Single-Participant Research Design

Bringing Science to Managed Care

David L. Morgan — *Spalding University*
Robin K. Morgan — *Indiana University Southeast*

The ongoing transition to managed health care continues to have repercussions for health care providers, perhaps the most important of which is an emphasis on accountability for demonstrating the usefulness of clinical interventions. This requirement places a premium on intervention research and highlights the historically strained relationship between psychological research and professional practice. In the midst of this challenge, researchers have increasingly criticized the logic and practice of traditional null hypothesis significance testing. This article describes the history, epistemology, and advantages of single-participant research designs for behavioral scientists and professionals in clinical settings. Although its lack of correspondence with the Fisherian tradition has precluded widespread adoption, the single-participant alternative features a design power and flexibility well suited to both basic science and applied research.

During the 1980s and 1990s, health care costs skyrocketed because of advances in technology, an abundance of health care providers in metropolitan regions, and traditional fee-for-service reimbursement mechanisms. Managed care programs were developed as a way to curb these costs. In the process, freedoms historically enjoyed by health care providers have been curtailed. In addition, greater demands have been placed on health care providers to deliver interventions that possess documented effectiveness and to do so in a cost-effective manner (Berman, 2000).

That the managed care revolution has changed the way professional health care providers operate is hardly contestable. The practices of mental health professionals have come under particular scrutiny in a contemporary climate characterized by conflict on many fronts, from confidentiality issues (Acuff et al., 1999), to disputes about treatment duration and reimbursement (Dean, 1998), to fundamental questions about accountability. To a considerable extent, many of the disputes about health care provision hinge on the concept of accountability, or whether the health care provider can document or attest to the clinical usefulness of a particular treatment regimen. Doing so necessarily requires that health care providers be knowledgeable about and make contact with research literature bearing specific relevance to their practice. Historically, this requirement has seldom been met. Although many reasons for this lack of documentation exist, chief among them is the often insurmountable chasm that separates basic research from applied practice (Hayes, 1981; Hilliard, 1993; Marten & Heimberg, 1995; Seligman & Levant, 1998). One concrete example of this dilemma was provided by Sanavio (1998), who argued that behavior therapy (in the form of exposure and response prevention) is a highly effective treatment for obsessive–compulsive disorder. Despite this treatment effectiveness research, research demonstrating the cost-effectiveness of this intervention is virtually nonexistent, and unless behavior therapy can be shown to be as cost-effective and as easily delivered as medication, reimbursement for behavioral treatments by managed care companies may disappear.

It is interesting that just as clinicians have long lamented the inaccessibility or irrelevance of basic research to clinical practice, so too have the basic logic and machinery of psychological research come under scrutiny. A burgeoning contemporary literature attests to a growing disposition on the part of methodologists to place time-honored research designs under the microscope (Abelson, 1997; Cohen, 1990, 1994; Estes, 1997; Harris, 1997; Loftus, 1993, 1996). The results often prove to be less than flattering, particularly in matters of data analysis. Indeed, at a symposium held during the 1996 convention of the American Psychological Association (APA), a group of methodologists and statisticians actually discussed the notion of banning significance tests from APA journals (Shrout et al., 1996)! APA subsequently convened a task force on statistical inference (Wilkinson & the Task Force on Statistical Inference, 1999) to assess and offer recommendations concerning the use of statistical methods in psychological research. The matter is, of course, not resolved, and the dialogue continues in earnest, as most recently manifested in the open peer commentary format of the journal *Behavioral and Brain Sciences* (Chow, 1998). Regardless of the eventual outcome of this debate, the APA meeting and its subsequent fallout would seem, at the very least, to reflect some systematic unrest among researchers concerning psychology's version of the scientific method.

David L. Morgan, Department of Psychology, Spalding University; Robin K. Morgan, School of Social Sciences, Indiana University Southeast.

Correspondence concerning this article should be addressed to David L. Morgan, Department of Psychology, Spalding University, 851 South Fourth Street, Louisville, KY 40203-2188. Electronic mail may be sent to dmorgan@spalding.edu.

Copyright 2001 by the American Psychological Association, Inc. 0003-066X/01/$5.00
Vol. 56, No. 2, 119–127 DOI: 10.1037//0003-066X.56.2.119

David L. Morgan

In joining the current dialogue concerning research methodology, in this article we argue that both basic and applied behavioral scientists have at their disposal a richer inventory of data collection and analysis techniques than circumscribed by the null hypothesis testing tradition. Among these techniques, single-participant experimentation boasts both a storied history and a largely unappreciated potential for basic science and professional practice. The single-participant alternative has produced solid and replicable empirical findings across a number of behavioral domains yet remains relatively obscure because of its disavowal of the statistical machinery that defines psychological research in the 21st century. At a time when psychology seems invested in examining how it conducts itself as a science, a critical evaluation of alternative research strategies seems prudent. This article demonstrates the unique epistemology and experimental strategies that distinguish single-participant research and the kinds of contributions that this research tradition can continue to make to the science and practice of psychology, especially for those professionals most affected by the exigencies of managed care.

A Brief History of Single-Participant Research

Single-participant research is hardly a newcomer to psychology. Indeed, early experimental psychology, borrowing heavily from physiological laboratory methods, usually entailed manipulations of independent variables at the level of the individual participant (Boring, 1929). The subject matter of interest, and the sometimes invasive independent variables manipulated, often mitigated against the use of large numbers of participants. Consequently, data were collected from and presented not in aggregate form but at the level of the individual participant. Of course, multiple participants actually took part in these studies, although this was done primarily as an exercise in interparticipant replication, not to enhance statistical power.

Even when experimentation turned to more psychological matters, the single-participant strategy served researchers well. Ebbinghaus (1885/1913), as the sole participant in his own research program, conducted the first systematic and thorough analysis of human memory. In doing so, he uncovered fundamental memory principles that remain, for the most part, unchallenged even today. In a similar manner, both classic and contemporary research in psychophysics has relied heavily on the intensive examination of individual perceptual processes, as represented, for example, in the receiver operating characteristics of signal-detection theory (Green & Swets, 1966; Swets, 1973). Finally, Pavlov's (1927) seminal research on the conditional reflex involved intrasubject comparisons of the dependent variable both before and after conditioning.

Perhaps the most ardent support for single-participant research has come from the operant research tradition, referred to by Skinner (1969) as the experimental analysis of behavior. To the behavior analyst, however, the method is more than a collection of experimental practices. Indeed, within the operant literature can be found a sizable and compelling epistemology supporting the single-participant method as a frequently superior alternative to the large-group hypothesis testing designs familiar to psychologists (Michael, 1974; Sidman, 1960; Skinner, 1953, 1956, 1966). In his autobiography, Skinner (1979) made no bones about his early intentions to reshape both the conceptual and methodological practices of scientific psychology. His revisions left little of the discipline untouched, and his rejection of group designs and inferential statistics was particularly notable for its intransigence. In addition to Skinner's writings, Sidman's groundbreaking text, *Tactics of Scientific Research,* still considered the bible of operant methodology, served both as a model for conducting research and as a convincing argument for the advantages of single-participant methodology. Finally, in 1958, operant researchers at Columbia University and Harvard University contemplated establishing a journal that would be more appropriate for their manuscripts:

> We had trouble getting our reports published in the regular journals. We used very small numbers of subjects, we did not "design our experiments" with matched groups, our cumulative records did not look like learning curves, and we were asking questions (for example, about schedules) that were not found in the "literature." (Skinner, 1987, p. 447)

Despite the unprecedented control that these researchers had shown over the behavior of their participants, journal editors were reluctant to accept the research, most often because the designs were not standard group designs and because the data analysis was not informed by inferential statistics. Thus, in 1958, in an effort spearheaded by Charles Ferster, this fledgling group of scientists founded *The Journal of the Experimental Analysis of Behavior,* a forum devoted exclusively to presenting experimental data

Robin K. Morgan

from individual organisms. It is interesting that this focus on individual-organism research remains the distinguishing feature of the journal today, regardless of the conceptual or theoretical framework characterizing the research.

The single-participant research design, then, has a rather lengthy history in the behavioral sciences, certainly more substantial than would be gleaned by perusing many contemporary textbooks on research methods in psychology. Moreover, the method was used by many of psychology's pioneers and across a rather diverse collection of research programs. Nevertheless, if one takes today's research design textbooks as reflective of the field, the single-participant method remains a relatively obscure option possessing little of the scientific rigor inherent in the more conventional group designs (Dermer & Hoch, 1999).

There are undoubtedly various reasons why single-participant methodology receives scant attention in serious discussions or writings on method in psychology. Perhaps paramount among these reasons is the uneasy fit between conventional inferential statistics and the data generated by single-participant studies. The issue is a complicated one, and little consensus has emerged regarding the appropriate method of analyzing sequential response data that, when obtained from a single organism, contain substantial serial dependency or autocorrelation (Bengali & Ottenbacher, 1997; Huitema, 1986). This serial dependency can increase the probability of both Type I and Type II errors in the interpretation of single-participant data (Sharpley & Alavosius, 1988). It is fortunate that researchers are aware of the problems inherent in analyzing such data, and a sizable literature has been devoted to exploring the role of both traditional and nontraditional inferential statistics in single-participant research (Franklin, Allison, & Gorman, 1997; R. R. Jones, Vaught, & Weinrott, 1977), including contemporary discussions of effect size estimation (Kromrey & Foster-Johnson, 1996) and meta-analysis (Baron & Derenne, 2000; Faith, Allison, & Gorman, 1997; Kollins, Newland, & Critchfield, 1999). In fact, at least one past associate editor of *The Journal of the Experimental Analysis of Behavior* reported an increase in the percentage of submitted manuscripts in which inferential statistics were used (Ator, 1999). At the very least, this may be interpreted as an increased awareness among researchers that inferential statistics may prove useful to single-participant researchers.

However, many single-participant researchers continue to justify the graphic display and visual analysis of such data, with little interest in formal null hypothesis testing (Hopkins, Cole, & Mason, 1998; Michael, 1974; Parsonson & Baer, 1986; Perone, 1999). These scholars argue that the data from an individual participant behaving under well-specified conditions should provide unequivocal evidence of an independent variable's effect and that such an effect should be visible to the naked eye. For Skinner (1969), the matter was hardly a contentious one:

> Unlike hypotheses, theories, and models, together with the statistical manipulations of data which support them, a smooth curve showing a change in probability of response as a function of a controlled variable is a fact in the bag, and there is no need to worry about it as one goes in search of others. (p. 84)

But it is not only the analysis and interpretation of data in single-participant studies that differ from traditional group designs. The method of data collection is similarly unorthodox and foreign to psychologists who have cut their methodological teeth on standard group designs. In an engaging and informative account of his development as a scientist, Skinner (1956) traced his growing frustration with large-*N* studies driven by statistical inference. Not only did group means obscure the orderly and systematic development of a behavioral repertoire in an experimental animal, but the requirement of manipulating variables in dozens of separate experimental cubicles proved impractical and stifled attempts at following up on interesting functional processes as they were identified. "No matter how significant might be the relations we actually demonstrated, our statistical Leviathan had swum aground. The art of the method had stuck at a particular stage of its development" (Skinner, 1956, p. 228).

Perhaps also because Skinner reported to graduate school at Harvard University without prior schooling in psychology, he was lacking any formal indoctrination in the discipline's growing enchantment with Fisherian designs. In fact, his early years were spent in William Crozier's physiology laboratory, where he came to appreciate the experimental control and precise measurement strategies common to the natural sciences. Skinner was generally unimpressed with psychology at Harvard, prompting the following proclamation in a letter to a close friend: "I have almost gone over to physiology, which I find fascinating. But my fundamental interests lie in the field of Psychology, and I shall probably continue therein, even, if necessary, by

making over the entire field to suit myself" (Skinner, 1979, p. 38). Nor would he waste much time redesigning the methods of experimental psychology that he found so wanting, in the process borrowing from Pavlov, whose impressive experimental achievements stemmed from a rigorous control over his subject matter and the observation of lawful behavioral principles in individual organisms. Pavlov's experimental method seemed to Skinner to be not only applicable to behavior as a subject matter but also consistent with his conceptual position that psychology was, in fact, a natural science.

Other sciences provided additional contributions to Skinner's developing methodology. From physical chemistry, he incorporated the concept of a steady state, which would prove invaluable to the experimental analysis of behavior. Skinner (1956) chose response rate as his primary dependent measure, and when behavior was observed repeatedly over long experimental sessions, it became possible to ascertain a steady state, as reflected in minimal variability and no discernible upward or downward trend in response rate. This strategy allowed for sensitive comparisons of steady-state responding under a preintervention or "baseline" phase and response rates during subsequent treatment phases. The steady-state strategy would eventually prove itself a powerful design feature not only in the context of both basic conditioning and psychophysical research but also in the then embryonic fields of behavioral pharmacology (Boren, 1966; Dews & Morse, 1961) and behavior therapy (M. C. Jones, 1924; Lindsley, Skinner, & Soloman, 1953; Wolpe, 1958). Of course, the strategy requires nearly continuous contact with one's subject matter and repeated measurement of dependent variables. Though this may at times be a tall order, the subject matter calls for nothing less. Both Skinner (1956) and Sidman (1960) were adamant about the importance of using research strategies that adequately captured the natural dimensions of the subject matter, and discrete group means seemed poorly suited to capturing the continuity of behavioral processes.

As a variant in research methodology, single-participant designs possess unambiguous advantages, if one is interested in the development of behavior in a single organism over time. Although this is not the explicit aim of all research in psychology, or the behavioral sciences in general, such a goal would seem to be endorsed, if perhaps unwittingly, more than might be acknowledged. In fact, at an applied level, it is seldom the case that treatment objectives center on how group averages respond to manipulated variables and interventions. For parents, teachers, therapists, and others charged with changing behavior, the individual ordinarily constitutes the unit of analysis, and change makes itself known only through multiple measures taken over prolonged observational periods. The single-participant design evolved because it allowed for a sensitive assessment of developing behavioral repertoires, which remains its primary advantage, whether realized in the basic laboratory or in clinical settings.

Features of Single-Participant Designs

The phrase *single participant* is, of course, misleading if it is interpreted to mean that such an experiment involves only one participant. The essence of single-participant experimentation is simply that all dependent measures are collected repeatedly over the course of the experiment, and these data are not combined with those from other participants to produce group averages for purposes of data analysis. In most such experiments, data are collected from a handful of participants, although the numbers clearly do not begin to approach the large sample sizes expected of group designs. With proper controls and experimental manipulations, numerous intra- and interparticipant replications can be conducted with a small number of participants, allowing for strong inferences concerning functional relationships between behavior and its controlling variables.

Repeated Measures

Loftus (1996) has argued that psychology's slow progress as a science can be traced in large measure to the way we analyze and interpret data:

> What we do, I sometimes think, is akin to trying to build a violin using a stone mallet and a chain saw. The tool-to-task fit is not very good, and, as a result, we wind up building a lot of poor quality violins. (p. 161)

It is not, however, the data analytic process alone that differentiates psychology from other sciences but the very conditions under which data are collected in the first place. Many observational strategies in psychology rely on a single observation of the dependent variable. The poor "goodness of fit" between this discrete measurement practice and a temporally dynamic subject matter is equivalent to underusing the resolving power of a microscope. In contrast, the single-participant design epitomizes the concept of repeated measures and, in so doing, inverts the measurement practices of group designs. Whereas the strategy in group research usually involves one or two dependent measures from a large sample of participants, single-participant research uses frequent and continuous measurement of the dependent variable from individual participants. This strategy is justified on two grounds. First, numerous measures of a participant's behavior increase the experimenter's confidence that the sample of behavior being measured is representative of that participant under those experimental conditions. This logic is not entirely unlike that attached to sampling issues in group designs, though the referents of such terms as *sample* and *population* are clearly markedly different in single-participant research. Second, repeated measures are considered a natural consequence of an epistemology that conceptualizes behavior as a continuously unfolding phenomenon. Behavior exhibits considerable serial dependence, and to be scientifically viable, observational and measurement schemes must make sufficient contact with this dimension of the subject matter. Computing measures of central tendency on dependent variables obtained from a single, discrete observation may be logistically convenient, but it compromises unnecessarily the natural dimensions of the subject matter.

Participants Serving as Their Own Controls

Similar to participants in within-participant group designs, participants in single-participant designs serve as their own controls, with comparisons being made across experimental conditions. But individual differences make no contribution to the variance in single-participant designs because no comparisons are made across participants. A participant's behavior in one phase of the experiment is compared with his or her own behavior under other phases, not with the behavior of other participants. This is viewed as the only relevant comparison, because in most natural settings, the question will be whether an individual's behavior has changed relative to his or her own baseline, not relative to that of another person. Indeed, the phrase *behavior change* assumes little meaning at the group level. In evaluating, for instance, whether a speech therapy intervention has been effective for a child, one needs to know how the child's speech after therapy compares with his or her speech before therapy, not how much the child's speech deviates from a statistically derived group average. It is difficult to envision a clinical application, at least within psychology, medicine, or other service-oriented disciplines, for which the purported goal is the alteration of group means. As a result, professionals in applied disciplines have come to recognize the limitations of standard group designs and the inherent logic and meaningfulness of single-participant designs to health science research (Elder, 1997; Ottenbacher, 1992; Perrin, 1998). This point was particularly well made by Lundervold and Bellwood (2000):

> Group experimental design methodology by definition is insensitive to the exigencies of everyday practice. Although group experimental design methodology is appropriate for technique testing, counseling practice is primarily concerned with the development of techniques that are effective for the individual case or technique building. Consequently, it is ironic that a research methodology, single-case (N = 1) design, developed for use in practice settings and capable of evaluating counseling process, evaluating counseling intervention outcomes, and demonstrating experimental control, continues to be the "best kept secret" in counseling. (p. 78)

Emphasis on Experimental Replication

The role of replication in scientific endeavors can hardly be overestimated. The self-corrective nature of science, often lauded as being one of the advantages of science as a method of inquiry, depends on the ability to establish the reliability of discovered functional relationships, and replication plays a crucial part in this agenda. Replication studies, however, are often hard to come by in the behavioral sciences, particularly certain areas of "soft psychology" (Meehl, 1978). The practical exigencies of repeating a study requiring large samples of participants may often preclude such work. Without substantial resources, including monetary support, a replication involving hundreds, or perhaps thousands, of participants is an unlikely event. This is unfortunate because it is precisely in such research areas, in which experimental controls are often lacking and variable definition and measurement exhibit considerable interstudy variability, that replication serves its most useful purpose.

Single-participant research, in contrast, not only allows for but also is in fact defined in part by its reliance on replication. The reliability of an independent variable manipulation often can be evaluated through simple intraparticipant replication across several phase changes in a single experiment. Replication is ordinarily done by alternating baseline (nontreatment) and treatment conditions. Of course, such reversal, or ABA, designs are not always feasible for ethical or logistical reasons. Many kinds of behavior, particularly changes produced through learning, do not simply go away or reverse once treatment is removed. When this is the case, interparticipant replication becomes a viable option, most commonly through a multiple-baseline strategy. Although there are several versions of the multiple-baseline design (see, e.g., Kazdin, 1994), perhaps the most common entails repeated implementation of treatment across several participants but after differing baseline durations. The staggered manner in which treatment is implemented controls for the kinds of threats to internal validity that are otherwise common in pretest–posttest designs. Thus, as a means of demonstrating interparticipant replication, the multiple-baseline design allows for strong causal inferences, especially as the number of replications increases. Moreover, such interparticipant replications are often quite manageable at little additional cost or effort when they are conducted within the context of single-participant research.

The programmatic manipulation of independent variables, both within and across participants, has long been the hallmark of research in the experimental analysis of behavior (Skinner, 1966). Moreover, when research is conducted at the level of an individual participant, replication can be flexible and maximally sensitive to previous experimental findings. Problems that arise, be they methodological or technical, can often be promptly dealt with, and serendipitous discoveries that lead in different directions can be expeditiously pursued. The method is unabashedly inductive and resembles logically the research strategy of the natural scientist. Moreover, the single-participant design is compatible with the idiographic decision making of what Stricker and Trierweiler (1995) called the "local clinical scientist." This flexibility of method is unheard of in large group designs in which hypothetico–deductive logic places substantial restraints on what will be observed during experimentation and what sorts of inferences can be drawn from the results. As Skinner (1956) suggested, this research mentality necessarily creates a kind of observational myopia whereby results that appear irrelevant to the experimental hypothesis garner little attention, despite the fact that they may bear substantial theoretical or empirical implications.

Graphic Presentation and Visual Analysis of Data

As we mentioned previously, data presentation and analysis in single-participant research differ markedly from data treatment in traditional group designs. The most conspic-

uous difference is, of course, the presentation of data from individual participants rather than summarized aggregate measures. Moreover, the conventional vehicle for data presentation in single-participant research is the real-time graph, in which dependent variable measures typically appear on the ordinate and independent variable conditions (often depicted across time) typically appear on the abscissa. Individual data points on such graphs usually depict such measures as response rate, percentage of correct responding, and so on.

As we previously discussed, interpretation of single-participant data is seldom informed by the statistical criteria associated with the Fisherian tradition, and the question of how best to fit such data into the Fisherian protocol remains quite contentious (Ator, 1999; Baron, 1999; Branch, 1999; Huitema, 1986; Michael, 1974; Perone, 1999). Single-participant researchers argue that meaningful effects of an independent variable ought to be noticeable on visual inspection, particularly when the full power of the steady-state strategy is used. Thus, visual inspection of dependent measures during independent variable conditions, relative to baseline measures, represents the standard treatment of single-participant data. Rather than endorsing the formal decision criteria of null hypothesis testing designs, single-participant research evaluates behavior change relative to benchmarks provided by participants themselves. In addition, the very process of data presentation and analysis is an ongoing effort in single-participant studies, as opposed to a process that "kicks in" only once the data have been collected, as is more common in group studies. Perone (1999) has argued that the single-participant researcher retains more intimate and continuous contact with the subject matter of interest because of this style of data presentation:

> Skinner, the consummate tinkerer, was quite willing to scout about for new ways to conduct experiments. He rejected group-statistical methods not because they collided with his radical behaviorist epistemology, but rather because his experience revealed that they insulated the investigator from the behavior of the subject. (p. 111)

Treatment of Variability in Single-Participant Research

All scientific pursuits can be conceptualized as attempts to account for variability in the phenomenon of interest. The manner in which this is done differs across disciplinary boundaries, and in its treatment of variability, the single-participant method both methodologically and epistemologically distinguishes itself from group designs in psychology. Among the functions of statistical techniques in group research is the "neutralization" of error variance, a significant portion of which is contributed by individual differences. The fact that such variance is often described in our textbooks as "noise" or "nuisance" variability does much to capture the spirit of the Fisherian approach. Indeed, many statistical procedures, such as analysis of covariance, have been developed for the express purpose of statistically managing variables not targeted for primary analysis. In addition, the very use of measures of central tendency, such as the mean, illustrate a certain discomfort or lack of patience with the variability inherent in natural phenomena.

There is ample justification, however, to be suspicious of attempts to represent complex natural phenomena by means of convenient mathematical abstractions. Gould (1996) offered a compelling argument that in reducing our subject matter to single aggregate measures, we end up neglecting the one irreducible property of all natural phenomena, variability:

> What can be more discombobulating than a full inversion, or "grand flip", in our concept of reality: in Plato's world, variation is accidental, while essences record a higher reality; in Darwin's reversal, we value variation as a defining (and concrete earthly) reality, while averages (our closest operational approach to "essences") become mental abstractions. (p. 41)

Gould's (1996) position is articulated within the context of a discussion of evolution, whose status as science's grand unifying theory emerged only when it was acknowledged that genetic variation served as the raw material on which selective pressures could operate over time. Much of his provocative *Full House: The Spread of Excellence From Plato to Darwin* (Gould, 1996) is a timely scolding of all researchers who forget that variability is nature's originally dealt hand and that mathematical summarization is simply an effort to impose some sense of order, albeit arbitrary, on our observations.

If one begins with the assumption that variability represents the core subject matter of science, rather than an inevitable nuisance to be sidestepped, certain consequences for methodology follow. For the single-participant researcher, variability, as evidenced in the data of a single participant during the course of an experimental phase, is pivotal information about the impact of independent variables over time. The ongoing measurement of dependent variables allows for a sensitive metric of the participant's behavior, often in response to prolonged exposure to experimental conditions. Moreover, such refined assessment may frequently reveal the unintentional effects of extraneous factors when, for instance, behavior shows abrupt and marked deviations from a steady state in the absence of independent variable manipulations. Such deviations from steady-state behavior are always informative about functional relationships and can, in applied settings, prove especially useful in the development and modification of a treatment program. The question of why such variability occurs is an empirical one, which is best pursued through refinements in experimental procedure and, if need be, observational strategies, not through statistical maneuverings that discourage further inquiry.

When one has statistically controlled for a variable, more powerful data analysis may result, but little has been learned about that variable's impact on the subject matter. Experimentally controlling for variables, when possible, requires the researcher to come into direct contact with the relationships that occur between extraneous variables and the behavior of interest. Because these variables do exist and often exert influence outside the confines of a formal

study, much can be gained in the process of attempting to control or eliminate such factors. What may be viewed as nuisance factors by the researcher may in fact be variables of unparalleled importance in the participant's natural setting. The single-participant method, owing largely to its flexible nature, renders assessment and control of such variables more feasible than is true of group designs.

The General Applicability of the Single-Participant Method

Perhaps because single-participant research received its most ardent endorsement from Sidman (1960) and Skinner (1966, 1969), the method and its supporting epistemology are viewed as idiosyncratic to the operant tradition. This is unfortunate, however, because the approach is theoretically neutral. Observing moment-to-moment interactions between an organism and its local environment does not commit one to any particular brand of theorizing or conceptual interpretation. Rather, it simply puts one in contact with a different dimension of the subject matter than offered by group designs and summarized measures.

A case in point is Newell and Simon's (1972) groundbreaking text entitled *Human Problem Solving*, which details a research program for which rich and informative data could not have been generated by standard group methods. In a manner remarkably similar to Skinner's rejection of group designs, Newell and Simon adopted a research strategy that logically addressed the inherent parameters of their subject matter. Their approach was to ask participants to verbalize every step taken in a problem-solving task—a sort of dialogue or self-narration describing the details of their ongoing strategy. The development of protocol analysis (Ericsson & Simon, 1984) was in fact an important milestone in the study of problem solving, and the data presented in *Human Problem Solving* are culled exclusively from this idiographic process. Curiously absent, however, are the statistical tests and hypothesis-testing conventions long thought necessary to research in the behavioral sciences. It is in fact remarkable how many of psychology's classic empirical contributions were derived from methodological approaches bearing no resemblance to the null hypothesis testing tradition. A veritable "Who's Who" of psychology's luminaries, including Pavlov, Piaget, Ebbinghaus, and Skinner, conducted their very substantial research programs with alarmingly little concern for sample size or alpha levels (Morgan, 1998).

The relevance of research designs targeting development and change in individual behavior has clearly not been lost on professionals who deliver one-on-one services to their clients. Among those currently singing the praises of single-participant methodology are nursing and occupational and physical therapy practitioners (e.g., Backman, Harris, Chisholm, & Monette, 1997; Blair, 1986; Bryson-Brockman & Roll, 1996; Elder, 1997; Ottenbacher, 1992; Sterling & McNally, 1992). For instance, Holm, Santangelo, Brown, and Walter (2000) implemented a single-participant design to investigate the impact of three occupation-based interventions for reducing the frequency of disruptive vocalizations, distraction of others, and withdrawal from appropriate social interaction. The study involved 2 participants, a 17-year-old woman diagnosed with bipolar disorder, intermittent explosive disorder, and mild mental retardation and a 19-year-old woman diagnosed with major depression with psychiatric features, borderline personality disorder, and moderate mental retardation. The occupation-based interventions were implemented across three separate settings (school, a sheltered workshop, and two variations of a community living arrangement) in a multiple-baseline design. Dysfunctional behaviors occurred with less frequency in the school and sheltered-workshop settings than in the community setting.

In a similar manner, Linderman and Stewart (1999) used a single-participant approach to examine the effects of sensory integrative-based occupational therapy on the functional behaviors of two 3-year-old boys with pervasive developmental disorder. Repeated measures were taken during both a two-week baseline and a subsequent treatment phase. Both boys displayed substantial improvement in the areas of social interaction, approach toward new activities, and response to holding or hugging. Conversely, disruptive behaviors decreased in frequency and duration. Both of these studies highlight the unique advantages of the single-participant method. Because behavioral disorders invariably manifest themselves in an idiosyncratic manner, the processes of reliable assessment, treatment development and implementation, and treatment evaluation take on similarly individualistic dimensions. Moreover, group designs intended to identify effective interventions necessarily produce "conditional" knowledge, in that treatment efficacy will ultimately depend on several separate factors, some of which pertain to idiosyncratic client features. These idiosyncratic features, ordinarily beyond the scope of a large group study, are the heart and soul of single-participant designs.

Limitations of Single-Participant Research Designs

There is, of course, no reason to suppose that all questions about human experience can be pursued efficaciously only through single-participant methods. The nature of one's subject matter and the particular goals of the research project must inform all design and measurement issues. As we mentioned previously, single-participant designs are most suited to projects in which the unfolding behavioral repertoire of an individual organism is of primary interest. When, instead, one is interested in population parameters for the purpose of establishing social policies or regulations affecting educational, political, or social institutions, then group designs may have considerable usefulness.

Single-participant research is decidedly experimental in its approach. That is, the most powerful use of the method is when independent variables can be manipulated and conspicuous extraneous variables can be effectively controlled. Although such design features are more easily realized in laboratory settings, the method is capable of surprisingly effective exportation to natural settings. In fact, the flexible nature of the strategy and its capacity for rapid adaptation to changes in participants' behavior or

setting features make single-participant research especially well suited to applied settings. In addition, the emphasis on ongoing dependent variable observation offers substantial benefits even when, for ethical or practical reasons, variables cannot be purposefully manipulated. The systematic collection of behavioral data under properly specified environmental conditions can provide invaluable information about a participant's behavior and its controlling variables, even in the absence of explicit interventions. Perhaps the most frequently mentioned shortcoming of single-participant (often called small-*N*) designs is their presumably minimal external validity. Their resemblance to clinical case studies in this regard is seldom ignored by methodology textbook authors who question the usefulness of data generated by a single participant. But the issue of generality is a complicated one, and the full measure of an experimental design cannot be adequately evaluated without a proper stocktaking of the phenomenon in question and the domains across which generality is being evaluated. Nor is the issue conveniently put to rest by the size of one's sample:

> We cannot dispose of the problem of subject generality by employing large groups of subjects and using statistical measures, such as the mean and variance of the groups. It is not true that the larger the group, the greater is the generality of the data. Representativeness is an actuarial problem to which the currently prevalent statistical design is not applicable. (Sidman, 1960, p. 47)

Ottenbacher (1990) has also argued that the formal statistical and probability requirements that are ordinarily considered necessary to ensure external validity almost never eventuate in clinical research. "Given the empirical exigencies associated with most clinical research in rehabilitation, generalizability judgments based on a statistical model are simply not possible or statistically legitimate" (Ottenbacher, 1990, p. 290).

As we previously described, single-participant designs are distinguished by their reliance on both intra- and inter-participant replications. The latter, in particular, serve not only as a reliability check on the particular functional relationship being pursued but also as an assessment of individual differences in its expression. Such replications, a staple feature of the method, are made possible by the ease with which changes in experimental conditions can be made across participants and on short notice. Thus, the generality of a behavioral phenomenon is seen not so much as an exercise in statistical inference but as an experimental practice in which replication allows for a thorough evaluation of generality across independent variable parameters, stimulus conditions, and participant variables.

Conclusion

The advent of managed care has led to significant changes in the landscape of health care practice, including a forceful mandate that practitioners be able to document the effectiveness of their clinical interventions. This emphasis on accountability places a premium on the conduct of research in applied settings. Yet the conventional null hypothesis machinery of psychological science is embarrassingly unwieldy in a practice environment, and health care providers consequently perceive themselves as disenfranchised from the business of demonstrating treatment effectiveness. Recently, however, both practitioners and methodologists representing divergent training and theoretical persuasions have questioned the continued uncritical acceptance of the Fisherian strategy within the behavioral sciences. If no alternative methods existed, such criticism would ring hollow. Alternatives, however, do exist, and among them single-participant research has enjoyed documented success in contributing to the empirical database of psychology. The method, rich in history, epistemology, and design power, has remained largely unappreciated because of its poor fit with the logic of statistical inference and deductive hypothesis testing. Its features include an unabashed interest in the development of behavioral repertoires, a staunch declaration that such development is obscured by group measures, and a flexibility of method reminiscent of the natural sciences. These features should be especially appealing to practicing clinicians who are delivering services to individual clients and whose professional responsibilities increasingly include documentation of treatment efficacy.

REFERENCES

Abelson, R. P. (1997). On the surprising longevity of flogged horses: Why there is a case for the significance test. *Psychological Science, 8,* 12–15.

Acuff, C., Bennett, B. E., Bricklin, P. M., Canter, M. B., Knapp, S. J., Moldawsky, S., & Phelps, R. (1999). Considerations for ethical practice in managed care. *Professional Psychology: Research and Practice, 30,* 563–575.

Ator, N. A. (1999). Statistical inference in behavior analysis: Environmental determinants? *The Behavior Analyst, 22,* 93–97.

Backman, C. L., Harris, S. R., Chisholm, J. M., & Monette, A. D. (1997). Single-subject research in rehabilitation: A review of studies using AB, withdrawal, multiple baseline, and alternating treatments designs. *Archives of Physical Medicine and Rehabilitation, 78,* 1145–1153.

Baron, A. (1999). Statistical inference in behavior analysis: Friend or foe? *The Behavior Analyst, 22,* 83–85.

Baron, A., & Derenne, A. (2000). Quantitative summaries of single-subject studies: What do group comparisons tell us about individual performances? *The Behavior Analyst, 23,* 101–106.

Bengali, M. K., & Ottenbacher, K. J. (1997). The effect of autocorrelation on the results of visually analyzing data from single-subject designs. *American Journal of Occupational Therapy, 52,* 650–655.

Berman, B. (2000). The academic children's hospital primary care clinic: Responding to the challenge of a changing health care environment. *Clinical Pediatrics, 39,* 473–478.

Blair, C. E. (1986, May). A case for single-subject research in nursing. *AARN Newsletter, 42,* 15–16.

Boren, J. J. (1966). The study of drugs with operant techniques. In W. K. Honig (Ed.), *Operant behavior: Areas of research and application* (pp. 531–564). Englewood Cliffs, NJ: Prentice Hall.

Boring, E. G. (1929). *A history of experimental psychology.* New York: Appleton-Century-Crofts.

Branch, M. N. (1999). Statistical inference in behavior analysis: Some things significance testing does and does not do. *The Behavior Analyst, 22,* 87–92.

Bryson-Brockman, W., & Roll, D. (1996). Single-case experimental designs in medical education: An innovative research method. *Academic Medicine, 71,* 78–85.

Chow, S. L. (1998). Precis of statistical significance: Rationale, validity, and utility. *Behavioral and Brain Sciences, 21,* 169–194.

Cohen, J. (1990). Some things I have learned (so far). *American Psychologist, 45,* 1304–1312.

Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist, 49*, 997–1003.
Dean, R. G. (1998). Postmodernism and brief treatment: A more inclusive model. *Crisis Intervention and Time-Limited Treatment, 4*, 101–112.
Dermer, M. L., & Hoch, T. A. (1999). Improving descriptions of single-subject experiments in research texts written for undergraduates. *The Psychological Record, 49*, 49–66.
Dews, P. B., & Morse, W. H. (1961). Behavioral pharmacology. *Annual Review of Pharmacology, 1*, 145–174.
Ebbinghaus, H. (1913). *Memory* (H. A. Rueger & C. E. Bussenius, Trans). New York: Teachers College. (Original work published 1885)
Elder, J. H. (1997). Single-subject experimentation for psychiatric nursing. *Archives of Psychiatric Nursing, 11*, 133–138.
Ericsson, K. A., & Simon, H. A. (1984). *Protocol analysis: Verbal reports as data.* Cambridge, MA: MIT Press.
Estes, W. K. (1997). Significance testing in psychological research: Some persisting issues. *Psychological Science, 8*, 18–20.
Faith, M. S., Allison, D. B., & Gorman, B. S. (1997). Meta-analysis of single-case research. In R. D. Franklin, D. B. Allison, & B. S. Gorman (Eds.), *Design and analysis of single-case research* (pp. 245–277). Mahwah, NJ: Erlbaum.
Franklin, R. D., Allison, D. B., & Gorman, B. S. (Eds.). (1997). *Design and analysis of single-case research.* Mahwah, NJ: Erlbaum.
Gould, S. J. (1996). *Full house: The spread of excellence from Plato to Darwin.* New York: Three Rivers Press.
Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics.* New York: Wiley.
Harris, R. J. (1997). Significance tests have their place. *Psychological Science, 8*, 8–11.
Hayes, S. C. (1981). Single-case experimental design and empirical clinical practice. *Journal of Consulting and Clinical Psychology, 49*, 193–211.
Hilliard, R. B. (1993). Single-case methodology in psychotherapy process and outcome research. *Journal of Consulting and Clinical Psychology, 61*, 373–380.
Holm, M. B., Santangelo, M. A., Brown, S. O., & Walter, H. (2000). Effectiveness of everyday occupations for changing client behaviors in a community living arrangement. *American Journal of Occupational Therapy, 54*, 361–371.
Hopkins, B. L., Cole, B. L., & Mason, T. L. (1998). A criticism of the usefulness of inferential statistics in applied behavior analysis. *The Behavior Analyst, 21*, 125–137.
Huitema, B. E. (1986). Autocorrelation in behavioral research: Wherefore art thou? In A. Poling & R. W. Fuqua (Eds.), *Research methods in applied behavior analysis: Issues and advances* (pp. 187–208). New York: Plenum.
Jones, M. C. (1924). The elimination of children's fears. *Journal of Experimental Psychology, 7*, 383–390.
Jones, R. R., Vaught, R. S., & Weinrott, M. (1977). Time-series analysis in operant research. *Journal of Applied Behavior Analysis, 4*, 45–49.
Kazdin, A. E. (1994). *Behavior modification in applied settings* (5th ed.). Pacific Grove, CA: Brooks/Cole.
Kollins, S. H., Newland, M. C., & Critchfield, T. S. (1999). Quantitative integration of single-subject studies: Methods and misinterpretations. *The Behavior Analyst, 22*, 149–157.
Kromrey, J. D., & Foster-Johnson, L. (1996). Determining the efficacy of intervention: The use of effect sizes for data analysis in single-subject research. *Journal of Experimental Education, 65*, 73–93.
Linderman, T. M., & Stewart, K. B. (1999). Sensory integrative-based occupational therapy and functional outcomes in young children with pervasive developmental disorders: A single-subject study. *American Journal of Occupational Therapy, 53*, 207–213.
Lindsley, O. R., Skinner, B. F., & Soloman, H. C. (1953). *Studies in behavior therapy: Status Report 1.* Waltham, MA: Metropolitan State Hospital.
Loftus, G. R. (1993). A picture is worth a thousand p values: On the irrelevance of hypothesis testing in the microcomputer age. *Behavior Research Methods, Instruments & Computers, 25*, 250–256.
Loftus, G. R. (1996). Psychology will be a much better science when we change the way we analyze data. *Current Directions in Psychological Science, 5*, 161–171.
Lundervold, D. A., & Bellwood, M. F. (2000). The best kept secret in counseling: Single-case (large $N = 1$) experimental designs. In M. R. Jalongo, G. J. Gerlach, & W. Yan (Eds.), *Annual editions: Research methods* (pp. 78–90). Guilford, CT: McGraw-Hill/Dushkin.
Marten, P. A., & Heimberg, R. G. (1995). Toward an integration of independent practice and clinical research. *Professional Psychology: Research and Practice, 26*, 48–53.
Meehl, P. E. (1978). Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. *Journal of Consulting and Clinical Psychology, 46*, 806–834.
Michael, J. (1974). Statistical inference for individual organism research: Mixed blessing or curse? *Journal of Applied Behavior Analysis, 7*, 647–653.
Morgan, D. L. (1998). Selectionist thought and methodological orthodoxy in psychological science. *The Psychological Record, 48*, 439–456.
Newell, A., & Simon, H. A. (1972). *Human problem solving.* Englewood Cliffs, NJ: Prentice Hall.
Ottenbacher, K. J. (1990). Clinically relevant designs for rehabilitation research: The idiographic model. *American Journal of Physical Medicine and Rehabilitation, 69*, 286–292.
Ottenbacher, K. J. (1992). Analysis of data in idiographic research: Issues and methods. *American Journal of Physical Medicine and Rehabilitation, 71*, 202–208.
Parsonson, B. S., & Baer, D. M. (1986). The graphic analysis of data. In A. Poling & R. W. Fuqua (Eds.), *Research methods in applied behavior analysis: Issues and advances* (pp. 157–186). New York: Plenum.
Pavlov, I. P. (1927). *Conditioned reflexes.* London: Clarendon Press.
Perone, M. (1999). Statistical inference in behavior analysis: Experimental control is better. *The Behavior Analyst, 22*, 109–116.
Perrin, T. (1998). Single-system methodology: A way forward in dementia care. *British Journal of Occupational Therapy, 61*, 448–452.
Sanavio, E. (1998). *Behavior and cognitive therapy today: Essays in honor of Hans J. Eysenck.* Oxford, England: Elsevier Science.
Seligman, M. E. P., & Levant, R. F. (1998). Managed care policies rely on inadequate science. *Professional Psychology: Research and Practice, 29*, 211–212.
Sharpley, C. F., & Alavosius, M. P. (1988). Autocorrelation in behavioral data: An update. *Behaviour Change, 4*, 40–45.
Shrout, P. E. (Chair), Hunter, J. E., Harris, R. J., Wilkinson, L., Strouss, M. E., Applebaum, M. I., Hunt, E., & Levin, J. R. (1996, August). *Significance tests—Should they be banned from APA journals?* Symposium conducted at the 104th Annual Convention of the American Psychological Association, Toronto, Ontario, Canada.
Sidman, M. (1960). *Tactics of scientific research.* New York: Basic Books.
Skinner, B. F. (1953). *Science and human behavior.* New York: Macmillan.
Skinner, B. F. (1956). A case history in scientific method. *American Psychologist, 11*, 221–233.
Skinner, B. F. (1966). What is the experimental analysis of behavior? *Journal of the Experimental Analysis of Behavior, 9*, 213–218.
Skinner, B. F. (1969). *Contingencies of reinforcement: A theoretical analysis.* Englewood Cliffs, NJ: Prentice Hall.
Skinner, B. F. (1979). *The shaping of a behaviorist: Part two of an autobiography.* New York: Knopf.
Skinner, B. F. (1987). Antecedents. *Journal of the Experimental Analysis of Behavior, 48*, 447–493.
Sterling, Y. M., & McNally, J. A. (1992). Single-subject research for nursing practice. *Clinical Nurse Specialist, 6*, 21–26.
Stricker, G., & Trierweiler, S. J. (1995). The local clinical scientist: A bridge between science and practice. *American Psychologist, 50*, 995–1002.
Swets, J. A. (1973, December 7). The relative operating characteristics in psychology. *Science, 182*, 990–1000.
Wilkinson, L., & the Task Force on Statistical Inference. (1999). Statistical methods in psychology: Guidelines and explanations. *American Psychologist, 54*, 594–604.
Wolpe, J. (1958). *Psychotherapy by reciprocal inhibition.* Stanford, CA: Stanford University Press.