

Chapter 4

DIMENSIONS OF SINGLE-CASE RESEARCH DESIGN AND DATA DISPLAY

As we saw in Chapter 2, single-case research has a rather long history in the behavioral sciences. Despite having been largely overshadowed for much of the past century by the large-group-hypothesis testing designs introduced by Fisher, single-case designs have emerged as an important alternative, particularly for researchers who conduct empirical studies of clinical interventions in the field. No longer is the single-case strategy relegated to basic studies of operant conditioning (Sidman, 1960) or psychophysics within the relatively artificial environs of the laboratory. Increasingly, clinical researchers are taking advantage of the natural flexibility of single-case designs to assess the effectiveness of psychological and physical interventions in applied settings (Borckardt, Nash, Murphy, Moore, Shaw, & O'Neil, 2008; Elder, 1997; Kuentzel, Henderson, Zambo, Stine, & Schuster, 2003; Nuehring & Pascone, 1986; Robey, Schultz, Crawford, & Sinner, 1999; Sterling & McNally, 1992). Thus, the single-case method finds itself today being applied to a much larger universe of natural phenomena than could have been envisioned by some of its earlier practitioners. Nevertheless, the method is characterized by a unique set of philosophical assumptions and strategic practices that set it apart from the Fisherian tradition. In this chapter, we explore these philosophical and strategic features of single-case research as well as the methods used by researchers for displaying single-case data.

AN IDIOGRAPHIC APPROACH TO BEHAVIOR

The most fundamental feature of single-case research design is an exclusive focus on the behavior of the individual organism. Research in this tradition is often referred to as **idiographic**, because the entire research enterprise, from conceptualization to observation and measurement, is directed toward drawing conclusions at the level of the individual case. This approach contrasts with a **nomothetic** perspective, which involves studying many cases for the purpose of ascertaining generalizable principles or laws. As we have already seen, the nomothetic approach is exemplified in the large group designs that have dominated research in the behavioral sciences for many years.

The distinction between idiographic and nomothetic approaches takes on special significance for scientists conducting research in applied settings. For example, a psychologist implementing a behavior analysis program for an autistic child engaged in self-injurious behavior would be expected to have a considerable interest in the effectiveness of the program. Although the psychologist might be curious as to whether the program would prove helpful to other clients exhibiting such behavior, this would not be the central question addressed by the study. The primary question is undoubtedly an idiographic one: Does the program as implemented bring about a clinically relevant reduction in the self-injurious behavior of *this particular client*? If it does, then the program will be considered successful, even if the same program fails to impact the same type of behavior in other clients. In fact, this kind of variability in treatment outcome is common. The same surgical technique, for instance, does not always prove successful, even for patients whose diagnoses are identical. Nor do chemical interventions, including standard antibiotics, produce the same outcome in all patients.

Inductive and Deductive Research

Of course, none of this implies that single-case researchers are entirely disinterested in the generality of a particular experimental or treatment effect. Obviously, a busy psychologist, nurse, or occupational therapist would be very appreciative of an intervention that proved useful to large numbers of clients presenting similar symptoms. In fact, it is nearly inevitable that once an effect is demonstrated for one research participant or client, the question will arise as to whether the effect will generalize to others. For the single-case researcher, answering this question requires a follow-up replication study, in which all elements of the research design are repeated systematically with another participant or client. Through a series of additional replications, the generality of the particular finding can be established, ultimately leading to the identification of a general principle, or law. Moving from

specific cases to general principles represents an **inductive** research strategy and remains a hallmark of single-case research design. In contrast, **deductive** strategies involve moving from general principles to specific cases. The large-group research tradition established by Fisher is, in fact, often called the *hypothetico-deductive method*, in part because the process of drawing inferences begins with a general principle, from which specific, testable hypotheses can be generated. Support for this generic principle is usually obtained through large-scale group designs, and it may then be used to make more specific predictions about individual cases.

It may strike you that both inductive and deductive approaches to science have the same ultimate objective. All scientific endeavors strive to make sense of the world, and this is best done when general, widely applicable principles can be identified. The laws of gravity, after all, would be of little use if they changed daily or if they were applicable in only isolated regions on earth. Similarly, we would have little use for principles of behavior or physiological functioning that were in need of substantial revision with every new client. A treatment regimen that proves effective for only one client will naturally be appreciated by that client, but it will have limited utility to the clinician if it is not more widely effective. The question, then, of generality is important to scientists working in applied settings. Both inductive and deductive strategies converge on the issue of generality, but in doing so they rely on different kinds of data (individual vs. group), and thus they draw inferences in different directions.

As we have seen, the single-case approach is decidedly idiographic in nature and supports an inductive method of drawing inferences. This philosophy is particularly “user friendly” to clinicians conducting research in applied settings because the concern in such cases is ordinarily with the effectiveness of a treatment regimen for a particular client and not establishment of a general principle or confirmation of a theory. All data collection, analysis, and decision making are understood to pertain only to the specific case under study. In a sense, applied research entails a conscientious effort to integrate scientific logic and professional practice. The general philosophy of this position was well articulated by Stricker and Trierweiler (1995) in their construct of the “local clinical scientist”:

The local clinical scientist brings the attitudes and knowledge base of the scientist to bear on the problems that must be addressed by the clinician in the consulting room. The problems of inadequate generalizability are reduced by a recognition of the value of local observations and local solutions to problems. (p. 995)

Idiographic and nomothetic perspectives on behavior should not be viewed as competing or mutually exclusive conceptual stances. Remember that all research represents an effort to reduce uncertainty about some natural phenomenon. The

74 SINGLE-CASE RESEARCH METHODS

kinds of questions asked by scientists approaching their subject matter from an idiographic perspective are often informed by data and theory generated by nomothetic research. A case in point is the concept of reinforcement in behavioral psychology. Psychologists refer to *reinforcers* as response-contingent stimuli that increase the frequency of the behavior that produces them. This definition is explicitly functional, meaning that reinforcers are identified solely on the basis of their ability to strengthen behavior, regardless of whether we would expect them to or whether we understand how they do so. Thus, at the conceptual level, reinforcement is viewed as a fairly generic process, believed to have relevance to all biological creatures.

Although the concept of reinforcement is considered a general principle of behavior, the actual study of reinforcement as a process becomes a much more specific enterprise. In many laboratory-based studies of operant conditioning, reinforcers include important biological resources, such as food and water. If an experimental animal is deprived of the reinforcer prior to the experiment, then that stimulus will tend to be effective in increasing almost any behavior on which it is made contingent. Because food and water are so vital to the survival of all animals, it hardly surprises us that they can be effective reinforcers in operant experiments.

In the applied human setting, however, the use of reinforcement principles becomes very idiosyncratic. Suppose you were a behavioral psychologist charged with enhancing the social skills of a shy and withdrawn adolescent. Your first order of business would be to conduct a proper assessment to identify possible reinforcers that might be effective in altering the child's behavior. Because depriving a human client of basic necessities such as food and water is an unethical practice, these particular stimuli would not be available to you to be used as reinforcers. However, by talking with the child or simply observing the child's natural behavior, you could probably obtain useful information about potential reinforcers for that client. Perhaps you would observe that the child likes to play video games, read books, go on hikes, shoot pool, and listen to country music. For this particular child, these represent preferred activities that might function as effective reinforcers; consequently, you might develop a program that allows the child to obtain a preferred activity by first engaging in some appropriate social behavior.

On the other hand, you would not necessarily expect these same activities to be preferred by other children for whom you might be developing an intervention. To construct a behavioral program for another child, you would need to do a separate assessment because events that act as reinforcers for one person may not function in the same way for another. Social attention is a good example. For many of us, there is nothing more powerful or self-validating than to have others, such as parents, teachers, or coworkers, draw attention to or praise our achievements. Thus, for most people, positive social attention and commendations are quite effective reinforcing

consequences. For some people, however, particularly socially withdrawn individuals, social attention is actually an aversive stimulus and would probably be ineffective if used as a reinforcer. This simply highlights the importance of conducting an idiographic assessment of possible reinforcers before constructing a program to alter behavior.

Notice that in the preceding example we are not claiming that the concept of reinforcement would not be expected to apply across many clients. In fact, reinforcement remains one of the most powerful and well-documented behavior principles in psychology, and its general applicability is for the most part uncontested. Reinforcement, as a nomothetic principle or law of behavior, enjoys rather rare status within the behavioral sciences. This fact, however, is of little use to the applied clinician who intends to use the principle idiographically to alter the behavior of a specific client. The specific stimuli that function as reinforcers differ from one person to the next, and a stimulus that reinforces behavior at one time may not reinforce it at another. Reading may be an effective reinforcer if one has not had the opportunity to read for some time, but if one has been reading uninterrupted for hours, the reinforcing capacity of reading may diminish. This is true simply because reinforcement, though a general principle, operates idiographically, or, as Stricker and Trierweiler (1995) would say, at the “local” level.

Replication in Single-Case Research

An important part of the logic of science is the establishment of important findings through experimental replication. The purpose of replication is straightforward. Scientists, like everybody else, are capable of making mistakes. When an especially important scientific finding is published, we want to be sure that it is reliable, particularly if large amounts of money are going to be invested in further research or technological development generated by the finding. But we all lose when a scientific finding receives substantial acclaim or attention, only to be proved incorrect by subsequent research. This is the reason that potentially significant findings are usually followed up immediately by attempts at replication, in which the procedural details of the experiment are repeated in order to discover whether the results remain the same. Only when a finding can be replicated over and over does it become an established part of the corpus of scientific knowledge. It is this continual process of experiment and replication that gives scientific knowledge its cumulative or progressive nature.

In single-case research, replication takes on added importance, especially in applied settings, where independent variables usually take the form of clinical treatments. In such circumstances, replications actually serve two functions, one of which is idiographic, the other nomothetic.

76 SINGLE-CASE RESEARCH METHODS

Intrasubject Replication. When an independent variable is manipulated in a single-case study, repeated measurement of the dependent variable allows the researcher to detect any change in behavior brought about by the independent variable. Remember, though, that an essential part of drawing conclusions from experiments is eliminating other possible explanations for the results. We cannot identify a causal relationship between the independent and dependent variables if some other extraneous variable intervened. If some other variable, unbeknownst to the researcher, happened to coincide with the independent variable, then the researcher could falsely identify the independent variable as the reason for the change in the dependent variable. In many applied studies, this amounts to declaring the treatment to have an effect when in fact it may not. We wish, quite obviously, to guard against such conclusions and the potentially dire implications they may provoke.

One method of ensuring that the change in behavior (dependent variable) observed in a study is, in fact, due to the independent variable (treatment) is to replicate the manipulation of the independent variable with the same subject. This ordinarily amounts to withdrawing or removing the treatment, allowing behavior to return to its pretreatment or baseline level, and then once again introducing the treatment. Because all of this is done at the level of the individual subject or client, we call this process **intrasubject replication**.

Suppose, for example, that a parent consults with a psychologist about a child who has developed the habit of crying, whining, and throwing tantrums each night at bedtime. The psychologist hypothesizes that the parent's tendency to comfort the child when this behavior emerges is actually reinforcing the problem behavior. As a result, the psychologist recommends that the parent no longer respond to the problem behavior at all, thus placing it on an extinction schedule. If the tantrum behavior diminishes or ceases altogether during the extinction phase, we would be tempted to attribute this behavior change to the extinction strategy. There are, however, other possibilities. Perhaps the child simply tired of the nightly tirades, or developed some precocious empathy for his or her frazzled parents. To evaluate whether the behavior change was due to the extinction procedure or some other variable, the parent could once again respond to inappropriate bedtime behavior with attention and comforting. Should the tantrum behavior return, the parent could implement extinction once again. This second application represents a replication of the first extinction phase. Should tantrum behavior reduce or extinguish once again during this phase, we would be unlikely to attribute it to an extraneous variable. It is simply improbable that some such variable just happened to coincide twice with the extinction component. In general, each successive replication of an independent variable's effect increases the researcher's confidence that behavior change was due to the manipulated variable, not some extraneous factor. Thus, intrasubject replications represent an idiographic strategy for strengthening causal

conclusions concerning the independent and dependent variables, thereby enhancing the internal validity of a study.

Intersubject Replication. Single-case researchers utilize **intersubject replication** as an inductive strategy to establish the generality or applicability of a manipulation or treatment beyond one subject or to other kinds of behavioral phenomena. For instance, having demonstrated that extinction of tantrum behavior in a child could be brought about by removing parental attention, we might want to know whether the same effect could be produced with a separate child. Successful replication of the treatment with several different children would demonstrate that the conclusions drawn from the first study seem to be applicable or generalizable to others, thus suggesting the **external validity** of the original findings.

Because they bear important consequences for the generality of particular findings, intersubject replications play an especially critical role in both basic and applied science. Results that are known to generalize, or apply across a wide variety of subjects, settings, or behaviors, clearly are likely to be taken quite seriously. Indeed, the more circumstances across which a particular principle seems to hold, the more status it will probably enjoy, either as a basic scientific principle or as an applied intervention. To discover, for instance, that withdrawal of parental attention successfully leads to the extinction of tantrum behavior in a particular child is certainly noteworthy. However, to demonstrate this principle across several children shows us that this relationship is not entirely idiosyncratic. Obviously, the more domains across which the finding can be shown to generalize, the more powerful the principle would appear to be. We might want to know, for example, whether a similar intervention might work in a classroom setting. We could also determine whether other kinds of behavior, besides tantrums, could be extinguished by simply removing attention. Would the same be true of adults as well? Notice that in each case we are asking whether a particular independent–dependent variable relationship will continue to hold up when applied to different types of subjects, behaviors, or settings than originally studied.

The use of intrasubject and intersubject replication often parallels the distinction between basic and applied research. As we have already seen, a clinical researcher evaluating a specific treatment in an applied setting is primarily interested in whether that intervention is effective for a particular client. Questions about treatment generality may simply not emerge. On the other hand, the question of whether the same treatment would be equally effective for a different kind of client population, or in a different institutional setting, is an inquiry on a larger scale. Some scientists believe that research of the latter sort is much needed in the behavioral and health sciences and that the esteem afforded applied interventions will prove to be proportionate to the amount of data attesting to their generaliz-

78 SINGLE-CASE RESEARCH METHODS

ability. This is a rather tall order, though, given that many applied researchers have neither the time nor the interest in establishing the conditions under which specific treatments will or will not prove effective. Research of this kind must usually be conducted programmatically, using a series of well-planned experiments intended to answer specific questions about generality. Johnston and Pennypacker (1993a) referred to this kind of research as *thematic* and argued that researchers pursuing generality must necessarily adopt a different kind of agenda than do most applied researchers: “In thematic research, when there is a conflict between experimental and service goals, the scales are tipped in favor of science so that the resulting interpretations may be unambiguous” (p. 179).

Tracking Behavior in Real Time

Most behavioral research from a nomothetic perspective places a major emphasis on collecting information from a large sample of research participants. Because time constraints place restrictions on how long the data collection phase of a study can take and because observations are being made of many participants (perhaps in the hundreds), it becomes impractical to make several observations of each participant over time. Thus, group designs often restrict themselves to one or two observations of behavior for each participant. In the case of experimental designs utilizing both experimental and control groups, this may amount to a single measure of behavior in each research participant after manipulation of the independent variable, as in a posttest-only control group design (Leary, 2004). Thus, behavior in large group studies is usually measured only once.

Behavior as a Continuous Subject Matter

Single-case researchers have long argued that observational and measurement strategies that treat behavior as a discrete subject matter fail to do justice to a phenomenon that is actually extended in time. For many scientists, **behavior** represents an ongoing, adaptive interplay between an organism and its environment; by its very nature, behavior as a subject matter resists efforts at discrete categorization. This point should not be underestimated, nor should it be viewed as merely esoteric, intellectual nit-picking. The position taken by single-case researchers is in fact quite consistent with commonsense perceptions of behavior. You need only consider examples from your own life to see why this is so. It is difficult to conceptualize even mundane instances of behavior without taking their temporal qualities into account. Even relatively unsophisticated behaviors, such as tying your shoes, cooking a meal, holding a phone conversation, or writing a letter, involve sometimes lengthy sequences of responses and fine adjustments to ongoing environmental feedback.

In fact, understanding behavior is often inconceivable without taking environmental feedback into account. Riding a bicycle is a particularly good example. Keeping a bicycle upright while riding involves moment-to-moment sensory feedback about body posture and shifts in one's center of gravity, and this information must be integrated with the ongoing activity of the arms (steering) and legs (pedaling). Suppose you were going to try to teach a child to ride a bike and your first order of business was to discover how far along the child was in acquiring this skill. Assume that two kinds of information were available to you prior to the first lesson: (1) a still photograph of the child and (2) a videotape of the child, both depicting the child attempting to ride the bike. Which kind of information would you prefer? Naturally, the answer to this question is pretty transparent. The still photograph represents a very impoverished and unrepresentative portrayal of the behavior of interest, whereas the videotape offers a rich database from which to draw conclusions about how best to help the child learn to ride. In the same vein, single-case research endorses the logic that observation and measurement of behavior should, whenever possible, be continuous, not discrete; that is, our research design should be developed in such a way as to accommodate the natural dimensions of the subject matter.

Behavior Change as Focus of Applied Research

An emphasis on behavioral continuity is especially apparent in applied research, primarily because this kind of research ordinarily concerns itself with some aspect of behavior change. Regardless of differences in training and credentials, mental health workers, nurses, and occupational and physical therapists all serve as behavior change agents. The interventions utilized by these professionals all entail helping clients address behavioral deficits or excesses that have implications for adaptive functioning. The psychologist providing relaxation training to a client paralyzed by social anxiety, the nurse teaching a diabetic patient how to assess his or her own blood sugar levels, the physical therapist developing an exercise regimen for an accident victim are all trying to bring about functional changes in the behavioral repertoires of their respective clients. Evaluating whether such efforts are effective requires a measurement strategy that is sensitive to changes in relevant dimensions of behavior, and such measurement presumes an ability to track the behavior over time.

Of course, the logic of continuous measurement extends to many nonclinical domains as well. Many human interactions have behavior change as their primary objective, although this fact is not always readily acknowledged. Both parenting and teaching, for instance, involve explicit attempts to alter behavioral repertoires. Learning to read, for instance, involves mastering incremental skills, from letter recognition and pronunciation to word and sentence recognition. The ability to

80 SINGLE-CASE RESEARCH METHODS

read fluently does not emerge instantaneously at any point, and it would be an odd claim to say that a child was not able to read on Wednesday but was able to do so on Thursday! Adequate assessment of a program designed to teach reading would require frequent, if not continuous, monitoring simply because the phenomenon itself unfolds incrementally over time.

INTERIM SUMMARY

Single-case research focuses on behavioral development within the individual, and this idiographic approach differs from the nomothetic approach typical of traditional group research. The primary vehicle for drawing conclusions is experimental replication. Intrasubject replication allows for rigorous conclusions at the level of the individual subject, and intersubject replication allows single-case research to be extended inductively to the more general case. Behavior change is the major emphasis in applied research, and single-case research is well adapted to studying behavior change because of its commitment to continuous observation and measurement of behavior over time.

INDEPENDENT AND DEPENDENT VARIABLES IN SINGLE-CASE RESEARCH

Single-case research design represents an experimental approach to the study of behavior. We have already seen that an experiment is a study in which the researcher controls or manipulates an independent variable while also controlling the effects of extraneous variables. When these two objectives can be achieved, alternative explanations can be ruled out, and the researcher is in a position to draw cause-and-effect conclusions about the variables in the study. Notice that this is the fundamental logic of experimentation, regardless of whether we are considering single-case or more traditional group designs. Moreover, the ability to manipulate the independent variable and rule out alternative explanations is also what distinguishes a single-case experiment from a case study. The latter involves no control over variables, usually because the relevant variables have already occurred for natural reasons. The resemblance between single-case designs and case studies is primarily due to the focus on the individual, not the procedural or design strategies or the kinds of inferences that each type of study supports.

Types of Independent Variables

The range of independent variables manipulated in single-case research is enormous, particularly if one surveys research endeavors across such varied disciplines

as psychology, nursing, physical therapy, and occupational therapy. The nature of the independent variable will depend on a number of factors, one being whether the study addresses a basic or an applied issue. For instance, an operant psychologist studying the effects of reinforcer magnitude on lever-pressing in rats might vary the number of food pellets (one vs. three) delivered contingent on responding on two separate levers. The research is probably being conducted to answer a specific question about reinforcement, or perhaps to evaluate a specific theory of reinforcement, not necessarily to solve a problem concerning rodent behavior in the wild.

On the other hand, a therapist providing a specific cognitive intervention for a stroke patient is not only manipulating a different kind of independent variable but is also doing so to bring about meaningful changes in the patient's functional behavior. In fact, in most clinical studies the independent variable is some kind of therapeutic or educational program, perhaps made up of several phases or components. A cognitive intervention for a stroke patient, for example, may include such specific tasks as learning rehearsal strategies, self-monitoring of behavior, and imagery enhancement. This *cognitive intervention*, as an independent variable, is a qualitatively and quantitatively more complex variable than is *reinforcer magnitude*, as used in the laboratory study just described.

This point is important because the scientific strength of any study, whether in the laboratory or the field, depends on the precision with which variables are conceptualized, manipulated, and measured. In applied settings, interventions must consist of thoroughly defined, concretely articulated activities, and the actual delivery of the intervention must be consistent with its verbal description. Indeed, one of the major sources of controversy surrounding therapy outcome research is the problem of defining and standardizing treatment methods (McGlinchey & Dobson, 2003; Nezu & Nezu, 2005; Schlosser, 2002). If one researcher's cognitive restructuring intervention is different from another therapist's cognitive restructuring therapy, then the two studies, even if otherwise well controlled, will yield uninterpretable results, at least with respect to this ambiguously defined independent variable. Moreover, the cumulative nature of science requires that experiments be replicated to establish the reliability or generality of any particular finding. This proves difficult when an independent variable, particularly in the form of a clinical intervention, is either poorly operationalized or is delivered in a manner that bears no relationship to its verbal description. Fortunately, the issue of *treatment integrity* has received a good deal of recent attention, and a resounding alarm has been sounded by several scientists (Carr, Bailey, Carr, & Coggin, 1996; Gresham, 1996; Gresham, Gansle, Noell, Cohen, & Rosenblum, 1993; Moncher & Prinz, 1991; Perepletchikova & Kazdin, 2005; Peterson, Homer, & Wonderlich, 1982). Treatment regimens, however complex, are the independent variables of applied research and are thus likely to be interpreted as causes of any observed improvements in clinical behavior. Consequently,

82 SINGLE-CASE RESEARCH METHODS

the proper identification and consistent manipulation of such interventions are essential to establishing both the scientific and applied validity of any particular research study.

Types of Dependent Variables

The range of variables that serve as dependent measures in behavioral and health science research is truly immense, as one would expect from a collection of such diverse disciplines as psychology, nursing, and physical and occupational therapy. Although defined and measured in different ways, variables in this kind of research usually represent behaviors having important adaptive or functional repercussions for the subjects or participants. The acquisition of simple self-care activities (proper hygiene or getting dressed), though taken for granted by most of us, may reflect an important milestone for a child with developmental disabilities. A home care patient may be required to learn how to monitor and operate an intravenous drip as a part of an ongoing medical regimen. A victim of an industrial accident, having lost a dominant limb, may have to train the nondominant limb to do the work of the lost limb. The following are some of the kinds of dependent variables encountered in applied behavioral and health care settings:

- Words read per minute by a person with visual impairment (Buning & Hanzlik, 1993)
- Weight distribution of an affected limb in a patient suffering from a form of paralysis known as *hemiplegia* (Wu, Huang, Lin, & Chen, 1996)
- Walking speed in meters per second in chronic stroke patients (Kollen, Rietberg, Kwakkel, & Emmelot, 2000)
- Number of observation intervals containing tics in children exhibiting different kinds of motor tics (Woods, Miltenberger, & Lumley, 1996)
- Caffeine intake in milligrams (Foxy & Rubinoff, 1979)
- Frequency of on-topic and off-topic conversational statements by nursing home residents with dementia (Hoerster, Hickey, & Bourgeois, 2001)
- Number of prompted signs given per minute in infants being taught American Sign Language (Thompson, McKerchar, & Dancho, 2004)
- Number of correct football moves (e.g., tackles) as a result of feedback via posted performance charts (Ward & Carnes, 2002)
- Correct pronunciation of Mandarin Chinese characters by college students (Wu & Miller, 2007)

Of course, it is difficult to assess the adaptive nature of a behavior unless it can be observed for a time course sufficient to exhibit change. This is the primary justification for the continuous, or at least repeated, measurement of dependent

variables that characterizes single-case research. Many behaviors, especially those undergoing treatment, exhibit gradual change over time, not abrupt, discrete shifts. We would not, for example, expect an accident victim to transfer a complex manual function from an amputated dominant limb to a nondominant limb in a few days. In fact, a successful response to therapy in such a case may not be noticeable for weeks or even months. Thus, a single measure of the behavior after treatment may tell us very little about the effects of treatment and absolutely nothing about the process itself. Repeated measurement provides not only an “after” treatment picture but also a more refined and informative “during” treatment picture.

Importance of Multiple Measures of Dependent Variables

The treatment of dependent variables in single-case research follows logically from the manner in which the subject matter itself is conceptualized. If, as we have argued, behavior represents a continuous interplay between an organism and its environment, then our measures of behavior must reflect this fact. It makes little sense to observe and measure a phenomenon one time only, in a discrete manner, if one is primarily interested in how the phenomenon changes or unfolds over a period of time. Because behavior exhibits this basic characteristic and because most behavioral and health care research, especially in applied settings, is conducted for the purpose of assessing changes in behavior, then continuous measurement emerges as a necessary strategy. Thus, instead of obtaining a single measure of the dependent variable of interest, single-case researchers utilize observation and measurement tactics that allow for multiple measures of the dependent variable over a period of time.

The process of measuring the dependent variable is frequently not entirely unique to single-case designs. Group researchers do, on occasion, conduct *repeated measures designs* in which subjects in all groups (experimental and control) are observed more than once after manipulation of the independent variable. Nevertheless, the number of dependent variable measures that can be taken when dealing with large numbers of subjects is likely to be constrained by practical considerations. For this reason, one seldom encounters a group study in which the dependent variable was measured more than two or three times following independent-variable manipulation. Single-case designs, on the other hand, ordinarily involve many more measures of the dependent variable. Indeed, the ideal strategy would be to continuously measure behavior throughout the entirety of the study. This is of course not always possible, particularly in applied settings, but the logic of single-case research quite clearly places a premium on collecting repeated measures of behavior.

The single-case reliance on multiple measures of the dependent variable extends not only to postintervention phases but also to behavioral measurement prior to the intervention. This feature also distinguishes single-case designs from group designs

84 SINGLE-CASE RESEARCH METHODS

that utilize some degree of repeated measures strategy. The purpose of obtaining multiple measures of the dependent variable prior to independent variable implementation is to establish a benchmark against which each subject's behavior during treatment can be evaluated. This benchmark is referred to as a **baseline**, and it represents a critical strategic feature of single-case designs. In essence, the subject's level and pattern of behavior during the baseline, or nontreatment, phase of the study serve as a comparison phase for that same subject's behavior during treatment. In this way, the subject in this type of study serves in both the control and experimental conditions, thus providing the same kind of rationale for a comparison seen in group designs that use separate control and experimental groups.

INTERIM SUMMARY

Single-case research represents an experimental approach to studying behavior in which independent variables, often in the form of clinical treatments, are directly controlled by the researcher. Dependent variables usually consist of behaviors deemed clinically or socially relevant to the subject, as identified by the relevant health science discipline (e.g., psychology, nursing, physical therapy). A key feature of single-case research is the repeated observation and measurement of the dependent variable both prior to and during or after the independent variable or intervention condition. Such repeated observation and measurement are viewed as necessary to tracking and effectively interpreting behavior change in individual subjects or clients.

DATA PRESENTATION IN SINGLE-CASE RESEARCH

In many ways, data serve the same function in single-case research as they do in traditional group designs. In applied health care settings, behavioral and/or medical data collected from subjects are used to draw conclusions about the relative effectiveness of a clinical intervention. As in group designs, data collected under different independent-variable conditions are compared, and such comparisons form the basis of conclusions drawn by the researcher. However, unlike group designs, single-case research designs involve no aggregating or summing of data across multiple research subjects, and the resulting data analysis usually does not consist of evaluating group means through tests of statistical significance. Data collected from individual subjects over prolonged periods of time are importantly different from single, discrete measures summed across many subjects; consequently, the collection, presentation, and analysis of single-case data differ markedly from the conventions of group designs.

We have made reference on occasion to the fact that single-case research tends to resemble more the methodological practices of the natural sciences than those of the behavioral and social sciences. This fact is evident as well in the manner in which single-case researchers present data collected during the course of a study. Because single-case researchers have no group data to present, the presentation of means or other measures of central tendency are not as paramount as in group designs. However, single-case researchers do routinely report central tendency measures, particularly the mean, as one descriptive statistic, to help illuminate an individual subject's performance within particular phases of a study, for example, the mean number of vocalizations during baseline and the mean number of vocalizations during intervention. What the single-case researcher has, on the other hand, are large amounts of data collected from an individual subject or participant, representing appropriate dimensions of the behavior over time and in response to various experimental conditions. Data are presented to describe individual performance for reasons that have already been detailed, most notably the fact that aggregating data across subjects usually misrepresents or obscures the behavior of individuals, and professionals providing health care are, almost without exception, interested in the responses of individuals to clinical interventions.

THE REAL TIME OR TIME-SERIES GRAPH

The prevalent vehicle and long-standing convention for presenting single-case data are the *real-time graph*, also referred to as a *time-series* or *trend graph*, and sometimes simply as a *line graph*. Tufte (2001) reported that 75% of all graphs published in journal articles and magazines are time series graphs, and Henry (1995) noted that such graphs were used to portray political and economic trends, such as importing and exporting, as early as the 1700s. Given the ingenuity of the researcher, the amount and type of information that can be conveyed in a time-series graph is actually quite astounding, as attested to by Figure 4.1, a graph depicting New York City's weather for the year 1980.

Graphing Conventions

Researchers benefit from practices of data presentation and analysis characterized by some degree of standardization, and although no such universal regulations govern graphing, some effort has been made to bring consistency to this process. Among the researchers whose efforts frequently result in graphed data, applied behavior analysts have exhibited an especially steadfast devotion to depicting behavior change in single-case data. Indeed, the *Journal of Applied*

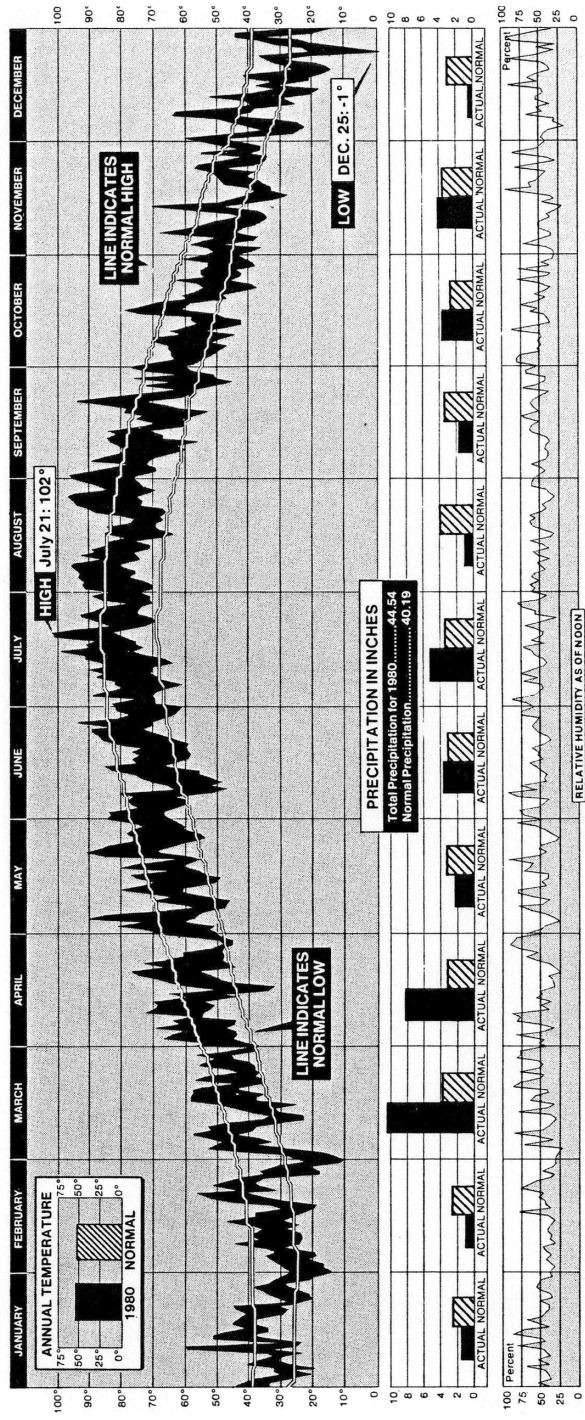


Figure 4.1 Time-series graph of New York's weather for 1980

Source: Tufte, E. R. (2001). *The Visual Display of Quantitative Information*, 2nd ed. Cheshire, CT: Graphics Press. Reprinted with permission.

Behavior Analysis (JABA) has been the primary outlet for research in this field since 1968, and it may be unique in the behavioral sciences in its commitment to publishing single-case data generated in applied settings. Although representing various dimensions of behavior observed and recorded under differing conditions, all graphs published in *JABA* are expected to meet the following guidelines:

- Individual data points represent the relevant dimension of behavior (rate, percentage correct, etc.) for the subject across all experimental conditions.
- Condition changes (nontreatment/baseline to treatment) are denoted by a vertical dotted line.
- The zero level on the y-axis (denoting the dependent variable) is raised above the horizontal line to allow for less confusing visual assessment.
- Data representing different experimental conditions are labeled with a descriptor centered above the data.
- A break in the y-axis is used to indicate that the axis scale is not continuous.
- Subject information (e.g., pseudonyms) is placed in a box in the graph's lower right-hand corner.
- All relevant information is placed within the boundaries of the graph.

These graphing conventions ensure that some degree of uniformity exists in data presented in *JABA*, even though the data may represent very different kinds of behaviors, measured in different dimensions, and across similarly varying independent variable conditions. The flexibility of time-series graphs allows for an accommodation of an extraordinarily wide assortment of dependent variables measured under varying independent variable conditions and, perhaps most important, from very disparate applied fields, such as psychology, nursing, and physical and occupational therapy.

Such graphs have become an indispensable tool for contemporary single-case researchers. An example of a real-time or time-series graph depicting the behavior of a single person is depicted in Figure 4.2. The data in the graph represent how often a client, John Doe, washes his hands each day over a 2-week period. John experiences a debilitating behavior problem known as obsessive-compulsive disorder, characterized by an obsessive fear of contamination. This fear leads John to wash his hands compulsively, with more frequency than is desirable, even to the point of washing away the natural oils produced by healthy skin, eventually leading to dry, chapped hands, covered with ulcerations. We will use the data from John Doe throughout this chapter to illustrate not only graphing conventions but also how a behavior therapist would go about evaluating the effects of a clinical intervention for this client.

88 SINGLE-CASE RESEARCH METHODS

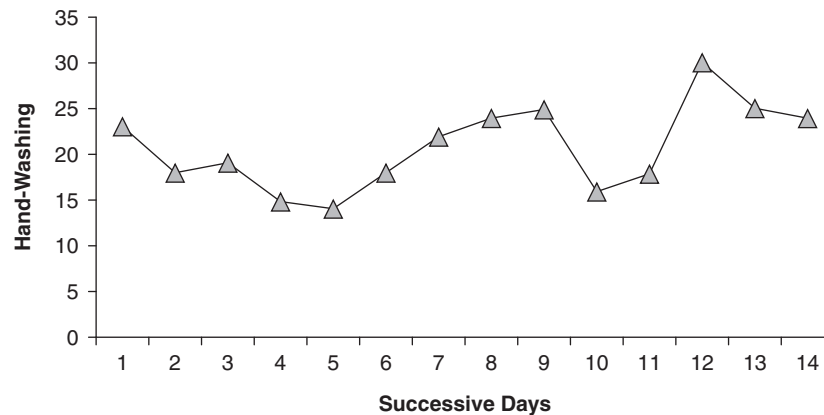


Figure 4.2 John Doe's hand-washing behavior

You'll notice that the graph in Figure 4.2 shares many features with other graphs you may encounter in textbooks, magazines, newspapers, and other sources. Two lines running perpendicular to one another meet or intersect at a point called the *origin* (lower left-hand corner). These lines, ordinarily referred to as *axes*, are labeled to represent the relevant quantitative and/or qualitative variables of the study. Although deviations from this convention are sometimes warranted, such graphs are ordinarily depicted in an aspect ratio of 1:2, meaning the horizontal axis—that is, the *x*-axis—tends to be twice as long as the vertical axis, the *y*-axis. In the behavioral and health sciences, graphs are usually presented to depict some aspect of the subject's behavior, recorded repeatedly over time, both prior to and during the clinical intervention or independent-variable manipulation. Ordinarily, the dimension of behavior (e.g., response rate, percentage correct responses, intensity) is represented on the vertical axis, and some dimension of time is represented on the horizontal axis. Of course, the passage of time is not itself a variable influencing behavior, but any intervention must necessarily occur over time, and only by representing behavior in this manner can we entertain questions about behavior change in relation to an intervention or independent variable. Also, the horizontal axis and the particular dimension of time that it depicts further illustrate the importance of repeated, or continuous, measurement which is a hallmark of single-case research. Repeated measures across time allow for a much more fine-tuned tracking and assessment of behavior change than discrete measures for groups, often collected at one or two points in time, giving single-case data collection and presentation a descriptive power unprecedented in conventional behavioral research.

Importance of Scale Decisions

As is always the case with graphing conventions, researchers must make important decisions about how to represent both independent and dependent variables on the graph. Perhaps the most important decision is the choice of how variables will be scaled on each axis. If it is true that a picture is worth a thousand words, then single-case researchers are in the business of “talking” with graphs. Thus, the researcher has the responsibility of presenting data in a manner that accurately portrays the behavior of interest and increases the likelihood of proper conclusions concerning the effectiveness of the intervention.

Figures 4.3 and 4.4 demonstrate how graphing conventions may result in misleading interpretations. The data representing John Doe’s hand-washing rate are identical but presented on graphs using different scales on the vertical axis. Despite the identical nature of the data, Figure 4.3 would seem to provoke the conclusion that John’s hand washing is fairly stable, exhibiting little variability from day to day. Figure 4.4, however, seems to suggest a good deal more variability in this behavior. It is important to be aware of how scale axes will affect interpretation, but at the same time, there are no hard-and-fast rules concerning which scale to use for a particular graph. Clearly, the scale needs to be able to depict the full range of the dependent variable. At the same time, the scale should be sensitive enough to render an accurate picture of trends and variability in the behavior. In the present example, Figure 4.3 uses a y-axis that exceeds the highest frequency of hand washing observed by nearly three times, so that most of the data are “squeezed” into the lower third of the graph. Consequently, this graph wastes

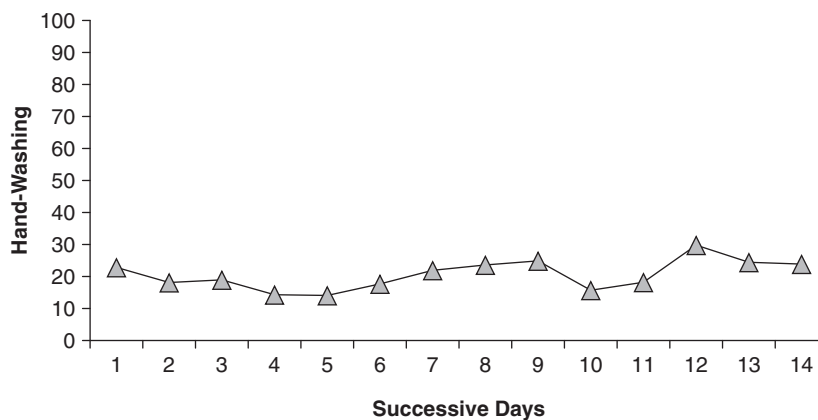


Figure 4.3 John Doe’s hand-washing behavior

90 SINGLE-CASE RESEARCH METHODS

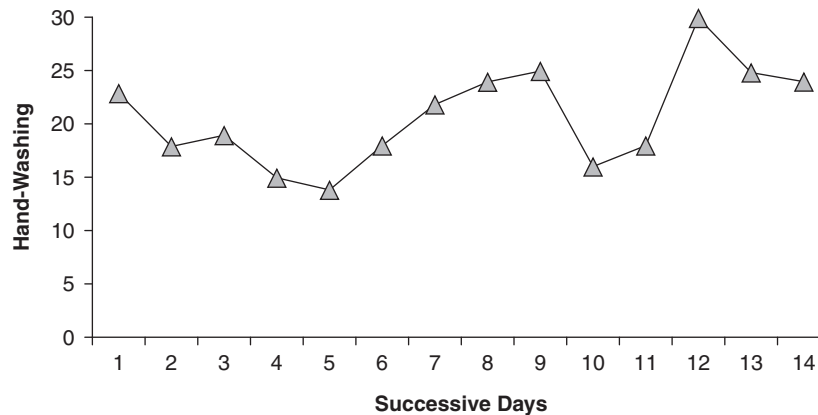


Figure 4.4 John Doe's hand-washing behavior

space and probably underrepresents the degree of variability occurring in John's hand washing over the 2-week period. The *y*-axis in graph 4.4, on the other hand, both utilizes space more effectively and demonstrates the considerable day-to-day variability in the dependent variable.

Of course, there is more information to be gleaned from the *y*-axis of a graph than just the measurement scale. The text found running alongside the axis is important because it identifies not only the behavior or dependent variable that was of interest to the researcher but also the relevant dimensions along which this variable was measured. In fact, the nature of the dependent variable and its measurable dimensions is often the most distinguishing factor about a study and what separates one study from the next, both within and across research disciplines.

In most published accounts of clinical interventions, single-case graphs are a bit busier than those depicted in Figures 4.3 and 4.4. This is because the function of such graphs is to show the effects of the intervention on the relevant target behavior. To show the effects of the intervention, some kind of comparison of behavior both prior to and during or after the intervention is necessary. An example of how this is done can be seen in Figure 4.5, which depicts our client with obsessive-compulsive disorder (John Doe) once again. The major difference between Figure 4.5 and the earlier figures (4.3 and 4.4), however, is that now we have an opportunity to compare John's frequency of hand washing prior to and during treatment. We will assume, for current purposes, that John has received some version of cognitive-behavior therapy, a fairly standard intervention for obsessive-compulsive disorder.

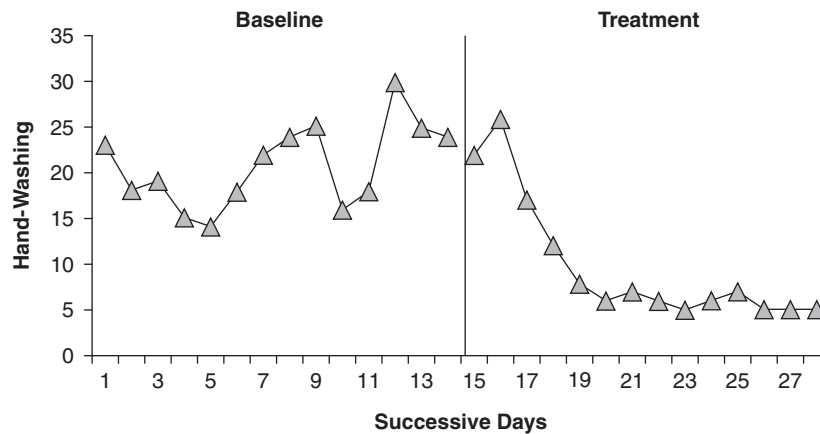


Figure 4.5 John Doe's hand washing before and after treatment

Figure 4.5 depicts John's hand-washing episodes both prior to the clinical intervention (cognitive-behavior therapy) and during the intervention. These two phases of the study (before and during intervention) are separated on the graph by a vertical line, and each phase is also identified by text at the top of the graph. The data represented on the left side of the graph are referred to as *baseline* data, because they represent the level of the target behavior before any clinical intervention has been delivered. It is the comparison of these data in the baseline phase with the data subsequently presented on the right side of the vertical line—that is, data in the treatment or intervention phase—that allows us to draw conclusions about the effects of the intervention. This baseline-treatment comparison is logically similar to the control group-experimental group comparison with which you are probably already familiar in group studies. In single-case research, each subject serves as its (or his or her) own control, with behavior being measured both prior to and after introducing the independent variable (clinical intervention).

The data in Figure 4.5 provide the critical information needed for the researcher to draw conclusions about the effects of the independent variable (intervention) on the dependent variable (target behavior) in a single-case study. However, the details of the research design and the criteria for drawing decisions about intervention effectiveness are themselves rather complex issues. We discuss the different kinds of research designs available to single-case researchers, as well as the decision criteria for evaluating intervention effectiveness, in subsequent chapters.

THE STANDARD CELERATION CHART

The real-time or time-series graph is a convenient and flexible device for depicting continuous behavioral data. Such graphs are able to accommodate dependent variables across numerous behavioral and health science disciplines and measured along varying dimensions. However, displaying data in this way requires considerable nuance when the measured dependent variable varies across a sizable number of values. For instance, we might be interested in assessing the rate at which a client experiences undesirable brief facial movements, or tics. Because such movements may occur quite frequently under conditions of anxiety but with very little frequency under less anxious conditions, measures of this dependent variable might vary from dozens of times a minute to only a few occurrences an hour. In fact, the sheer variability of this behavior might lead the researcher to question the appropriate response rate cycle in which to measure the behavior: minutes or hours. Thus in depicting the behavior graphically, the researcher must decide on a scale axis that both accommodates the substantial range of the behavior while also being sensitive enough to demonstrate small amounts of variability that may be of some clinical importance.

Fortunately, data that evidence substantial variability even more than exhibited in Figures 4.3 and 4.4 can be adequately portrayed using the Standard Celeration Chart (see Figure 4.6). Ogden Lindsley and colleagues developed this graphing format in the 1950s and 1960s, primarily within the context of educational research. Lindsley had been a graduate student of B. F. Skinner's during the 1950s and consequently had developed an appreciation for collecting and interpreting data at the level of the individual subject. Lindsley was, in fact, among the first researchers to employ laboratory-based principles of behavior in an applied setting. His pioneering efforts to utilize fundamental conditioning principles to change behavioral symptoms in institutionalized psychiatric patients (Lindsley, 1956, 1960) were early contributions to the development of behavior therapy and behavior modification (Kazdin, 1982a).

By the 1960s, Lindsley had become interested in the application of behavioral interventions to educational practice. His background in experimental methodology and the study of operant behavior led Lindsley to question not only the instructional methods that characterized American schools at the time but also the methods for evaluating their relative effectiveness. Lindsley argued, not surprisingly, given his scientific training, that assessment of learning had to take place at the level of the individual student and that the development of objective and standard measures of learning was a necessary first step in this process. Lindsley and his colleagues suggested that the rate at which an academic behavior occurs (e.g., number of words read per minute) represents not only a universal dimension of

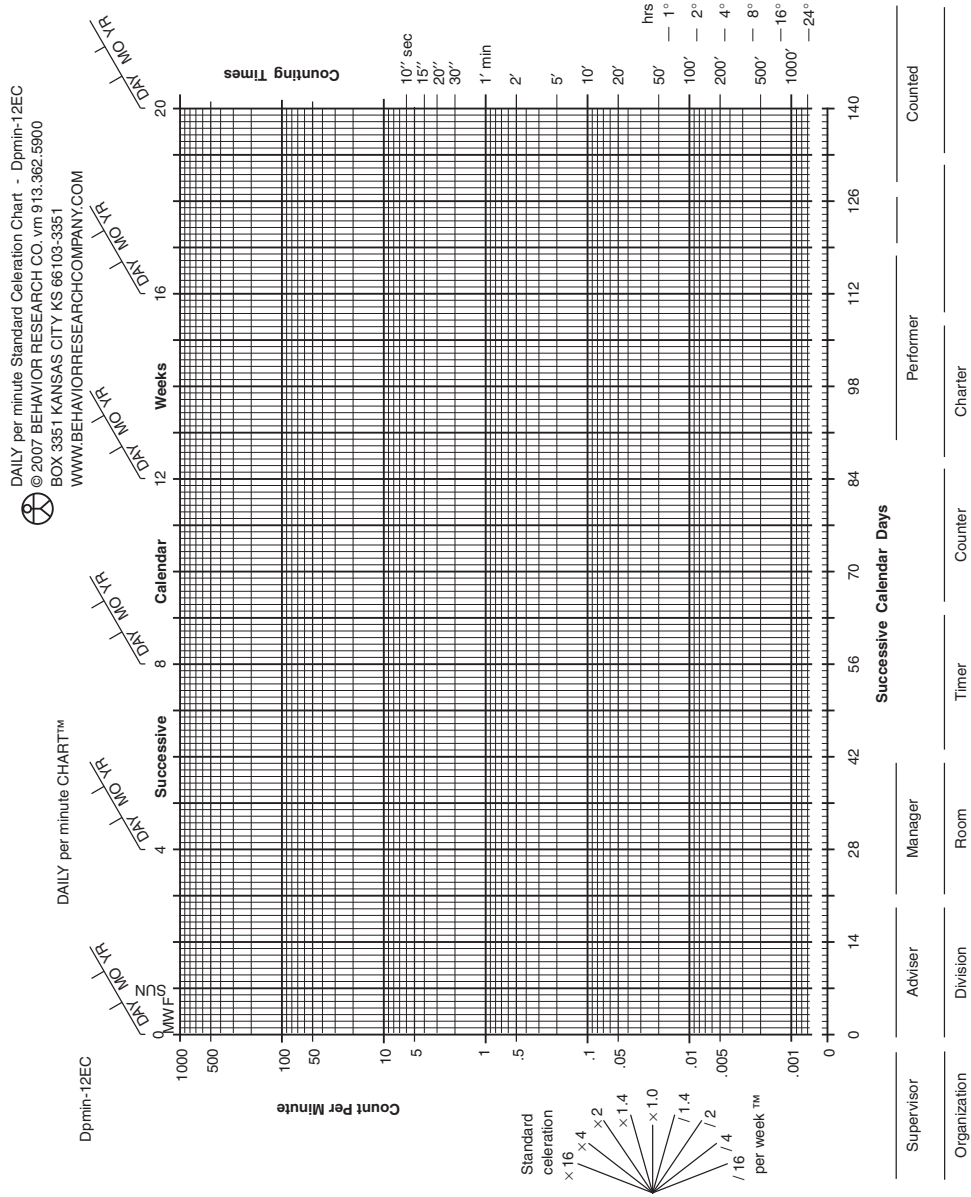


Figure 4.6 The Standard Celeration Chart

94 SINGLE-CASE RESEARCH METHODS

behavior but also a highly sensitive metric of skill development. This may have seemed an odd idea at the time because many educators then, and even now, tended to favor percentage of correct responses as the best indication of academic progress. One problem with percentage correct as a dependent measure is that once a skill has reached a certain level of proficiency, a percentage measure can no longer depict continued progress. For example, two children who both answer 10 out of 10 math problems correctly would be considered equivalent in amount of learning, because the percentage of correct responses, 100%, is the same for both students. However, one child may be able to answer these 10 math problems correctly in less than 1 minute, whereas the other child may require 3 minutes to achieve this same percentage level of correct responding. A percentage-correct measure fails to depict the differences in *fluency* (correct responses per unit of time) that characterize the two students' performances (Binder, 1996). Consequently, Lindsley and his colleagues chose response rate as their primary dimension for measuring behavior change in academic settings. As we saw in chapter 3, rate is an important dimension of many kinds of behavior; indeed, the objective of many clinical or applied interventions is to bring about marked increases or decreases in the frequency of a target behavior.

Over the course of several years, Lindsley and his colleagues experimented on various graphing strategies for depicting academic learning. Having observed that academic progress often occurs rapidly (at least under optimal instructional conditions) and that many types of behavior change occur proportionally, not arithmetically, these researchers opted against using a standard equal interval or add-subtract graph to depict learning; instead, they spent several years developing a graph that would more properly display the proportional or multiplicative changes characteristic of behavioral processes. The result of these efforts is the *Standard Celeration Chart*, an example of which is shown in Figure 4.6.

Although the chart shares many properties of any Cartesian graph, including *x*- and *y*-axes, the Standard Celeration Chart utilizes a logarithmic progression rather than an arithmetic progression to depict behavior change on the *y*-axis. Notice that in Figure 4.6 tick marks on the *y*-axis represent multiples of 10, ranging from .001 to 1,000. Distances between multiples (e.g., 1 to 10, 10 to 100, 100 to 1,000) are equal, indicating proportionally equivalent changes in the dependent variable. On such a graph, a behavior that increases from once per minute to 10 times per minute would demonstrate the same change in slope as would a behavior changing from 10 times a minute to 100 times a minute. On a more conventional equal-interval or add-subtract graph a change from 1 to 10 responses per minute will pale in comparison to a change from 10 to 100 responses per minute. Even though both increases represent the same ratio, or proportion, of behavior change, an equal-interval graph produces a

considerable distortion that may lead researchers to misinterpret the amount of behavior change. Moreover, it would be challenging to represent a range of behavior change from once per minute to 100 times per minute on an equal-interval graph without severely “squeezing” the y-axis in such a way as to deflate visually the apparent variability in the behavior.

Although there are other characteristics of the Standard Celeration Chart that recommend its use to applied professionals, its strongest feature is the extent to which it allows for standardized depiction of behavior change in individual subjects. The chart allows researchers and clinicians to monitor subjects’ progress on a continuous basis and to follow changes in trend, such as target behaviors that accelerate and decelerate, hence the name Standard Celeration Chart. The Standard Celeration Chart has become a vital component of educational assessment among a growing number of teachers who utilize the principles of precision teaching developed by Ogden Lindsley and his students in the 1960s and 1970s. Relying heavily on the basic behavior principles studied by behavior analysts for many decades, precision teaching represents a decidedly science-based instructional strategy characterized by objective measurement, continuous behavioral monitoring, and development and modification of instructional practices informed by empirical data. A major source of data in precision teaching are the Standard Celeration Charts maintained by both teachers and students alike in precision teaching classrooms. In response to the question “What does our Standard Celeration Chart do?” Ogden Lindsley offered the following answer:

It simplifies things. It simplifies charting so that six year olds can learn it and teach it to others. It simplifies chart reading, making it so fast that we can share charts at 2 minutes each. It simplifies chart checking so much that you can check for x2 learning on 60 charts posted on a ten foot stretch of wall as you walk past without slowing your pace. It simplifies understanding of all growth and decay. (See <http://www.celeration.org/faqs>)

The Standard Celeration Chart has for many years been enthusiastically endorsed and used by precision teaching educators, but it has not been widely adopted by most researchers and clinicians in the applied health sciences, or by most schoolteachers. This reluctance may be due to many people’s lack of familiarity with measuring proportional changes in dependent variables and the corresponding logarithmic scale used by the Standard Celeration Chart. Interested readers are encouraged to contact the Standard Celeration Society (<http://www.celeration.org>) to learn more about the history of the chart’s development, charting tips, and the many applications of the chart to studies of behavior change.

INTERIM SUMMARY

The typical manner of presenting single-case data for purposes of interpretation is as continuous behavior on a time-series or real-time graph. This form of display takes full advantage of the continuous nature of single-case data collection and allows for meaningful interpretation of behavior change in response to independent-variable manipulation. Interpretation of behavior change is influenced by a number of graphing strategies, perhaps most important being the scale dimension chosen to depict the dependent variable. In an effort to standardize both display and interpretation of single-case data, Ogden Lindsley and his colleagues developed the Standard Celeration Chart, which uses a logarithmic (multiplicative) rather than traditional additive scale for plotting behavior on the y-axis. This multiplicative scale not only allows for a standardized method of depicting proportional change in behavior but also makes it possible to portray behaviors that may show excessive variability over time. In the next chapter, we begin discussing the specific kinds of design strategies used by single-case researchers to evaluate data collected during actual applied studies.

KEY TERMS GLOSSARY

Idiographic research Research that utilizes observation and measurement at the level of the individual subject with no intention of establishing general laws.

Nomothetic research Research that collects and aggregates data across subjects for the purpose of establishing general laws.

Inductive strategy Research in which general principles are derived from specific observations.

Deductive strategy Research in which specific observations or hypotheses are derived from general principles or theories.

Intrasubject replication Research in which replications of experimental conditions occur at the level of the individual subject.

Intersubject replication Research in which replications of experimental conditions occur across two or more subjects.

External validity The extent to which the results of a particular study are applicable or generalizable to other populations or settings.

Behavior An adaptive, ongoing interplay between an organism and its immediate environment.

Baseline A period of observation of the target behavior in its natural state prior to manipulation of the independent variable or intervention.

SUPPLEMENTS

Review Questions

1. What is the difference between an idiographic and a nomothetic approach to studying behavior? Which one of these approaches is endorsed by single-subject researchers, and why?
2. What is the difference between intrasubject and intersubject replication? Which kind of replication study would be conducted if a researcher were interested in discovering generally applicable principles of behavior?
3. Suppose a physical therapist is conducting an intervention study in which a power-walking regimen is used to try and build leg strength in a patient recovering from leg surgery. In this study, what are the independent and dependent variables, respectively?
4. What is a time-series graph, and how does it help exhibit the continuous nature of behavioral data? Also, on which axis is time usually represented on such a graph?
5. What is the primary difference between conventional time-series graphs and the Standard Celeration Chart? Why did Lindsley and his colleagues recommend use of the Standard Celeration Chart?

SUGGESTED READINGS/HELPFUL WEB SITES

Tufte, E. R. (2001). *The visual display of quantitative information* (2nd ed.). Cheshire, CT: Graphics Press.

Edward Tufte has been called the “Galileo” of graphic presentation. In several elegant books, including this one, Tufte has described the powerful and flexible use of graphics to portray evidence, as well as the common mistakes made by scientists and others in the presentation of data. Tufte’s books are both extremely attractive and fascinating explorations of the use of imagery in conveying information.

98 SINGLE-CASE RESEARCH METHODS

<http://www.edwardtufte.com>

This is Edward Tufte's Web site. You can purchase any of Tufte's highly acclaimed books on graphic display, read articles by Tufte about display strategies, including the problems with PowerPoint, and even ask Tufte questions about display issues.

<http://www.celeration.org>

This is the Web site of the Standard Celeration Society, which publishes the Standard Celeration Chart and data gathered by scientists and educators using the chart. There are links to articles describing the history of the Standard Celeration Chart, its advantages over equal-interval charts for behavioral research, and examples of the chart used in applied settings, especially by precision teachers.

<http://seab.envmed.rochester.edu/jaba/>

This is the Web site for the *Journal of Applied Behavior Analysis*, a scientific journal devoted to research involving basic behavioral interventions in applied settings, including schools, workplaces, and the home. *Journal of Applied Behavior Analysis* researchers are devoted to the single-case method of data collection, analysis, and display. The journal is an especially good resource for individuals who want to learn how to display single-case data in graphic form.