CHAPTER 7

Analyzing Behavior Change: Basic Assumptions and Strategies

Key Terms

A-B design
affirmation of the consequent
ascending baseline
baseline
baseline logic
confounding variable
dependent variable
descending baseline
experimental control
experimental design
experimental question
external validity
extraneous variable
independent variable
internal validity
parametric analysis
practice effects
prediction
replication
single-subject designs
stable baseline
steady state responding
steady state strategy
variable baseline
verification

Behavior Analyst Certification Board® BCBA® & BCABA® Behavior Analyst Task List©, Third Edition

Content Area 3: Principles, Processes, and Concepts	
3–10	Define and provide examples of functional relations.
Content Area 5: Experimental Evaluation of Interventions	
5–1	Systematically manipulate independent variables to analyze their effects on treatment.
5–2	Identify and address practical and ethical considerations in using various experimental designs.
5–4	Conduct a parametric analysis (i.e., determining effective parametric values of consequences, such as duration or magnitude).

Measurement can show whether, when, and how much behavior has changed, but measurement alone cannot reveal why, or more accurately *how,* behavior change occurred. A useful technology of behavior change requires an understanding of the specific arrangements of environmental variables that will produce desired behavior change. Without this knowledge, efforts to change behavior could only be considered one-shot affairs consisting of procedures selected randomly from a bag of tricks with little or no generality from one situation to the next.

The search for and demonstration of functional and reliable relations between socially important behavior and its controlling variables is a defining characteristic of applied behavior analysis. A major strength of applied behavior analysis is its insistence on experimentation as its method of proof, which enables and demands an ongoing, self-correcting search for effectiveness.

> Our technology of behavior change is also a technology of behavior measurement and of experimental design; it developed as that package, and as long as it stays in that package, it is a self-evaluating enterprise. Its successes are successes of known magnitude; its failures are almost immediately detected as failures; and whatever its outcomes, they are attributable to known inputs and procedures rather than to chance events or coincidences. (D. M. Baer, personal communication, October 21, 1982)

An experimental analysis must be accomplished to determine if and how a given behavior functions in relation to specific changes in the environment. This chapter introduces the basic concepts and strategies that underlie the *analysis* in applied behavior analysis.[1] The chapter begins with a brief review of some general conceptions of science, followed by a discussion of two defining features and two assumptions about the nature of behavior that dictate the experimental methods most conducive to the subject matter. The chapter then describes the necessary components of any experiment in applied behavior analysis and concludes by explaining the basic logic that guides the experimental methods used by applied behavior analysts.

[1]The analysis of behavior has benefited immensely from two particularly noteworthy contributions to the literature on experimental method: Sidman's *Tactics of Scientific Research* (1960/1988) and Johnston and Pennypacker's *Strategies and Tactics of Human Behavioral Research* (1980, 1993a). Both books are essential reading and working references for any serious student or practitioner of behavior analysis, and we acknowledge the significant part each has played in the preparation of this chapter.

Concepts and Assumptions Underlying the Analysis of Behavior

As was discussed in Chapter 1, scientists share a set of common perspectives that include assumptions about the nature of the phenomena they study (determinism), the kind of information that should be gathered on the phenomena of interest (empiricism), the way questions about the workings of nature are most effectively examined (experimentation), and how the results of experiments should be judged (with parsimony and philosophic doubt). These attitudes apply to all scientific disciplines, including the scientific study of behavior. "The basic characteristics of science are not restricted to any particular subject matter" (Skinner, 1953, p. 11).

The overall goal of science is to achieve an understanding of the phenomena under study—socially significant behavior, in the case of applied behavior analysis. Science enables various degrees of understanding at three levels: description, prediction, and control. First, systematic observation enhances the understanding of natural phenomena by enabling scientists to describe them accurately. Descriptive knowledge of this type yields a collection of facts about the observed events—facts that can be quantified and classified, a necessary and important element of any scientific discipline.

A second level of scientific understanding occurs when repeated observation discovers that two events consistently covary. That is, the occurrence of one event (e.g., marriage) is associated with the occurrence of another event at some reliable degree of probability (e.g., longer life expectancy). The systematic covariation between two events—termed a *correlation*—can be used to predict the probability that one event will occur based on the presence of the other event.

The ability to predict successfully is a useful result of science; prediction allows preparation. However, the greatest potential benefits of science are derived from the third, and highest, level of scientific understanding, which comes from establishing experimental control. "The experimental method is a method for isolating the relevant variables within a pattern of events. Methods that depend merely on observed correlations, without experimental intervention, are inherently ambiguous" (Dinsmoor, 2003, p. 152).

Experimental Control: The Path to and Goal of Behavior Analysis

Behavior is the interaction between an organism and its environment and is best analyzed by measuring changes in behavior that result from imposed variations on the environment. This statement embodies the general strategy and the goal of behavioral research: to demonstrate that measured changes in the target behavior occur because of experimentally manipulated changes in the environment.

Experimental control is achieved when a predictable change in behavior (the dependent variable) can be reliably produced by the systematic manipulation of some aspect of the person's environment (the independent variable). Experimentally determining the effects of environmental manipulation on behavior and demonstrating that those effects can be reliably produced constitute the *analysis* in applied behavior analysis. An analysis of a behavior has been achieved when a reliable functional relation between the behavior and some specified aspect of the environment has been demonstrated convincingly. Knowledge of functional relations enables the behavior analyst to reliably alter behavior in meaningful ways.

An analysis of behavior "requires a believable demonstration of the events that can be responsible for the occurrence or nonoccurrence of that behavior. An experimenter has achieved an analysis of a behavior when he can exercise control over it" (Baer, Wolf, & Risley, 1968, p. 94).[2] Baer and colleague's original definition of analysis highlights an important point. Behavior analysis' seeking and valuing of the experimental isolation of a given environmental variable of which a behavior is shown to be a function has often been misinterpreted as support for a simplistic conception of the causes of behavior. The fact that a behavior varies as a function of a given variable does not preclude its varying as a function of other variables. Thus, Baer and colleagues described an experimental analysis as a convincing demonstration that a variable *can* be responsible for the observed behavior change. Even though a complete analysis (i.e., understanding) of a behavior has not been achieved until all of its multiple causes have been accounted for, an *applied* (i.e., technologically useful) analysis has been accomplished when the investigator has isolated an environmental variable (or group of variables that operate together as a treatment package) that reliably produces socially significant behavior change. An *applied* analysis of behavior also requires that the target behavior be a function of an environmental event that can be practically and ethically manipulated.

Experiments that show convincingly that changes in behavior are a function of the independent variable and are not the result of uncontrolled or unknown variables are said to have a high degree of **internal validity.** A study without internal validity can yield no meaningful statements regarding functional relations between the variables examined in the experiment, nor can it be used as the basis for any statements regarding the generality of the findings to other persons, settings, and/or behaviors.[3]

When initially planning an experiment and later when examining the actual data from an ongoing study, the investigator must always be on the lookout for threats to internal validity. Uncontrolled variables known or suspected to exert an influence on the dependent variable are called **confounding variables.** For example, suppose a researcher wants to analyze the effects of guided lecture notes on high school biology students' learning as measured by their scores on next-day quizzes. One potential confounding variable that the researcher would need to take into account would be each student's changing level of interest in and background knowledge about the specific curriculum content (e.g., a student's high score on a quiz following a lecture on sea life may be due to his prior knowledge about fishing, not the guided notes provided during that lecture).

A primary factor in evaluating the internal validity of an experiment is the extent to which it eliminates or controls the effects of confounding variables while still investigating the research questions of interest. It is impossible to eliminate all sources of uncontrolled variability in an experiment, although the researcher always strives for that ideal. In reality, the goal of experimental design is to eliminate as many uncontrolled variables as possible and to hold constant the influence of all other variables except the independent variable, which is purposefully manipulated to determine its effects.

Behavior: Defining Features and Assumptions that Guide Its Analysis

> Behavior is a difficult subject matter, not because it is inaccessible, but because it is extremely complex. Since it is a process, rather than a thing, it cannot easily be held still for observation. It is changing, fluid, evanescent, and for this reason it makes great technical demands upon the ingenuity and energy of the scientist.
>
> —*B. F. Skinner (1953, p. 15)*

How a science defines its subject matter exerts profound influence and imposes certain constraints on the experimental strategies that will be most effective in an understanding of it. "In order for the scientific study of behavior to be as effective as possible, it is necessary for the methods of the science to accommodate the characteristics of its subject matter" (Johnston & Pennypacker, 1993a,

[2]The researcher's audience ultimately determines whether a claimed functional relation is believable, or convincing. We will explore the believability of research findings further in Chapter 10.

[3]**External validity** commonly refers to the degree to which a study's results are generalizable to other subjects, settings, and/or behaviors. Strategies for assessing and extending the external validity of experimentally demonstrated functional relations are discussed in Chapter 10.

p. 117). The experimental methods of behavior analysis are guided by two defining features of behavior: (a) the fact that behavior is an individual phenomenon and (b) the fact that behavior is a continuous phenomenon; and by two assumptions about its nature: (a) that behavior is determined and (b) that behavioral variability is extrinsic to the organism.

Behavior Is an Individual Phenomenon

If behavior is defined as a person's interaction with the environment, it follows that a science seeking to discover general principles or laws that govern behavior must study the behavior of individuals. Groups of people do not behave; individual people do. Thus, the experimental strategy of behavior analysis is based on within-subject (or single-subject) methods of analysis.

The average performance of a group of individuals is often interesting and useful information and may, depending on the methods by which individuals were selected to be in the group, enable probability statements about the average performance within the larger population represented by the group. However, "group data" provide no information about the behavior of any individual or how any individual might perform in the future. For example, although administrators and taxpayers may be justifiably interested in the average increase in students' reading comprehension from grade level to grade level, such information is of little use to the classroom teacher who must decide how to improve a given student's comprehension skills.

Nonetheless, learning how behavior–environment relations work with many individuals is vital. A science of behavior contributes to a useful technology of behavior change only to the extent that it discovers functional relations with generality across individuals. The issue is how to achieve that generality. Behavior analysts have found that discovery of behavioral principles with generality across persons is best accomplished by replicating the already demonstrated functional relations with additional subjects.

Behavior Is a Dynamic, Continuous Phenomenon

Just as behavior cannot take place in an environmental void (it must happen somewhere), so must behavior occur at particular points in time. Behavior is not a static event; it takes place in and changes over time. Therefore, single measures, or even multiple measures sporadically dispersed over time, cannot provide an adequate description of behavior. Only continuous measurement over time yields a complete record of behavior as it occurs in context with its environmental influences. Because true continuous measurement is seldom feasible in applied settings, the systematic repeated measurement of behavior (as described in Chapters 4 and 5) has become the hallmark of applied behavior analysis.

Behavior Is Determined

As discussed in Chapter 1, all scientists hold the assumption that the universe is a lawful and orderly place and that natural phenomena occur in relation to other natural events.

> The touchstone of all scientific research is order. In the experimental analysis of behavior, the orderliness of relations between environmental variables and the subject's behavior is at once the operating assumption upon which the experimenter proceeds, the observed fact that permits doing so, and the goal that continuously focuses experimental decisions. That is, the experimenter begins with the assumption that the subject's behavior is the result of variables in the environment (as opposed to having no causes at all). (Johnston & Pennypacker, 1993a, p. 238)

In other words, the occurrence of any event is determined by the functional relations it holds with other events. Behavior analysts consider behavior to be a natural phenomenon that, like all natural phenomena, is determined. Although determinism must always remain an assumption—it cannot be proven—it is an assumption with strong empirical support.

> Data gathered from all scientific fields indicate that *determinism* holds throughout nature. It has become clear that the *law of determinism,* that is, that all things are determined, holds for the behavioral area also. . . . When looking at actual behavior we've found that in situation 1, behavior is caused; in situation 2, behavior is caused; in situation 3, behavior is caused; . . . and in situation 1001, behavior is caused. Every time an experimenter introduces an independent variable that produces some behavior or some change in behavior, we have further *empirical* evidence that behavior is caused or deterministic. (Malott, General, & Snapper, 1973, pp. 170, 175)

Behavioral Variability Is Extrinsic to the Organism

When all conditions during a given phase of an experiment are held constant and repeated measures of the behavior result in a great deal of "bounce" in the data (i.e., the subject is not responding in a consistent fashion), the behavior is said to display variability.

The experimental approach most commonly used in psychology and other social and behavioral sciences (e.g., education, sociology, political science) makes two assumptions about such variability: (a) Behavioral variability is an intrinsic characteristic of the organism, and

(b) behavioral variability is distributed randomly among individuals in any given population. These two assumptions have critical methodological implications: (a) Attempting to experimentally control or investigate variability is a waste of time—it simply exists, it's a given; and (b) by averaging the performance of individual subjects within large groups, the random nature of variability can be statistically controlled or canceled out. Both of these assumptions about variability are likely false (empirical evidence points in the opposite direction), and the methods they encourage are detrimental to a science of behavior. "Variables are not canceled statistically. They are simply buried so their effects are not seen" (Sidman, 1960/1988, p. 162).[4]

Behavior analysts approach variability in their data quite differently. A fundamental assumption underlying the design and guiding the conduct of experiments in behavior analysis is that, rather than being an intrinsic characteristic of the organism, behavioral variability is the result of environmental influence: the independent variable with which the investigator seeks to produce change, some uncontrolled aspect of the experiment itself, and/or an uncontrolled or unknown factor outside of the experiment.

The assumption of extrinsic variability yields the following methodological implication: Instead of averaging the performance of many subjects in an attempt to mask variability (and as a result forfeiting the opportunity to understand and control it), the behavior analyst experimentally manipulates factors suspected of causing the variability. Searching for the causal factors contributes to the understanding of behavior, because experimental demonstration of a source of variability implies experimental control and thus another functional relation. In fact, "tracking down these answers may even turn out to be more rewarding than answering the original experimental question" (Johnston & Pennypacker, 1980 p. 226).

From a purely scientific viewpoint, experimentally tracking down sources of variability is always the preferred approach. However, the applied behavior analyst, with a problem to solve, must often take variability as it presents itself (Sidman, 1960/1988). Sometimes the applied researcher has neither the time nor the resources to experimentally manipulate even suspected and likely sources of variability (e.g., a teacher who interacts with a student for only part of the day has no hope of controlling the many variables outside the classroom). In most settings the applied behavior analyst seeks a treatment variable robust enough to overcome the variability induced by uncontrolled variables and produce the desired effects on the target behavior (Baer, 1977b).

Components of Experiments in Applied Behavior Analysis

> Nature to be commanded must be obeyed. . . . But, that coin has another face. Once obeyed, nature can be commanded.
>
> —*B. F. Skinner (1956, p. 232)*

Experimentation is the scientist's way of discovering nature's rules. Discoveries that prove valid and reliable can contribute to a technology of effective behavior change. All experiments in applied behavior analysis include these essential components:

- At least one participant (subject)
- At least one behavior (dependent variable)
- At least one setting
- A system for measuring the behavior and ongoing visual analysis of the data
- At least one treatment or intervention condition (independent variable)
- Manipulations of the independent variable so that its effects on the dependent variable, if any, can be detected (experimental design)

Because the reason for conducting any experiment is to learn something from nature, a well-planned experiment begins with a specific question for nature.

Experimental Question

> We conduct experiments to find out something we do not know.
>
> —*Murray Sidman (1960/1988, p. 214)*

For the applied behavior analyst, Sidman's "something we do not know" is cast in the form of a question about the existence and/or specific nature of a functional relation between meaningful improvement in socially significant behavior and one or more of its controlling variables. An **experimental question** is "a brief but specific statement of what the researcher wants to learn from conducting the experiment" (Johnston & Pennypacker (1993b, p. 366). In published reports of applied behavior analysis studies, the experimental (or research) question is sometimes stated explicitly in the form of a question, as in these examples:

- Which method of self-correction, after attempting each of 10 words or after attempting a list of 10

[4]Some investigators use group comparison designs not just to cancel randomly distributed variability but also to produce results that they believe will have more external validity. Group comparison and within subject experimental methods are compared in Chapter 10.

words, will produce better effects on (a) the acquisition of new spelling words as measured by end-of-the-week tests, and (b) the maintenance of practiced spelling words as measured by 1-week maintenance tests by elementary school students with learning disabilities? (Morton, Heward, & Alber, 1998)

- What are the effects of training middle school students with learning disabilities to recruit teacher attention in the special education classroom on (a) the number of recruiting responses they emit in the general education classroom, (b) the number of teacher praise statements received by the students in the general education classroom, (c) the number of instructional feedback statements received by the students in the general education classroom, and (d) the students' academic productivity and accuracy in the general education classroom? (Alber, Heward, & Hippler, 1999, p. 255)

More often, however, the research question examined by the experiment is implicit within a statement of the study's purpose. For example:

- The purpose of the present study was to compare the relative effectiveness of nonremoval of the spoon and physical guidance as treatments for food refusal and to assess the occurrence of corollary behaviors produced by each procedure. (Ahearn, Kerwin, Eicher, Shantz, & Swearingin, 1996, p. 322)
- The present study was conducted to determine if habit reversal is effective in treating verbal tics in children with Tourette syndrome. (Woods, Twohig, Flessner, & Roloff, 2003, p. 109)
- The purpose of this study was to determine if observed SIB during the tangible condition was confounded by the simultaneous delivery of therapist attention. (Moore, Mueller, Dubard, Roberts, & Sterling-Turner, 2002, p. 283)
- The purpose of this study was to determine whether naturally occurring meals would affect performance adversely during postmeal sessions in which highly preferred food was used as reinforcement. (Zhou, Iwata, & Shore, 2002, pp. 411–412)

Whether an experimental question is stated explicitly in the form of a question or implicit within a statement of purpose, all aspects of an experiment's design and conduct should follow from it.

> A good design is one that answers the question convincingly, and as such needs to be constructed in reaction to the question and then tested through arguments in that context (sometimes called, "thinking through"), rather than imitated from a textbook. (Baer, Wolf, & Risley, 1987, p. 319)

Subject

Experiments in applied behavior analysis are most often referred to as **single-subject** (or *single-case*) **designs.** This is not because behavior analysis studies are necessarily conducted with only one subject (though some are), but because the experimental logic or reasoning for analyzing behavior changes often employs the subject as her own control.[5] In other words, repeated measures of each subject's behavior are obtained as she is exposed to each condition of the study (e.g., the presence and absence of the independent variable). A subject is often exposed to each condition several times over the course of an experiment. Measures of the subject's behavior during each phase of the study provide the basis for comparing the effects of experimental variables as they are presented or withdrawn in subsequent conditions.

Although most applied behavior analysis studies involve more than one subject (four to eight is common), each subject's data are graphed and analyzed separately.[6] Instead of using *single-subject design* to refer to the experiments in which each subject serves as his or her own control, some authors use more aptly descriptive terms such as *within-subject design* or *intra-subject design.*

Sometimes the behavior analyst is interested in assessing the total effect of a treatment variable within a group of subjects—for example, the number of homework assignments completed by members of a class of

[5]It has become commonplace in the behavior analysis literature to refer to the person(s) whose behavior is the dependent variable in an experiment as a *participant,* instead of the more traditional term, *subject.* We use both terms in this text and urge readers to consider Sidman's (2002) perspective on the issue: "[W]e are no longer permitted to call our subjects 'subjects.' The term is supposed to be dehumanizing, and so we are supposed to call them 'participants.' I think this is completely misguided. Experimenters, too, are participants in their experiments. What does making them nonparticipants do to our perception of science and of scientists? Are experimenters merely robots who follow prescribed and unbreakable scientific rules? Are they supposed just to manipulate variables and coldly record the results of their manipulations? Separating them as nonparticipating manipulators and recorders of the behavior of participants really dehumanizes not only experimenters but, along with them, the whole scientific process." (p. 9)

[6]An experiment by Rindfuss, Al-Attrash, Morrison, and Heward (1998) provides a good example of the extent to which the term *single-subject research* can be a misnomer. A within-subject reversal design was used to evaluate the effects of response cards on the quiz and exam scores of 85 students in five eighth-grade American history classes. Although a large group of subjects participated in this study, it actually consisted of 85 individual experiments; or 1 experiment and 84 replications!

fifth-grade students. In such cases the total number of assignments completed may be measured, graphed, and analyzed as a dependent variable within a "single-subject" design. However, it must be remembered that unless each student's data are individually graphed and interpreted, no individual student's behavior has been analyzed, and the data for the group may not be representative of any individual subject.

Use of a single participant, or a small number of participants, each of whom is considered an intact experiment, stands in sharp contrast to the group comparison designs traditionally used in psychology and the other social sciences that employ large numbers of subjects.[7] Proponents of group comparison designs believe that large numbers of subjects control for the variability discussed earlier and increase the generality (or external validity) of any findings to the population from which the subjects were selected. The advantages and disadvantages of an experimental approach based on within-subject comparisons of the behavior of individual subjects versus comparisons of the average performance of different groups of subjects will be discussed in Chapter 10. For now, we will leave this issue with Johnston and Pennypacker's (1993b) astute observation:

> When well done, the procedures of within-subject designs preserve the pure characteristics of behavior, uncontaminated by intersubject variability. In contrast, the best between groups design practices obfuscate the representation of behavior in various ways, particularly by mixing intersubject variability with treatment-induced variability. (p. 188)

Behavior: Dependent Variable

The target behavior in an applied behavior analysis experiment, or more precisely a measurable dimensional quantity of that behavior (e.g., rate, duration), is called the **dependent variable.** It is so labeled because the experiment is designed precisely to determine whether the behavior is, in fact, *dependent on* (i.e., a function of) the independent variable(s) manipulated by the investigator. (The criteria and procedures for selecting and defining response classes that meet the *applied* requirements for an applied behavior analysis were described in Chapter 3.)

In some studies more than one behavior is measured. One reason for measuring multiple behaviors is to provide data patterns that can serve as controls for evaluating and replicating the effects of an independent variable as it is sequentially applied to each of the behaviors.[8] A second reason for multiple dependent measures is to assess the presence and extent of the independent variable's effects on behaviors other than the response class to which it was applied directly. This strategy is used to determine whether the independent variable had any collateral effects—either desired or undesired—on other behaviors of interest. Such behaviors are referred to as secondary dependent variables. The experimenter obtains regular measures of their rate of occurrence, though perhaps not with the same frequency with which measures of the primary dependent variable are recorded.

Still another reason for measuring multiple behaviors is to determine whether changes in the behavior of a person other than the subject occur during the course of an experiment and whether such changes might in turn explain observed changes in the subject's behavior. This strategy is implemented primarily as a control strategy in assessing the effects of a suspected confounding variable: The extra behavior(s) measured are not true dependent variables in the sense of undergoing analysis. For example, in a classic study analyzing the effects of the self-recording by a junior high school girl on her classroom study behavior, Broden, Hall, and Mitts (1971) observed and recorded the number of times the girl's teacher paid attention to her throughout the experiment. If teacher attention had been found to covary with changes in study behavior, a functional relation between self-recording and study behavior would not have been demonstrated. In that case, teacher attention would likely have been identified as a potential confounding variable, and the focus of the investigation would likely have shifted to include efforts to experimentally control it (i.e., to hold teacher attention constant) or to systematically manipulate and analyze its effects. However, the data revealed no functional relation between teacher attention and study behavior during the first four phases of the experiment, when concern was highest that teacher attention may have been a confounding variable.

Setting

> Control the environment and you will see order in behavior.
> —*B. F. Skinner (1967, p. 399)*

Functional relations are demonstrated when observed variations in behavior can be attributed to specific operations imposed on the environment. **Experimental control** is

[7]For a history of single-case research, see Kennedy (2005).

[8]This is the distinguishing feature of the *multiple baseline across behaviors design,* an experimental tactic used widely in applied behavior analysis. Multiple baseline designs are presented in Chapter 9.

achieved when a predictable change in behavior (the dependent variable) can be reliably and repeatedly produced by the systematic manipulation of some aspect of the subject's environment (the independent variable). To make such attributions properly, the investigator must, among other things, control two sets of environmental variables. First, the investigator must control the independent variable by presenting it, withdrawing it, and/or varying its value. Second, the investigator must control, by holding constant, all other aspects of the experimental setting—**extraneous variables**—to prevent unplanned environmental variation. These two operations—precisely manipulating the independent variable and maintaining the constancy of every other relevant aspect of the experimental setting—define the second meaning of *experimental control.*

In basic laboratory research, experimental space is designed and furnished to maximize experimental control. Lighting, temperature, and sound, for example, are all held constant, and programmed apparatus virtually guarantee the presentation of antecedent stimuli and the delivery of consequences as planned. Applied behavior analysts, however, conduct their studies in the settings where socially important behaviors naturally occur—the classroom, home, and workplace. It is impossible to control every feature of an applied environment; and to add to the difficulty, subjects are typically in the experimental setting for only part of each day, bringing with them the influences of events and contingencies operating in other settings.

In spite of the complexity and ever-changing nature of applied settings, the behavior analyst must make every effort to hold constant all seemingly relevant aspects of the environment. When unplanned variations take place, the investigator must either wait out their effects or try to incorporate them into the design of the experiment. In any event, repeated measures of the subject's behavior are the barometer for assessing whether unplanned environmental changes are of concern.

Applied studies are often conducted in more than one setting. Researchers sometimes use concurrent measures of the same behavior obtained in multiple settings as controls for analyzing the effects of an independent variable that is sequentially applied to the behavior in each setting.[9] In addition, data are often collected in multiple settings to assess the extent to which behavior changes observed in the primary setting have also occurred in the other setting(s). (Strategies for promoting the generalization of behavior change across settings are described in Chapter 28.)

[9]This analytic tactic is known as a *multiple-baseline across settings design.* Multiple-baseline designs are presented in Chapter 9.

Measurement System and Ongoing Visual Analysis

Beginning students of behavior analysis sometimes believe that the discipline is preoccupied with issues and procedures related to the observation and measurement of behavior. They want to get on with the analysis. However, the results of any experiment can be presented and interpreted only in terms of what was measured, and the observation and recording procedures used in the study determine not only what was measured, but also how well it was measured (i.e., how representative of the subject's actual behavior is the estimate provided by the experimental data—all measurements of behavior, no matter how frequent and technically precise, are estimates of true values). It is critical that observation and recording procedures be conducted in a completely standardized manner throughout each session of an experiment. Standardization involves every aspect of the measurement system, from the definition of the target behavior (dependent variable) to the scheduling of observations to the manner in which the raw data are transposed from recording sheets to session summary sheets to the way the data are graphed. As detailed in Chapter 5, an adventitious change in measurement tactics can result in unwanted variability or confounded treatment effects.

The previous chapter outlined the advantages that accrue to the behavioral researcher who maintains direct contact with the experimental data by ongoing visual inspection of graphic displays. The behavior analyst must become skilled at recognizing changes in level, trend, and degree of variability as these changes develop in the data. Because behavior is a continuous, dynamic phenomenon, experiments designed to discover its controlling variables must enable the investigator to inspect and respond to the data continuously as the study progresses. Only in this way can the behavior analyst be ready to manipulate features of the environment at the time and in the manner that will best reveal functional relations and minimize the effects of confounding variables.

Intervention or Treatment: Independent Variable

Behavior analysts seek reliable relations between behavior and the environmental variables of which it is a function. The particular aspect of the environment that the experimenter manipulates to find out whether it affects the subject's behavior is called the **independent variable.** Sometimes called the *intervention, treatment,* or *experimental variable,* this component of an experiment is called the independent variable because the researcher can control or manipulate it independent of the subject's behavior or any other event. (Though, as we will soon see, manipulating the independent variable without regard to what is happening with the dependent variable is unwise.) Whereas any changes that must be made in the experimental setting to conduct the study (e.g., the addition of observers to measure behavior) are made with the goal of minimizing their effects on the dependent variable, "changes in the independent variable are arranged by the experimenter in order to maximize . . . its influence on responding" (Johnston & Pennypacker, 1980, p. 260).

Manipulations of the Independent Variable: Experimental Design

Experimental design refers to the particular arrangement of conditions in a study so that meaningful comparisons of the effects of the presence, absence, or different values of the independent variable can be made. Independent variables can be introduced, withdrawn, increased or decreased in value, or combined across behaviors, settings, and/or subjects in an infinite number of ways.[10] However, there are only two basic kinds of independent variable changes that can be made with respect to the behavior of a given subject in a given setting.

> A new condition can be introduced or an old condition can be reintroduced. . . . Experimental designs are merely temporal arrangements of various new and old conditions across behaviors and settings in ways that produce data that are convincing to the investigator and the audience. (Johnston & Pennypacker, 1980, p. 270)

In the simplest case—from an analytic perspective, but not necessarily a practical point of view—an independent variable can be manipulated so that it is either present or absent during each time period or phase of the study. When the independent variable is in either of these conditions during a study, the experiment is termed a nonparametric study. In contrast, a **parametric analysis** seeks to discover the differential effects of a range of values of the independent variable. For example, Lerman, Kelley, Vorndran, Kuhn, and LaRue (2002) conducted a parametric study when they assessed the effects of different reinforcer magnitudes (i.e., 20 seconds, 60 seconds, or 300 seconds of access to toys or escape from demands) on the duration of postreinforcement pause and resistance to extinction. Parametric experiments are sometimes used because a functional relation may have more generality if it is based on several values of the independent variable.

Sometimes the investigator is interested in comparing the effects of several treatment alternatives. In this case multiple independent variables become part of the experiment. For example, perhaps two separate treatments are evaluated as well as the effects of a third treatment, which represents a combination of both variables. However, even in experiments with multiple independent variables, the researcher must heed a simple but fundamental rule of experimentation: *Change only one variable at a time.* Only in this manner can the behavior analyst attribute any measured changes in behavior to a specific independent variable. If two or more variables are altered simultaneously and changes in the dependent variable are noted, no conclusions can be made with regard to the contribution of any one of the altered variables to the behavior change. If two variables changed together, both could have contributed equally to the resultant behavior change; one variable could have been solely, or mostly, responsible for the change; or one variable may have had a negative or counterproductive effect, but the other independent variable was sufficiently strong enough to overcome this effect, resulting in a net gain. Any of these explanations or combinations may have accounted for the change.

As stated previously, applied behavior analysts often conduct their experiments in "noisy" environments where effective treatments are required for reasons related to personal safety or exigent circumstances. In such cases, applied behavior analysts sometimes "package" multiple and well-documented and effective treatments, knowing that multiple independent variables are being introduced. As implied earlier, a package intervention is one in which multiple independent variables are being combined or bundled into one program (e.g., token reinforcement + praise, + self-recording + time out). However, from the perspective of experimental analysis, the rule still holds. When manipulating a treatment package, the experimenter must ensure that the entire package is presented or withdrawn each time a manipulation occurs. In this situation, it is important to understand that the entire package is being evaluated, not the discrete components that make up the package. If at a later time, the analyst wishes to determine the relative contributions of each part of the package, a component analysis would need to be carried out. Chapters 8 and 9 describe experimental tactics for component analyses.

There are no off-the-shelf experimental designs available for a given research problem (Baer et al., 1987; Johnston & Pennypacker, 1980, 1993a; Sidman, 1960/1988; Skinner, 1956, 1966). The investigator must not get locked into textbook "designs" that (a) require a priori assumptions about the nature of the functional relations

[10]How many different experimental designs are there? Because an experiment's design includes careful selection and consideration of all of the components discussed here (i.e., subject, setting, behavior, etc.), not counting the direct replication of experiments, one could say that there are as many experimental designs as there are experiments.

one seeks to investigate and (b) may be insensitive to unanticipated changes in behavior. Instead, the behavior analyst should select and combine experimental tactics that best fit the research question, while standing ever ready to "explore relevant variables by manipulating them in an improvised and rapidly changing design" (Skinner, 1966, p. 21).

Simultaneous with, and to a large degree responsible for, the growth and success of applied behavior analysis have been the development and refinement of a powerful group of extremely flexible experimental tactics for analyzing behavior-environment relations. The most widely used of these tactics will be described in detail in Chapters 8 and 9. However, to most effectively select, modify, and combine these tactics into convincing experiments, the behavior analyst must first fully understand the experimental reasoning, or logic, that provides the foundation for within-subject experimental comparisons.

Steady State Strategy and Baseline Logic

Steady or **stable state responding**—which "may be defined as a pattern of responding that exhibits relatively little variation in its measured dimensional quantities over a period of time" (Johnston & Pennypacker, 1993a, p. 199)—provides the basis for a powerful form of experimental reasoning commonly used in behavior analysis, called baseline logic. **Baseline logic** entails three elements—prediction, verification, and replication—each of which depends on an overall experimental approach called steady state strategy. **Steady state strategy** entails repeatedly exposing a subject to a given condition while trying to eliminate or control any extraneous influences on the behavior and obtaining a stable pattern of responding before introducing the next condition.

Nature and Function of Baseline Data

Behavior analysts discover behavior–environment relations by comparing data generated by repeated measures of a subject's behavior under the different environmental conditions of the experiment. The most common method of evaluating the effects of a given variable is to impose it on an ongoing measure of behavior obtained in its absence. These original data serve as the **baseline** against which any observed changes in behavior when the independent variable is applied can be compared. A baseline serves as a control condition and does not necessarily mean the absence of instruction or treatment as such, only the absence of a specific independent variable of experimental interest.

Why Establish a Baseline?

From a purely scientific or analytic perspective, the primary purpose for establishing a baseline level of responding is to use the subject's performance in the absence of the independent variable as an objective basis for detecting the effects of the independent variable when it is introduced in the future. However, obtaining baseline data can yield a number of applied benefits. For one, systematic observation of the target behavior before a treatment variable is introduced provides the opportunity to look for and note environmental events that occur just before and just after the behavior. Such empirically obtained descriptions of antecedent-behavior-consequent correlations are often invaluable in planning an effective intervention (see Chapter 24). For example, baseline observations revealing that a child's disruptive outbursts are consistently followed by parent or teacher attention can be used in designing an intervention of ignoring outbursts and contingent attention following desired behavior.

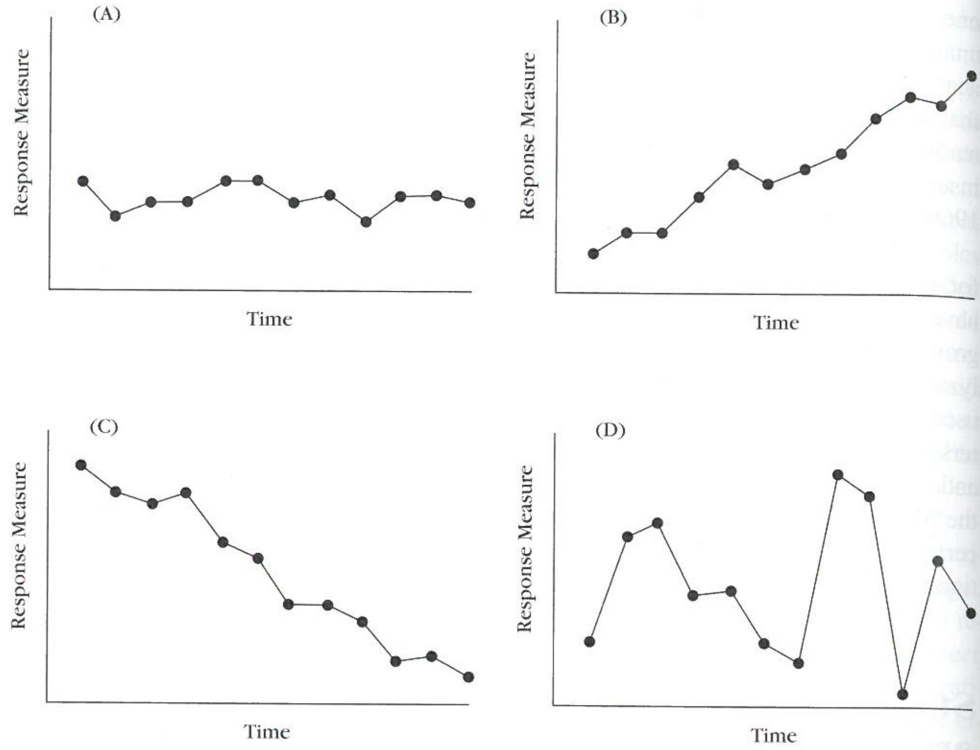
Second, baseline data can provide valuable guidance in setting initial criteria for reinforcement, a particularly important step when a contingency is first put into effect (see Chapter 11). If the criteria are too high, the subject never comes into contact with the contingency; if they are too low, little or no improvement can be expected.

From a practical perspective, a third reason for collecting baseline data concerns the merits of objective measurement versus subjective opinion. Sometimes the results of systematic baseline measurement convince the behavior analyst or significant others to alter their perspectives on the necessity and value of attempting to change the behavior. For example, a behavior being considered for intervention because of several recent and extreme instances is no longer targeted because baseline data show it is decreasing. Or, perhaps a behavior's topography attracted undue attention from teachers or parents, but objective baseline measurement over several days reveals that the behavior is not occurring at a frequency that warrants an intervention.

Types of Baseline Data Patterns

Examples of four data patterns sometimes generated by baseline measurement are shown in Figure 7.1. It must be stressed that these hypothetical baselines represent only four examples of the wide variety of baseline data patterns an experimenter or practitioner will encounter. The potential combinations of different levels, trends, and degrees of variability are, of course, infinite. Nevertheless, in an effort to provide guidance to the beginning behavior analyst, some general statements will be given about the experimental decisions that might be warranted by the data patterns shown in Figure 7.1.

Figure 7.1 Data patterns illustrating stable (A), ascending (B), descending (C), and variable (D) baselines.



Graph A shows a relatively **stable baseline.** The data show no evidence of an upward or downward trend, and all of the measures fall within a small range of values. A stable baseline provides the most desirable basis, or context, against which to look for effects of an independent variable. If changes in level, trend, and/or variability coincide with the introduction of an independent variable on a baseline as stable as that shown in Graph A, one can reasonably suspect that those changes may be related to the independent variable.

The data in Graphs B and C represent an **ascending baseline** and a **descending baseline,** respectively. The data path in Graph B shows an increasing trend in the behavior over time, whereas the data path in Graph C shows a decreasing trend. The applied behavior analyst must treat ascending and descending baseline data cautiously. By definition, dependent variables in applied behavior analysis are selected because they represent target behaviors that need to be changed. But ascending and descending baselines reveal behaviors currently in the process of changing. The effects of an independent variable introduced at this point are likely to be obscured or confounded by the variables responsible for the already-occurring change. But what if the applied investigator needs to change the behavior immediately? The applied perspective can help solve the dilemma.

Whether a treatment variable should be introduced depends on whether the trending baseline data represent improving or deteriorating performance. When an ascending or descending baseline represents behavior change in the therapeutically desired direction, the investigator should withhold treatment and continue to monitor the dependent variable under baseline conditions. When the behavior ceases to improve (as evidenced by stable responding) or begins to deteriorate, the independent variable can be applied. If the trend does not level off and the behavior continues to improve, the original problem may no longer be present, leaving no reason for introducing the treatment as planned (although the investigator might be motivated to isolate and analyze the variables responsible for the "spontaneous" improvement). Introducing an independent variable to an already-improving behavior makes it difficult, and often impossible, to claim any continued improvement as a function of the independent variable.

An ascending or descending baseline that represents significantly deteriorating performance signals an immediate application of the independent variable. From an applied perspective the decision to intervene is obvious: The subject's behavior is deteriorating, and a treatment designed to improve it should be introduced. An independent variable capable of affecting desired behavior change in spite of other variables "pushing" the behavior in the opposite direction is most likely a robust variable, one that will be a welcome addition to the behavior analyst's list of effective treatments. The decision to introduce a treatment variable on a deteriorating baseline is also a sound one from an analytic perspective, which will be discussed in the next section.

Graph D in Figure 7.1 shows a highly unstable or **variable baseline.** The data in Graph D show just one of many possible patterns of unstable responding. The data points do not consistently fall within a narrow range of values, nor do they suggest any clear trend. Introducing the independent variable in the presence of such variability is unwise from an experimental standpoint. Variability is assumed to be the result of environmental variables, which in the case shown by Graph D, seem to be operating in an uncontrolled fashion. Before the researcher can analyze the effects of an independent variable effectively, these uncontrolled sources of variability must be isolated and controlled.

Stable baseline responding provides an index of the degree of experimental control the researcher has established. Johnston and Pennypacker stressed this point in both editions of *Strategies and Tactics of Human Behavioral Research:*

> If unacceptably variable responding occurs under baseline conditions, this is a statement that the researcher is probably not ready to introduce the treatment conditions, which involves adding an independent variable whose effects are in question. (1993a, p. 201)

These authors were more direct and blunt in the first edition of their text:

> If sufficiently stable responding cannot be obtained, the experimenter is in no position to add an independent variable of suspected but unknown influence. To do so would be to compound confusion and lead to further ignorance. (1980, p. 229)

Again, however, applied considerations must be balanced against purely scientific pursuits. The applied problem may be one that cannot wait to be solved (e.g., severe self-injurious behavior). Or, confounding variables in the subject's environment and the setting(s) of the investigation may simply be beyond the experimenter's control.[11] In such situations the independent variable is introduced with the hope of producing stable responding in its presence. Sidman (1960/1988) agreed that "the behavioral engineer must ordinarily take variability as he finds it, and deal with it as an unavoidable fact of life" (p. 192).

[11]The applied researcher must guard very carefully against assuming automatically that unwanted variability is a function of variables beyond his capability or resources to isolate and control, and thus fail to pursue the investigation of potentially important functional relations.

Prediction

Prediction can be defined as "the anticipated outcome of a presently unknown or future measurement. It is the most elegant use of quantification upon which validation of all scientific and technological activity rests" (Johnston & Pennypacker, 1980, p. 120). Figure 7.2 shows a series of hypothetical measures representing a stable pattern of baseline responding. The consistency of the first five data points in the series encourage the prediction that—if no changes occur in the subject's environment—subsequent measures will fall within the range of values obtained thus far. Indeed, a sixth measure is taken that gives credence to this prediction. The same prediction is made again, this time with more confidence, and another measure of behavior shows it to be correct. Throughout a baseline (or any other experimental condition), an ongoing prediction is made and confirmed until the investigator has every reason to believe that the response measure will not change appreciably under the present conditions. The data within the shaded portion of Figure 7.2 represent unobtained but predicted measures of future responding under "relatively constant environmental conditions."[12] Given the stability of the obtained measures, few experienced scientists would quarrel with the prediction.

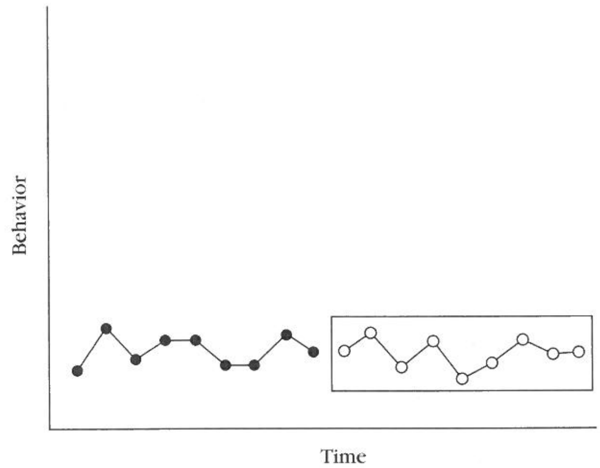


Figure 7.2 Solid data points represent actual measures of behavior that might be generated in a stable baseline; open data points within the box represent the level of responding that would be predicted on the basis of the obtained measures, should the environment remain constant.

[12]"The above reference to 'relatively constant environmental conditions' means only that the experimenter is not knowingly producing uncontrolled variations in functionally related environmental events" (Johnston & Pennypacker, 1980, p. 228).

How many measures must be taken before an experimenter can use a series of data points to predict future behavior with confidence? Baer and colleagues (1968) recommended continuing baseline measurement until "its stability is clear." Even though there are no set answers, some general statements can be made about the predictive power of steady states. All things being equal, many measurements are better than a few; and the longer the period of time in which stable responding is obtained, the better the predictive power of those measures. Also, if the experimenter is not sure whether measurement has produced stable responding, in all likelihood it has not, and more data should be collected before the independent variable is introduced. Finally, the investigator's knowledge of the characteristics of the behavior being studied under constant conditions is invaluable in deciding when to terminate baseline measurement and introduce the independent variable. That knowledge can be drawn from personal experience in obtaining stable baselines on similar response classes and from familiarity with patterns of baseline responding found in the published literature.

It should be clear that guidelines such as "collect baseline data for at least five sessions" or "obtain baseline measures over two consecutive weeks" are misguided or naive. Depending on the situation, five data points obtained over one or two weeks of baseline conditions may or may not provide a convincing picture of steady state responding. The question that must be addressed is: Are the data sufficiently stable to serve as the basis for experimental comparison? This question can be answered only by ongoing prediction and confirmation using repeated measures in an environment in which all relevant conditions are held constant.

Behavior analysts are often interested in analyzing functional relations between an instructional variable and the acquisition of new skills. In such situations it is sometimes assumed that baseline measures are zero. For example, one would expect repeated observations of a child who has never tied her shoes to yield a perfectly stable baseline of zero correct responses. However, casual observations that have never shown a child to use a particular skill do not constitute a scientifically valid baseline and should not be used to justify any claims about the effects of instruction. It could be that if given repeated opportunities to respond, the child would begin to emit the target behavior at a nonzero rate. The term **practice effects** refers to improvements in performance resulting from repeated opportunities to emit the behavior so that baseline measurements can be obtained. For example, attempting to obtain stable baseline data for students performing arithmetic problems can result in improved levels of responding simply because of the repeated practice inherent in the measurement process. Practice effects confound a study, making it impossible to separate and account for the effects of practice and instruction on the student's final performance. Repeated baseline measures should be used either to reveal the existence or to demonstrate the nonexistence of practice effects. When practice effects are suspected or found, baseline data collection should be continued until steady state responding is attained.

The necessity to demonstrate a stable baseline and to control for practice effects empirically does not require applied behavior analysts to withhold needed treatment or intervention. Nothing is gained by collecting unduly long baselines of behaviors that cannot reasonably be expected to be in the subject's repertoire. For example, many behaviors cannot be emitted unless the subject is competent in certain prerequisite behaviors; there is no legitimate possibility of a child's tying his shoes if he currently does not pick up the laces, or of a student's solving division problems if she cannot subtract and multiply. Obtaining extended baseline data in such cases is unnecessary pro forma measurement. Such measures would "not so much represent zero behavior as zero opportunity for behavior to occur, and there is no need to document at the level of well-measured data that behavior does not occur when it cannot" (Horner & Baer, 1978, p. 190).

Fortunately, applied behavior analysts need neither abandon the use of steady state strategy nor repeatedly measure nonexistent behavior at the expense of beginning treatment. The multiple probe design, described in Chapter 9, is an experimental tactic that enables the use of steady state logic to analyze functional relations between instruction and the acquisition of behaviors shown to be nonexistent in the subject's repertoire prior to the introduction of the independent variable.

Affirmation of the Consequent

The predictive power of steady state responding enables the behavior analyst to employ a kind of inductive logic known as **affirmation of the consequent** (Johnston & Pennypacker, 1980). When an experimenter introduces an independent variable on a stable baseline, an explicit assumption has been made: If the independent variable were not applied, the behavior, as indicated by the baseline data path, would not change. The experimenter is also predicting that (or more precisely, questioning whether) the independent variable will result in a change in the behavior.

The logical reasoning behind affirmation of the consequent begins with a true antecedent–consequent (if-A-then-B) statement and proceeds as follows:

1. If A is true, then B is true.
2. B is found to be true.
3. Therefore, A is true.

The behavior analyst's version goes like this:

1. If the independent variable is a controlling factor for the behavior (A), then the data obtained in the presence of the independent variable will show that the behavior has changed (B).
2. When the independent variable is present, the data show that the behavior has changed (B is true).
3. Therefore, the independent variable is a controlling variable for the behavior (therefore, A is true).

The logic, of course, is flawed; other factors could be responsible for the truthfulness of A. But, as will be shown, a successful (i.e., convincing) experiment affirms several if-A-then-B possibilities, each one reducing the likelihood of factors other than the independent variable being responsible for the observed changes in behavior.

Data shown in Figures 7.3 to 7.5 illustrate how prediction, verification, and replication are employed in a hypothetical experiment using the reversal design, one of the most common and powerful analytic tactics used by behavior analysts (see Chapter 8). Figure 7.3 shows a successful affirmation of the consequent. Steady state responding during baseline enabled the prediction that, if no changes were made in the environment, continued measurement would yield data similar to those in the shaded portion of the graph. The independent variable was then introduced, and repeated measures of the dependent variable during this treatment condition showed that the behavior did indeed change. This enables two comparisons, one real and one hypothetical. First, the real difference between the obtained measures in the presence of the independent variable and the baseline level of responding represents the extent of a possible effect of the independent variable and supports the prediction that treatment would change the behavior.

The second, hypothetical, comparison is between the data obtained in the treatment condition with the predicted measures had the treatment variable not been introduced (i.e., the open data points within the boxed area of Figure 7.3). This comparison represents the behavior analyst's hypothetical approximation of the ideal but impossible-to-achieve experimental design: the simultaneous measurement and comparison of the behavior of an individual subject in both the presence and absence of the treatment variable (Risley, 1969).

Although the data in Figure 7.3 affirm the initial antecedent–consequent statement—a change in the behavior was observed in the presence of the independent

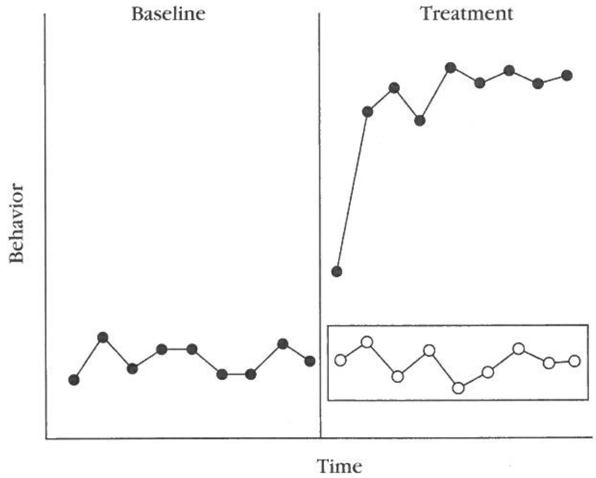


Figure 7.3 Affirmation of the consequent supporting the possibility of a functional relation between the behavior and treatment variable. The measures obtained in the presence of the treatment variable differ from the predicted level of responding in the absence of the treatment variable (open data points within boxed area).

variable—asserting a functional relation between the independent and dependent variables at this point is unwarranted. The experiment has not yet ruled out the possibility of other variables being responsible for the change in behavior. For example, perhaps some other event that is responsible for the change in behavior occurred at the same time that the independent variable was introduced.[13]

A firmer statement about the relation between the treatment and the behavior can be made at this point, however, if changes in the dependent variable are *not* observed in the presence of the independent variable. Assuming accurate measures of the behavior and a measurement system sensitive to changes in the behavior, then no behavior change in the presence of the independent variable constitutes a disconfirmation of the consequent (B was shown not to be true), and the independent variable is eliminated as a controlling variable. However, eliminating a treatment from the ranks of controlling variables on the basis of no observed effects presupposes experimental control of the highest order (Johnston & Pennypacker, 1993a).

[13]Although two-phase experiments consisting of a pretreatment baseline condition followed by a treatment condition (called ***A–B design***) enable neither verification of the prediction of continued responding at baseline levels nor replication of the effects of the independent variable, studies using A–B designs can nevertheless contribute important and useful findings (e.g., Azrin & Wesolowski, 1974; Reid, Parsons, Phillips, & Green, 1993).

However, in the situation illustrated in Figure 7.3, a change in behavior was observed in the presence of the independent variable, revealing a correlation between the independent variable and the behavior change. To what extent was the observed behavior change a function of the independent variable? To pursue this question, the behavior analyst employs the next component of baseline logic: verification.

Verification

The experimenter can increase the probability that an observed change in behavior was functionally related to the introduction of the independent variable by verifying the original prediction of unchanging baseline measures. **Verification** can be accomplished by demonstrating that the prior level of baseline responding would have remained unchanged had the independent variable not been introduced (Risley, 1969). If that can be demonstrated, this operation verifies the accuracy of the original prediction of continued stable baseline responding and reduces the probability that some uncontrolled (confounding) variable was responsible for the observed change in behavior. Again, the reasoning behind affirmation of the consequent is the logic that underlies the experimental strategy.

Figure 7.4 illustrates the verification of effect in our hypothetical experiment. When steady state responding has been established in the presence of the independent variable, the investigator removes the treatment variable, thereby returning to the previous baseline conditions. This tactic allows the possibility of affirming two different antecedent–consequent statements. The first statement and its affirmation follows this pattern:

1. If the independent variable is a controlling factor for the behavior (A), then its removal will coincide with changes in the response measure (B).
2. Removal of the independent variable is accompanied by changes in the behavior (B is true).
3. Therefore, the independent variable controls responding (therefore, A is true).

The second statement and affirmation follows this pattern:

1. If the original baseline condition controlled the behavior (A), then a return to baseline conditions will result in similar levels of responding (B).
2. The baseline condition is reinstated and levels of responding similar to those obtained during the original baseline phase are observed (B is true).
3. Therefore, the baseline condition controlled the behavior both then and now (therefore, A is true).

The six measures within the shaded area obtained during Baseline 2 of our hypothetical experiment in Figure 7.4 verify the prediction made for Baseline 1. The open data points in the shaded area in Baseline 2 represent the predicted level of responding if the independent variable had not been removed. (The prediction component of baseline logic applies to steady state responding obtained during any phase of an experiment, baseline and treatment conditions alike.) The difference between the data actually obtained during Treatment (solid data points) and the data obtained during Base-line 2 (solid data points) affirms the first if-A-then-B statement: If the treatment is a controlling variable, then its removal will result in changes in behavior. The similarity between measures obtained during Baseline 2 and those obtained during Baseline 1 confirms the second if-A-then-B statement: If baseline conditions controlled the behavior before, reinstating baseline conditions will result in similar levels of responding.

Again, of course, the observed changes in behavior associated with the application and withdrawal of the independent variable are subject to interpretations other

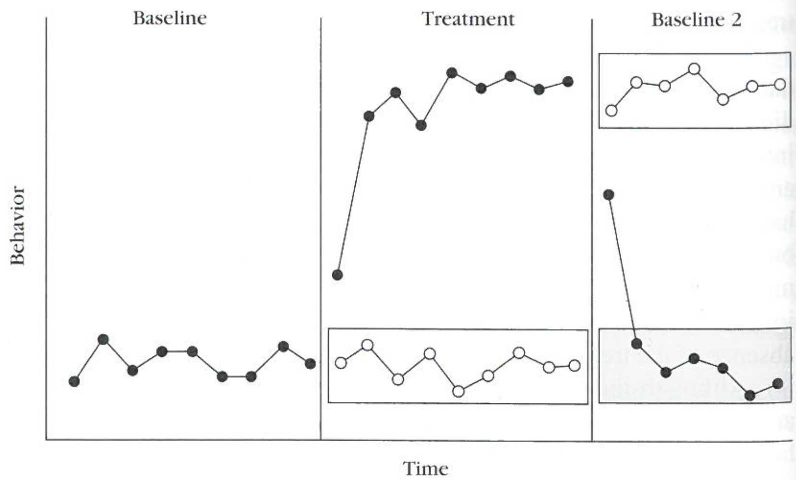


Figure 7.4 Verification of a previously predicted level of baseline responding by termination or withdrawal of the treatment variable. The measures obtained during Baseline 2 (solid data points within shaded area) show a successful verification and a second affirmation of the consequent based on a comparison with the predicted level of responding (open dots in Baseline 2) in the continued presence of the treatment variable.

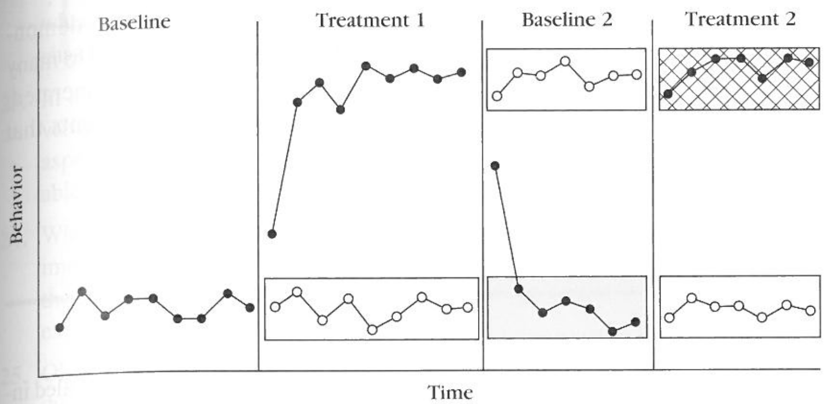


Figure 7.5 Replication of experimental effect accomplished by reintroducing the treatment variable. The measures obtained during Treatment 2 (data points within area shaded with cross hatching) enhance the case for a functional relation between the treatment variable and the target behavior.

than a claim of a functional relation between the two events. However, the case for the existence of a functional relation is becoming stronger. When the independent variable was applied, behavior change was observed; when the independent variable was withdrawn, behavior again changed and responding returned to baseline levels. To the extent that the experimenter effectively controls the presence and absence of the independent variable and holds constant all other variables in the experimental setting that might influence the behavior, a functional relation appears likely: An important behavior change has been produced and reversed by the introduction and withdrawal of the independent variable. The process of verification reduces the likelihood that a variable other than the independent variable was responsible for the observed behavior changes.

Does this two-step strategy of prediction and verification constitute sufficient demonstration of a functional relation? What if some uncontrolled variable covaried with the independent variable as it was presented and withdrawn and this uncontrolled variable was actually responsible for the observed changes in behavior? If such was the case, claiming a functional relation between the target behavior and the independent variable would at best be inaccurate and at the worst perhaps end a search for the actual controlling variables whose identification and control would contribute to an effective and reliable technology of behavior change.

The appropriately skeptical investigator (and research consumer) will also question the reliability of the obtained effect. How reliable is this verified behavior change? Was the apparent functional relation a fleeting, one-time-only phenomenon, or will repeated application of the independent variable reliably (i.e., consistently) produce a similar pattern of behavior change? An effective (i.e., convincing) experimental design yields data that are responsive to these important questions. To investigate uncertain reliability, the behavior analyst employs the final, and perhaps the most important, component of baseline logic and experimental design: replication.

Replication

Replication is the essence of believability.

—*Baer, Wolf, and Risley (1968, p. 95)*

Within the context of any given experiment, **replication** means repeating independent variable manipulations conducted previously in the study and obtaining similar outcomes.[14] Replication within an experiment has two important purposes. First, replicating a previously observed behavior change reduces the probability that a variable other than the independent variable was responsible for the now twice-observed behavior change. Second, replication demonstrates the reliability of the behavior change; it can be made to happen again.

Figure 7.5 adds the component of replication to our hypothetical experiment. After steady state responding was obtained during Baseline 2, the independent variable is reintroduced; this is the Treatment 2 phase. To the extent that the data obtained during the second application of the treatment (data points within area shaded with cross-hatched lines) resemble the level of responding observed during Treatment 1, replication has occurred. Our hypothetical experiment has now produced powerful evidence of a functional relation exists between the independent and the dependent variable. The extent to which one has confidence in the assertion of a functional relation rests on numerous factors, some of the most important of which are the accuracy and sensitivity of the measurement system, the degree of control the experimenter maintained over all relevant variables, the duration of experimental phases, the stability of responding within each phase, and the speed, magnitude, and consistency of behavior change between conditions. If each of these

[14]Replication also refers to the repeating of experiments to determine the reliability of functional relations found in previous experiments and the extent to which those findings can be extended to other subjects, settings, and/or behaviors (i.e., generality or external validity). The replication of experiments is examined in Chapter 10.

considerations is satisfied by the experimental design and is supported by the data as displayed within the design, then replication of effect becomes perhaps the most critical factor in claiming a functional relation.

An independent variable can be manipulated in an effort to replicate an effect many times within an experiment. The number of replications required to demonstrate a functional relation convincingly is related to many considerations, including all of those just enumerated, and to the existence of other similar experiments that have produced the same effects.

Summary

Introduction

1. Measurement can show whether and when behavior changes, but measurement alone cannot reveal how the change has come about.
2. Knowledge of specific functional relations between behavior and environment is necessary if a systematic and useful technology of behavior change is to develop.
3. An experimental analysis must be performed to determine how a given behavior functions in relation to specific environmental events.

Concepts and Assumptions Underlying the Analysis of Behavior

4. The overall goal of science is to achieve an understanding of the phenomena under study—socially important behaviors, in the case of applied behavior analysis.
5. Science produces understanding at three levels: description, prediction, and control.
6. Descriptive research yields a collection of facts about the observed events—facts that can be quantified and classified.
7. A correlation exists when two events systematically covary with one another. Predictions can be made about the probability that one event will occur based on the occurrence of the other event.
8. The greatest potential benefits of science are derived from the third, and highest, level of scientific understanding, which comes from establishing experimental control.
9. Experimental control is achieved when a predictable change in behavior (the dependent variable) can be reliably produced by the systematic manipulation of some aspect of the person's environment (the independent variable).
10. A functional analysis does not eliminate the possibility that the behavior under investigation is also a function of other variables.
11. An experiment that shows convincingly that changes in behavior are a function of the independent variable and not the result of uncontrolled or unknown variables has internal validity.
12. External validity refers to the degree to which a study's results are generalizable to other subjects, settings, and/or behaviors.
13. Confounding variables exert unknown or uncontrolled influences on the dependent variable.
14. Because behavior is an individual phenomenon, the experimental strategy of behavior analysis is based on within-subject (or single-subject) methods of analysis.
15. Because behavior is a continuous phenomenon that occurs in and changes through time, the repeated measurement of behavior is a hallmark of applied behavior analysis.
16. The assumption of determinism guides the methodology of behavior analysis.
17. Experimental methods in behavior analysis are based on the assumption that variability is extrinsic to the organism; that is, variability is imposed by environmental variables and is not an inherent trait of the organism.
18. Instead of masking variability by averaging the performance of many subjects, behavior analysts attempt to isolate and experimentally manipulate the environmental factors responsible for the variability.

Components of Experiments in Applied Behavior Analysis

19. The experimental question is a statement of what the researcher seeks to learn by conducting the experiment and should guide and be reflected in all aspects of the experiment's design.
20. Experiments in applied behavior analysis are most often referred to as *single-subject* (or *single-case*) research designs because the experimental logic or reasoning for analyzing behavior change often employs the subject as her own control.
21. The dependent variable in an applied behavior analysis experiment is a measurable dimensional quantity of the target behavior.
22. Three major reasons behavior analysts use multiple-response measures (dependent variables) in some studies are (a) to provide additional data paths that serve as controls for evaluating and replicating the effects of an independent variable that is sequentially applied to each behavior, (b) to assess the generality of treatment effects to behaviors other than the response class to which the independent variable was applied, and (c) to determine whether changes in the behavior of a person other than the subject occur during the course of an experiment and

whether such changes might in turn explain observed changes in the subject's behavior.

23. In addition to precise manipulation of the independent variable, the behavior analyst must hold constant all other aspects of the experimental setting—extraneous variables—to prevent unplanned environmental variation.

24. When unplanned events or variations occur in the experimental setting, the behavior analyst must either wait out their effects or incorporate them into the design of the experiment.

25. Observation and measurement procedures must be conducted in a standardized manner throughout an experiment.

26. Because behavior is a continuous and dynamic phenomenon, ongoing visual inspection of the data during the course of an experiment is necessary to identify changes in level, trend, and/or variability as they develop.

27. Changes in the independent variable are made in an effort to maximize its effect on the target behavior.

28. The term *experimental design* refers to the way the independent variable is manipulated in a study.

29. Although an infinite number of experimental designs are possible as a result of the many ways independent variables can be manipulated and combined, there are only two basic kinds of changes in independent variables: introducing a new condition or reintroducing an old condition.

30. A parametric study compares the differential effects of a range of different values of the independent variable.

31. The fundamental rule of experimental design is to change only one variable at a time.

32. Rather than follow rigid, pro forma experimental designs, the behavior analyst should select experimental tactics suited to the original research questions, while standing ready to "explore relevant variables by manipulating them in an improvised and rapidly changing design" (Skinner, 1966, p. 21).

Steady State Strategy and Baseline Logic

33. Stable, or steady state, responding enables the behavior analyst to employ a powerful form of inductive reasoning, sometimes called baseline logic. Baseline logic entails three elements: prediction, verification, and replication.

34. The most common method for evaluating the effects of a given variable is to impose it on an ongoing measure of behavior obtained in its absence. These preintervention data serve as the baseline by which to determine and evaluate any subsequent changes in behavior.

35. A baseline condition does not necessarily mean the absence of instruction or treatment per se, only the absence of the specific independent variable of experimental interest.

36. In addition to the primary purpose of establishing a baseline as an objective basis for evaluating the effects of the independent variable, three other reasons for baseline data collection are as follows: (a) Systematic observation of the target behavior prior to intervention sometimes yields information about antecedent-behavior-consequent correlations that may be useful in planning an effective intervention; (b) baseline data can provide valuable guidance in setting initial criteria for reinforcement; and (c) sometimes baseline data reveal that the behavior targeted for change does not warrant intervention.

37. Four types of baseline data patterns are stable, ascending, descending, and variable.

38. The independent variable should be introduced when stable baseline responding has been achieved.

39. The independent variable should not be introduced if either an ascending or descending baseline indicates improving performance.

40. The independent variable should be introduced if either an ascending or descending baseline indicates deteriorating performance.

41. The independent variable should not be imposed on a highly variable, unstable baseline.

42. Prediction of future behavior under relatively constant environmental conditions can be made on the basis of repeated measures of behavior showing little or no variation.

43. In general, given stable responding, the more data points there are and the longer the time period in which they were obtained, the more accurate the prediction will likely be.

44. Practice effects refer to improvements in performance resulting from opportunities to emit the behavior that must be provided to obtain repeated measures.

45. Extended baseline measurement is not necessary for behaviors that have no logical opportunity to occur.

46. The inductive reasoning called affirmation of the consequent lies at the heart of baseline logic.

47. Although the logic of affirming the consequent is not completely sound (some other event may have caused the change in behavior), an effective experimental design confirms several if-A-then-B possibilities, thereby eliminating certain other factors as responsible for the observed changes in behavior.

48. Verification of prediction is accomplished by demonstrating that the prior level of baseline responding would have remained unchanged if the independent variable had not been introduced.

49. Replication within an experiment means reproducing a previously observed behavior change by reintroducing the independent variable. Replication within an experiment reduces the probability that a variable other than the independent variable was responsible for the behavior change and demonstrates the reliability of the behavior change.