Using a Single-Subject Research Design to Evaluate the Effectiveness of Treatment

Philippa H. Campbell

Key Words: pediatrics • program evaluation • single-subject research

Single-subject designs allow therapists to evaluate the effects of treatment on individual students, the effects of individualized occupational therapy services across a group of students, and the effects of a specific treatment procedure on a group of students. The purpose of this article is to provide information to enable therapists to use single-subject designs effectively in public school occupational therapy programs.

The use of a single-subject research design offers a mechanism for easily combining research with clinical practice (Ottenbacher & York, 1984). Randomization, experimental and control groups, and high numbers of similar subjects are not necessary when single-subject procedures are used. In addition, these procedures are easily combined with more traditional methods used to document progress in therapy (Campbell, Clegg, & McFarland, 1982). Single-subject research is derived from the field of applied behavior analysis and is used to study individual subjects who are not easily combined into groups. The use of these designs requires only one subject, who functions as his or her own control (subject) while receiving specifically designed treatment (Tawney & Gast, 1986). Complicated statistical procedures are not required but can be used if subjects are combined or if sophisticated single-subject designs are used.

Each pediatric occupational therapist who works in a school is responsible for documenting the effects of treatment by determining the extent to which the treatment is helping the infants and children who are receiving therapy. Most therapists use traditional methods, including initial evaluations of functioning, written treatment plans and goals, and progress records. The initial assessment measures include standardized instruments, such as developmental evaluations or tests of sensory integration, and less formal measures, such as clinical observation or various therapist-designed checklists. Readministration of instruments assesses change in performance. These methods may be adequate for the basic documentation of a child’s performance but are not sufficient as measures of the specific effects of treatment. Children may progress (or not progress) for a variety of reasons—one of which may be the type, quality, or quantity of the therapeutic methods used. Pediatric therapists must take responsibility not only for documenting children’s behavior but also for demonstrating that changes in behavior are the direct result of treatment. Using a single-subject design enables the therapist to document progress while also assessing the efficacy of therapeutic practices (Campbell & Stewart, 1986).

Essential Features of Single-Subject Designs

A number of different designs are used in single-subject research, some of which are more appropriate for particular situations than others (e.g., Ottenbacher & York, 1984; Worley & Harris, 1982). All share certain features, including (a) a dependent variable, which will be measured repeatedly; and (b) an independent variable, which is the treatment approach that will be used. Table 1 lists some basic designs and their essential features.
Table 1
Single-Subject Designs

<table>
<thead>
<tr>
<th>Design</th>
<th>Component</th>
<th>Distinguishing Feature(s)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reversal</td>
<td>ABA; ABBA; ABAB</td>
<td>Treatment is reversed, which should result in a return to baseline performance.</td>
<td>Tawney &amp; Gast, 1986</td>
</tr>
<tr>
<td>Multielement</td>
<td>ABCD ...</td>
<td>Different types of treatment procedures are used as required, or the number of treatment procedures is systematically decreased over time.</td>
<td>Murphey, Doughty, &amp; Nunes, 1979; Ullman &amp; Sulzewski- Azeroff, 1975</td>
</tr>
<tr>
<td>Changing criterion</td>
<td>ABCD ...</td>
<td>The criterion for performance changes over time.</td>
<td>Hartmann &amp; Hall, 1976</td>
</tr>
<tr>
<td>Multiple baseline</td>
<td>A</td>
<td>Data are simultaneously taken on several types of behavior, but only one behavior is the target of treatment at any time.</td>
<td>Hall, Cristler, Cranston, &amp; Tucker, 1970</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alternating treatment</td>
<td>AB; ABC</td>
<td>Several treatment procedures are used to address the same behavior; treatments may be alternating or randomly provided (e.g., AACBACB).</td>
<td>Kazdin &amp; Kopel, 1975</td>
</tr>
</tbody>
</table>

Baseline Phase

Most single-subject designs begin with a baseline phase, a period during which quantifiable levels of performance without any treatment are obtained on the dependent measure(s). Stable baseline is a term used to describe the level of baseline performance necessary before any treatment is initiated. Continuing the baseline phase until performance is stable is important for two primary reasons. The first is that the baseline phase, in most designs, represents the control (without treatment) measure of a subject's performance. These data establish that a subject is unable to perform on the selected dependent measure(s). The effects of treatment will be judged against performance in this no-treatment condition. The second reason is that the baseline performance defines the extent to which a subject requires treatment. The baseline phase, in this sense, functions as a precise assessment of a subject's performance on the identified dependent measure(s). The number of required baseline sessions will vary depending on performance conditions. The baseline phase, in general, must consist of a minimum of at least three to five data points to be interpretable (Alberto & Troutman, 1986). Continuing the baseline phase for at least three to five sessions can be difficult for adults who know that a child is unable to perform on the dependent measure. A stable and consistent baseline, however, assists in the interpretation of treatment data. Intervention effects can be judged more accurately when consistent and reliable baseline data have been obtained.

Reliability

Obtaining reliability is an often difficult aspect of single-subject research for therapists working in school-based programs. Data collected by two or more individuals at the same time and with similar results are considered reliable. Reliability measures are collected on at least 20% to 33% of data points in each design phase (baseline and intervention phases). Various methods can be used to calculate reliability, but the measure most typically used is to take the number of agreements divided by the number of agreements plus disagreements and multiply this figure by 100. The result is a percentage of agreement. Most researchers require a minimum of 85% to above 90% agreement, depending on the type of data collected.

Single-Subject Design in School-Based Therapy

Therapists in most school districts are required to establish goals and objectives for students included on direct-service, consultation, and monitoring case-loads. The annual review of the individualized education program (IEP) requires therapists to determine, through objective measures, the extent to which a student has achieved established goals and objectives. The single-subject design is an ideal approach for this purpose (e.g., Ottenbacher & York, 1984; Campbell & Stewart, 1986).

A second use for single-subject design is to evaluate the effects of specific treatment procedures with individual students. Occupational therapy methods
A program is effective when the majority of students are responsible for serving children with a variety of factors to measure overall progress in each student as an receiving programming benefit from the services provided. Occupational therapists in school settings and therapists, document the effectiveness of an occupational therapy program. Objective measures can be used to measure the effects of treatment with all students on therapy caseloads, but are probably best used for students for whom the use of repeated standardized assessments, such as the Miller Assessment Profile (Miller, 1982), is inappropriate.

**Evaluating Programs**

A program is effective when the majority of students receiving programming benefit from the services provided. Occupational therapists in school settings are responsible for serving children with a variety of unique needs. A single-subject design allows therapists to measure overall progress in each student as an individual while grouping results to assess overall program effectiveness (Campbell & Stewart, 1986). Single-subject measures, summarized across students and therapists, document the effectiveness of an occupational therapy program. Objective measures can be reported by groups of students in terms of a number of factors, such as (a) diagnostic category, (b) degree of severity of students’ disability, (c) age of students, (d) specific treatment used, or (e) model of service used.

Single-subject design was the basis of a program of motor therapy provided for 24 children with a variety of disorders who received individual therapy from one of three therapists on a twice-a-week basis. Therapists identified a goal (or goals) for each child, selected the dependent measure that would be used to measure performance, and developed a written plan that was followed during treatment sessions. Baseline measures were taken to establish behavior levels without treatment. Treatment procedures were begun after a stable baseline condition had been achieved. Data were summarized after 1 semester of treatment. Each child had the opportunity to receive therapy for a total of 30 treatment sessions during this time period. The results of therapeutic programming are presented in Table 2, with children grouped by major diagnosis.

The effectiveness of the treatment provided was assessed individually for each child. These results were grouped to determine how effective the therapy program as a whole was for all the children enrolled. A total of 29 goals was attempted across all 24 children. Baseline assessments (without intervention) were completed for 8 (27.5%) of the goals. The data for 14 of the remaining 21 goals showed acceleration. No change was reflected in the data for 7 goals. The specific treatment goals were those that showed no change on therapy goals were those...
Determining the Effects of a Particular Treatment With an Individual Child

The data from one child included in the total analysis presented in Table 2 are illustrated in Figure 1. This 2½-year-old child with mild spasticity in the lower extremities demonstrated delays in all areas of development. She was able to walk with assistance, but was unable to walk independently with normally coordinated movement in the lower extremities. The objective for therapy was to work on independent walking with emphasis on a neutrally aligned pelvis, neutral hip rotation, keeping the feet flat, and equal diagonal weight shifts. A program of facilitation was used to establish these motor components. Measurement of independent walking followed each therapy session and consisted of 10 trials where the child walked for a distance of 2 ft (0.61 m) in each trial. The therapist observed the walking pattern and indicated the first instance where any component of the pattern was marked incorrectly. As can be seen in Figure 1, an improved pattern was noted in neutral rotation of the hips, keeping the feet flat, and diagonal weight shift. Ability to maintain the pelvis in neutral showed almost no change. However, the child improved enough to be able to walk with all components correct for 30% of the trials on the final 2 days of therapy in this semester.

Determining the Effects of the Same Treatment With Groups of Children

The integration of clinical therapy and applied research is achieved when single-subject design is used to determine the effects of the same treatment with several different students. Single-subject studies of this type can be found in the professional literature, providing therapists with examples of designs and procedures that can be used in clinical application.

Single-subject designs have been employed to investigate the efficacy of various procedures used to enhance oral-motor skills in eating and drinking (Sobsey & Orelave, 1984). In this study, the number of bites with lip closure, rotary chewing, and the number of spills per bite were used as dependent measures to assess the effects of a series of oral-motor facilitation techniques that were provided for approximately 15 minutes before lunch. Single-subject design has also been used to study changes in postural reactions of students with cerebral palsy receiving a specific treatment designed to improve these reactions (Noonan, 1984) and to determine the effects of sensory integration procedures on changes in postrotatory nystagmus scores (Ottenbacher, 1982). Single-subject designs have been used to investigate the effects of specific procedures in improving the rates of switch activation of students with severe and multiple disabilities (Campbell & Mulhauser, 1986; Everson & Goodwyn, 1987; Giangreco, 1986) and to measure the effects of switch activation training with infants with disabilities (Brinker & Lewis, 1982). Detailed procedures for measuring quantitative changes in various types of motor behavior have been described by Campbell, Clegg, and McFarland (1982) and by Guess, Rues, Warren, and Lyon (1980), who edited a series of volumes that provide guidelines for measuring various aspects of motor behavior. Each of these studies, as well as others, provides examples of the application of single-subject methodology.

Implementing a Single-Subject Design

A number of useful references are available to help therapists plan and implement studies using single-subject designs (Campbell, Clegg, & McFarland, 1982; Ottenbacher & York, 1984; Tawney & Gast, 1986). Each of these references as well as studies that have used single-subject methodology can provide ideas concerning the selection of an appropriate design and appropriate measurement strategies.

Setting up a single-subject study requires five steps. The first is to determine a question that can be answered appropriately under single-subject procedures. Questions are stated to identify specifically both the independent (treatment/intervention) and dependent (measurement of outcome) variables under consideration. The way in which the question is phrased, however, guides the remaining steps of a study. Poorly written questions (such as, “Is NDT
[neurodevelopmental treatment] effective?" result in data collection that may be unnecessary, inappropriate, or insufficient to answer the stated question.

The second step is to determine the measurement strategies that will be used to assess outcome. Basic types of measurement include (a) frequency (counting the number of times a behavior occurs); (b) duration (timing the length of time a behavior occurs); and (c) rate of behavior (frequency divided by duration). A therapist might count the number of mouthfuls of food that a child eats using appropriate lip closure, the number of times that a child successfully produces a switch closure (to activate an object), or measure the length of time taken to complete a specific task (such as a vocational task). Measurement strategies range from simple to complex. For example, a therapist might count the number of times a child reaches and contacts an object while also recording the extent to which required movement components, such as forward humeral flexion or wrist extension, are correctly performed (e.g., Campbell & Stewart, 1986; Valvano & DeGangi, 1986).

The third step is to select the most appropriate treatment design. Some questions are better addressed by particular design structures, and some designs are more easily implemented within specific situations. For example, an alternating treatment design (see Table 1) would be quite difficult to implement with a child who received intervention on a once-a-week basis. A multiple baseline or changing criterion design would be easier to implement in this situation because of the relatively infrequent opportunities for treatment and measurement.

Describing the treatment procedures that will be used is the fourth step. The length of a treatment session (e.g., 30 minutes) and the number of sessions (e.g., 3 times a week; a total of 15 sessions) to be provided are specified. The treatment methods and procedures that will be used in sessions should be described in sufficient detail to allow the same treatment to be provided in each session. A therapist might describe treatment procedures generally as "NDT procedures" or "SI [sensory integration] therapy." The more specifically described the treatment procedures are, however, the easier it will be to implement consistent practice across a number of treatment sessions. Moreover, the more precisely the procedures are described, the easier it will be to determine their effects on a child's performance. For example, a therapist, who uses and describes specific facilitation procedures to help the student stabilize the shoulder to allow for a normal pattern of upper extremity movement for bringing the hand to the mouth in self-feeding will have a better idea of what procedures were effective than a therapist who describes the procedures simply as "NDT."

Therapy treatment techniques can be assessed as a "package." However, such an approach weakens the conclusions that can be drawn after the study is completed. For example, a therapist might decide to treat a child with a package of procedures designed to increase swallowing and achieve lip and jaw closure. Techniques such as jaw control, deep pressure with the finger under the jaw to facilitate swallowing, and oral sensitization procedures carried out before eating could be combined and defined as a package. However, it would not be possible to determine the specific procedure that influenced a child's swallowing and lip and jaw closure abilities (assuming that progress is demonstrated).

The fifth and final step is data collection and interpretation. Data are collected on the specific measures (defined in Step 2 above) and in accord with a preestablished schedule. These data are recorded on individual data sheets and data summary sheets. Data may also be graphed. Data are the basis for interpreting the effects of treatment. Essentially, a child's performance changes in the desired direction (shows an increase or a decrease) when treatment is effective and is either variable, unchanging, or changed in the wrong direction (e.g., shows an increase when a decrease was desired) when treatment is not effective. The extent to which changes in performance can be attributed reasonably to the treatment procedures implemented and not to some other factor (e.g., medication changes; new adaptive equipment) is inferred through a visual inspection of the graphed data and, in specific instances, through statistical analysis.

**Interpretation and Use of Results**

A visual inspection of the graphed data and inferences about the effects of treatment are made in accord with the particular design structure used. For example, in an ABA or reversal design that contrasts a no-treatment condition with a treatment condition, the probability that treatment was the important factor in changing a child's behavior is increased when the child demonstrates different levels of performance in the A and B conditions. Using a reversal design, a therapist who wished to measure the rate at which a child activated a particular switch would provide no facilitation of upper extremity movement before the child hit the switch during the A (no-treatment) condition but would provide facilitation in the B (treatment) condition. If the rate of activation were higher in the treatment than in the no-treatment condition, the therapist could conclude that the treatment was effective. On the other hand, if the rates of activation were similar in both conditions, the therapist could conclude that the treatment was ineffective.

One of the most important reasons for collecting data on individual students is to provide a data base
for decision making. Data that are summarized and
graphed regularly (e.g., weekly) can be analyzed by a
therapist to monitor the ongoing effects of treatment.
A particular treatment procedure that has been imple­
menced for a reasonable period of time (e.g., five
treatment sessions) and that is not altering a child's
performance can be modified quickly, thereby poten­
tially ensuring greater levels of progress. An ongoing
data base is helpful, particularly for students for whom
changes in performance are difficult to determine on
the basis of clinical observation alone (e.g., students
with severe and multiple disabilities). Regular review
of a student's data can prevent ineffective treatment
from being continued indefinitely (or until the IEP is
reviewed). Figure 2 provides an example of a deci­
sion-making chart for ongoing data analysis.

Many factors potentially influence the interpreta­
tion of the results of a single-subject study, including
(a) the type of design, (b) the measurement strategy,
(c) consistency in the delivery of treatment, and (d)
reliability of measurement. The potential impact of
each of these factors must be considered both before
and after the implementation of the study. A therapist
can eliminate or minimize the effects of each of these
factors by planning for their potential influence when
designing the study. Realistically, it may not be possi­
ble to minimize the effects of all factors, no matter
what design is selected or how carefully the interven­
tion is provided and monitored. Single-subject design
provides valuable information for therapists—even
when all potentially influencing factors are not well
controlled—and offers an easily implemented pro­
cess for integrating research with clinical practice.

References
analysis for teachers. Columbus, OH: Merrill.
work with microcomputers: A learning prosthesis for handi­
capped infants. Exceptional Children, 49, 163–172.
Measuring motor behavior. In K. Stremel-Campbell (Ed.),
Data collection in the classroom. (pp. 59–89) Seattle: WES­
TAR.
Campbell, P. H., McInerney, W. F., & Cooper, M. A.


