CHAPTER 8

Reversal and Alternating Treatments Designs

Key Terms

A-B-A design
A-B-A-B design
alternating treatments design
B-A-B design
DRI/DRA reversal technique
DRO reversal technique
irreversibility
multielement design
multiple treatment interference
multiple treatment reversal design
(NCR) reversal technique
reversal design
sequence effects
withdrawal design

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

	Content Area 5: Experimental Evaluation of Interventions
5-1	Systematically manipulate independent variables to analyze their effects on treatment.
(a)	Use withdrawal designs.
(b)	Use reversal designs.
(c)	Use alternating treatments (i.e., multielement, simultaneous treatment, multiple or concurrent schedule) designs.
5-2	Identify and address practical and ethical considerations in using various experimental designs.

This chapter describes the reversal and alternating treatments designs, two types of experimental analysis tactics widely used by applied behavior analysts. In a reversal design, the effects of introducing, withdrawing (or "reversing" the focus of), and reintroducing an independent variable are observed on the target behavior. In an alternating treatments analysis, two or more experimental conditions are rapidly alternated, and the differential effects on behavior are noted. We explain how each design incorporates the three elements of steady state strategy—prediction, verification, and replication—and present representative examples illustrating the major variations of each. Considerations for selecting and using reversal and alternating treatments designs are also presented.

Reversal Design

An experiment using a **reversal design** entails repeated measures of behavior in a given setting that requires at least three consecutive phases: (a) an initial baseline phase in which the independent variable is absent, (b) an intervention phase during which the independent variable is introduced and remains in contact with the behavior, and (c) a return to baseline conditions accomplished by withdrawal of the independent variable. In the widely used notation system for describing experimental designs in applied behavior analysis, the capital letters *A* and *B* denote the first and second conditions, respectively, that are introduced in a study. Typically baseline (A) data are collected until steady state responding is achieved. Next, an intervention (B) condition is applied that signifies the presence of a treatment—the independent variable. An experiment entailing one reversal is described as an **A-B-A design.** Although studies using an A-B-A design are reported in the literature (e.g., Christle & Schuster, 2003; Geller, Paterson, & Talbott, 1982; Jacobson, Bushell, & Risley, 1969; Stitzer, Bigelow, Liebson, & Hawthorne, 1982), an **A-B-A-B design** is preferred because reintroducing the B condition enables the replication of treatment effects, which strengthens the demonstration of experimental control (see Figure 8.1).[1]

[1]Some authors use the term ***withdrawal design*** to describe experiments based on an A-B-A-B analysis and reserve the term *reversal design* for studies in which the behavioral focus of the treatment variable is reversed (or switched to another behavior), as in the DRO and DRI/DRA reversal techniques described later in this chapter (e.g., Leitenberg, 1973; Poling, Method, & LeSage, 1995). However, *reversal design,* as the term is used most often in the behavior analysis literature, encompasses both withdrawals and reversals of the independent variable, signifying the researcher's attempt to demonstrate "behavioral reversibility" (Baer, Wolf, & Risley, 1968; Thompson & Iwata, 2005). Also, *withdrawal design* is sometimes used to describe an experiment in which the treatment variable(s) are sequentially or partially withdrawn after their effects have been analyzed in an effort to promote maintenance of the target behavior (Rusch & Kazdin, 1981).

The A-B-A-B reversal is the most straightforward and generally most powerful within-subject design for demonstrating a functional relation between an environmental manipulation and a behavior. When a functional relation is revealed with a reversal design, the data show how the behavior works.

> As explanations go, the one offered by the reversal design was not at all a bad one. In answer to the question, "How does this response work?" we could point out demonstrably that it worked like so [e.g., see Figure 8.1]. Of course, it might also work in other ways; but, we would wait until we had seen the appropriate graphs before agreeing to any other way. (Baer, 1975, p. 19)

Baer's point must not be overlooked: Showing that a behavior works in a predictable and reliable way in the presence and absence of a given variable provides only one answer to the question, How does this behavior work? There may be (and quite likely are) other controlling variables for the targeted response class. Whether additional experimentation is needed to explore those other possibilities depends on the social and scientific importance of obtaining a more complete analysis.

Operation and Logic of the Reversal Design

Risley (2005) described the rationale and operation of the reversal design as follows:

> The reversal or ABAB design that Wolf reinvented from Claude Bernard's early examples in experimental medicine entailed establishing a baseline of repeated quantified observations sufficient to see a trend and forecast that trend into the near future (A); to then alter conditions and see if the repeated observations become different than they were forecast to be (B); to then change back and see if the repeated observations return to confirm the original forecast (A); and finally, to reintroduce the altered conditions and see if the repeated observations again become different than forecast (B). (pp. 280–281)[2]

Because the reversal design was used in Chapter 7 to illustrate baseline logic, a brief review here of the roles of prediction, verification, and replication in the reversal design will suffice. Figure 8.2 shows the same data from Figure 8.1 with the addition of the open data points representing predicted measures of behavior if conditions in the previous phase had remained unchanged. After a stable pattern of responding, or a countertherapeutic trend, is obtained during Baseline 1, the independent variable is

[2]Risley (1997, 2005) credits Montrose Wolf with designing the first experiments using the reversal and multiple baseline designs. "The research methods that Wolf pioneered in these studies were groundbreaking. That methodology came to define applied behavior analysis" (pp. 280–281).

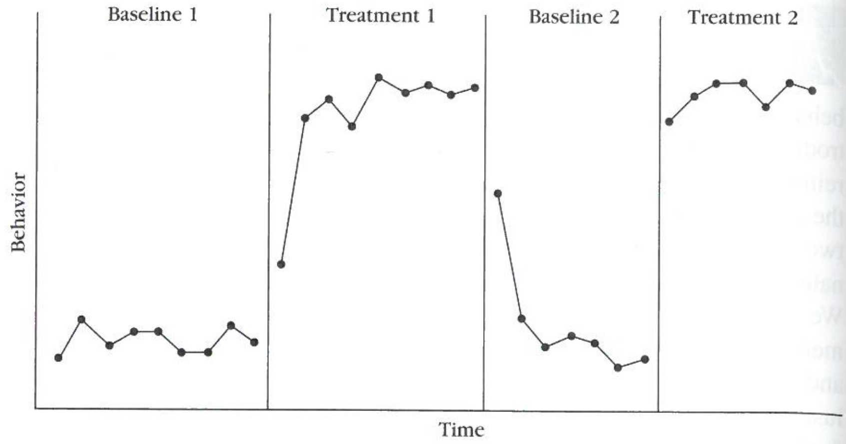


Figure 8.1 Graphic prototype of the A-B-A-B reversal design.

introduced. In our hypothetical experiment the measures obtained during Treatment 1, when compared with those from Baseline 1 and with the measures *predicted* by Baseline 1, show that behavior change occurred and that the change in behavior coincided with the intervention. After steady state responding is attained in Treatment 1, the independent variable is withdrawn and baseline conditions are reestablished. If the level of responding in Baseline 2 is the same as or closely approximates the measures obtained during Baseline 1, *verification* of the prediction made for Baseline 1 data is obtained. Stated otherwise, had the intervention not been introduced and had the initial baseline condition remained in effect, the predicted data path would have appeared as shown in Baseline 2. When withdrawal of the independent variable results in a reversal of the behavior change associated with its introduction, a strong case builds that the intervention is responsible for the observed behavior change. If reintroduction of the independent variable in Treatment 2 reproduces the behavior change observed during Treatment 1, *replication* of effect has been achieved, and a functional relation has been demonstrated. Again stated in other terms, had the intervention continued and had the second baseline condition not been introduced, the predicted data path of the treatment would have appeared as shown in Treatment 2.

Romaniuk and colleagues (2002) provided an excellent example of the A-B-A-B design. Three students with developmental disabilities who frequently displayed problem behaviors (e.g., hitting, biting, whining, crying, getting out of seat, inappropriate gestures, noises, and comments) when given academic tasks participated in the study. Prior to the experiment a functional analysis (see Chapter 24) had shown that each student's problem behaviors were maintained by escape from working on the task (i.e., problem behavior occurred most often when followed by being allowed to take a break from the task). The researchers wanted to determine whether providing students with a choice of which task to work on would reduce the frequency of their problem behavior, even though problem behavior, when it occurred, would still result in a break. The experiment consisted of two conditions: no choice (A) and choice (B). The same set of teacher-nominated tasks was used in both conditions.

Each session during the no-choice condition began with the experimenter providing the student with a task and saying, "This is the assignment you will be working on today" or "It's time to work on _____" (p. 353). Dur-

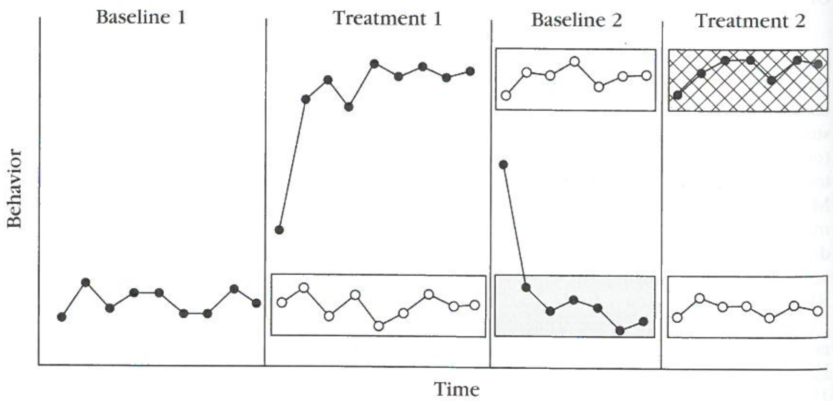


Figure 8.2 Illustration of A-B-A-B reversal design. Open data points represent data predicted if conditions from previous phase remained in effect. Data collected during Baseline 2 (within shaded box) verify the prediction from Baseline 1. Treatment 2 data (cross-hatched shading) replicate the experimental effect.

ing the choice condition (B), the experimenter placed the materials for four to six tasks on the table before the student and said, "Which assignment would you like to work on today?" (p. 353). The student was also told that he or she could switch tasks at any time during the session by requesting to do so. Occurrences of problem behavior in both conditions resulted in the experimenter stating, "You can take a break now" and giving a 10-second break.

Figure 8.3 shows the results of the experiment. The data reveal a clear functional relation between the opportunity to choose which tasks to work on and reduced occurrence of problem behavior by all three students. The percentage of session time in which each student exhibited problem behavior (reported as a total duration measure obtained by recording the second of onset and offset as shown by the VCR timer) decreased sharply from the no-choice (baseline) levels when the choice condition was implemented, returned (reversed) to baseline levels when choice was withdrawn, and decreased again when choice was reinstated. The A-B-A-B design enabled Romaniuk and colleagues to conduct a straightforward, unambiguous demonstration that significant reductions in problem behavior exhibited by each student were a function of being given a choice of tasks.

In the 1960s and early 1970s applied behavior analysts relied almost exclusively on the A-B-A-B reversal design. The straightforward A-B-A-B design played such a dominant role in the early years of applied behavior analysis that it came to symbolize the field (Baer, 1975). This was no doubt due, at least in part, to the reversal design's ability to expose variables for what they are—strong and reliable or weak and unstable. Another reason for the reversal design's dominance may have been that few alternative analytic tactics were available at that time that effectively combined the intrasubject experimental elements of prediction, verification, and replication. Although the reversal design is just one of many experimental designs available to applied behavior analysts today, the simple, unadorned A-B-A-B design continues to play a major role in the behavior analysis literature (e.g., Anderson & Long, 2002 [see Figure 21.2]; Ashbaugh & Peck, 1998 [see Figure 15.7]; Cowdery, Iwata, & Pace, 1990 [see

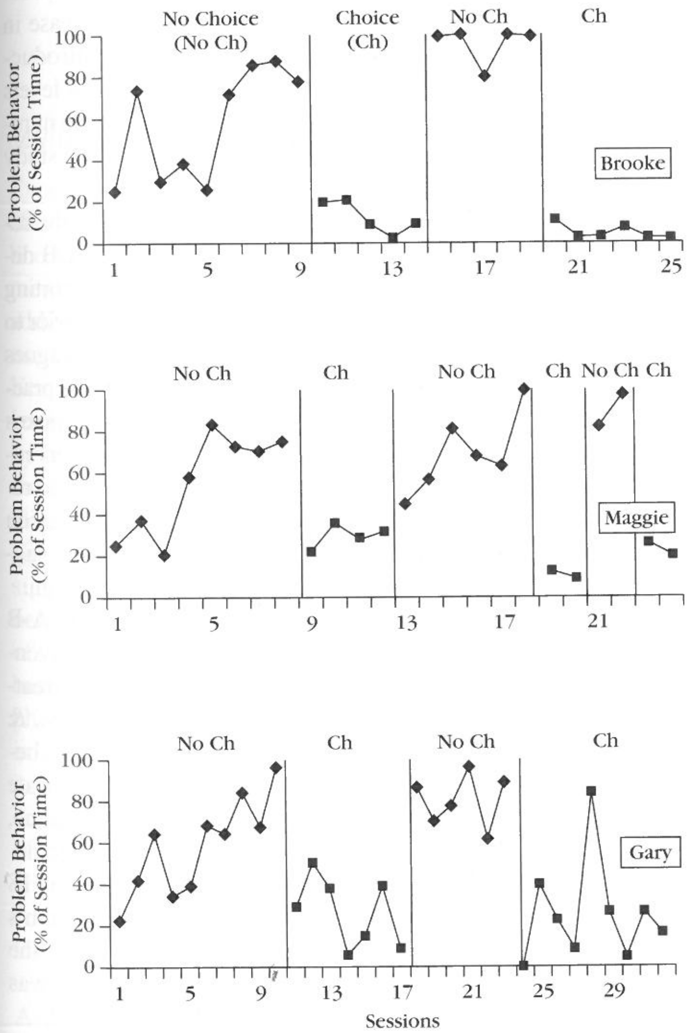


Figure 8.3 An A-B-A-B reversal design.

From "The Influence of Activity Choice on Problems Behaviors Maintained by Escape versus Attention" by C. Romaniuk, R. Miltenberger, C. Conyers, N. Jenner, M. Jurgens, and C. Ringenberg, 2002, *Journal of Applied Behavior Analysis, 35,* p. 357. Copyright 2002 by the Society for the Experimental Analysis of Behavior, Inc. Reprinted by permission.

Figure 22.7]; Deaver, Miltenberger, & Stricker, 2001 [see Figure 21.3]; Gardner, Heward, & Grossi, 1994; Levondoski & Cartledge, 2000; Lindberg, Iwata, Kahng, & DeLeon, 1999 [see Figure 22.7]; Mazaleski, Iwata, Rodgers, Vollmer, & Zarcone, 1994; Taylor & Alber, 2003; Umbreit, Lane, & Dejud, 2004).

Variations of the A-B-A-B Design

Many applied behavior analysis studies use variations or extensions of the A-B-A-B design.

Repeated Reversals

Perhaps the most obvious variation of the A-B-A-B reversal design is a simple extension in which the independent variable is withdrawn and reintroduced a second time; A-B-A-B-A-B (see the graph for Maggie in Figure 8.3). Each additional presentation and withdrawal that reproduces the previously observed effects on behavior increases the likelihood that the behavior changes are the result of manipulating the independent variable. All other things being equal, an experiment that incorporates multiple reversals presents a more convincing and compelling demonstration of a functional relation than does an experiment with one reversal (e.g., Fisher, Lindauer, Alterson, & Thompson, 1998 [Figure 6.2]; Steege et al., 1990). That said, it is also possible to reach a point of redundancy beyond which the findings of a given analysis are no longer enhanced significantly by additional reversals.

B-A-B Design

The **B-A-B design** begins with the application of the independent variable: the treatment. After stable responding has been achieved during the initial treatment phase (B), the independent variable is withdrawn. If the behavior worsens in the absence of the independent variable (the A condition), the treatment variable is reintroduced in an attempt to recapture the level of responding obtained during the first treatment phase, which would verify the prediction based on the data path obtained during the initial treatment phase.

Compared to the A-B-A design, the B-A-B design is preferable from an applied sense in that the study ends with the treatment variable in effect. However, in terms of demonstrating a functional relation between the independent variable and dependent variable, the B-A-B design is the weaker of the two because it does not enable an assessment of the effects of the independent variable on the preintervention level of responding. The nonintervention (A) condition in a B-A-B design cannot verify a prediction of a previous nonexistent baseline. This weakness can be remedied by withdrawing and then reintroducing the independent variable, as in a B-A-B-A-B design (e.g., Dixon, Benedict, & Larson, 2001 [see Figure 22.1]).

Because the B-A-B design provides no data to determine whether the measures of behavior taken during the A condition represent preintervention performance, sequence effects cannot be ruled out: The level of behavior observed during the A condition may have been influenced by the fact that the treatment condition preceded it. Nevertheless, there are exigent situations in which initial baseline data cannot be collected. For instance, the B-A-B design may be appropriate with target behaviors that result in physical harm or danger to the participant or to others. In such instances, withholding a possibly effective treatment until a stable pattern of baseline responding can be obtained may present ethical problems. For example, Murphy, Ruprecht, Baggio, and Nunes (1979) used a B-A-B design to evaluate the effectiveness of mild punishment combined with reinforcement on the number of self-choking responses by a 24-year-old man with profound mental retardation. After the treatment was in effect for 24 sessions, it was withdrawn for three sessions, during which an immediate and large increase in self-choking was recorded (see Figure 8.4). Reintroduction of the treatment package reproduced behavior levels noted during the first treatment phase. The average number of self-chokes during each phase of the B-A-B study was 22, 265, and 24, respectively.

Despite the impressive reduction of behavior, the results of Murphy and colleagues' study using a B-A-B design may have been enhanced by gathering and reporting objectively measured data on the level of behavior prior to the first intervention. Presumably, Murphy and colleagues chose not to collect an initial baseline for ethical and practical reasons. They reported anecdotally that self-chokes averaged 434 per day immediately prior to their intervention when school staff had used a different procedure to reduce the self-injurious behavior. This anecdotal information increased the believability of the functional relation suggested by the experimental data from the B-A-B design.

At least two other situations exist in which a B-A-B design might be warranted instead of the more conventional A-B-A-B design. These include (a) when a treatment is already in place (e.g., Marholin, Touchette, & Stuart, 1979; Pace & Troyer, 2000) and (b) when the behavior analyst has limited time in which to demonstrate practical and socially significant results. For instance, Robinson, Newby, and Ganzell (1981) were asked to develop a behavior management system for a class of 18 hyperactive boys with the stipulation that the program's effectiveness be demonstrated within 4 weeks. Given "the stipulation of success in 4 weeks, a B-A-B design was used" (pp. 310–311).

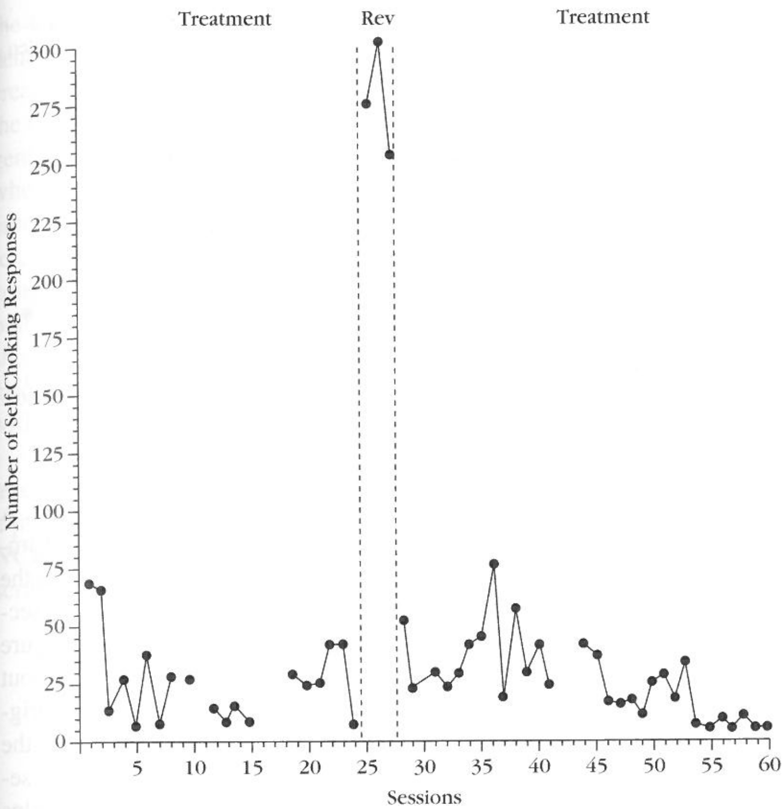


Figure 8.4 A B-A-B reversal design.
From "The Use of Mild Punishment in Combination with Reinforcement of Alternate Behaviors to Reduce the Self-Injurious Behavior of a Profoundly Retarded Individual" by R. J. Murphy, M. J. Ruprecht, P. Baggio, and D. L. Nunes, 1979, *AAESPH Review, 4,* p. 191. Copyright 1979 by the *AAESPH Review.* Reprinted by permission.

Multiple Treatment Reversal Designs

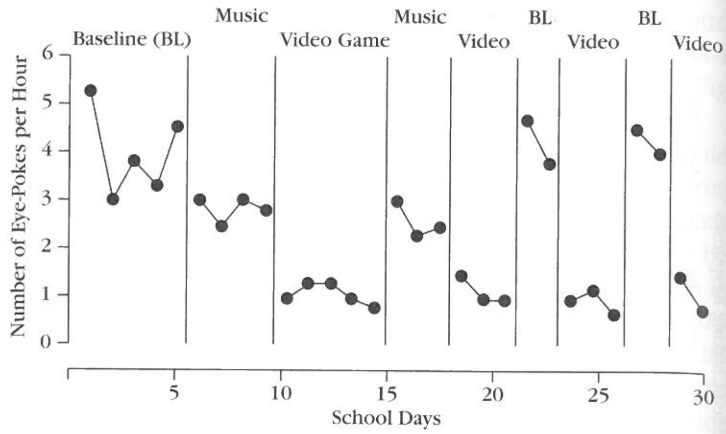
Experiments that use the reversal design to compare the effects of two or more experimental conditions to baseline and/or to one another are said to use a **multiple treatment reversal design.** The letters *C, D,* and so on, denote additional conditions, as in the A-B-C-A-C-B-C design used by Falcomata, Roane, Hovanetz, Kettering, and Keeney (2004); the A-B-A-B-C-B-C design used by Freeland and Noell (1999); the A-B-C-B-C-B-C design used by Lerman, Kelley, Vorndran, Kuhn, and LaRue (2002); the A-B-A-C-A-D-A-C-A-D design used by Weeks and Gaylord-Ross (1981); and the A-B-A-B-B+C-B-B+C design of Jason and Liotta (1982). As a whole, these designs are considered variations of the reversal design because they embody the experimental method and logic of the reversal tactic: Responding in each phase provides baseline (or control condition) data for the subsequent phase (prediction), independent variables are withdrawn in an attempt to reproduce levels of behavior observed in a previous condition (verification), and each independent variable that contributes fully to the analysis is introduced at least twice (replication). Independent variables can be introduced, withdrawn, changed in value, combined, and otherwise manipulated to produce an endless variety of experimental designs.

For example, Kennedy and Souza (1995) used an A-B-C-B-C-A-C-A-C design to analyze and compare the effects of two kinds of competing sources of stimulation on eye poking by a 19-year-old student with profound disabilities. Geoff had a 12-year history of poking his forefinger into his eyes during periods of inactivity, such as after lunch or while waiting for the bus. The two treatment conditions were music (B) and a video game (C). During the music condition, Geoff was given a Sony Walkman radio with headphones. The radio was tuned to a station that his teacher and family thought he preferred. Geoff had continuous access to the music during this condition, and he could remove the headphones at any time. During the video game condition, Geoff was given a small handheld video game on which he could observe a variety of visual patterns and images on the screen with no sound. As with the music condition, Geoff had continuous access to the video game and could discontinue using it at any time.

Figure 8.5 shows the results of the study. Following an initial baseline phase (A) in which Geoff averaged 4 eye pokes per hour, the music condition (B) was introduced and eye pokes decreased to a mean of 2.8 per hour. The video game (C) was implemented next, and eye pokes decreased further to 1.1 per hour. Measures obtained during the next two phases—a reintroduction of music (B) followed by a second phase of the video game (C)—replicated previous levels of responding under each condition. This B-C-B-C portion of the experiment

Figure 8.5 Example of multiple-treatment reversal design (A-B-C-B-C-A-C-A-C).

From "Functional Analysis and Treatment of Eye Poking" by C. H. Kennedy and G. Souza, 1995, *Journal of Applied Behavior Analysis, 28,* p. 33. Copyright 1995 by the Society for the Experimental Analysis of Behavior, Inc. Reprinted by permission.



revealed a functional relation between video game condition and lower frequency of eye pokes compared to music. The final five phases of the experiment (C-A-C-A-C) provided an experimental comparison of the video game and baseline (no-treatment) condition.

In most instances, extended designs involving multiple independent variables are not preplanned. Instead of following a predetermined, rigid structure that dictates when and how experimental manipulations must be made, the applied behavior analyst makes design decisions based on ongoing assessments of the data.

> In this sense, a single experiment may be viewed as a number of successive designs that are collectively necessary to clarify relations between independent and dependent variables. Thus, some design decisions might be made in response to the data unfolding as the investigation progresses. This sense of design encourages the experimenter to pursue in more dynamic fashion the solutions to problems of experimental control immediately upon their emergence. (Johnston & Pennypacker, 1980, pp. 250–251)

Students of applied behavior analysis should not interpret this description of experimental design as a recommendation for a completely free-form approach to the manipulation of independent variables. The researcher must always pay close attention to the rule of changing only one variable at a time and must understand the opportunities for legitimate comparisons and the limitations that a given sequence of manipulations places on the conclusions that can be drawn from the results.

Experiments that use the reversal design to compare two or more treatments are vulnerable to confounding by sequence effects. **Sequence effects** are the effects on a subject's behavior in a given condition that are the result of the subject's experience with a prior condition. For example, caution must be used in interpreting the results from the A-B-C-B-C design that results from the following fairly common sequence of events in practice: After baseline (A), an initial treatment (B) is implemented and little or no behavioral improvements are noted. A second treatment (C) is then tried, and the behavior improves. A reversal is then conducted by reintroducing the first treatment (B), followed by reinstatement of the second treatment (C) (e.g., Foxx & Shapiro, 1978 [Figure 15.3). In this case, we can only speak knowingly about the effects of C when it follows B. Recapturing the original baseline levels of responding before introducing the second treatment condition (i.e., an A-B-A-C-A-C sequence) reduces the threat of sequence effects (or helps to expose them for what they are).

An A-B-A-B-C-B-C design, for instance, enables direct comparisons of B to A and C to B, but not of C to A. An experimental design consisting of A-B-A-B-B+C-B-B+C (e.g., Jason & Liotta, 1982) permits an evaluation of the additive or interactive effects of B+C, but does not reveal the independent contribution of C. And in both of these examples, it impossible to determine what effects, if any, C may have had on the behavior if it had been implemented prior to B. Manipulating each condition so that it precedes and follows every other condition in the experiment (e.g., A-B-A-B-C-B-C-A-C-A-C) is the only way to know for sure. However, manipulating multiple conditions requires a large amount of time and resources, and such extended designs become more susceptible to confounding by maturation and other historical variables not controlled by the experimenter.

NCR Reversal Technique

With interventions based on positive reinforcement, it can be hypothesized that observed changes in behavior are the result of the participant's feeling better about himself because of the improved environment created by the reinforcement, not because a specific response class has been immediately followed by contingent reinforcement. This hypothesis is most often advanced when interventions consisting of social reinforcement are involved. For example, a person may claim that it doesn't matter *how* the teacher's praise and attention were given; the student's behavior improved because the praise and attention created a warm and supporting environment. If, however, the behavioral improvements observed during a contingent reinforcement condition are lost during a condition when equal amounts of the same consequence are delivered independent of the occurrence of the target behavior, a functional relation between the reinforcement contingency and behavior change is demonstrated. In other words, such an experimental control technique can show that behavior change is the result of *contingent* reinforcement, not simply the presentation of or contact with the stimulus event (Thompson & Iwata, 2005).

A study by Baer and Wolf (1970a) on the effects of teachers' social reinforcement on the cooperative play of a preschool child provides an excellent example of the **NCR reversal technique** (Figure 8.6). The authors described the use and purpose of the design as follows:

> [The teachers first collected] baselines of cooperative and other related behaviors of the child, and of their own interaction with the child. Ten days of observation indicated that the child spent about 50% of each day in proximity with other children (meaning within 3 feet of them indoors, or 6 feet outdoors). Despite this frequent proximity, however, the child spent only about 2% of her day in cooperative play with these children. The teachers, it was found, interacted with this girl about 20% of the day, not all of it pleasant. The teachers, therefore, set up a period of intense social reinforcement, offered not for cooperative play but free of any response requirement at all: the teachers took turns standing near the girl, attending closely to her activities, offering her materials, and smiling and laughing with her in a happy and admiring manner. The results of 7 days of this noncontingent extravagance of social reinforcement were straightforward: the child's cooperative play changed not at all, despite the fact that the other children of the group were greatly attracted to the scene, offering the child nearly double the chance to interact with them cooperatively. These 7 days having produced no useful change, the teachers then began their planned reinforcement of cooperative behavior. . . . Contingent social reinforcement, used in amounts less than half that given during the noncontingent period, increased the child's cooperative play from its usual 2% to a high of 40% in the course of 12 days of reinforcement. At that point, in the interests of certainty, the teachers discontinued contingent reinforcement in favor of noncontingent. In the course of 4 days, they lost virtually all of the cooperative behavior they had gained during the reinforcement period of the study, the child showing about a 5% average of cooperative play over that period of time. Naturally, the study concluded with a return to the contingent use of social reinforcement, a recovery of desirable levels of cooperative play, and a gradual reduction of the teacher's role in maintaining that behavior. (pp. 14–15)

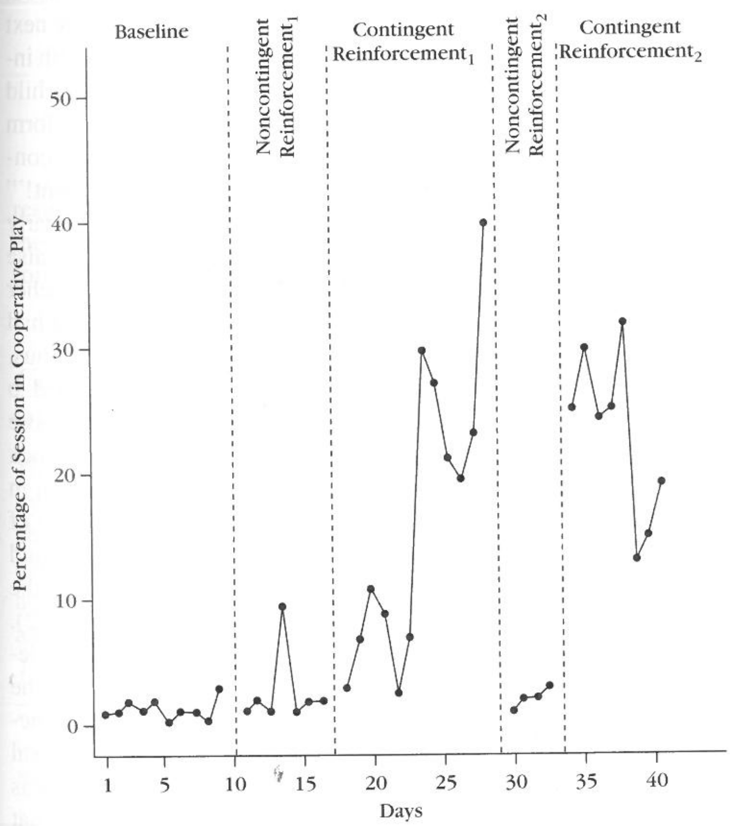


Figure 8.6 Reversal design using noncontingent reinforcement (NCR) as a control technique.

From "Recent Examples of Behavior Modification in Pre-School Settings" by D. M. Baer and M. M. Wolf in *Behavior Modification in Clinical Psychology*, pp. 14–15, edited by C. Neuringer and J. L. Michael, 1970, Upper Saddle River, NJ: Prentice Hall. Copyright 1970 by Prentice Hall. Adapted by permission.

Using NCR as a control conditions to demonstrate a functional relation is advantageous when it is not possible or appropriate to eliminate completely the event or activity used as a contingent reinforcement. For example, Lattal (1969) employed NCR as a control condition to "reverse" the effects of swimming as reinforcement for tooth brushing by children in a summer camp. In the contingent reinforcement condition, the campers could go swimming only if they had brushed their teeth; in the NCR condition, swimming was available whether or not tooth brushing occurred. The campers brushed their teeth more often in the contingent reinforcement condition.

The usual procedure is to deliver NCR on a fixed or variable time schedule independent of the subject's behavior. A potential weakness of the NCR control procedure becomes apparent when a high rate of the desired behavior has been produced during the preceding contingent reinforcement phase. It is probable in such situations that at least some instances of NCR, delivered according to a predetermined time schedule, will follow occurrences of the target behavior closely in time, and thereby function as adventitious, or "accidental reinforcement" (Thompson & Iwata, 2005). In fact, an intermittent schedule of reinforcement might be created inadvertently that results in even higher levels of performance than those obtained under contingent reinforcement. (Intermittent schedules of reinforcement and their effects are described in Chapter 13). In such cases the investigator might consider using one of the two control techniques described next, both of which involve "reversing" the behavioral focus of the contingency.[3]

DRO Reversal Technique

One way to ensure that reinforcement will not immediately follow the target behavior is to deliver reinforcement immediately following the subject's performance of *any behavior other than the target behavior.* With a **DRO reversal technique**, the control condition consists of delivering the event suspected of functioning as reinforcement following the emission of any behavior other than the target behavior (e.g., Baer, Peterson, & Sherman, 1967; Osbourne, 1969; Poulson, 1983). For example, Reynolds and Risley (1968) used contingent teacher attention to increase the frequency of talking in a 4-year-old girl enrolled in a preschool program for disadvantaged children. After a period of teacher attention contingent on verbalization, in which the girl's talking increased from a baseline average of 11% of the intervals observed to 75%, a DRO condition was implemented during which the teachers attended to the girl for any behavior except talking. During the 6 days of DRO, the girl's verbalization dropped to 6%. Teacher attention was then delivered contingent on talking, and the girl's verbalization "immediately increased to an average of 51%" (p. 259).

DRI/DRA Reversal Technique

During the control condition in a **DRI/DRA reversal technique**, occurrences of a specified behavior that is either incompatible with the target behavior (i.e., the two behaviors cannot possibly be emitted at the same time) or an alternative to the target behavior are immediately followed by the same consequence previously delivered as contingent reinforcement for the target behavior. Goetz and Baer's (1973) investigation of the effects of teacher praise on preschool children's creative play with building blocks illustrates the use of a DRI control condition. Figure 8.7 shows the number of different block forms (e.g., arch, tower, roof, ramp) constructed by the three children who participated in the study. During baseline (data points indicated by the letter *N*), "the teacher sat by the child as she built with the blocks, watching closely but quietly, displaying neither criticism nor enthusiasm about any particular use of the blocks" (p. 212). During the next phase (the *D* data points), "the teacher remarked with interest, enthusiasm, and delight every time that the child placed and/or rearranged the blocks so as to create a form that had not appeared previously in that session's construction(s). . . . 'Oh, that's very nice—that's different!'" (p. 212). Then, after increasing form diversity was clearly established, instead of merely withdrawing verbal praise and returning to the initial baseline condition, the teacher provided descriptive praise only when the children had constructed the same forms (the *S* data points). "Thus, for the next two to four sessions, the teacher continued to display interest, enthusiasm, and delight, but only at those times when the child placed and/or rearranged a block so as to create a repetition of a form already apparent in that session's construction(s). . . . Thus, no first usage of a form in a session was reinforced, but every second usage of that form and every usage thereafter within the session was. . . . 'How nice—another arch!'" (p. 212). The final phase of the experiment entailed a return to descriptive praise for different forms. Results show that the form diversity of children's block building was a function of teacher praise and comments. The DRI reversal tactic allowed Goetz and Baer to determine that it was not just the delivery of teacher praise and comment that

[3]Strictly speaking, using NCR as an experimental control technique to demonstrate that the contingent application of reinforcement is requisite to its effectiveness is not a separate variation of the A-B-A reversal design. Technically, the NCR reversal technique, as well as the DRO and DRI/DRA reversal techniques described next, is a multiple treatment design. For example, the Baer and Wolf (1970a) study of social reinforcement shown in Figure 8.6 used an A-B-C-B-C design, with B representing the NCR conditions and C representing the contingent reinforcement conditions.

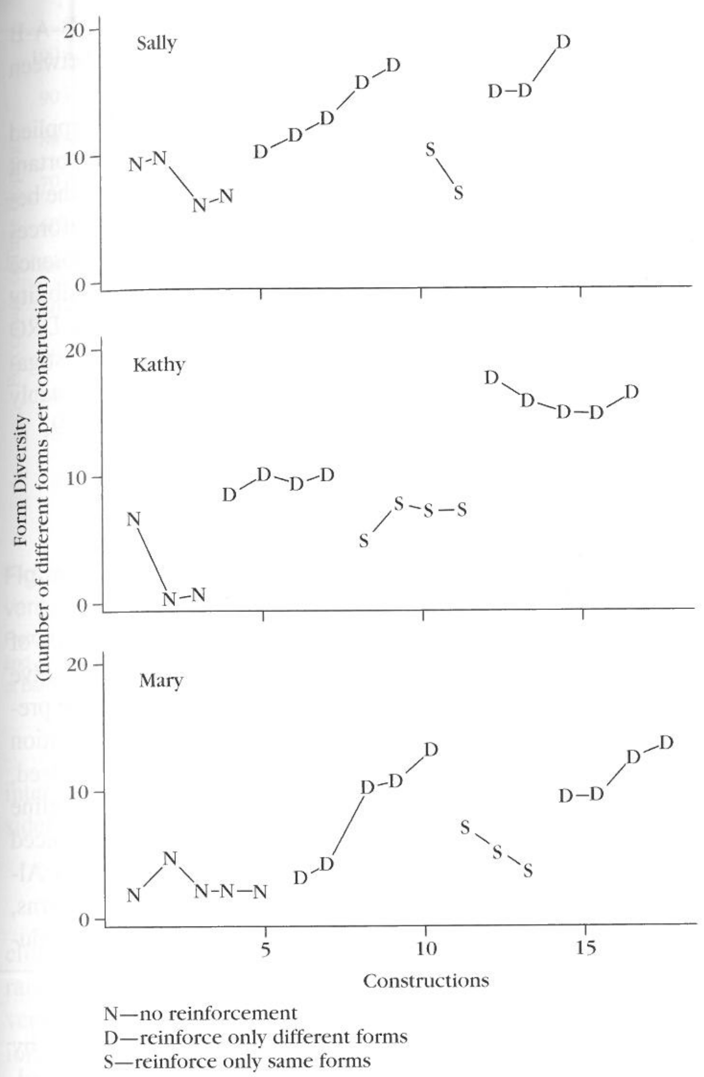


Figure 8.7 Reversal design using a DRI control technique.

From "Social Control of Form Diversity and the Emergence of New Forms in Children's Blockbuilding" by E. M. Goetz and D. M. Baer, 1973, *Journal of Applied Behavior Analysis, 6,* p. 213. Copyright 1973 by the Society for the Experimental Analysis of Behavior, Inc. Reprinted by permission.

resulted in more creative block building by the children; the praise and attention had to be contingent on different forms to produce increasing form diversity.[4]

Considering the Appropriateness of the Reversal Design

The primary advantage of the reversal design is its ability to provide a clear demonstration of the existence (or absence) of a functional relation between the independent and dependent variables. An investigator who reliably turns the target behavior on and off by presenting and withdrawing a specific variable makes a clear and convincing demonstration of experimental control. In addition, the reversal design enables quantification of the amount of behavior change over the preintervention level of responding. And the return to baseline provides information on the need to program for maintenance. Furthermore, a complete A-B-A-B design ends with the treatment condition in place.[5]

In spite of its strengths as a tool for analysis, the reversal design entails some potential scientific and social disadvantages that should be considered prior to its use. The considerations are of two types: irreversibility, which affects the scientific utility of the design; and the social, educational, and ethical concerns related to withdrawing a seemingly effective intervention.

[4]The extent to which the increased diversity of the children's block building can be attributed to the attention and praise ("That's nice") or the descriptive feedback (". . . that's different") in the teacher's comments cannot be determined from this study because social attention and descriptive feedback were delivered as a package.

[5]Additional manipulations in the form of the partial or sequential withdrawal of intervention components are made when it is necessary or desirable for the behavior to continue at its improved level in the absence of the complete intervention (cf., Rusch & Kazdin, 1981).

Irreversibility: A Scientific Consideration

A reversal design is not appropriate in evaluating the effects of a treatment variable that, by its very nature, cannot be withdrawn once it has been presented. Although independent variables involving reinforcement and punishment contingencies can be manipulated with some certainty—the experimenter either presents or withholds the contingency—an independent variable such as providing information or modeling, once presented, cannot simply be removed. For example, a reversal design would not be an effective element of an experiment investigating the effects of attending an in-service training workshop for teachers during which participants observed a master teacher use contingent praise and attention with students. After the participants have listened to the rationale for using contingent praise and attention and observed the master teacher model it, the exposure provided by that experience could not be withdrawn. Such interventions are said to be irreversible.

Irreversibility of the dependent variable must also be considered in determining whether a reversal would be an effective analytic tactic. Behavioral **irreversibility** means that a level of behavior observed in an earlier phase cannot be reproduced even though the experimental conditions are the same as they were during the earlier phase (Sidman, 1960). Once improved, many target behaviors of interest to the applied behavior analyst remain at their newly enhanced level even when the intervention responsible for the behavior change is removed. From a clinical or educational standpoint, such a state of affairs is desirable: The behavior change is shown to be durable, persisting even in the absence of continued treatment. However, irreversibility is a problem if demonstration of the independent variable's role in the behavior change depends on verification by recapturing baseline levels of responding.

For example, baseline observations might reveal very low, almost nonexistent, rates of talking and social interaction for a young child. An intervention consisting of teacher-delivered social reinforcement for talking and interacting could be implemented, and after some time the girl might talk to and interact with her peers at a frequency and in a manner similar to that of her classmates. The independent variable, teacher-delivered reinforcement, could be terminated in an effort to recapture baseline rates of talking and interacting. But the girl might continue to talk to and interact with her classmates even though the intervention, which may have been responsible for the initial change in her behavior, is withdrawn. In this case a source of reinforcement uncontrolled by the experimenter—the girl's classmates talking to and playing with her as a consequence of her increased talking and interacting with them—could maintain high rates of behavior after the teacher-delivered reinforcement is no longer provided. In such instances of irreversibility, an A-B-A-B design would fail to reveal a functional relation between the independent variable and the target behavior.

Nonetheless, one of the major objectives of applied behavior analysis is establishing socially important behavior through experimental treatments so that the behavior will contact natural "communities of reinforcement" to maintain behavioral improvements in the absence of treatment (Baer & Wolf, 1970b). When irreversibility is suspected or apparent, in addition to considering DRO or DRI/DRA conditions as control techniques, investigators can consider other experimental tactics, most notably the multiple baseline designs described in Chapter 9.

Withdrawing an Effective Intervention: A Social, Educational, and Ethical Consideration

Although it can yield an unambiguous demonstration of experimental control, withdrawing a seemingly effective intervention to evaluate its role in behavior change presents a legitimate cause for concern. One must question the appropriateness of any procedure that allows (indeed, seeks) an improved behavior to deteriorate to baseline levels of responding. Various concerns have been voiced over this fundamental feature of the reversal design. Although there is considerable overlap among the concerns, they can be classified as having primarily a social, educational, or ethical basis.

Social Concerns. Applied behavior analysis is, by definition, a social enterprise. Behaviors are selected, defined, observed, measured, and modified by and for people. Sometimes the people involved in an applied behavior analysis—administrators, teachers, parents, and participants—object to the withdrawal of an intervention they associate with desirable behavior change. Even though a reversal may provide the most unqualified picture of the behavior–environment relation under study, it may not be the analytic tactic of choice because key participants do not want the intervention to be withdrawn. When a reversal design offers the best experimental approach scientifically and poses no ethical problems, the behavior analyst may choose to explain the operation and purpose of the tactic to those who do not favor it. But it is unwise to attempt a reversal without the full support of the people involved, especially those who will be responsible for withdrawing the intervention (Tawney & Gast, 1984). Without their cooperation the procedural integrity of the experiment could easily be compromised. For example, people who are against the withdrawal of treatment might sabotage the return to baseline conditions by implementing the

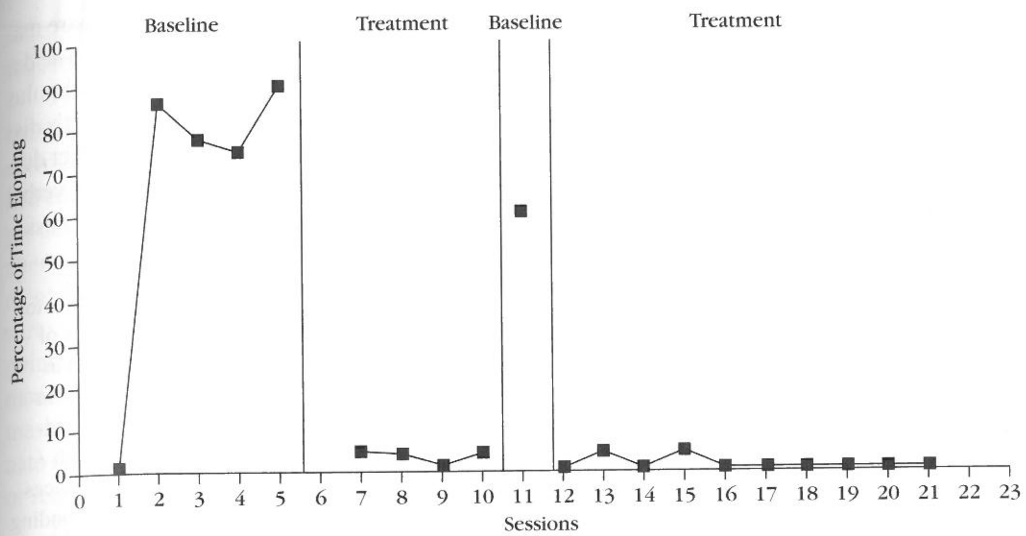


Figure 8.8 Reversal design with a single-session return-to-baseline probe to evaluate and verify effects of treatment for a potentially dangerous behavior.

From "Functional Analysis and Treatment of Elopement for a Child with Attention Deficit Hyperactivity Disorder" by T. Kodak, L. Grow, and J. Northrup, 2004, *Journal of Applied Behavior Analysis, 37,* p. 231. Copyright 2004 by the Society for the Experimental Analysis of Behavior, Inc. Reprinted by permission.

intervention, or at least those parts of it that they consider the most important.

Educational and Clinical Issues. Educational or clinical issues concerning the reversal design are often raised in terms of instructional time lost during the reversal phases, as well as the possibility that the behavioral improvements observed during intervention may not be recaptured when treatment is resumed after a return to baseline conditions. We agree with Stolz (1978) that "extended reversals are indefensible." If preintervention levels of responding are reached quickly, reversal phases can be quite short in duration. Sometimes only three or four sessions are needed to show that initial baseline rates have been reproduced (e.g., Ashbaugh & Peck, 1998 [Figure 15.7]; Cowdery Iwata, & Pace, 1990 [Figure 22.6]). Two or three brief reversals can provide an extremely convincing demonstration of experimental control. Concern that the improved levels of behavior will not return when the treatment variable is reintroduced, while understandable, has not been supported by empirical evidence. Hundreds of published studies have shown that behavior acquired under a given set of environmental conditions can be reacquired rapidly during subsequent reapplication of those conditions.

Ethical Concerns. A serious ethical concern must be addressed when the use of a reversal design is considered for evaluating a treatment for self-injurious or dangerous behaviors. With mild self-injurious or aggressive behaviors, short reversal phases consisting of one or two baseline probes can sometimes provide the empirical evidence needed to reveal a functional relation (e.g., Kelley, Jarvie, Middlebrook, McNeer, & Drabman, 1984; Luce, Delquadri, & Hall, 1980; Murphy et al., 1979 [Figure 8.4]). For example, in their study evaluating a treatment for elopement (i.e., running away from supervision) by a child with attention-deficient/hyperactivity disorder, Kodak, Grow, and Northrup (2004) returned to baseline conditions for a single session (see Figure 8.8).

Nonetheless, with some behaviors it may be determined that withdrawing an intervention associated with improvement for even a few one-session probes would be inappropriate for ethical reasons. In such cases experimental designs that do not rely on the reversal tactic must be used.

Alternating Treatments Design

An important and frequently asked question by teachers, therapists, and others who are responsible for changing behavior is, Which of these treatments will be most effective with this student or client? In many situations, the research literature, the analyst's experience, and/or logical extensions of the principles of behavior point to several possible interventions. Determining which of several possible treatments or combination of treatments will produce the greatest improvement in behavior is a primary task for applied behavior analysts. As described earlier, although a multiple treatment reversal design (e.g., A-B-C-B-C) can be used to compare the effects of two or more treatments, such designs have some inherent limitations. Because the different treatments in a multiple treatment reversal design are implemented during separate phases that occur in a particular order, the design is particularly vulnerable to confounding because of sequence effects (e.g., Treatment C may have produced its effect only because it followed Treatment B, not because it was more robust in its own right). A second disadvantage of comparing multiple treatments with the reversal tactic is the extended time required to demonstrate differential effects. Most behaviors targeted for change by teachers and therapists are selected because they need immediate improvement. An experimental design that will quickly reveal the most effective treatment among several possible approaches is important for the applied behavior analyst.

The alternating treatments design provides an experimentally sound and efficient method for comparing the effects of two or more treatments. The term *alternating treatments design,* proposed by Barlow and Hayes (1979), accurately communicates the operation of the design. Other terms used in the applied behavior analysis literature to refer to this analytic tactic include ***multielement design*** (Ulman & Sulzer-Azaroff, 1975), *multiple schedule design* (Hersen & Barlow, 1976), *concurrent schedule design* (Hersen & Barlow, 1976), and *simultaneous treatment design* (Kazdin & Hartmann, 1978).[6]

Operation and Logic of the Alternating Treatments Design

The **alternating treatments design** is characterized by the rapid alternation of two or more distinct treatments (i.e., independent variables) while their effects on the target behavior (i.e., dependent variable) are measured. In contrast to the reversal design in which experimental manipulations are made after steady state responding is achieved in a given phase of an experiment, the different interventions in an alternating treatments design are manipulated independent of the level of responding. The design is predicated on the behavioral principle of stimulus discrimination (see Chapter 17). To aid the subject's discrimination of which treatment condition is in effect during a given session, a distinct stimulus (e.g., a sign, verbal instructions, different colored worksheets) is often associated with each treatment.

> The data are plotted separately for each intervention to provide a ready visual representation of the effects of each treatment. Because confounding factors such as time of administration have been neutralized (presumably) by counterbalancing, and because the two treatments are readily discriminable by subjects through instructions or other discriminative stimuli, differences in the individual plots of behavior change corresponding with each treatment should be attributable to the treatment itself, allowing a direct comparison between two (or more) treatments. (Barlow & Hayes, 1979, p. 200)

Figure 8.9 shows a graphic prototype of an alternating treatments design comparing the effects of two treatments, A and B, on some response measure. In an alternating treatments design, the different treatments can be alternated in a variety of ways. For example, the treatments might be (a) alternated across daily sessions, one treatment in effect each day; (b) administered in separate sessions occurring within the same day; or (c) implemented each during a portion of the same session. Counterbalancing the days of the week, times of day, sequence in which the different treatments occur (e.g., first or second each day), persons delivering the different treatments, and so forth, reduces the probability that any observed differences in behavior are the result of variables other than the treatments themselves. For example, assume that Treatments A and B in Figure 8.9 were each administered

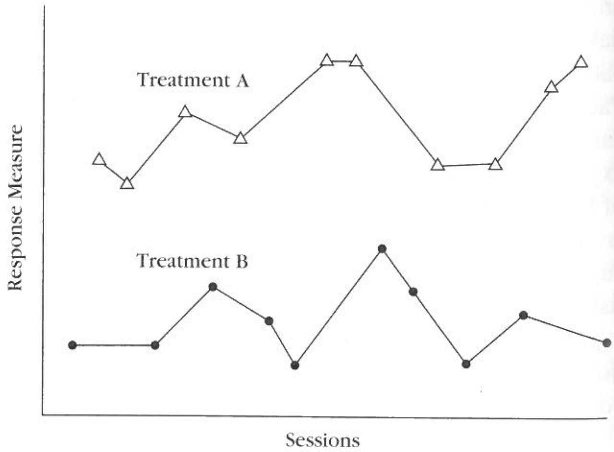


Figure 8.9 Graphic prototype of an alternating treatments design comparing the differential effects of two treatments (A and B).

[6] A design in which two or more treatments are concurrently or simultaneously presented, and in which the subject chooses between treatments, is correctly termed a concurrent schedule or simultaneous treatment design. Some published studies described by their authors as using a simultaneous treatment design have, in fact, employed an alternating treatments design. Barlow and Hayes (1979) could find only one true example of a simultaneous treatment design in the applied literature: a study by Browning (1967) in which three techniques for reducing the bragging of a 10-year-old boy were compared.

for a single 30-minute session each day, with the daily sequence of the two treatments determined by a coin flip.

The data points in Figure 8.9 are plotted on the horizontal axis to reflect the actual sequence of treatments each day. Thus, the horizontal axis is labeled *Sessions,* and each consecutive pair of sessions occurred on a single day. Some published reports of experiments using an alternating treatments design in which two or more treatments were presented each day (or session) plot the measures obtained during each treatment above the same point on the horizontal axis, thus implying that the treatments were administered simultaneously. This practice masks the temporal order of events and has the unfortunate consequence of making it difficult for the researcher or reader to discover potential sequence effects.

The three components of steady state strategy—prediction, verification, and replication—are found in the alternating treatments design. However, each component is not readily identified with a separate phase of the design. In an alternating treatments design, each successive data point for a specific treatment plays all three roles: It provides (a) a basis for the *prediction* of future levels of responding under that treatment, (b) potential *verification* of the previous prediction of performance under that treatment, and (c) the opportunity for *replication* of previous effects produced by that treatment.

To see this logic unfold, the reader should place a piece of paper over all the data points in Figure 8.9 except those for the first five sessions of each treatment. The visible portions of the data paths provide the basis for predicting future performance under each respective treatment. Moving the paper to the right reveals the two data points for the next day, each of which provides a degree of verification of the previous predictions. As more data are recorded, the predictions of given levels of responding within each treatment are further strengthened by continued verification (if those additional data conform to the same level and/or trend as their predecessors). Replication occurs each time Treatment A is reinstated and measurement reveals responding similar to previous Treatment A measures and different from those obtained when Treatment B is in effect. Likewise, another mini-replication is achieved each time a reintroduction of Treatment B results in measures similar to previous Treatment B measures and different from Treatment A levels of responding. A consistent sequence of verification and replication is evidence of experimental control and strengthens the investigator's confidence of a functional relation between the two treatments and different levels of responding.

The presence and degree of experimental control in an alternating treatments design is determined by visual inspection of the differences between (or among) the data paths representing the different treatments. Experimental control is defined in this instance as objective, believable evidence that different levels of responding are predictably and reliably produced by the presence of the different treatments. When the data paths for two treatments show no overlap with each other and either stable levels or opposing trends, a clear demonstration of experimental control has been made. Such is the case in Figure 8.9, in which there is no overlap of data paths and the picture of differential effects is clear. When some overlap of data paths occurs, a degree of experimental control over the target behavior can still be demonstrated if the majority of data points for a given treatment fall outside the range of values of the majority of data points for the contrasting treatment.

The extent of any differential effects produced by two treatments is determined by the vertical distance—or fractionation—between their respective data paths and quantified by the vertical axis scale. The greater the vertical distance, the greater the differential effect of the two treatments on the response measure. It is possible for experimental control to be shown between two treatments but for the amount of behavior change to be socially insignificant. For instance, experimental control may be demonstrated for a treatment that reduces a person's severe self-injurious behavior from 10 occurrences per hour to 2 per hour, but the participant is still engaged in self-mutilation. However, if the vertical axis is scaled meaningfully, the greater the separation of data paths on the vertical axis, the higher the likelihood that the difference represents a socially significant effect.

Data from an experiment that compared the effects of two types of group-contingent rewards on the spelling accuracy of fourth-grade underachievers (Morgan, 1978) illustrate how the alternating treatments design reveals experimental control and the quantification of differential effects. The six children in the study were divided into two equally skilled teams of three on the basis of pretest scores. Each day during the study the students took a five-word spelling test. The students received a list of the words the day before, and a 5-minute study period was provided just prior to the test. Three different conditions were used in the alternating treatments design: (a) *no game,* in which the spelling tests were graded immediately and returned to the students, and the next scheduled activity in the school day was begun; (b) *game,* in which test papers were graded immediately, and each member of the team who had attained the highest total score received a mimeographed Certificate of Achievement and was allowed to stand up and cheer; and (c) *game plus,* consisting of the same procedure as the game condition, plus each student on the winning team also received a small trinket (e.g., a sticker or pencil).

The results for Student 3 (see Figure 8.10) show that experimental control over spelling accuracy was obtained

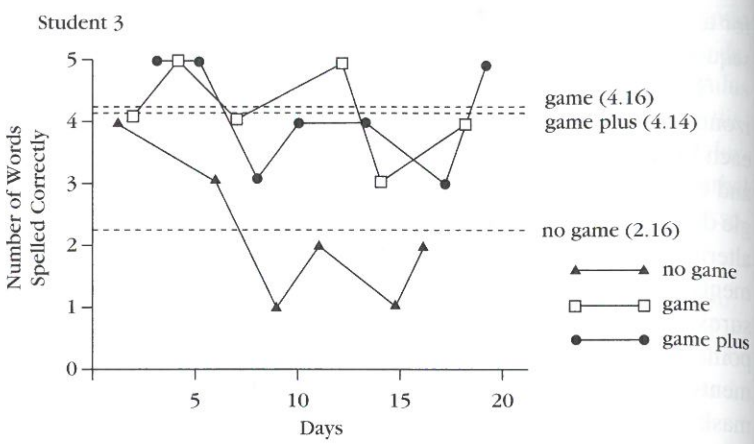


Figure 8.10 Alternating treatments design comparing the effects of three different treatments on the spelling accuracy of a fourth-grade student.

From *Comparison of Two "Good Behavior Game" Group Contingencies on the Spelling Accuracy of Fourth-Grade Students* by Q. E. Morgan, 1978, unpublished master's thesis, The Ohio State University. Reprinted by permission.

between the no-game condition and both the game and the game-plus conditions. Only the first two no-game data points overlap the lower range of scores obtained during the game or the game-plus conditions. However, the data paths for the game and game-plus conditions overlap completely and continuously throughout the study, revealing no difference in spelling accuracy between the two treatments. The vertical distance between the data paths represents the amount of improvement in spelling accuracy between the no-game condition and the game and the game-plus conditions. The mean difference between the two game conditions and the no-game condition was two words per test. Whether such a difference represents a significant improvement is an educational question, not a mathematical or statistical one, but most educators and parents would agree that an increase of two words spelled correctly out of five is socially significant, especially if that gain can be sustained from week to week. The cumulative effect over a 180-day school year would be impressive. There was virtually no difference in Student 3's spelling performance between the game and game-plus conditions. However, even a larger mean difference would not have contributed to the conclusions of the study because of the lack of experimental control between the game and the game-plus treatments.

Student 6 earned consistently higher spelling scores in the game-plus condition than he did in the game or no-game conditions (see Figure 8.11). Experimental control was demonstrated between the game-plus and the other two treatments for Student 6, but not between the no-game and game conditions. Again, the difference in responding between treatments is quantified by the vertical distance between the data paths. In this case there was a mean difference of 1.55 correctly spelled words per test between the game-plus and no-game conditions.

Figures 8.10 and 8.11 illustrate two other important points about the alternating treatments design. First, the two graphs show how an alternating treatments design enables a quick comparison of interventions. Although the study would have been strengthened by the collection of additional data, after 20 sessions the teacher had sufficient empirical evidence for selecting the most effective consequences for each student. If only two conditions had been compared, even fewer sessions may have been required to identify the most effective intervention. Second, these data underscore the importance of evaluating treatment effects at the level of the individual subject. All six

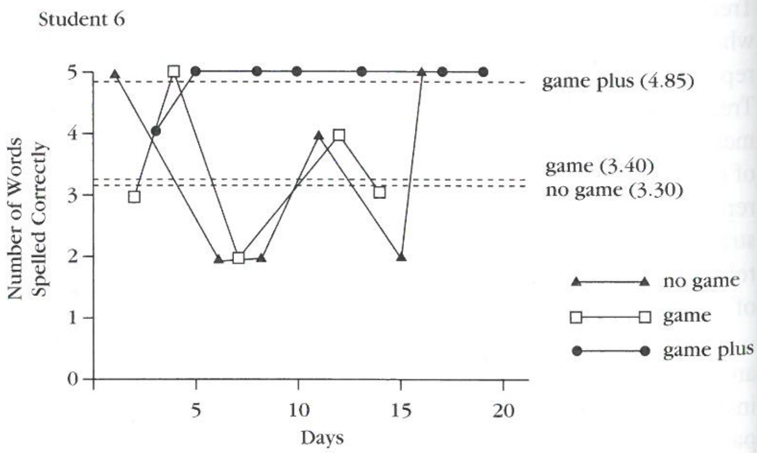


Figure 8.11 Alternating treatments design comparing the effects of three different treatments on the spelling accuracy of a fourth-grade student.

From *Comparison of Two "Good Behavior Game" Group Contingencies on the Spelling Accuracy of Fourth-Grade Students* by Q. E. Morgan, 1978, unpublished master's thesis, The Ohio State University. Reprinted by permission.

children spelled more words correctly under one or both of the game conditions than they did under the no-game condition. However, Student 3's spelling accuracy was equally enhanced by either the game or the game-plus contingency, whereas Student 6's spelling scores improved only when a tangible reward was available.

Variations of the Alternating Treatments Design

The alternating treatments design can be used to compare one or more treatments to a no-treatment or baseline condition, assess the relative contributions of individual components of a package intervention, and perform parametric investigations in which different values of an independent variable are alternated to determine differential effects on behavior change. Among the most common variations of the alternating treatments design are the following:

- Single-phase alternating treatments design without a no-treatment control condition
- Single-phase design in which two or more conditions, one of which is a no-treatment control condition, are alternated
- Two-phase design consisting of an initial baseline phase followed by a phase in which two or more conditions (one of which may be a no-treatment control condition) are alternated
- Three-phase design consisting of an initial baseline, a second phase in which two or more conditions (one of which may be a no-treatment control condition) are alternated, and a final phase in which only the treatment that proved most effective is implemented

Alternating Treatments Design without a No-Treatment Control Condition

One application of the alternating treatments design consists of a single-phase experiment in which the effects of two or more treatment conditions are compared (e.g., Barbetta, Heron, & Heward, 1993; McNeish, Heron, & Okyere, 1992; Morton, Heward, & Alber, 1998). A study by Belfiore, Skinner, and Ferkis (1995) provides an excellent example of this design. They compared the effects of two instructional procedures—trial-repetition and response-repetition—on the acquisition of sight words by three elementary students with learning disabilities in reading. An initial training list of five words for each condition was created by random selection from a pool of unknown words (determined by pretesting each student). Each session began with a noninstructional assessment of unknown and training words, followed by both conditions. The order of instructional conditions was counterbalanced across sessions. Words spoken correctly on three consecutive noninstructional assessments were considered mastered and replaced as training words with unknown words.

The trial-repetition condition consisted of one response opportunity within each of five interspersed practice trials per word. The experimenter placed a word card on the table and said, "Look at the word, and say the word." If the student made a correct response within 3 seconds, the experimenter said, "Yes, the word is _____." (p. 347). If the student's initial response was incorrect, or the student made no response within 3 seconds, the experimenter said, "No, the word is _____," and the student repeated the word. The experimenter then presented the next word card and the procedure repeated until five practice trials (antecedent-response-feedback) were provided with each word.

The response-repetition condition also consisted of five response opportunities per word, but all five responses occurred within a single practice trial for each word. The experimenter placed a word card on the table and said, "Look at the word, and say the word." If the student made a correct response within 3 seconds, the experimenter said, "Yes, the word is _____, please repeat the word four more times" (p. 347). If the student made an incorrect response or no response within 3 seconds, the experimenter said, "No, the word is _____." The student then repeated the word and was instructed to repeat it four more times.

Figure 8.12 shows the cumulative number of words mastered by each student under both conditions. Even though the number of correct responses per word during instruction was identical in both conditions, all three students had higher rates of learning new words in the trial-repetition condition than in the response-repetition condition. These results obtained within the simple alternating treatments design enabled Belfiore and colleagues (1995) to conclude that, "Response repetition outside the context of the learning trial (i.e., of the three-term contingency) was not as effective as repetition that included antecedent and consequent stimuli in relation to the accurate response" (p. 348).

Alternating Treatments Design with No-Treatment Control Condition

Although not a requirement of the design, a no-treatment condition is often incorporated into the alternating treatments design as one of the treatments to be compared. For example, the no-game condition in the Morgan (1978) study served as a no-treatment control condition against which the students' spelling scores in the game

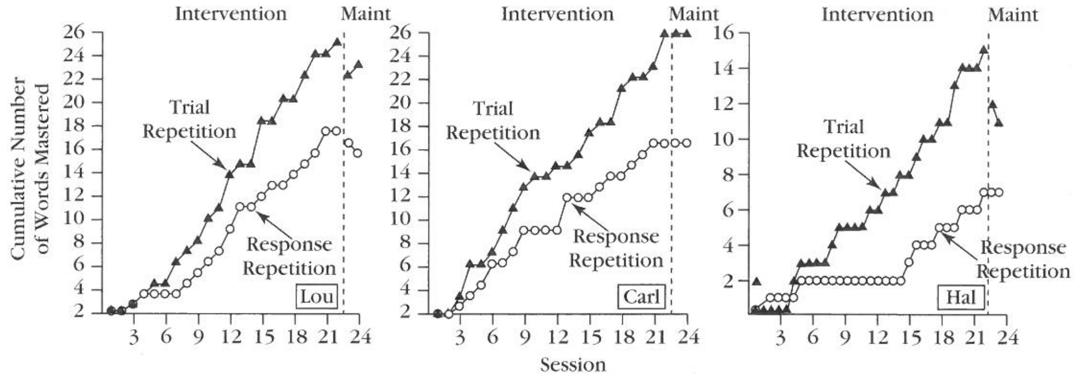


Figure 8.12 Single-phase alternating treatments design without a no-treatment control condition.

From "Effects of Response and Trial Repetition on Sight-Word Training for Students with Learning Disabilities" by P. J. Belfiore, C. H. Skinner, and M. A. Ferkis, 1995, *Journal of Applied Behavior Analysis, 28,* p. 348. Copyright 1995 by the Society for the Experimental Analysis of Behavior, Inc. Reprinted by permission.

and game-plus conditions were compared (see Figures 8.10 and 8.11).

Including a no-treatment control condition as one of the experimental conditions in an alternating treatments design provides valuable information on any differences in responding under the intervention treatment(s) and no treatment. However, the measures obtained during the no-treatment control condition should not be considered representative of an unknown preintervention level of responding. It may be that the measures obtained in the no-treatment condition represent only the level of behavior under a no-treatment condition when it is interspersed within an ongoing series of treatment condition(s), and do not represent the level of behavior that existed before the alternating treatments design was begun.

Alternating Treatments Design with Initial Baseline

Investigators using the alternating treatments tactic often use a two-phase experimental design in which baseline measures are collected until a stable level of responding or countertherapeutic trend is obtained prior to the alternating treatments phase (e.g., Martens, Lochner, & Kelly, 1992 [see Figure 13.6]). Sometimes the baseline condition is continued during the alternating treatments phase as a no-treatment control condition.

A study by J. Singh and N. Singh (1985) provides an excellent example of an alternating treatments design incorporating an initial baseline phase. The experiment evaluated the relative effectiveness of two procedures for reducing the number of oral reading errors by students with mental retardation. The first phase of the study consisted of a 10-day baseline condition in which each student was given a new 100-word passage three times each day and told, "Here is the story for this session. I want you to read it. Try your best not to make any errors" (p. 66). The experimenter sat nearby but did not assist the student, correct any errors, or attend to self-corrections. If a student requested help with new or difficult words, he was prompted to continue reading.

During the alternating treatments phase of the study, three different conditions were presented each day in separate sessions of about 5 minutes each: control (the same procedures as during baseline), word supply, and word analysis. To minimize any sequence or carryover effects from one condition to another, the three conditions were presented in random order each day, each condition was preceded with specific instructions identifying the procedure to be implemented, and an interval of at least 5 minutes separated consecutive sessions. During the word-supply condition, each student was instructed, "Here is the story for this session. I want you to read it. I will help you if you make a mistake. I will tell you the correct word while you listen and point to the word in the book. After that, I want you to repeat the word. Try your best not to make any errors" (p. 67). The experimenter supplied the correct word when an oral reading error was made, had the child repeat the correct word once, and instructed the child to continue reading. During the word-analysis condition, each student was instructed, "Here is the story for this session. I want you to read it. I will help you if you make a mistake. I will help you sound out the word and then you can read the word correctly before you carry on reading the rest of the story. Try your best not to make any errors" (p. 67). When errors were made in this condition, the experimenter directed the child's attention to the phonetic elements of the word and coaxed the child to sound out correctly each part of the word. Then the experimenter had the student read the entire word at the normal speed and instructed him or her to continue reading the passage.

The results for the four students who participated in the study are shown in Figure 8.13. Each baseline data

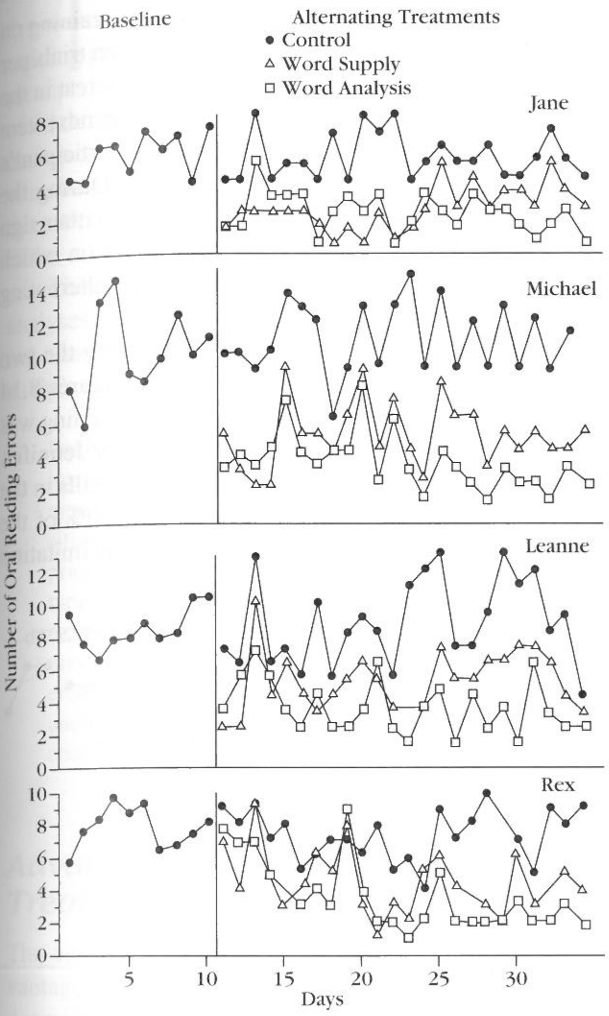


Figure 8.13 Alternating treatments design with an initial baseline.

From "Comparison of Word-Supply and Word-Analysis Error-Correction Procedures on Oral Reading by Mentally Retarded Children" by J. Singh and N. Singh, 1985, *American Journal of Mental Deficiency, 90,* p. 67. Copyright 1985 by the *American Journal of Mental Deficiency.* Reprinted by permission.

point is the mean number of errors for the three daily sessions. Although the data in each condition are highly variable (perhaps because of the varied difficulty of the different passages used), experimental control is evident. All four students committed fewer errors during the word-supply and the word-analysis conditions than they did during the control condition. Experimental control of oral reading errors, although not complete because of some overlap of the data paths, is also demonstrated between the word-supply and word-analysis conditions, with all four students making fewer errors during the word-analysis condition.

By beginning the study with a baseline phase, J. Singh and N. Singh (1985) were able to compare the level of responding obtained during each of the treatments to the natural level of performance uncontaminated by the introduction of either error-correction intervention. In addition, the initial baseline served as the basis for predicting and assessing the measures obtained during the control sessions of the alternating treatments phase of the study. The measures obtained in the alternating control condition matched the relatively high frequency of errors observed during the initial baseline phase, providing evidence that (a) the vertical distance between the data paths for the word-supply and word-analysis conditions and the data path for the control condition represents the true amount of improvement produced by each treatment and (b) the frequency of errors during the control condition was not influenced by reduced errors during the other two treatments (i.e., no generalized reduction in oral reading errors from the treated passages to untreated passages occurred).

Alternating Treatments Design with Initial Baseline and Final Best Treatment Phase

A widely used variation of the alternating treatments design consists of three sequential phases: an initial baseline phase, a second phase comparing alternating treatments, and a final phase in which only the most effective treatment is administered (e.g., Heckaman, Alber, Hooper, & Heward, 1998; Kennedy & Souza, 1995, Study 4; Ollendick, Matson, Esvelt-Dawson, & Shapiro, 1980; N. Singh, 1990; N. Singh & J. Singh, 1984; N. Singh & Winton, 1985). Tincani (2004) used an alternating treatments design with an initial baseline and final best treatment phase to investigate the relative effectiveness of sign language and picture exchange training on the acquisition of mands (requests for preferred items) by two children with autism.[7] A related research question was whether a relation existed between students' preexisting motor imitation skills and their abilities to learn mands through sign language or by picture exchange. Two assessments were conducted for each student prior to baseline. A stimulus preference assessment (Pace, Ivancic, Edwards, Iwata, & Page, 1985) was conducted to identify a list of 10 to 12 preferred items (e.g., drinks, edibles, toys), and each student's ability to imitate 27 hand, arm, and finger movements similar to those required for sign language was assessed.[8]

The purpose of baseline was to ensure that the participants were not able to request preferred items with picture exchange, sign language, or speech prior to training. Baseline trials consisted of giving the student 10 to 20 seconds of noncontingent access to a preferred item, removing the item briefly, and then placing it out of the

[7]The mand is one of six types of elementary verbal operants identified by Skinner (1957). Chapter 25 describes Skinner's analysis of verbal behavior and its importance to applied behavior analysis.

[8]Stimulus preference assessment procedures are described in Chapter 11.

student's reach. A laminated 2-inch-by-2-inch picture of the item was placed in front of the student. If the student placed the picture symbol in the experimenter's hand, signed the name of the item, or spoke the name of the item within 10 seconds, the experimenter provided access to the item. If not, the item was removed and the next item on the list was presented. Following a three-session baseline, during which neither participant emitted an independent mand in any modality, the alternating treatments phase was begun.

The sign language training procedures were adapted from Sundberg and Partington's (1998) *Teaching Language to Children with Autism or Other Developmental Disabilities.* The simplest sign from American Sign Language for each item was taught. Procedures used in the PECS training condition were adapted from Bondy and Frost's (2002) *The Picture Exchange Communication System Training Manual.* In both conditions, training on each preferred item continued for five to seven trials per session, or until the participant showed no interest in the item. At that time training then began on the next item and continued until all 10 or 12 items on the participant's list of preferred items had been presented. During the study's final phase, each participant received either sign language or PECS training only, depending on which method had been most successful during the alternating treatments phase.

The percentage of independent mands by the two students throughout the study is shown in Figures 8.14 (Jennifer) and 8.15 (Carl). Picture exchange training was clearly more effective than sign language for Jennifer. Jennifer demonstrated weak motor imitation skills in the prebaseline assessment, correctly imitating 20% of the motor movements attempted in the prebaseline imitation

Figure 8.14 Alternating treatments design with an initial baseline and a final best-treatment-only condition.

From "Comparing the Picture Exchange Communication System and Sign Language Training for Children with Autism" by M. Tincani, 2004, *Focus on Autism and Other Developmental Disabilities, 19,* p. 160. Copyright 2004 by Pro-Ed. Used by permission.

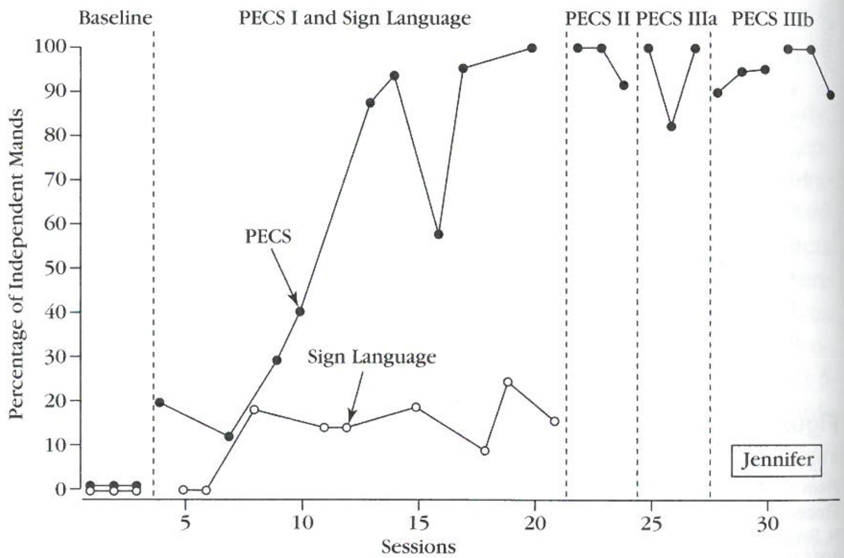
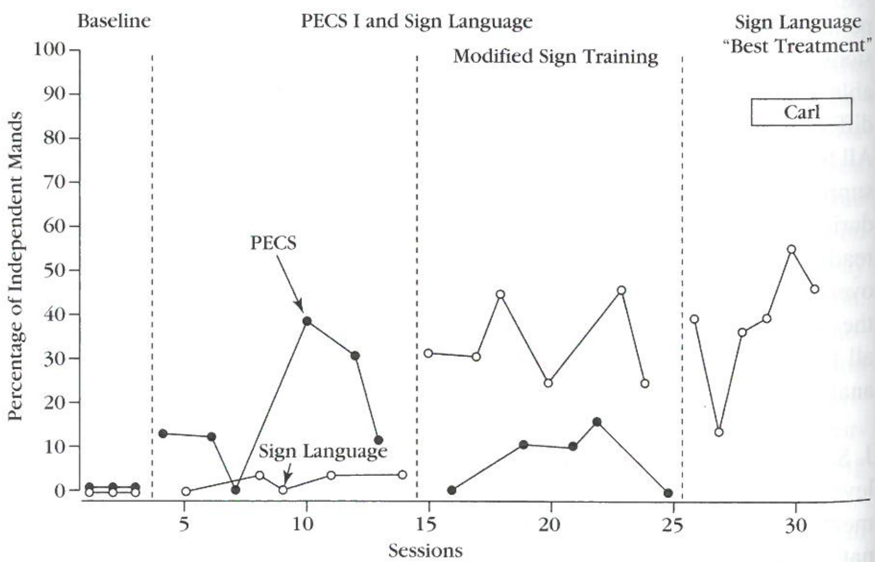


Figure 8.15 Alternating treatments design with an initial baseline and a final best-treatment-only condition.

From "Comparing the Picture Exchange Communication System and Sign Language Training for Children with Autism" by M. Tincani, 2004, *Focus on Autism and Other Developmental Disabilities, 19,* p. 159. Copyright 2004 by Pro-Ed. Used by permission.



assessment. After a slight modification in sign language training procedures was implemented to eliminate Carl's prompt dependency, he emitted independent mands more often during sign language training than with picture exchange training. Carl's preexisting motor imitation skills were better than Jennifer's. He imitated correctly 43% of the attempted motor movements in the prebaseline imitation assessment.

This study highlights the importance of individual analyses and exploring the possible influence of variables not manipulated during the study. In discussing the study's results, Tincani (2004) noted that

> For learners without hand-motor imitation skills, including many children with autism, PECS training may be more appropriate, at least in terms of initial mand acquisition. Jennifer had weak hand-motor imitation skills prior to intervention and learned picture exchange more rapidly than sign language. For learners who have moderate hand-motor imitation skills, sign language training may be equally, if not more, appropriate. Carl had moderate hand-motor imitation skills prior to intervention and learned sign language more rapidly than picture exchange. (p. 160)

Advantages of the Alternating Treatments Design

The alternating treatments design offers numerous advantages for evaluating and comparing two or more independent variables. Most of the benefits cited here were described by Ulman and Sulzer-Azaroff (1975), who are credited with first bringing the rationale and possibilities of the alternating treatments design to the attention of the applied behavior analysis community.

Does Not Require Treatment Withdrawal

A major advantage of the alternating treatments design is that it does not require the investigator to withdraw a seemingly effective treatment to demonstrate a functional relation. Reversing behavioral improvements raises ethical issues that can be avoided with the alternating treatments design. Regardless of ethical concerns, however, administrators and teachers may be more likely to accept an alternating treatments design over a reversal design even when one of the alternating treatments is a no-treatment control condition. "It would appear that a return to baseline conditions every other day or every third day is not as disagreeable to a teacher as is first establishing a high level of desirable behavior for a prolonged period, and then reinstating the baseline behaviors" (Ulman & Sulzer-Azaroff, 1975, p. 385).

Speed of Comparison

The experimental comparison of two or more treatments can often be made quickly with the alternating treatments design. In one study an alternating treatments design enabled the superiority of one treatment over another in increasing the cooperative behavior of a 6-year-old boy to be determined after only 4 days (McCullough, Cornell, McDaniel, & Mueller, 1974). The alternating treatments design's ability to produce useful results quickly is a major reason that it is the basic experimental tactic used in functional behavior analysis (see Chapter 24 and Figures 24.4, 24.5, 24.6, and 24.9).

When the effects of different treatments become apparent early in an alternating treatments design, the investigator can then switch to programming only the most effective treatment. The efficiency of the alternating treatments design can leave a researcher with meaningful data even when an experiment must be terminated early (Ulman & Sulzer-Azaroff, 1975). A reversal or multiple baseline design, on the other hand, must be carried through to completion to show a functional relation.

Minimizes Irreversibility Problem

Some behaviors, even though they have been brought about or modified by the application of the intervention, do not return to baseline levels when the intervention is withdrawn and thereby resist analysis with an A-B-A-B design. However, rapidly alternating treatment and no-treatment (baseline) conditions may reveal differences in responding between the two conditions, especially early in an experiment before responding in the no-treatment condition begins to approximate the level of responding in the treatment condition.

Minimizes Sequence Effects

An alternating treatments design, when properly conducted, minimizes the extent to which an experiment's results are confounded by sequence effects. Sequence effects pose a threat to the internal validity of any experiment, but especially to those involving multiple treatments. The concern over sequence effects can be summed up by this simple question: Would the results have been the same if the sequence of treatments had been different? Sequence effects can be extremely difficult to control in experiments using reversal or multiple tactics (see Chapter 9) to compare two or more independent variables because each experimental condition must remain in effect for a fairly long period of time, thereby producing a specific sequence of events. However, in an alternating treatments design, the independent variables are rapidly alternated with one another in a random fashion that produces no particular sequence. Also, each treatment is in effect for short periods of time, reducing the likelihood of carryover effects (O'Brien, 1968). The ability to minimize sequence effects makes the alternating treatments design a powerful tool for achieving complex behavior analyses.

Can Be Used with Unstable Data

Determining functional behavior–environment relations in the presence of unstable data presents a serious problem for the applied behavior analyst. Using steady state responding to predict, verify, and replicate behavioral changes is the foundation of experimental reasoning in behavior analysis (Sidman, 1960). Obtaining stable baseline responding, however, is extremely difficult with many socially important behaviors of interest to applied behavior analysts. Merely providing a subject with repeated opportunities to emit a target response can result in gradually improved performance. Although practice effects are worthy of empirical investigation because of their applied and scientific importance (Greenwood, Delquadri, & Hall, 1984; Johnston & Pennypacker, 1993a), the unstable baselines they create pose problems for the analysis of intervention variables. The changing levels of task difficulty inherent in moving through a curriculum of progressively more complex material also make obtaining steady state responding for many academic behaviors difficult.

Because the different treatment conditions are alternated rapidly in an alternating treatments design, because each treatment is presented many times throughout each time period encompassed by the study, and because no single condition is present for any considerable length of time, it can be presumed that any effects of practice, change in task difficulty, maturation, or other historical variables will be equally represented in each treatment condition and therefore will not differentially affect any one condition more or less than the others. For example, even though each of two data paths representing a student's reading performance under two different teaching procedures shows variable and ascending trends that might be due to practice effects and uneven curriculum materials, any consistent separation and vertical distance between the data paths can be attributed to differences in the teaching procedures.

Can Be Used to Assess Generalization of Effects

By alternating various conditions of interest, an experimenter can continually assess the degree of generalization of behavior change from an effective treatment to other conditions of interest. For example, by alternating different therapists in the final phase of their study of pica behavior, N. Singh and Winton (1985) were able to determine the extent to which the overcorrection treatment was effective when presented by different persons.

Intervention Can Begin Immediately

Although determining the preintervention level of responding is generally preferred, the clinical necessity of immediately attempting to change some behaviors precludes repeated measurement in the absence of intervention. When necessary, an alternating treatments design can be used without an initial baseline phase.

Considering the Appropriateness of the Alternating Treatments Design

The advantages of the alternating treatments design are significant. As with any experimental tactic, however, the alternating treatments design presents certain disadvantages and leaves unanswered certain questions that can be addressed only by additional experimentation.

Multiple Treatment Interference

The fundamental feature of the alternating treatments design is the rapid alternation of two or more independent variables irrespective of the behavioral measures obtained under each treatment. Although the rapid alternation minimizes sequence effects and reduces the time required to compare treatments, it raises the important question of whether the effects observed under any of the alternated treatments would be the same if each treatment were implemented alone. **Multiple treatment interference** refers to the confounding effects of one treatment on a subject's behavior being influenced by the effects of another treatment administered in the same study.

Multiple treatment interference must always be suspected in the alternating treatments design (Barlow & Hayes, 1979; McGonigle, Rojahn, Dixon, & Strain, 1987). However, by following the alternating treatments phase with a phase in which only the most effective treatment condition is in effect, the experimenter can assess the effects of that treatment when administered in isolation.

Unnatural Nature of Rapidly Alternating Treatments

The rapid back-and-forth switching of treatments does not reflect the typical manner in which clinical and educational interventions are applied. From an instructional perspective, rapid switching of treatments can be viewed as artificial and undesirable. In most instances, however, the quick comparison of treatments offered by the alternating treatments design compensates for concerns about its contrived nature. The concern of whether participants might suffer detrimental effects from the rapid alternation of conditions is an empirical question that can be determined only by experimentation. Also, it is helpful for practitioners to remember that one purpose of the alternating treatments design is to identify an effective intervention as quickly as possible so that the participant does not have to endure ineffective instructional approaches or treatments that would delay progress toward educational goals. On balance, the advantages of rapidly switching treatments to identify an efficacious intervention outweigh any undesirable effects that such manipulation may cause.

Limited Capacity

Although the alternating treatments design enables an elegant, scientifically sound method for comparing the differential effects of two or more treatments, it is not an open-ended design in which an unlimited number of treatments can be compared. Although alternating treatments designs with up to five conditions have been reported (e.g., Didden, Prinson, & Sigafoos, 2000), in most situations a maximum of four different conditions (one of which may be a no-treatment control condition) can be compared effectively within a single phase of an alternating treatments design, and in many instances only two different treatments can be accommodated. To separate the effects of each treatment condition from any effects that may be caused by aspects of the alternating treatments design, each treatment must be carefully counterbalanced across all potentially relevant aspects of its administration (e.g., time of day, order of presentation, settings, therapists). In many applied settings the logistics of counterbalancing and delivering more than two or three treatments would be cumbersome and would cause the experiment to require too many sessions to complete. Also, too many competing treatments can decrease the subject's ability to discriminate between treatments, thereby reducing the design's effectiveness.

Selection of Treatments

Theoretically, although an alternating treatments design can be used to compare the effects of any two discrete treatments, in reality the design is more limited. To enhance the probability of discrimination between conditions (i.e., obtaining reliable, measurable differences in behavior), the treatments should embody significant differences from one to the other. For example, an investigator using an alternating treatments design to study the effects of group size on students' academic performance during instruction might include conditions of 4, 10, and 20 students. Alternating conditions of 6, 7, and 8 students, however, is less likely to reveal a functional relation between group size and performance. However, a treatment condition should not be selected for inclusion in an alternating treatments design only because it might yield a data path that is easily differentiated from that of another condition. The *applied* in applied behavior analysis encompasses the nature of treatment conditions as well as the nature of behaviors investigated (Wolf, 1978). An important consideration in selecting treatment conditions should be the extent to which they are representative of current practices or practices that could conceivably be implemented. For example, although an experiment comparing the effects of 5 minutes, 10 minutes, and 30 minutes of math homework per school night on math achievement might be useful, a study comparing the effects of 5 minutes, 10 minutes, and 3 hours of math homework per night probably would not be. Even if such a study found 3 hours of nightly math homework extremely effective in raising students' achievement in math, few teachers, parents, administrators, or students would carry out a program of 3 hours of nightly homework for a single content area.

Another consideration is that some interventions may not produce important behavior change unless and until they have been implemented consistently over a continuous period of time.

> When a multielement baseline design is employed, overlapping data do not necessarily rule out the possible efficacy of an experimental procedure. The session-by-session alternation of conditions might obscure effects that could be observed if the same condition was presented during several consecutive sessions. It is therefore possible that a given treatment may prove to be effective with a reversal or multiple baseline design, but not with a multielement baseline design. (Ulman & Sulzer-Azaroff, 1975, p. 382)

The suspicion that a given treatment may be effective if it is presented in isolation for an extended period is an empirical question that can be explored properly only through experimentation. At one level, if extended application of a single treatment results in behavioral improvement, the practitioner might be satisfied, and no further action would be needed. However, the practitioner-researcher who is interested in determining experimental control might return to an alternating treatments design and compare the performance of the single treatment with that of another intervention.

Summary

Reversal Design

1. The reversal tactic (A-B-A) entails repeated measurement of behavior in a given setting during three consecutive phases: (a) a baseline phase (absence of the independent variable), (b) a treatment phase (introduction of the independent variable), and (c) a return to baseline conditions (withdrawal of the independent variable).
2. The reversal design is strengthened tremendously by reintroducing the independent variable in the form of an A-B-A-B design. The A-B-A-B design is the most straightforward and generally most powerful intrasubject design for demonstrating functional relations.

Variations of the A-B-A-B Design

3. Extending the A-B-A-B design with repeated reversals may provide a more convincing demonstration of a functional relation than a design with one reversal.
4. The B-A-B reversal design can be used with target behaviors for which an initial baseline phase is inappropriate or not possible for ethical or practical reasons.
5. Multiple treatment reversal designs use the reversal tactic to compare the effects of two or more experimental conditions to baseline and/or to one another.
6. Multiple treatment reversal designs are particularly susceptible to confounding by sequence effects.
7. The NCR reversal technique enables the isolation and analysis of the contingent aspect of reinforcement.
8. Reversal techniques incorporating DRO and DRI/DRA control conditions can also be used to demonstrate the effects of contingent reinforcement.

Considering the Appropriateness of the Reversal Design

9. An experimental design based on the reversal tactic is ineffective in evaluating the effects of a treatment variable that, by its very nature, cannot be withdrawn once it has been presented (e.g., instruction, modeling).
10. Once improved, some behaviors will not reverse to baseline levels even though the independent variable has been withdrawn. Such behavioral irreversibility precludes effective use of the reversal design.
11. Legitimate social, educational, and ethical concerns are often raised over withdrawing a seemingly effective treatment variable to provide scientific verification of its function in changing behavior.
12. Sometimes very brief reversal phases, or even one-session baseline probes, can demonstrate believable experimental control.

Alternating Treatments Design

13. The alternating treatments design compares two or more distinct treatments (i.e., independent variables) while their effects on the target behavior (i.e., dependent variable) are measured.
14. In an alternating treatments design, each successive data point for a specific treatment plays three roles: it provides (a) a basis for the *prediction* of future levels of responding under that treatment, (b) potential *verification* of the previous prediction of performance under that treatment, and (c) the opportunity for *replication* of previous effects produced by that treatment.
15. Experimental control is demonstrated in the alternating treatments design when the data paths for two different treatments show little or no overlap.
16. The extent of any differential effects produced by two treatments is determined by the vertical distance between their respective data paths and quantified by the vertical axis scale.

Variations of the Alternating Treatments Design

17. Common variations of the alternating treatments design include the following:
    - Single-phase alternating treatments design without a no-treatment control condition
    - Single-phase design with a no-treatment control condition
    - Two-phase design: initial baseline phase followed by the alternating treatments phase
    - Three-phase design: initial baseline phase followed by the alternating treatments phase and a final best treatment phase

Advantages of the Alternating Treatments Design

18. Advantages of the alternating treatments design include the following:
    - Does not require treatment withdrawal.
    - Quickly compares the relative effectiveness of treatments.
    - Minimizes the problem of irreversibility.
    - Minimizes sequence effects.
    - Can be used with unstable data patterns.
    - Can be used to assess generalization of effects.
    - Intervention can begin immediately.

Considering the Appropriateness of the Alternating Treatments Design

19. The alternating treatments design is susceptible to multiple treatment interference. However, by following the alternating treatments phase with a phase in which only one

treatment is administered, the experimenter can assess the effects of that treatment in isolation.

20. The rapid back-and-forth switching of treatments does not reflect the typical manner in which interventions are applied and may be viewed as artificial and undesirable.

21. An alternating treatments phase is usually limited to a maximum of four different treatment conditions.

22. The alternating treatments design is most effective in revealing the differential effects of treatment conditions that differ significantly from one another.

23. The alternating treatments design is not effective for assessing the effects of an independent variable that produces important changes in behavior only when it is consistently administered over a continuous period of time.