The Clinical Application Of Withdrawal, Single-Case Research Designs

G. Albyn Davis

Memphis State University
Memphis, Tennessee

Single subject experimental designs provide tools by which we can obtain scientifically sound data regarding treatment effectiveness while still providing services to our clients. Being a clinician/researcher need not seem overwhelming if we keep in mind Campbell and Stanley's (1963) statement that,

> "basic to scientific evidence...is the process of comparison.... (and that) securing scientific evidence involves making at least one comparison (p. 6)."

During the course of clinical events we can make at least one comparison involving a question about treatment of a single subject. However, careful planning is required so that data from the comparison will be as valid and reliable as possible.

In the case of withdrawal designs, our comparison would be between the levels of a target behavior when a treatment procedure is present versus when that procedure is absent or withdrawn. Comparing alternating periods defined by the presence and absence of a treatment procedure is often referred to as the ABAB single subject design. Today I shall try to familiarize you with several aspects of the ABAB design, using Hersen and Barlow's (1976) text as my main reference and the Journal of Applied Behavior Analysis (JABA) as a source of some examples (see also Kratochwill, 1978). As in any single subject design we might employ in clinical aphasiology, the ABAB design uses the subject as his own control, to control for the heterogeneity of aphasic subjects, and entails taking repeated measures over a period of time.

Familiar Designs. I want to work my way gradually into explaining the ABAB design by reviewing more familiar forms of single subject investigation. Hopefully, this background will enhance our understanding of withdrawal procedures.

Before and After Design. In his book on research in speech pathology, Silverman (1977) described the "before and after" design or the "pretest-posttest" design as the most basic single subject design. This is a traditional strategy used by clinicians to report the effectiveness of their treatment, and it is the model for the traditional case study (see Table 1). The comparison in this design is between an observation (O_1) of behavior before treatment (X) and a second observation (O_2) of behavior after treatment. You should note that there is no comparison between the presence and absence of treatment. Silverman was unhappy with this design because it does not minimize the influence of variables other than treatment that could account for a change between O_1 and O_2. However, his main recommendation

for improving this design was to make several "before" observations and several "after" observations to increase the reliability of this particular comparison.

Table 1. Before and After Design ("The B").

O_1	X	O_2	where O_1 = Pretest observation of target behavior
			X = Treatment
			O_2 = Posttest observation

Baseline-Treatment Design (A-B). A second design seems to resemble the structure of LaPointe's (1977) Base-10 forms and gives us a chance to begin to form a clearer image of the ABAB design. On the Base-10 form we record three observations of a behavior before we apply a treatment procedure to it. We usually refer to these three observations as the baseline (see Table 2). This strategy departs from the before- and after-model when we continue taking measures of the baseline behavior while it is being treated. This baseline-treatment strategy can be called an A-B design, where we have two phases of repeated measurement over time that begin to define the basic elements of the ABAB or withdrawal design.

Table 2. Baseline-Treatment Design (A-B) with Unequal Length of Periods.

			X_1	X_2	X_3	.....	X_{10}		
O_1	O_2	O_3	O_1	O_2	O_3	.....	O_{10}		
Baseline (A)			Treatment (B)						

Each letter in the A-B design designates a phase in which a series of behavioral measures is taken. The A phase involves a series of observations of the natural frequency of the target behavior under study (Hersen and Barlow, 1976). This behavior may be conversation in a natural setting or it may be a specified response to a set of discrete stimuli, as in a diagnostic subtest. The B phase contains the treatment procedure, while the behavior of the phase A continues to be observed. With the A-B design, we have a comparison between the presence and absence of a treatment variable. An effect of treatment on the target behavior would be implied if we observe a stable level of behavior during the A or baseline phase and an improvement of that behavior during the B phase. Unfortunately, with an A-B design a stable baseline is often difficult to obtain, and we cannot be sure that improvement during B would not have occurred anyway or that it did not coincide with some other event causing the change. The controlling effect of treatment might be more clearly observed if we withdraw it for a period of time.

Withdrawal Design. Withdrawing a treatment procedure extends the A-B design to an A-B-A design, in which we would repeat the conditions of the initial baseline period in a second A phase following the B phase. The A-B-A design is the simplest use of the withdrawal strategy and can be quite successful in demonstrating that treatment is the agent of behavioral change. In traditional application of this design, treatment effectiveness would be demonstrated if the treatment during B leads to improvement in the target behavior and removal of treatment in A_2 results in deterioration of the target behavior. In the case of aphasia treatment during B, we would not expect deterioration of the target behavior during A_2, which raises an issue to be addressed shortly. Furthermore, a major drawback of the A-B-A design as a withdrawal procedure is that the paradigm ends on a baseline or no-treatment phase, denying the patient full benefit of the experimental treatment (Hersen and Barlow, 1976). The A-B-A-B design, as shown in Table 3, does end on a treatment phase (B) which can be extended beyond the requirements for a given study and introduces some symmetry to the withdrawal design, in that the number of A and B phrases are equated. Also, this extended withdrawal design provides two occasions, rather than one, for demonstrating effects of the treatment variable: B_1 to A_2 and A_2 to B_2.

Table 3. Withdrawal Design (A-B-A-B) with Equal Length of Periods.

A_1			B_1			A_2			B_2		
			X_1	X_2	X_3				X_4	X_5	X_6
O_1	O_2	O_3	O_4	O_5	O_6	O_7	O_8	O_9	O_{10}	O_{11}	O_{12}

Now, I want to continue with some general features of the A-B-A-B design by discussing aphasia treatment as an irreversible procedure, then by showing that the withdrawal design need not entail the complete removal of treatment during the A phase, and by reviewing a basic decision strategy for constructing a single case experiment.

Irreversible Procedures. The withdrawal design is best suited for evaluating the controlling effects of a reversible procedure. A reversible procedure is one which would not produce a permanent change in the target behavior and, upon removal of the procedure for A_2, would result in a return toward baseline levels of performance for the target behavior. In typical A-B-A-B studies, treatment introduced in B involves only a reinforcement procedure, and its absence in A_2 represents a return to baseline conditions. On the other hand, aphasia treatment is thought to be irreversible; and the second A phase may not be a complete return to baseline conditions, since the subject may have permanently changed during the previous B phase. Hersen and Barlow (1976) suggested that the A-B-A-B design is not appropriate for irreversible procedures.

Let us consider a hypothetical experiment in order to see if we can use the withdrawal design. As in any experiment, we want to answer a well-defined research question. In terms of the A-B-A-B design, our question

should fit within the following framework: Does treatment introduced in the B phase have an effect on the target behavior observed in the A phase? Table 4 illustrates components of a design defined to answer the following specific question:

1. Does confrontation naming practice have an effect on word retrieval ability?

Table 4. Components of an A-B-A-B Sequence Examining the Effect of Naming Practice.

Target behavior = spoken word retrieval
O = naming 10 objects
X = naming practice with 30 pictures repeated twice
A = O_1 O_2 O_3 and O_7 O_8 O_9
B = X_1 X_2 X_3 and X_4 X_5 X_6
O_4 O_5 O_6 and O_{10} O_{11} O_{12}

First, we would develop a measure of the target behavior which could be a task of naming 10 objects. We observe this behavior for three days during the baseline period or initial A phase. Then we introduce naming practice with 30 pictures presented twice for three days during the first B phase. We continue to measure the target behavior with a test of naming 10 objects. Subsequently we would withdraw the treatment during A_2 and reinstate it for B_2. Since we hope that the treatment produces some improvement in the target behavior, we would expect a stairstep effect in the results (LaPointe, 1977) rather than the return to baseline level expected with a reversible procedure.

LaPointe (1977) suggested that this stairstep result, or an alternating flat and increasing configuration, could be used to demonstrate the role of aphasia treatment with an A-B-A-B design. Hersen and Barlow (1976) cited an exception to the use of A-B-A-B with only reversible procedures. In fact, they sometimes applied a post hoc evaluation of certain designs by giving a favorable recommendation for the design if it produced change that varied systematically as a function of the alternating phases. I recommend that we evaluate this design in the aphasia clinic, especially to see whether our clinical procedures are truly irreversible.

Maintaining Treatment. Up to this point, I have been presenting the comparison between B and A as a comparison between the presence and absence of treatment, respectively. However, most aphasia treatment procedures involve several variables, and the B and A comparison can be a comparison between the presence and absence of one treatment variable. Also, many single case experiments appear in JABA where two methods of treatment are alternated between adjacent phases. Therefore, using the A-B-A-B strategy does not necessarily deprive the patient of treatment.

The maintenance of treatment while comparing the presence and absence o one variable is illustrated with the following research question:

2. Do response cues have an effect on word retrieval during confrontation naming practice?

Here we would be examining a specific treatment variable by introducing it in B and withdrawing it in the subsequent A phase. Confrontation naming practice would continue throughout the experiment. This particular experiment might be conducted during a single session.

Decision Strategy. Before getting into more detailed design considerations and some more examples of A-B-A-B studies, I would like to review the basic steps one would use to develop a single subject study:

1. Formulate a specific research question about the effect of a treatment procedure on a target behavior. The target behavior would define phas A, and the treatment procedure or variable would be applied to the target behavior to define phase B.
2. Develop a valid and reliable measure to quantify observations of the target behavior during each phase of the experiment. Many single case experiments include at least a second observer as a reliability check.
3. Plan for the length of each phase and the frequency of measurement within each phase. Ideally, each phase should be the same period of time, determined, in part, by the time needed to establish a stable baseline during the initial A phase. However, clinical considerations sometimes dictate that the B phases be longer than A phases, especially when A involv absence of treatment. An A-B-A design by Walker and Buckley (1968) was use by Hersen and Barlow (1976) to show that a longer B phase may permit valid conclusions concerning the effects of a treatment variable. Similarly the number of measures during each phase should be equal among phases and frequent enough to detect trends in performance. Unfortunately, patient attendance does not always conform to the symmetrical beauty of an experimental design. We should plan our experiments as tightly as possible and then consider variations from the ideal in interpreting results.

Designing Answers to More Questions. The rest of my discussion deals with a few more rules for A-B-A-B single subject designs and a few ways in which the basic design might be extended to answer a variety of clinical questions I shall cite some examples from JABA and shall make up some others. Some of these examples may push A-B-A-B to its limits, if not beyond its limits; but I have already been somewhat pushy by suggesting that the withdrawal design be used at all to evaluate the effectiveness of aphasia treatment. Even though most aphasia treatment procedures are irreversible, I think we should see if these design principles can be applied or modified so that they can work for us.

Generality. First, let us consider the generality of findings from a single subject. Single case experiments have been sneezed at because an "N of one" has limited generality to a larger population. On the other hand, many studies of groups of aphasic subjects have limited applicability to individual cases in the clinic, because we cannot be sure that a group

mean reflects the performance of any individual with aphasia. We can infer some generalizations from single case studies if the subject is described completely and clearly. I have had the opportunity to evaluate single case studies where the single subject was not clearly described, thus making it difficult for me to determine whether the procedures in the study would be applicable to a similar client in my clinic. I want us to consider two areas in which generality of findings can be examined with an A-B-A-B design. These are generality across subjects and generality across behaviors.

Generality Across Subjects. Generalizing findings from a single case to other aphasic subjects simply entails replicating the A-B-A-B design with additional cases. Hersen and Barlow's (1976) review of replication studies showed that most investigators used replication to determine the reliability or consistency of their findings. However, replication can be used to establish whether a treatment procedure that was effective for one aphasic person will be effective for other aphasic persons. If we want to establish generality and reliability, additional subjects must be as similar as possible to the first subject on all relevant characteristics (Darley, 1972). Also, we could attempt to generalize findings with one procedure to different types of aphasia, thereby creating an additional comparison - a comparison between subjects.

Generality Across Behaviors. It is very important that we establish generality of findings with a treatment procedure across behaviors. We are usually interested in the effects of a rather circumscribed procedure on a variety of different communicative functions. My previous examples examined the effect of a confrontation naming procedure on the ability to name objects. However, would it not be more useful to examine that procedure's effectiveness in improving functional communicative abilities as well? In a study of agrammatic aphasic subjects, Weigl-Crump (1976) looked at the effect of repetition on the ability to describe pictures. Would it not have been useful to have examined that procedure's effectiveness in improving the ability to converse, also? We can examine the generality of a procedure's appropriateness to multiple behaviors by monitoring concurrent or non-targeted behavior during A and B phases. My first example of naming treatment would possess the additional features shown in Table 5:

3. Does confrontation naming practice have an effect on spoken word retrieval, graphic word retrieval, and amount of verbal output during conversation?

In this example, three behaviors would be monitored, one behavior being the target behavior for treatment with two additional behaviors used to study generalization. Like the notations used in Silverman's (1977) book, the observations are designated first with a behavior type number and second with the temporal position of the observation in the time series.

Comparison of Two Clinicians. I found an interesting twist on the A-B-A-B design in JABA, where the performances of two teachers were compared by extending the design with an additional A-B phase. In this study, (by Carnine, 1976), slow and fast task presentation rates were alternated between A and B phases respectively by a certified teacher during the

$A_1B_1A_2B_2$ phases; and then A_3B_3 phases were carried out by a student teacher. Perhaps Carnine's study provides a model for evaluating the effectiveness of student clinicians (or of their supervisors).

Table 5. Components of an A-B-A-B Sequence Examining Generality Across Behaviors.

Target behavior = spoken word retrieval							
O_1= spoken naming of 10 objects							
O_2= graphic naming of 10 objects							
O_3= number of nouns produced per opportunity in conversation							
X = naming practice with 30 pictures repeated twice							
A =	O_{11}	O_{12}	O_{13}		O_{17}	O_{18}	O_{19}
	O_{21}	O_{22}	O_{23}	and	O_{27}	O_{28}	O_{29}
	O_{31}	O_{32}	O_{33}		O_{37}	O_{38}	O_{39}
B =	X_1	X_2	X_3		X_4	X_5	X_6
	O_{14}	O_{14}	O_{16}	and	O_{110}	O_{111}	O_{112}
	O_{24}	O_{25}	O_{26}		O_{210}	O_{211}	O_{212}
	O_{34}	O_{35}	O_{36}		O_{310}	O_{311}	O_{312}

One Variable at a Time. There is another important consideration for designing A-B-A-B experiments which enable us to put a particular treatment method under a microscope to scrutinize specific aspects of it. You may recall my second example where the effect of a single treatment variable - response cues - was assessed by introducing that variable in B_1 and withdrawing it in A_2. However, a particular aphasia treatment procedure often involves several variables. If we wanted to compare the effectiveness of two variables within a treatment procedure, we must follow a design rule which requires us to manipulate one variable at a time between adjacent phases in a withdrawal design (Hersen and Barlow, 1976). My illustration of this principle is simply an extension of previous examples. We might ask the fourth question:

4. What are the relative contributions of response cues and reinforcement of correct responses to word retrieval during confrontation naming practice?

In the third example we looked at multiple observation measures. In this fourth example we are looking at more than one treatment variable. The letter B designates the response cue variable as in second example and the letter C designates the other treatment variable - reinforcement. Following the rule of manipulating one variable at a time between adjacent phases, we arrive at the appropriate design or phase sequence shown in Table 6.

Table 6. Abbreviated Illustration of Components in a Design Comparing the Effects of Two Treatment Variables.

Target behavior = spoken word retrieval

O = spoken naming of 10 objects

X_1 = naming practice with response cues.

X_2 = naming practice with reinforcement

$A_1 = O_1 \quad O_2 \quad O_3$

$B_1 = X_{11} \quad X_{12} \quad X_{13}$

$O_4 \quad O_5 \quad O_6$

$C = X_2 \quad X_2 \quad X_2$

PHASE SEQUENCE: A - B - A - B - BC - B - BC

Where $BC_1 = X_{17} \quad X_{18} \quad X_{19}$

$X_{21} \quad X_{22} \quad X_{23}$

$X_{13} \quad X_{14} \quad X_{15}$

We would begin by comparing response cues to baseline performance by alternating B and A phases as in the second example. After the second B phase, we would add the second variable of reinforcement and alternate the presence and absence of that variable with the B phase.

Comparison of Different Treatment Procedures. Finally, if we wanted to compare two different treatment procedures applied to a particular type of aphasia, we might consider alternating those procedures, with one procedure as phase B and another as phase C deriving a B-C-B-C design. Doing this would certainly violate the one-variable-at-a-time rule, since several variables would change between phases. In this case, we would simply recognize that we could address our comparison only to the treatments as a whole and not to the specific attributes of each procedure. Some simple comparisons between procedures have appeared in JABA. Quilitch and Risley (1973) compared two types of training materials by alternating their use between what they designated as A and B phases. At Memphis State we tried the B-C-B-C design when Pat Larkins (1978) recently compared two methods of eliciting meaningful speech from a Wernicke's aphasic patient by alternating these procedures on a weekly basis. The B-C-B-C model was avoided by Bondy and Erickson (1976), who compared modeling and token reinforcement in increasing question-asking of retarded children. They assigned the modeling procedure to one group of four children and the reinforcement procedure to another group of four children. A third group received both procedures, and a fourth group received none of the procedures. Each group was examined with its assigned procedure administered during the B phase of an A-B-A-B design which alternated the presence and absence of the assigned treatment for each of the three treatment groups.

Conclusions. Today I have tried to provide you with a basic introduction to the nature of withdrawal single subject research designs. Even within the restrictions of the several rules for constructing the ABAB design, we can create a wide variety of paradigms to answer many very specific questions about aphasia treatment methodology. Because of the nature of our subjects and settings in clinical aphasiology, some of our applications of this design may in fact be distoring it beyond recognition. But as we try to achieve some degree of experimental control while administering treatment to our clients, we may have to invent new rules for a new science of clinical aphasiology. Furthermore, the flexibility that is possible with single case designs can be extended to combining ABAB paradigms with the multiple baseline strategy to be discussed later by Chick LaPointe.

Though the realization of ABAB designs are potentially limitless, we should keep in mind a certain basic discipline. We should consult the new texts on single subject designs, and we should search journals like JABA for ideas or precedents related to the questions we might want to answer concerning aphasia treatment. Our studies should conform to at least two minimal requirements: 1) we make at least one comparison, and 2) we examine a treatment's effect on functionally relevant communicative behaviors with observation measures that are valid and reliable. If we do these things as clinicians, all of us can have a part in injecting a lot of sicence into the art of treating speech and language disorders.

References:

Bondy, A.S., and Erickson, M.T., Comparison of modeling and reinforcement procedures in increasing question-asking of mildly retarded children. JABA, 9, 108, 1976.

Campbell, D.T., and Stanley, J.C., Experimental and Quasi-Experimental Designs for Research. Chicago, Rand McNally, 1963.

Carnine, D.W., Effects of two teacher-presentation rates on off-task behavior, answering correctly, and participation. JABA, 9, 199-206, 1976.

Darley, F.L., The efficacy of language rehabilitation in aphasia. JSHD, 37, 3-21, 1972.

Hersen, M., and Barlow, D.H., Single Case Experimental Designs: Strategies for Studying Behavior Change. New York: Pergamon Press, 1976.

Kratochwill, T.R., (ed.), Single Subject Research Strategies for Evaluating Change. New York: Academic Press, 1978.

LaPointe, L.L., Base-10 programmed stimulation: Task specification, scoring, and plotting performance in aphasia therapy. JSHD, 42, 90-105, 1977.

Larkins, P.G., and Prance, C.A., Comparison of procedures to improve speech for a Wernicke's aphasic. Unpublished manuscript, 1978.

Quilitch, H.R., and Risley, T.R., The effects of play materials on social play. JABA, 6, 573-578, 1973.

Silverman, F.H., Research Design in Speech Pathology and Audiology. Englewood Cliffs, New Jersey: Prentice-Hall, 1977.

Walker, H.M., and Buckley, N.K., The use of positive reinforcement in conditioning attending behavior. JABA, 1, 245-250, 1968.

Weigl-Crump, C., Agrammatism and aphasia. In Y. Lebrun and R. Hoops (eds.), Recovery in Aphasics. Amsterdam: Swets & Zietlinger, 1976.