

# The Application of Single-Subject Research Designs to the Field of Learning Disabilities

Michael J. Guralnick, PhD

**This article discusses the acknowledged gap that exists between educational research and practice, specifically the limited relevance of the former for the latter. The author describes, illustrates, and advocates the multiple baseline, single-subject research method as an approach that produces applicable results. — G.M.S.**

---

*Arguments are presented suggesting that single-subject research designs can form a basis for establishing a working relationship between researchers and clinicians. At the heart of the arguments is an emphasis on focusing activities at the level of the individual child. Issues regarding variability, generality, sampling, and the "believability" of multiple-baseline single-subject designs are discussed from the researcher's perspective. The potential benefits of such a relationship to the clinician as well as to the learning disabilities field in general are also explored.*

---

It seems almost axiomatic in education that a substantial gap exist between research and practice. Blackman (1972) has commented, "Much of what is wrong with educational research today has to do with its acknowledged inability to make a major impact on the educational enterprise" (p. 181). In part, this reflects the differing goals, attitudes, training,

and skills of researchers and clinicians, as pointed out by Keogh (1977) in a recent guest editorial in the *Journal*, although the problem is probably also related to the relatively small number of investigations that have been conducted in the brief history of the field (Bryan 1974).

In her editorial, Keogh argued that in order for the learning disabilities field to progress substantially, cooperative relationships between researchers and clinicians must develop. Although clearly recognizing the many incompatibilities that exist between these two groups, she urged that one basis for a rapprochement could occur through a shared concern for empirically demonstrating the effectiveness of their activities. To accomplish this, it was suggested that researchers focus on applied problems in naturalistic or field settings and use a more descriptive methodology to identify critical variables. Furthermore, clinicians were urged to extend their data-recording efforts and to use this information for decision making.

Additionally, I would suggest that in order for such a rapprochement to occur, it is essential that researchers and clinicians agree to focus their short- and long-range objectives on solving the educational problems of a specific child, in a specific setting, at a specific time. Blackman has stated this most clearly: "The educational enterprise . . . can be fully subsumed under the concrete image of a teacher helping a child learn a needed skill or exhibit a behavior that was not present before instruction" (p. 181). Unfortunately, researchers working on applied problems tended to analyze group behavior, and little predictability to individual children has resulted.

This article will describe a methodology based on single-subject designs that is responsive to the issues raised by Keogh and Blackman. Although these designs can only be viewed as one alternative among a host of strategies, I believe they are capable of meeting the diverse needs of both groups. As such, they have the potential for serving as a basis for establishing a working relationship between researchers and clinicians.

## COMPARATIVE APPROACHES

Single-subject research designs can be contrasted with research conducted within the framework of multiple-group comparative analysis designs. It is well known that there are numerous instances in which the conventional designs are useful and even necessary (see Paul 1969), but it appears that these designs are not well suited to teasing out the complex variables and relationships that exist in developing fields such as learning disabilities (Hersen & Barlow 1976, Sidman 1960). Although it is not my purpose here to elaborate on these deficiencies, conventional designs are often very cumbersome and cost ineffective, require too many assumptions, often fail to solve the problems inherent in sampling from deviant populations, and provide group data that can obscure rather than clarify lawful relationships (Edgar & Billingsley 1974, Hersen & Barlow 1976, Sidman 1960). In addition, the structural and procedural aspects related to these designs

tend to restrict application of the results to educational and clinical settings. This is primarily due to the lack of flexibility in modifying the variables during the course of an experiment as well as to many other procedural constraints inherent in applying multiple-group principles. In contrast, the procedural requirements for single-subject designs readily lend themselves to the solution of problems relevant to teacher-child interactions (Guralnick 1973a). Given the current state of the art, these designs are more likely to assist in the identification of educationally relevant variables and the conditions under which they are effective; they are intended for natural settings and contain data collection procedures that can assist both clinicians and researchers in their decision making.

## MULTIPLE-BASELINE DESIGNS

Although there are many variations of single-subject designs (Baer, Wolfe, & Risley 1968, Hersen & Barlow 1976), multiple-baseline designs seem to be the most likely candidates for encouraging joint efforts between clinicians and researchers. Essentially, this procedure requires that baseline data be obtained on several behaviors concurrently (A, B, and C in Figure 1). Once stability is achieved, the treatment factor (i.e., an educational or therapeutic intervention) is selectively applied to one behavior. Following an appropriate change in the behavior, this treatment is then applied to the second behavior, with additional replications occurring as needed. Figure 1 illustrates this sequence of events. The causal relationship between the treatment factor and the change in behavior is inferred from the fact that change occurred *only* at the time the treatment was applied. Notice also that, for an unambiguous interpretation to occur, it is necessary for the untreated behaviors to remain unchanged until the intervention period.

It is possible to identify at least three types of multiple-baseline designs that are appropriate for an analysis of data similar to that noted in

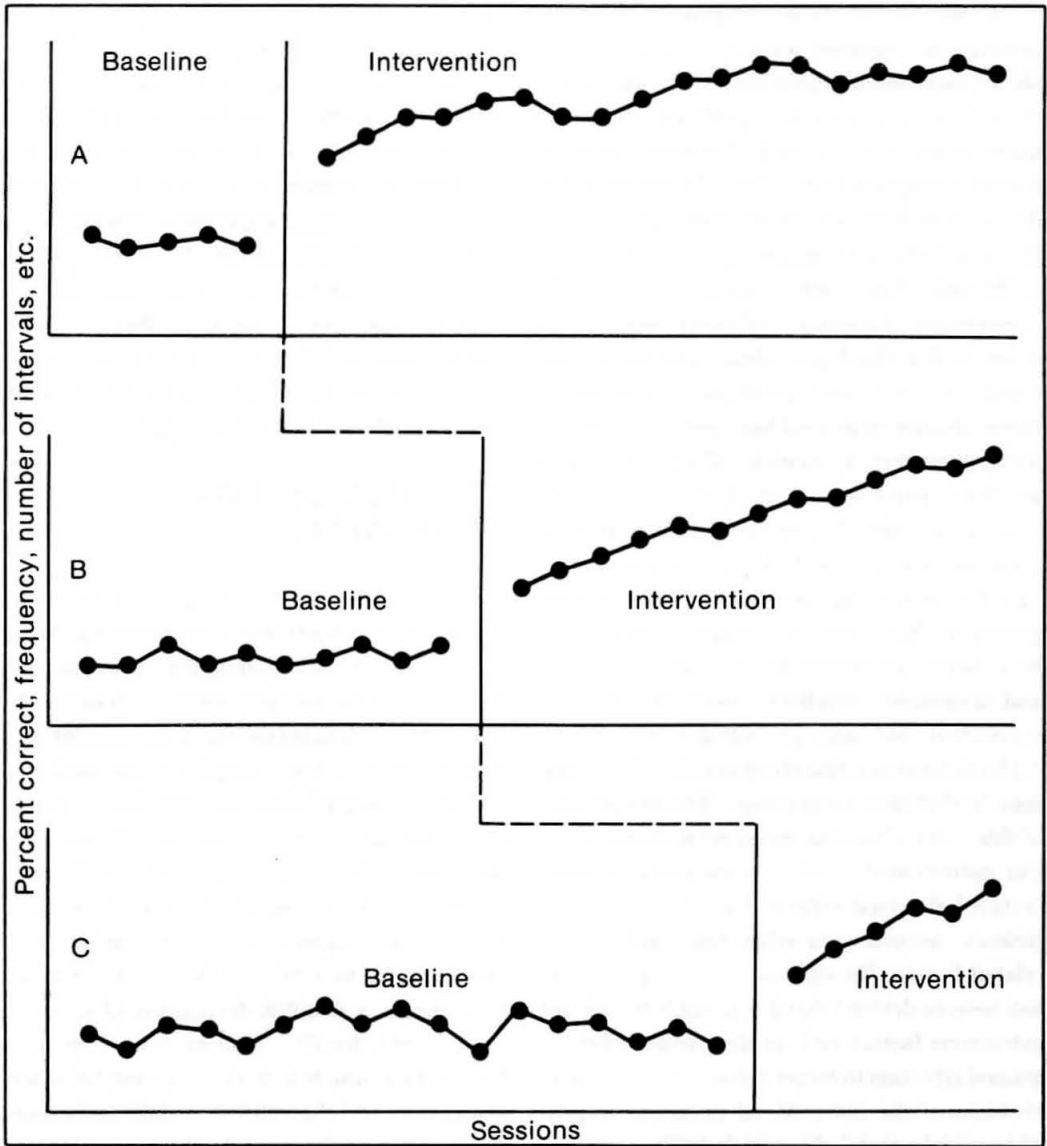


FIGURE 1. Illustration of the procedure and hypothetical data for multiple-baseline designs.

Figure 1 (Hall, Cristler, Cranston, & Tucker 1970). The first type looks at the behavior of one individual, with that same behavior measured in three different situations. For example, if we wanted to assess the effects of a reinforcement technique on the frequency of attentive behavior, this behavior might be measured during the math, reading, and play periods. These situations (periods) would correspond to graphs A, B, and

C in Figure 1. The effects of the treatment would be evaluated by detecting any changes that occur as the treatment variable is systematically applied to each situation in a sequential manner. Moreover, it is necessary to replicate this treatment with other individuals to establish the degree of generality of the technique. The issue of generality of results as well as variability during treatment will be discussed later.

In the second type of design, the same behavior is measured across different individuals. Its most useful application is the analysis of curriculum components, although there are many other uses as well. Referring again to Figure 1, graphs A, B, and C would correspond to different individuals, each being measured on the same behavior (e.g., percent correct on math problems). Intervention might refer to the presentation of a segment of a curriculum or any other well-defined procedure. The use of this design for curriculum development and formative evaluation in general have been described in detail elsewhere (Guralnick 1973a). Referred to as "the research-service model," the procedure includes a series of systematic instructional-unit presentations across individuals with a methodology for evaluating revisions of instructional programs. Moreover, it contains a process for both direct (different subjects, same conditions) and systematic (conditions permitted to vary) replication (Sidman), providing a firm empirical basis for evaluating the generality of the findings. At that time it was noted, "The end product of this process is a clear statement that this particular instructional unit does work, that is, it meets certain behavioral criteria that may include efficiency, accuracy, or other time and context related factors. [In addition] . . . the procedure has been so devised that it is possible to rule out extraneous factors such as the passage of time, general attention to target behaviors, or previous revisions of the instructional programs as possible causal agents" (Guralnick 1973a, p. 280).

The third common type of multiple-baseline design calls for the application of the treatment to different behaviors of one individual, but the situation remains essentially the same. For example, a child might display a variety of disruptive behaviors such as kicking, knocking over toys, or negative verbalizations in an unstructured play situation. To determine the effectiveness of a given treatment, each of the behaviors would be continuously recorded (A, B, and C in Figure 1 would refer to different behaviors, and the unit of measurement can

vary), with the treatment program applied sequentially to the behaviors in the usual way.

Recently, Hartmann and Hall (1976) have described a multiple-baseline design referred to as "the changing criterion design." The criterion for behavior change is adjusted in a stepwise fashion, with each adjustment constituting a potential replication of the effectiveness of the treatment procedure and a demonstration of experimental control. Although this design has many limitations, it provides one more potentially useful experimental strategy that is compatible with educational and clinical goals.

## VARIABILITY AND GENERALITY

One of the most significant aspects of the single-subject approach is that new experimental phases are usually not attempted until the previous ones have met some predetermined criterion. The failure to reach criterion for a given child on a curriculum unit, for example, is not used as a basis for assessing "error" or intersubject variability as in group designs but rather sets the occasion for a search for the causes of that variability.

Accordingly, when a failure occurs, different intervention strategies are systematically explored to try to achieve criterion. Throughout this process, a detailed description of program components, baseline contingencies (see Birnbrauer, Peterson, & Solnick 1974), and behavioral and historical characteristics of the individuals involved are maintained. In this way, although the original experiment may have been substantially modified, it may be possible, through this "rich" data base, to identify certain characteristics that individuals share in common, or certain elements of the curriculum that are correlated with the "failures." This process will also generate information on the possible effectiveness of the revisions that were developed. Moreover, once relationships between the type of intervention and the characteristics of individuals begin to emerge, more formalized procedures similar to the original multiple-baseline design

that prompted the search would be applied. Accordingly, we are beginning to approach the type of interactive model (aptitude-by-treatment) such as that described by Reynolds and Balow (1972) and advocated by Blackman — but at the level of the individual child. Since this research strategy parallels a teacher's normal instructional process, we are establishing a common ground for researchers and clinicians.

A related problem is variability in the baselines of behaviors that have not yet been subjected to the experimental intervention (i.e., response generalization). For example, the effects of a motivational system that is first applied to negative verbalizations as in the second design may immediately generalize to the behavior of knocking over toys and thus vitiate the experimental analysis. In this instance, as in the initial case of variability where the treatment was effective only for some subjects or settings, these data are accepted and are utilized in assessing the nature of stimulus control and the organization of response classes. In fact, recording multiple behaviors in this manner through direct and continuous measurement and analyzing the covariations that may occur among various behaviors is currently a very significant area of investigation (Guralnick 1973b, Nordquist 1971, 1978, Wahler 1975, Wahler & Nordquist 1973) and has many implications for learning disabled individuals.

This search process itself speaks to the question of the generality of findings. As Sidman has noted, generality is related to the control of experimental variables and the specification of functional relationships. In the final analysis, replicability is the key to establishing generality, and successful replications depend on our knowledge and control of relevant variables. In fact, in some multiple-baseline applications, systematic replication is a natural part of the experimental procedure and is a very powerful process related to generalizability. In this technique, certain variables are left free to vary while experimental control is demonstrated through replication. For example, success with a

curriculum unit administered by different teachers strengthens our confidence in the value of that particular unit. Replication of this sort is certainly not possible without a thorough understanding of the subject matter. Consequently, given the complex and uncertain situations in learning disability research, the fine-grained experimental analysis required by single-subject designs should substantially contribute in the long run to increasing the generality of the findings.

## GENERALIZATION AND SAMPLING

It is important to note here as well that a failure to use inferential statistical procedures, such as those used in conventional designs, does not appear to be a primary factor in evaluating the value of research. As Edgar and Billingsley (1974) have pointed out, most psychological and educational research does not, in fact, provide true random samples. They state:

In the absence of random samples, hypotheses may still be tested, but statements of significance are limited to the effect of the treatment of Ss actually utilized, generalizations being based on logical considerations of a nonstatistical nature (Edgington 1967). This logical rather than statistical basis for generalization is applied in most large sample studies and is equally applicable when  $N = 1$  whether tests of significance are applied to the data or not (as in the usual analysis of data obtained from reversal or multiple baseline investigations). *In many cases generalization may, in fact, be more readily made from  $N = 1$  studies than from large  $N$  studies due to the opportunity for more accurate delineation and precise control of relevant S characteristics*" (p. 153) (emphasis supplied).

Moreover, a functional, intensive approach to research common to single-case experiments, with the accompanying detailed statements of the characteristics of the subjects' behaviors and descriptions of situational characteristics, may well contribute to clarifying and refining the nature of the population of individuals referred to as having a learning disability.

## IS THIS REALLY RESEARCH?

Researchers ask questions that relate to establishing causal relationships between treatment



factors and measured behaviors. A working relationship between researchers and clinicians can only develop if this fact remains unaltered. Conceptually, the experimental designs described above do permit an investigator to draw these conclusions, and the confidence one has in the findings are subject to the same principles that affect confidence for large-group designs. Edgar and Billingsley (1974) have identified four dimensions that influence the "believability" of research findings and their value. They point out, first of all, that the logic of single-subject designs conforms to the logic of traditional research with regard to testing hypotheses and determining lawful relationships. The second dimension refers to the generation of data, with attention focused on the accuracy of these data. Single-subject designs collect data on a continuous basis, often in natural settings, and the method is usually one of direct observation. Although there are certain problems associated with this procedure, extensive, reliable, and sophisticated methods for defining variables and collecting these data are currently available (see Johnson & Bolstad 1973). Furthermore, as Keogh (1977) suggests, data generated through naturalistic observations are needed to clarify some basic issues in the learning disabilities field.

Internal validity (Campbell & Stanley 1963) is the third dimension and relates to establishing the causal relationship between the independent and dependent variables. Through repeated and systematic applications of the treatment factor using the processes of direct and systematic replication (Sidman), such relationships can be established within a single-subject methodology.

Finally, the relevance of statistical analysis has been of concern. Interestingly, Michael (1974) has identified a number of *detrimental* effects that arise from the use of statistical inference methodology (the fourth dimension), many of which are related to the fact that these procedures place an emphasis on statistical as opposed to experimental control. In most instances for applied problems, with the ultimate focus on a single subject, absolute criteria for

evaluating treatment outcomes are often more appropriate than relative comparisons among groups, with the group mean as the best estimate of treatment effectiveness (see Kazdin 1976). Of course, when one is interested in answering, for example, questions about cost effectiveness or related "engineering" problems for various treatments, a type of group design would be more appropriate. There have been recent advances in statistical analyses for single-subject designs, and some of the concerns of researchers have been reduced. Probably the most appropriate analysis is the time series design as described by Gottman (1975), with particular reference to the time-lagged control method, which is applicable to multiple-baseline techniques.

## WHAT'S IN IT FOR THE CLINICIAN?

At the outset of this article, I referred to Blackman's observations on the lack of impact of educational research and noted its relationship to existing differences between clinicians and researchers. For a variety of reasons, I have argued that interactions between researchers and clinicians are more likely to occur if efforts are focused at the level of the individual child in more natural settings. I have also argued that the type of research and the corresponding single-subject methodology advocated here are, in fact, directed at problem solving at this level. Given that most of the fundamental scientific needs of researchers can be met by this methodology, and the fact that the field probably could benefit from this or similar approaches, what motivates the clinician to become involved? Clearly, any participation by clinicians in an experiment, including those using single-subject designs, requires additional effort and compromising on various issues. In particular, more careful data collection and record keeping in general seem essential, and it is likely that schedule adjustments would be required as well.

Perhaps the most persuasive answer to this question is the fact that, due to the structure of

single-subject designs, the selection of research issues tends to be governed to a greater extent by the *teacher's priorities*. It is very difficult for researchers not to be more responsive to the learning and behavior problems that are so apparent yet so puzzling to the clinician, especially when the research domain is the child's natural setting. In addition to generating needs, clinicians can become active participants in the research enterprise (Hall et al. 1970), an activity that is likely to benefit both the research and service components (Guralnick 1973a).

Finally, it should be reiterated in this context that the form of the data and the degree of procedural specification required for single-subject designs are completely compatible with instructional purposes. In instances where instructional objectives are unclear, the greater precision required to answer research questions may well be of assistance in clarifying critical points. Moreover, the inherently flexible nature of the single-subject approach, the corresponding search to identify sources of variability, and the requirements for data-based feedback for decision making are all components that are more likely to enhance the instructional process (Guralnick 1975).

#### ABOUT THE AUTHOR

**Michael Guralnick, PhD**, received his doctorate in psychology and is currently director of The Nisonger Center at The Ohio State University and an associate professor in the Departments of Communication and Psychology. Requests for reprints should be sent to him at The Nisonger Center, The Ohio State University, 1580 Cannon Dr., Columbus, Ohio 43210.

#### REFERENCES

Baer, D.M., Wolfe, M.M., Risley, T.R.: Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis*, 1968, 1, 91-97.  
 Birnbrauer, J.S., Peterson, C.R., Solnick, J.V.: Design and interpretation of studies of single subjects. *American Journal of Mental Deficiency*, 1974, 79, 191-203.

Blackman, L.S.: Research and the classroom: Mahomet and the mountain revisited. *Exceptional Children*, 1972, 39, 181-191.  
 Bryan, T.H.: Learning disabilities: A new stereotype. *Journal of Learning Disabilities*, 1974, 7, 304-309.  
 Campbell, D.T., Stanley, J.C.: *Experimental and Quasi-experimental Designs for Research*. Chicago: Rand McNally & Co., 1963.  
 Edgar, E., Billingsley, F.: Believability when  $N = 1$ . *The Psychological Record*, 1974, 24, 147-160.  
 Edington, E.S.: Statistical inference from  $N = 1$  experiments. *The Journal of Psychology*, 1967, 65, 195-199.  
 Gottman, J.M.: N-of-one and N-of-two research in psychotherapy. *Psychological Bulletin*, 1975, 80, 93-105.  
 Guralnick, M.J.: A research-service model for support of handicapped children. *Exceptional Children*, 1973, 39, 277-282. (a)  
 Guralnick, M.J.: Behavioral techniques and preschool intervention programs. Paper presented at the annual meeting of the American Association on Mental Deficiency, Atlanta, Georgia, 1973. (b).  
 Guralnick, M.J.: Early classroom-based intervention and the role of organizational structure. *Exceptional Children*, 1975, 42, 25-31.  
 Hall, R.V., Cristler, C., Cranston, S.S., Tucker, B.: Teachers and parents as researchers using multiple baseline designs. *Journal of Applied Behavior Analysis*, 1970, 3, 247-255.  
 Hartmann, D.P., Hall, R.V.: The changing criterion design. *Journal of Applied Behavior Analysis*, 1976, 9, 527-532.  
 Hersen, M., Barlow, D.: Single case experimental designs: Strategies for studying behavior change. Elmsford, N.Y.: Pergamon Press, 1976.  
 Kazdin, A.: Statistical analyses for single-case experimental designs. In M. Hersen, D. Barlow (Eds.): *Single Case Experimental Designs: Strategies for Studying Behavior Change*. Elmsford, N.Y.: Pergamon Press, 1976.  
 Keogh, B.: Working together: A new direction. *Journal of Learning Disabilities*, 1977, 10, 478-482.  
 Johnson, S.M., Bolstad, O.D.: Methodological issues in naturalistic observation: Some problems and solutions for field research. In L.A. Hamerlynk, L.C. Handy, E.J. Mash (Eds.): *Behavior Change: Methodology, Concepts, and Practice*. Champaign, Ill.: Research Press, 1973.  
 Michael, J.: Statistical inference for individual organism research: Mixed blessing or curse? *Journal of Applied Behavior Analysis*, 1974, 7, 647-653.  
 Nordquist, V.M.: The modification of a child's enuresis: Some response-response relationships. *Journal of Applied Behavior Analysis*, 1971, 3, 241-247.  
 Nordquist, V.M.: A behavioral approach to the analysis of peer interactions. In M.J. Guralnick (Ed.): *Early Intervention and the Integration of Handicapped and Nonhandicapped Children*. Baltimore: University Park Press, 1978.  
 Paul, G.L.: Behavior modification research: Design and tactics. In C.M. Franks (Ed.): *Behavior Therapy: Appraisal and Status*. New York: McGraw-Hill, 1969.  
 Reynolds, M.C., Balow, B.: Categories and variables in special education. *Exceptional Children*, 1972, 38, 357-366.  
 Sidman, M.: *Tactics of Scientific Research*. New York: Basic Books, 1960.  
 Wahler, R.G.: Some structural aspects of deviant child behavior. *Journal of Applied Behavior Analysis*, 1975, 8, 27-42.  
 Wahler, R.G., Nordquist, V.M.: Adult discipline as a factor in childhood imitation. *Journal of Abnormal Child Psychology*, 1973, 1, 40-56.