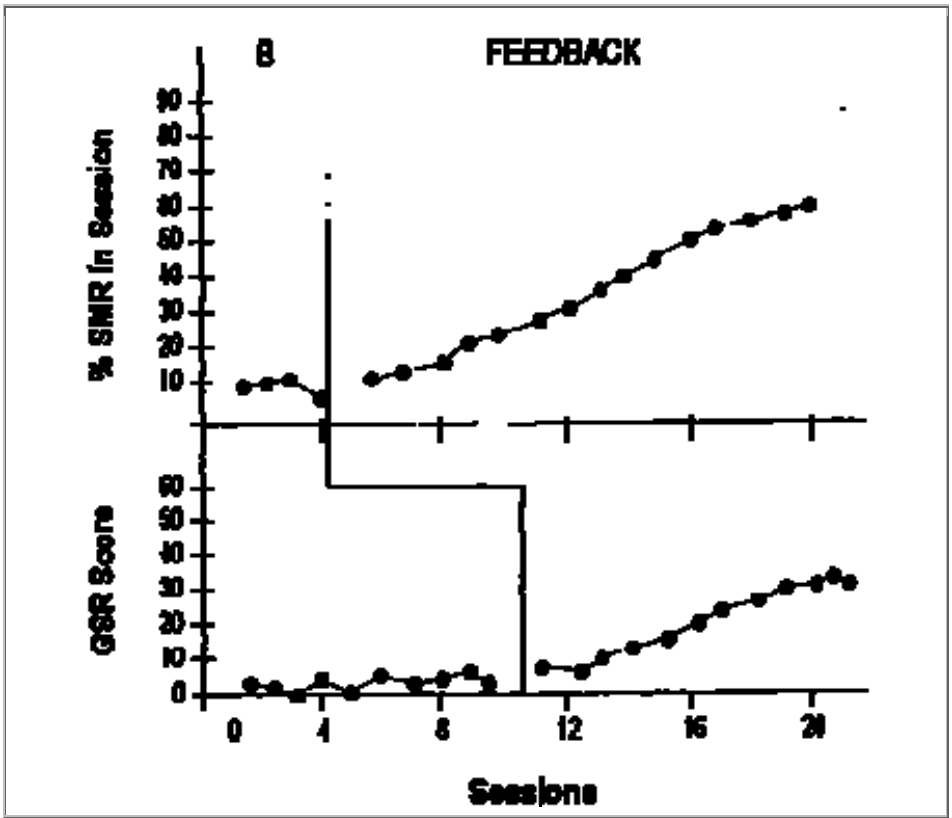

Fig. 4. Synthetic data illustrating a Multiple Baseline Across Behaviors Design. The top graph displays the percent of SMR recorded in each session before (Baseline) and during Neuro- therapy (Feedback). The second graph displays changes in skin resistance (GSR) in Baseline (B) and during biofeedback treatment (Feedback). GSR is scored in arbitrary units (After Quirk, 1995).

2. Procedures involving between-person replications

Every occasion on which an investigator uses a standard treatment protocol with yet another client is an instance of between-person replication, and clearly the more often the treatment is successful the more confident we may be that the treatment is effective. Bringing together a large series of single cases, and matching degrees of treatment success or failure against personal attributes and situational contexts permits us to establish the general effectiveness and the limitations of the treatment protocol. The multiple-baseline-across-subjects (or participants) permits between-person replication to be used in a more systematic way.

In the concurrent multiple-baseline procedure, two or more individuals with similar problems enter investigation at approximately the same time and undergo baseline assessment (e.g., a 19 electrode site multitask neurometric assessment). The critical feature of this design is that one individual then begins treatment, while the other(s) remain on baseline. As in other forms of multiple-baseline, it is possible to use probes (e.g., QEEG measures of theta/beta ratios from single sites) rather than repeated comprehensive assessments, for reasons of practicality, expense, or intrusiveness. After a lag of sufficient time to permit the first person to demonstrate a treatment effect, the second person enters treatment. Baseline measures continue for any additional

participants. They too then sequentially enter treatment, but each with a lag. Again, successive replications over the participants showing that their behavior changed when and only when treatment" wais introduced support the hypothesis that the treatment causes the change.

Obviously, this multiple-baseline design variant lends itself ideally to the requirements and inherent limitations of clinical practice. Even more, clinical work may lend itself to using the non-concurrent multiple baseline design (Hayes, 1981; Watson & Workman, 1981). This relaxes the requirement that the participants enter assessment at approximately the same time, and makes it possible to deploy the data from several clients seen at different times, exploiting the fact that their baseline assessment phases were of different lengths. When treatment effects are replicated in different cases, seen at different times, and following different durations of baseline assessment, arguments that treatment effects were due to common extratherapy events experienced by the patients or to cumulative impacts of non-specific, placebo effects derived from the experience of assessment and therapy may be confidently rejected as implausible.

Conclusion

Neurotherapy investigations are typically innovative, data-rich investigations involving the induction of performance changes in psycho physiologic response systems. The focus is always on the individual client, and supplementary corroborative data are routinely taken to assess the impact of therapy on the pathology that brought the client to treatment. These features render single-case research designs and neurotherapy very congenial partners. Our recommendation is that neurotherapy scientist-practitioners seriously consider incorporating single-case designs in their work. This will ensure that all their cases contribute valid knowledge to the body of scientific knowledge about the brain-behavior interrelationships, and help them break free of the shackles of the conventional, between-groups research paradigm.

References

American Psychiatric Association. (1994). **Diagnostic and Statistical Manual for Mental Disorders** ***(4th ed.). Washington, DC: Author.***

Baer, D. M., Wolf, M. M., & Risley, T R. (1968). Some current dimensions of applied behavior analysis. **Journal** ***of Applied Behavior Analysis, 1, 91-97.***

Baer, D. M., Wolf, M. M., & Risley, T R. (1987). Some still-current dimensions of applied behavior analysis. **Journal** ***of Applied Behavior Analysis, 20, 313-327.***

Barabasz, A. F., & Barabasz, M. (1992). Research designs and considerations. In E. Fromm & M. R. Nash (Eds.), **Contemporary Hypnosis Research (pp.** ***173-200). New York, NY. Guilford.***

Barabasz, A., & Barabasz, M. (1995). Attention Deficit Hyperactivity Disorder: Neurological basis and treatment alternatives. **Journal of Neurotherapy, 1(1),** ***1-10.***

Barabasz, A., & Barabasz, M. (1996). Neurotherapy and alert hypnosis in the treatment of Attention-Deficit Hyperactivity Disorder. In S. Lynn, I. Kirsch, & J. Ruh, (Eds.), **Clinical Hypnotherapy Casebook.** *Washington, D.C.: American Psychological Association.*

Barabasz, M., Barabasz, A., & Blampied, N. (1996). A primer of case study research in neurotherapy **Journal** *of Neurotherapy, 1(4), 12-14.*

Barlow, D. H., Hayes, S. C., & Nelson, R. 0. (1984). **The Scientist-practitioner: Research and Accountability in Clinical and Educational Settings.** *Oxford: Pergamon.*

Berkson, J. (1942). The test of significance considered as evidence. **Journal** *of the American Statistical Association, 37, 325-333.*

Boakes, R. A. (1984). **From Darwin to Behaviourism: Psychology and the Minds of Animals.** *New York: Cambridge University Press.*

Brown, V W, (1995). Neurofeedback and Lyme's Disease: A clinical application of the five phase model of CNS functional transformation and integration. **Journal** *of Neurotherapy, 1(2), 60-73.*

Byers, A. P. (1995). Neurofeedback for a mild head injury **Journal** *of Neurotherapy, 1(1), 22-37.*

Cohen, J. (1969). **Statistical Power Analysis for the Behavioral Sciences.** *San Diego, CA: Academic Press.*

Cohen, J. (1990). Things I have learned (so far). **American Psychologist, 45,** *13041312.*

Cooper, J. 0., Heron, T. E., & Heward, W L. (1987). **Applied Behavior Analysis.** *Columbus, OH: Merrill.*

Dar, R., Serlin, R. C., & Omer, H. (1994). Misuse of statistical tests in three decades of psychotherapy research. **Journal** *of Consulting & Clinical Psychology, 62, 75-82.*

Hayes, S. C. (1981). Single-case research designs and empirical clinical practice. **Journal** *of Consulting & Clinical Psychology, 49, 193-211.*

Hersen, M., & Barlow, D. H. (1976). **Single case Experimental Designs: Strategies for Studying Behavior Change.** *Oxford: Pergamon.*

Horner, R. D., & Baer, D. M. (1978). Multiple-probe technique: A variation on the multiple-baseline design. **Journal** *of Applied Behavior Analysis, 11, 189-196.*

Johnston, J. M., & Pennypacker, H. S. (1980). **Strategies and Tactics of Human Behavioral Research.** *Hillsdale, NJ: Erlbaum.*

Lexicor Medical Technology, Inc. (1992). **NeuroSearch-24: User's Manual.** *Boulder, CO: Author.*

Meehl, P. E. (1978). Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. **Journal of Consulting & Clinical Psychology, 46,** *808-834.*

Quirk, D. A. (1995). Composite biofeedback conditioning and dangerous offenders. **Journal of Neurotherapy,** *1(2), 44-54.*

Packard, R. C., & Ham, L. P. (1996). EEG biofeedback in the treatment of Lyme Disease: A case study. **Journal of Neurotherapy,** *1(3), 22-31.*

Parsonson, B., & Baer, D. M. (1978). The analysis and presentation of graphic data. In T R. Kratochwill (Ed.), **Single subject Research: Strategies** *for* **Evaluating Change (pp.** *101-165). New York, NY: Academic Press.*

Roseboom, W W. (1960). The fallacy of the null-hypothesis significance test. **Psychological Bulletin, 57,** *416-428.*

Sedlmeier, P., & Gigerenzer, G. (1989). Do studies of statistical power have an effect on the power of studies? **Psychological Bulletin, 105,** *309- 306.*

Sidman, M. (1960). **Tactics of Scientific Research.** *New York: Basic Books.*

Skinner, B. F. (1956). A case history in scientific method. **American Psychologist,** *11, 221-233.*

Thomson, T. (1984). The Examining Magistrate for Nature: A retrospective review of Claude Bernard's An introduction to the Study of Experimental Medicine. **Journal of the Experimental Analysis of Behavior, 41,** *211-216.*

Valsiner, J. (1986). **The Individual Subject and Scientific Psychology.** *New York: Plenum.*

Watson, P. J., & Workman, E. A. (1981). The non-current multiple-baseline across individuals design: An extension of the traditional multiple baseline design. **Journal of Behavior Therapy and Experimental Psychiatry, 12,** *257-*

About the Authors:

This paper was written while N. M. Blampied was on Study Leave. Grateful thanks are expressed to the Attentional Processes Lab, Counseling Psychology program and to the Department of Psychology, Washington State University for their hospitality and provision of facilities, and to the University of Canterbury for making leave available.

Neville Blampied, M.SC., Senior Lecturer in Psychology, University of Canterbury, Christchurch, New Zealand, Fellow, New Zealand Psychological Association.

Marianne Barabasz, EdD., Associate Professor, Washington State University, Fellow, American Psychological Association, Society of Clinical and Experimental Hypnosis, American Association of Applied and Preventive Psychology.

Arreed Barabasz, EdD., Ph.D., ABPP, Professor, Washington State University, Board Certified Psychologist, Fellow, American Psychological Association, American Psychological Society, Society of Clinical and Experimental Hypnosis.

Address correspondence to either N. M. Blampied, Department of Psychology, University of Canterbury, PB 4800, Christchurch, New Zealand [n.blampied@psyc.canterbury.ac.nz] or to A. Barabasz, Attentional Processes Laboratory, Cleveland Hall, Washington State University, Pullman, WA 99164-2136 [arreed@mail.wsu.edu].