

The Use of Single-Subject Research to Identify Evidence-Based Practice in Special Education

ROBERT H. HORNER
University of Oregon

EDWARD G. CARR
State University of New York at Stony Brook

JAMES HALLE
University of Illinois

GAIL MCGEE
Emory University

SAMUEL ODOM
Indiana University

MARK WOLERY
Vanderbilt University

ABSTRACT: *Single-subject research plays an important role in the development of evidence-based practice in special education. The defining features of single-subject research are presented, the contributions of single-subject research for special education are reviewed, and a specific proposal is offered for using single-subject research to document evidence-based practice. This article allows readers to determine if a specific study is a credible example of single-subject research and if a specific practice or procedure has been validated as “evidence-based” via single-subject research.*

Single-subject research is a rigorous, scientific methodology used to define basic principles of behavior and establish evidence-based practices. A long and productive history exists in which single-subject research has provided useful information for

the field of special education (Kennedy, in press; Odom & Strain, 2002; Tawney & Gast, 1984; Wolery & Dunlap, 2001). Since the methodology was first operationalized over 40 years ago (Sidman, 1960), single-subject research has proven particularly relevant for defining educational practices at the level of the individual

learner. Educators building individualized educational and support plans have benefited from the systematic form of experimental analysis single-subject research permits (Dunlap & Kern, 1997). Of special value has been the ability of single-subject research methods to provide a level of experimental rigor beyond that found in traditional case studies. Because single-subject research documents experimental control, it is an approach, like randomized control-group designs (Shavelson & Towne, 2002), that may be used to establish evidence-based practices.

The systematic and detailed analysis of individuals that is provided through single-subject research methods has drawn researchers not only from special education, but also from a growing array of scholarly disciplines, with over 45 professional journals now reporting single-subject research (American Psychological Association, 2002; Anderson, 2001). Further, an array of effective interventions is now in use that emerged through single-subject research methods. Reinforcement theory or operant psychology has been the substantive area that has benefited most from single-case research methodology. In fact, operant principles of behavior have been empirically demonstrated and replicated within the context of single-subject experiments for more than 70 years. However, the close association between operant analysis of human behavior and single-subject experimental research is not exclusionary. That is, many procedures based on diverse theoretical approaches to human behavior can be evaluated within the confines of single-subject research. Interventions derived from social-learning theory, medicine, social psychology, social work, and communication disorders are but a sample of procedures that have been analyzed by single-subject designs and methods (cf., Hersen & Barlow, 1976; Jayaratne & Levy, 1979; McReynolds & Kearns, 1983).

The specific goals of this article are to (a) present the defining features of single-subject research methodology, (b) clarify the relevance of single-subject research methods for special education, and (c) offer objective criteria for determining when single-subject research results are sufficient for documenting evidence-based practices. Excellent introductions to single-subject research exist (Hersen & Barlow, 1976; Kazdin,

1982; Kratochwill & Levin, 1992; Richard, Taylor, Ramasamy & Richards, 1999; Tawney & Gast, 1984), and our goal here is not to provide an introduction to the single-subject research, but to clarify how single-subject research is used to establish knowledge within special education and define the empirical support needed to document evidence-based practices.

SINGLE-SUBJECT RESEARCH METHODOLOGY

Single-subject research is experimental rather than correlational or descriptive, and its purpose is to document causal, or functional, relationships between independent and dependent variables. Single-subject research employs within- and between-subjects comparisons to control for major threats to internal validity and requires systematic replication to enhance external validity (Martella, Nelson, & Marchand-Martella, 1999). Several critical features define this methodology. Each feature is described in the following sections and organized later in a table of quality indicators that may be used to assess if an individual study is an acceptable exemplar of single-subject research.

THE INDIVIDUAL PARTICIPANT IS THE UNIT OF ANALYSIS

Single-subject designs may involve only one participant, but typically include multiple participants (e.g., 3 to 8) in a single study. Each participant serves as his or her own control. Performance prior to intervention is compared to performance during and/or after intervention. In most cases a research participant is an individual, but it is possible for each participant to be a group whose performance generates a single score per measurement period (e.g., the rate of problem behavior performed by all children within a classroom during a 20 min period).

PARTICIPANT AND SETTING DESCRIPTION

Single-subject research requires operational descriptions of the participants, setting, and the process by which participants were selected (Wolery & Ezell, 1993). Another researcher should be able to use the description of participants and setting to recruit similar participants who inhabit similar

settings. For example, operational participant descriptions of individuals with a disability would require that the specific disability (e.g., autism spectrum disorder, Williams syndrome) and the specific instrument and process used to determine their disability (e.g., the Autism Diagnostic Interview–Revised) be identified. Global descriptions such as identifying participants as having developmental disabilities would be insufficient.

DEPENDENT VARIABLE

Single-subject research employs one or more dependent variables that are defined and measured. In most cases the dependent variable in single-subject educational research is a form of observable behavior. Appropriate application of single-subject methodology requires dependent variables to have the following features:

- *Dependent variables are operationally defined* to allow (a) valid and consistent assessment of the variable and (b) replication of the assessment process. Dependent variables that allow direct observation and empirical summary (e.g., words read correctly per min; frequency of head hits per min; number of s between request and initiation of compliance) are desirable. Dependent variables that are defined subjectively (e.g., frequency of helping behaviors, with no definition of “helping” provided) or too globally (e.g., frequency of “aggressive” behavior) would not be acceptable.
- *Dependent variables are measured repeatedly* within and across controlled conditions to allow (a) identification of performance patterns prior to intervention and (b) comparison of performance patterns across conditions/phases. The repeated measurement of individual behaviors is critical for comparing the performance of each participant with his or her own prior performance. Within an experimental phase or condition, sufficient assessment occasions are needed to establish the overall pattern of performance under that condition (e.g., level, trend, variability). Measurement of the behavior of the same individual across phases or conditions allows comparison of performance patterns under different environmental conditions.

Single-subject research is experimental rather than correlational or descriptive, and its purpose is to document causal, or functional, relationships between independent and dependent variables.

- *Dependent variable recording is assessed for consistency* throughout the experiment by frequent monitoring of interobserver agreement (e.g., the percentage of observational units in which independent observers agree) or an equivalent. The measurement of interobserver agreement should allow assessment for each variable across each participant in each condition of the study. Reporting interobserver agreement only for the baseline condition or only as one score across all measures in a study would not be appropriate.
- *Dependent variables are selected for their social significance.* A dependent variable is chosen not only because it may allow assessment of a conceptual theory, but also because it is perceived as important for the individual participant, those who come in contact with the individual, or for society.

INDEPENDENT VARIABLE

The independent variable in single-subject research typically is the practice, intervention, or behavioral mechanism under investigation. Independent variables in single-subject research are operationally defined to allow both valid interpretation of results and accurate replication of the procedures. Specific descriptions of procedures typically include documentation of materials (e.g., 7.5 cm × 12.5 cm card) as well as actions (e.g., peer tutors implemented the reading curriculum in a 1:1 context, 30 min per day, 3 days per week). General descriptions of an intervention procedure (e.g., cooperative play) that are prone to high variability in implementation would not meet the expectation for operational description of the independent variable.

To document experimental control, *the independent variable in single-subject research is actively, rather than passively, manipulated.* The

researcher must determine when and how the independent variable will change. For example, if a researcher examines the effects of hard versus easy school work (independent variable) on rates of problem behavior (dependent variable), the researcher would be expected to operationally define, and systematically introduce, hard and easy work rather than simply observe behavior across the day as work of varying difficulty was naturally introduced.

In single-subject research the *fidelity of independent variable implementation* is documented. Fidelity of implementation is a significant concern within single-subject research because the independent variable is applied over time. As a result, documentation of adequate implementation fidelity is expected either through continuous direct measurement of the independent variable, or an equivalent (Gresham, Gansel, & Kurtz, 1993).

BASELINE/COMPARISON CONDITION

Single-subject research designs typically compare the effects of an intervention with performance during a *baseline*, or comparison, condition. The baseline condition is similar to a *treatment as usual* condition in group designs. Single-subject research designs compare performance during the baseline condition, and then contrast this pattern with performance under an intervention condition. The emphasis on comparison across conditions requires measurement during, and detailed description of, the baseline (or comparison) condition. Description of the baseline condition should be sufficiently precise to allow replication of the condition by other researchers.

Measurement of the dependent variable during a baseline should occur until the observed pattern of responding is sufficiently consistent to allow prediction of future responding. Documentation of a predictable pattern during baseline

To document experimental control, the independent variable in single-subject research is actively, rather than passively, manipulated.

typically requires multiple data points (five or more, although fewer data points are acceptable in specific cases) without substantive trend, or with a trend in the direction opposite that predicted by the intervention. Note that if the data in a baseline documents a trend in the direction predicted by the intervention, then the ability to document an effect following intervention is compromised.

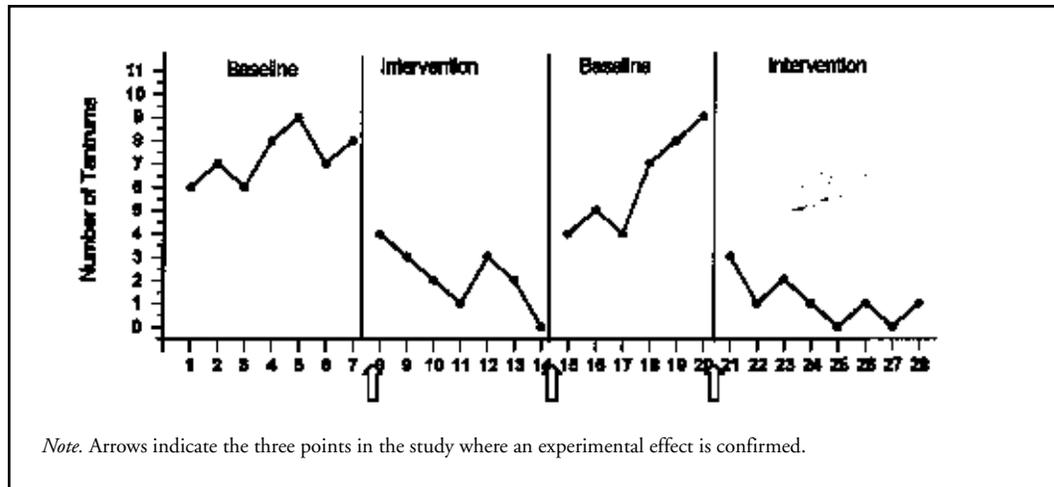
EXPERIMENTAL CONTROL

Single-subject research designs provide experimental control for most threats to internal validity and, thereby, allow confirmation of a functional relationship between manipulation of the independent variable and change in the dependent variable. In most cases *experimental control is demonstrated when the design documents three demonstrations of the experimental effect at three different points in time with a single participant (within-subject replication), or across different participants (inter-subject replication)*. An experimental effect is demonstrated when predicted change in the dependent variable covaries with manipulation of the independent variable (e.g., the level, and/or variability of the dataset in a phase decreases when a behavior-reduction intervention is implemented, or the level and/or variability of the dataset in a phase increases when the behavior-reduction intervention is withdrawn). Documentation of experimental control is achieved through (a) the introduction and withdrawal (or reversal) of the independent variable; (b) the staggered introduction of the independent variable at different points in time (e.g., multiple baseline); or (c) the iterative manipulation of the independent variable (or levels of the independent variable) across observation periods (e.g., alternating treatments designs).

For example, Figure 1 presents a typical A (Baseline)–B (Intervention)–A (Baseline 2)–B (Intervention 2) single-subject research design that establishes three demonstrations of the experimental effect at three points in time through demonstration that behavior change covaries with manipulation (introduction and removal) of the independent variable between Baseline and Intervention phases. Three demonstrations of an experimental effect are documented at the three arrows in Figure 1 by (a) an initial reduction in

FIGURE 1

Example of a Single-Subject Reversal Design Demonstrating Experimental Control



tantrums between the first A phase (Baseline) and the first B phase (Intervention); (b) a second change in response patterns (e.g., return to Baseline patterns) with re-introduction of the Baseline conditions in the second A phase; and (c) a third change in response patterns (e.g., reduction in tantrums) with re-introduction of the intervention in the second B phase.

A similar logic for documenting experimental control exists for multiple baseline designs with three or more data series. The staggered introduction of the intervention within a multiple baseline design allows demonstration of the experimental effect not only within each data series, but also across data series at the staggered times of intervention. Figure 2 presents a design that includes three series, with introduction of the intervention at a different point in time for each series. The results document experimental control by demonstrating a covariation between change in behavior patterns and introduction of the intervention within three different series at three different points in time.

Excellent sources exist describing the growing array of single-subject designs that allow documentation of experimental control (Hersen & Barlow, 1976; Kazdin, 1982, 1998; Kennedy, in press; Kratochwill & Levin, 1992; McReynolds & Kearns, 1983; Richard, et al., 1999; Tawney & Gast, 1984). Single-subject designs provide experimental documentation of unequivocal relation-

ships between manipulation of independent variables and change in dependent variables. Rival hypotheses (e.g., passage of time, measurement effects, uncontrolled variables) must be discarded to document experimental control. Traditional *case study* descriptions, or studies with only a baseline followed by an intervention, may provide useful information for the field, but do not provide adequate experimental control to qualify as single-subject research.

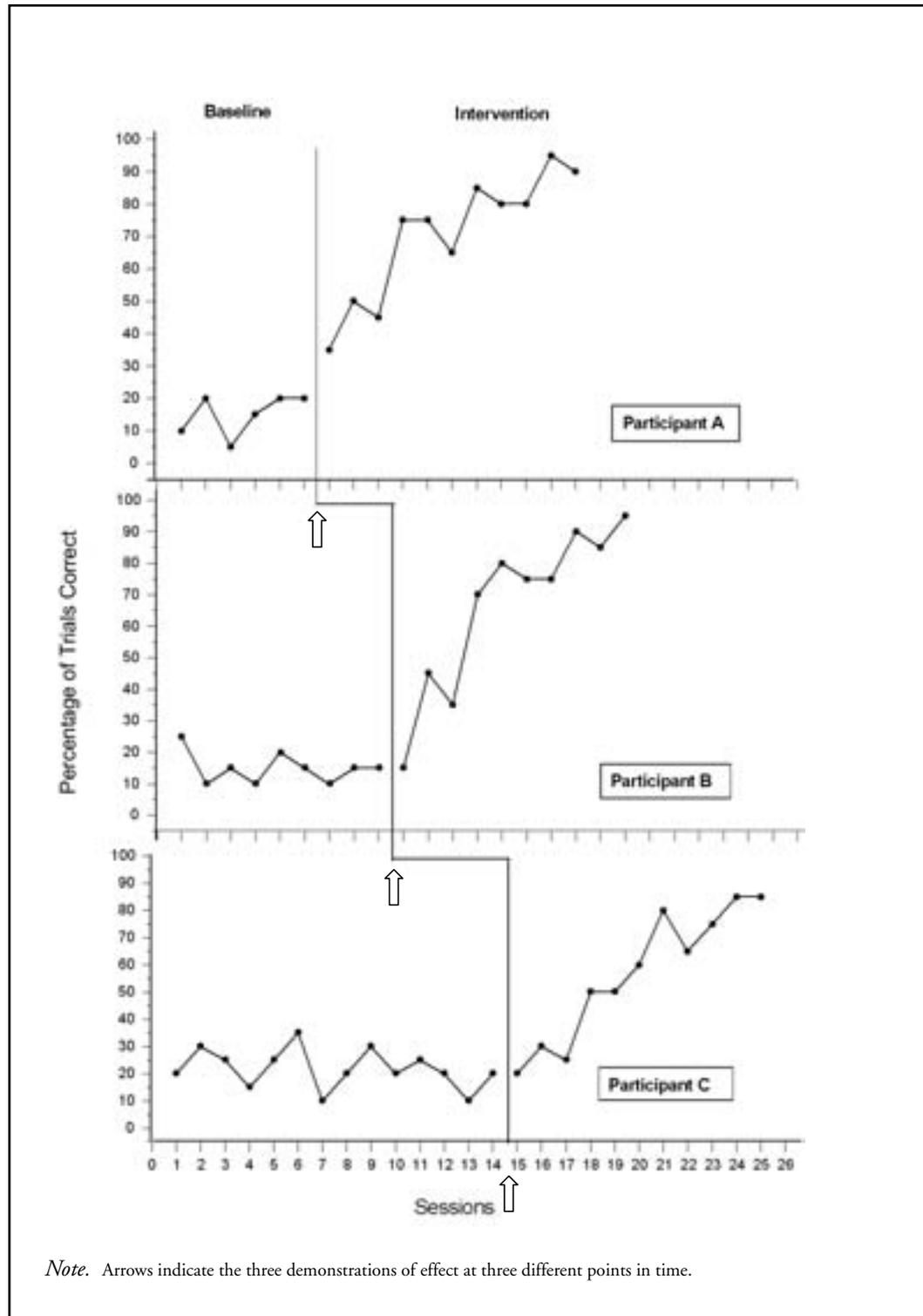
VISUAL ANALYSIS

Single-subject research results may be interpreted with the use of statistical analyses (Todman & Dugard, 2001); however, the traditional approach to analysis of single-subject research data involves systematic visual comparison of responding within and across conditions of a study (Parsonson & Baer, 1978). Documentation of experimental control requires assessment of all conditions within the design. Each design (e.g., reversal, multiple baseline, changing criterion, al-

Single-subject designs provide experimental documentation of unequivocal relationships between manipulation of independent variables and change in dependent variables.

FIGURE 2

Example of a Multiple Baseline Design Across Participants That Demonstrates Experimental Control



Note. Arrows indicate the three demonstrations of effect at three different points in time.

ternating treatments) requires a specific data pattern for the researcher to claim that change in the dependent variable is, and only is, a function of manipulating the independent variable.

Visual analysis involves interpretation of the level, trend, and variability of performance occurring during baseline and intervention conditions. Level refers to the mean performance during a condition (i.e., phase) of the study. Trend references the rate of increase or decrease of the best-fit straight line for the dependent variable within a condition (i.e., slope). Variability refers to the degree to which performance fluctuates around a mean or slope during a phase. In visual analysis, the reader also judges (a) the immediacy of effects following the onset and/or withdrawal of the intervention, (b) the proportion of data points in adjacent phases that overlap in level, (c) the magnitude of changes in the dependent variable, and (d) the consistency of data patterns across multiple presentations of intervention and nonintervention conditions. The integration of information from these multiple assessments and comparisons is used to determine if a functional relationship exists between the independent and dependent variables.

Documentation of a functional relationship requires compelling demonstration of an effect (Parsonson & Baer, 1992). Demonstration of a functional relationship is compromised when (a) there is a long latency between manipulation of the independent variable and change in the dependent variable, (b) mean changes across conditions are small and/or similar to changes within conditions, and (c) trends do not conform to those predicted following introduction or manipulation of the independent variable.

A growing set of models also exists for conducting meta-analysis of single-subject research (Busk & Serlin, 1992; Didden, Duker, & Korzilius, 1997; Faith, Allison, & Gorman, 1996; Hershberger, Wallace, Green, & Marquis, 1999; Marquis et al., 2000). This approach to analysis is of special value in documentation of comparative trends in a field.

EXTERNAL VALIDITY

Single-subject designs are used to (a) test conceptual theory and (b) identify and validate effective clinical interventions. A central concern is the ex-

tent to which an effect documented by one study has relevance for participants, locations, materials, and behaviors beyond those defined in the study. External validity of results from single-subject research is enhanced through replication of the effects across different participants, different conditions, and/or different measures of the dependent variable.

Although a study may involve only one participant, features of external validity of a single study are improved if the study includes multiple participants, settings, materials, and/or behaviors. It is typical for single-subject studies to demonstrate effects with at least three different participants. It also is expected that the generality and/or “boundaries” of an intervention will be established not by a single study, but through systematic replication of effects across multiple studies conducted in multiple locations and across multiple researchers (Birnbrauer, 1981). External validity in single-subject research also is enhanced through operational description of (a) the participants, (b) the context in which the study is conducted, and (c) the factors influencing a participant’s behavior prior to intervention (e.g., assessment and baseline response patterns).

The external validity for a program of single-subject studies is narrowed when selection and attrition bias (e.g., the selection of only certain participants, or the publication of only successful examples) limit the range of examples available for analysis (Durand & Rost, in press). Having and reporting specific selection criteria, however, assist in defining for whom, and under what conditions a given independent variable is likely to result in defined changes in the dependent measures. Attrition is a potent threat to both the internal and external validity of single-subject studies, and any participant who experienced

External validity of results from single-subject research is enhanced through replication of the effects across different participants, different conditions, and/or different measures of the dependent variable.

both conditions (i.e., baseline and intervention) of a study should be included in reports of that study.

SOCIAL VALIDITY

Within education, single-subject research has been used not only to identify basic principles of behavior (e.g., theory), but also to document interventions (independent variables) that are functionally related to change in socially important outcomes (dependent variables; Wolf, 1978). The emphasis on intervention has resulted in substantial concern about the social validity, or practicality, of research procedures and findings. The social validity of single-subject research goals, procedures and findings is enhanced by:

- Emphasis on the selection of dependent variables that have high social importance.
- Demonstration that the independent variables can be applied with fidelity by typical intervention agents (e.g., teachers, parents) in typical intervention contexts across meaningful periods of time.
- Demonstration that typical intervention agents (a) report the procedures to be acceptable, (b) report the procedures to be feasible within available resources, (c) report the procedure to be effective, and (d) choose to continue use of the intervention procedures after formal support/expectation of use is removed. For example, an effective procedure designed for use by young parents where the procedure fits within the daily family routines would have good social validity, whereas an intervention that disrupted family routines and compromised the ability of a family to function

Within education, single-subject research has been used not only to identify basic principles of behavior (e.g., theory), but also to document interventions (independent variables) that are functionally related to change in socially important outcomes (dependent variables).

normatively would *not* have good social validity.

- Demonstration that the intervention produced an effect that met the defined, clinical need.

Within special education, single-subject research has been used to examine strategies for building academic achievement (Greenwood, Tapia, Abbott & Walton, 2003; Miller, Gunter, Venn, Hummel, & Wiley, 2003; Rohena, Jitendra & Browder, 2002); improving social behavior and reducing problem behavior (Carr et al., 1999; Koegel & Koegel, 1986, 1990); and enhancing the skills of teachers (Moore et al., 2002) or families who implement interventions (Cooper, Wacker, Sasso, Reimers, & Donn, 1990; Hall et al., 1972).

Single-subject research also can be used to emphasize important distinctions between, and integration of, efficacy research (documentation that an experimental effect can be obtained under carefully controlled conditions) and effectiveness research (documentation that an experimental effect can be obtained under typical conditions) that may affect large-scale implementation of a procedure (Flay, 1986).

RESEARCH QUESTIONS APPROPRIATE FOR SINGLE-SUBJECT RESEARCH METHODS

The selection of any research methodology should be guided, in part, by the research question(s) under consideration. No research approach is appropriate for all research questions, and it is important to clarify the types of research questions that any research method is organized to address. Single-subject research designs are organized to provide fine-grained, time-series analysis of change in a dependent variable(s) across systematic introduction or manipulations of an independent variable. They are particularly appropriate when one wishes to understand the performance of a specific individual under a given set of conditions.

Research questions appropriately addressed with single-subject methods (a) examine causal, or functional, relations by examining the effects that introducing or manipulating an independent variable (e.g., an intervention) has on change in one or more dependent variables; (b) focus on the effects that altering a component of a multicomponent independent variable (e.g., an interven-

tion package) has on one or more dependent variables; or (c) focus on the relative effects of two or more independent variable manipulations (e.g., alternative interventions) on one or more dependent variables. Examples of research questions appropriately addressed by single-subject methods include

- Does functional communication training reduce problem behavior?
- Do incidental teaching procedures increase social initiations by young children with autism?
- Is time delay prompting or least-to-most prompt hierarchy more effective in promoting self-help skills of young children with severe disabilities?
- Does pacing of reading instruction increase the rate of acquisition of reading skills by third graders?
- Does the use of a new drug for children with AD/HD result in an increase in sustained attention?

QUALITY INDICATORS FOR SINGLE-SUBJECT RESEARCH

By its very nature, research is a process of approximations. The features listed previously define the core elements of single-subject research methodology, but we recognize that these features will be met with differing levels of precision. We also recognize that there are conditions in which exceptions are appropriate. It is important, therefore, to offer guidance for assessing the degree to which single-subject research methods have been applied adequately within a study, and an objective standard for determining if a particular study meets the minimally acceptable levels that permit interpretation.

Impressive efforts exist for quantifying the methodological rigor of specific single-subject studies (Busk & Serlin, 1992; Kratochwill, & Stoiber, 2002). In combination with the previous descriptions, we offer the information in Table 1 as content for determining if a study meets the “acceptable” methodological rigor needed to be a credible example of single-subject research.

Single-subject research designs are organized to provide fine-grained, time-series analysis of change in a dependent variable(s) across systematic introduction or manipulations of an independent variable.

IMPORTANCE OF SINGLE-SUBJECT RESEARCH METHODS FOR RESEARCH IN SPECIAL EDUCATION

Single-subject research methods offer a number of features that make them particularly appropriate for use in special education research. Special education is a field that emphasizes (a) the individual student as the unit of concern, (b) active intervention, and (c) practical procedures that can be used in typical school, home, and community contexts. Special education is a problem-solving discipline, in which ongoing research in applied settings is needed. Single-subject research matches well with the needs of special education in the following ways.

- *Single-subject research focuses on the individual.* Causal, or functional, relationships can be identified without requiring the assumptions needed for parametric analysis (e.g., normal distribution). Research questions in special education often focus on low-incidence or heterogeneous populations. Information about mean performance of these groups may be of less value for application to individuals. Single-subject methods allow targeted analysis at the unit of the “individual,” the same unit at which the intervention will be delivered.
- *Single-subject research allows detailed analysis of “nonresponders” as well as “responders.”* Control group designs produce conclusions about the generality of treatment effects as they relate to group means, not as they relate to specific individuals. Even in the most successful group designs, there are individuals whose behavior remains unaffected, or is made worse, by the treatment (e.g., “nonresponders”). Single-subject designs provide an empirically rigorous method for analyzing the characteristics of these nonresponders, thereby advancing

TABLE 1*Quality Indicators Within Single-Subject Research**Description of Participants and Settings*

- Participants are described with sufficient detail to allow others to select individuals with similar characteristics (e.g., age, gender, disability, diagnosis).
- The process for selecting participants is described with replicable precision.
- Critical features of the physical setting are described with sufficient precision to allow replication.

Dependent Variable

- Dependent variables are described with operational precision.
- Each dependent variable is measured with a procedure that generates a quantifiable index.
- Measurement of the dependent variable is valid and described with replicable precision.
- Dependent variables are measured repeatedly over time.
- Data are collected on the reliability or interobserver agreement associated with each dependent variable, and IOA levels meet minimal standards (e.g., IOA = 80%; Kappa = 60%).

Independent Variable

- Independent variable is described with replicable precision.
- Independent variable is systematically manipulated and under the control of the experimenter.
- Overt measurement of the fidelity of implementation for the independent variable is highly desirable.

Baseline

- The majority of single-subject research studies will include a baseline phase that provides repeated measurement of a dependent variable and establishes a pattern of responding that can be used to predict the pattern of future performance, if introduction or manipulation of the independent variable did not occur.
- Baseline conditions are described with replicable precision.

Experimental Control/Internal Validity

- The design provides at least three demonstrations of experimental effect at three different points in time.
- The design controls for common threats to internal validity (e.g., permits elimination of rival hypotheses).
- The results document a pattern that demonstrates experimental control.

External Validity

- Experimental effects are replicated across participants, settings, or materials to establish external validity.

Social Validity

- The dependent variable is socially important.
- The magnitude of change in the dependent variable resulting from the intervention is socially important.
- Implementation of the independent variable is practical and cost effective.
- Social validity is enhanced by implementation of the independent variable over extended time periods, by typical intervention agents, in typical physical and social contexts.

knowledge about the possible existence of subgroups and subject-by-treatment interactions. Analysis of nonresponders also allows identification of intervention adaptations needed to produce intended outcomes with a wider range of participants.

Single-subject research methods offer a number of features that make them particularly appropriate for use in special education research.

- *Single-subject research provides a practical methodology for testing educational and behavioral interventions.* Single-subject methods allow unequivocal analysis of the relationship between individualized interventions and change in valued outcomes. Through replication, the methodology also allows testing of the breadth, or external validity, of findings.
- *Single-subject research provides a practical research methodology for assessing experimental effects under typical educational conditions.* Single-subject designs evaluate interventions

(independent variables) under conditions similar to those recommended for special educators, such as repeated applications of a procedure over time. This allows assessment of the *process* of change as well as the *product* of change, and facilitates analysis of maintenance as well as initial effects.

- *Single-subject research designs allow testing of conceptual theory.* Single-subject designs can be used to test the validity of theories of behavior that predict conditions under which behavior change (e.g., learning) should and should not occur.
- *Single-subject research methods are a cost-effective approach to identifying educational and behavioral interventions that are appropriate for large-scale analysis.* Single-subject research methods, when applied across multiple studies, can be used to guide large-scale policy directives. Single-subject research also can be used cost effectively to produce a body of reliable, persuasive evidence that justifies investment in large, often expensive, randomized control group designs. The control group designs, in turn, can be used to further demonstrate external validity of findings established via single-subject methodology.

THE IDENTIFICATION OF EVIDENCE-BASED PRACTICES USING SINGLE-SUBJECT RESEARCH

Current legislation and policy within education emphasize commitment to, and dissemination of, evidence-based (or research-validated) practices (Shavelson & Towne, 2002). Appropriate concern exists that investment in practices that lack adequate empirical support may drain limited educational resources and, in some cases, may result in the use of practices that are not in the best interest of children (Beutler, 1998; Nelson, Roberts, Mathur, & Rutherford, 1999; Whitehurst, 2003). To support the investment in evidence-based practices, it is appropriate for any research method to define objective criteria that local, state or federal decision makers may use to determine if a practice is evidence based (Chambless & Hollon, 1998; Chambless & Ollendick, 2001; Odom

& Strain, 2002; Shernoff, Kratochwill, & Stoiber, 2002). This is a logical, but not easy, task (Christenson, Carlson, & Valdez, 2002). We provide here a context for using single-subject research to document evidence-based practices in special education that draws directly from recommendations by the Task Force on Evidence-Based Interventions in School Psychology (Kratochwill & Stoiber, 2002), and the Committee on Science and Practice, Division 12, American Psychological Association (Weisz & Hawley, 2002).

A practice refers to a curriculum, behavioral intervention, systems change, or educational approach designed for use by families, educators, or students with the express expectation that implementation will result in measurable educational, social, behavioral, or physical benefit. A practice may be a precise intervention (e.g., functional communication training; Carr & Durand, 1985), a procedure for documenting a controlling mechanism (e.g., the use of high-probability requests to create behavioral momentum; Mace et al., 1988), or a larger program with multiple components (e.g., direct instruction; Gettinger, 1993).

Within single-subject research methods, as with other research methods, the field is just beginning the process of determining the professional standards that allow demonstration of an evidence-based practice (Kratochwill & Stoiber, 2002). It is prudent to propose initial standards that are conservative and draw from existing application in the field (e.g., build from examples of practices such as functional communication training that are generally accepted as evidence based). We propose five standards that may be applied to assess if single-subject research results document a practice as evidence based. The standards were drawn from the conceptual logic for single-subject methods (Kratochwill & Stoiber), and from standards proposed for identifying evidence-based practices using group designs (Shavelson & Towne, 2002).

Single-subject research documents a practice as evidence based when (a) the practice is operationally defined; (b) the context in which the practice is to be used is defined; (c) the practice is implemented with fidelity; (d) results from single-subject research document the practice to be functionally related to change in dependent measures; and (e) the experimental effects are repli-

Appropriate concern exists that investment in practices that lack adequate empirical support may drain limited educational resources and, in some cases, may result in the use of practices that are not in the best interest of children.

cated across a sufficient number of studies, researchers, and participants to allow confidence in the findings. Each of these standards is elaborated in the following list.

- *The practice is operationally defined.* A practice must be described with sufficient precision so that individuals other than the developers can replicate it with fidelity.
 - *The context and outcomes associated with a practice are clearly defined.* Practices seldom are expected to produce all possible benefits for all individuals under all conditions. For a practice to be considered evidence based it must be defined in a context. This means operational description of (a) the specific conditions where the practice should be used, (b) the individuals qualified to apply the practice, (c) the population(s) of individuals (and their functional characteristics) for whom the practice is expected to be effective, and (d) the specific outcomes (dependent variables) affected by the practice. Practices that are effective in typical performance settings such as the home, school, community, and workplace are of special value.
 - *The practice is implemented with documented fidelity.* Single-subject research studies should provide adequate documentation that the practice was implemented with fidelity.
 - *The practice is functionally related to change in valued outcomes.* Single-subject research studies should document a causal, or functional, relationship between use of the practice and change in a socially important dependent variable by controlling for the effects of extraneous variables.
- *Experimental control is demonstrated across a sufficient range of studies, researchers, and participants to allow confidence in the effect.* Documentation of an evidence-based practice typically requires multiple single-subject studies. We propose the following standard: A practice may be considered evidence based when (a) a minimum of five single-subject studies that meet minimally acceptable methodological criteria and document experimental control have been published in peer-reviewed journals, (b) the studies are conducted by at least three different researchers across at least three different geographical locations, and (c) the five or more studies include a total of at least 20 participants.

An example of applying these criteria is provided by the literature assessing functional communication training (FCT). As a practice, FCT involves (a) using functional assessment procedures to define the consequences that function as reinforcers for undesirable behavior, (b) teaching a socially acceptable, and equally efficient, alternative behavior that produces the same consequence as the undesirable behavior, and (c) minimizing reinforcement of the undesirable behavior. Documentation of this practice as evidence-based is provided by the following citations, which demonstrate experimental effects in eight peer-reviewed articles across five major research groups and 42 participants (Bird, Dores, Moniz, & Robinson, 1989; Brown et al., 2000; Carr & Durand, 1985; Durand & Carr, 1987, 1991; Hagopian, Fisher, Sullivan, Acquistio, & LeBlanc, 1998; Mildon, Moore, & Dixon, 2004; Wacker et al., 1990).

CONCLUSION

We offer a concise description of the features that define single-subject research, the indicators that can be used to judge quality of single-subject research, and the standards for determining if an intervention, or practice, is validated as evidence based via single-subject methods. Single-subject research offers a powerful and useful methodology for improving the practices that benefit individuals with disabilities and their families. Any systematic policy for promoting the development

and/or dissemination of evidence-based practices in education should include single-subject research as an encouraged methodology.

REFERENCES

- American Psychological Association. (2002). Criteria for evaluating treatment guidelines. *American Psychologist*, *57*, 1052-1059.
- Anderson, N. (2001). *Design and analysis: A new approach*. Mahwah, NJ: Erlbaum.
- Beutler, L. (1998). Identifying empirical supported treatments: What if we didn't? *Journal of Consulting and Clinical Psychology*, *66*, 113-120.
- Bird, F., Dores, P. A., Moniz, D., & Robinson, J. (1989). Reducing severe aggressive and self-injurious behaviors with functional communication training: Direct, collateral, and generalized results. *American Journal on Mental Retardation*, *94*, 37-48.
- Birnbrauer, J. S. (1981). External validity and experimental investigation of individual behavior. *Analysis and Intervention in Developmental Disabilities*, *1*, 117-132.
- Brown, K. A., Wacker, D. P., Derby, K. M., Peck, S. M., Richman, D. M., Sasso, G. M. et al. (2000). Evaluating the effects of functional communication training in the presence and absence of establishing operations. *Journal of Applied Behavior Analysis*, *33*, 53-71.
- Busk, P., & Serlin, R. (1992). Meta-analysis for single-participant research. In T. R. Kratochwill & J. R. Levin (Eds.), *Single-case research design and analysis: New directions for psychology and education* (pp. 187-212). Mahwah, NJ: Erlbaum.
- Carr, E. G., & Durand, V. M. (1985). Reducing behavior problems through functional communication training. *Journal of Applied Behavior Analysis*, *18*, 111-126.
- Carr, E. G., Levin, L., McConnachie, G., Carlson, J. I., Kemp, D. C., Smith, C. E. et al. (1999). Comprehensive multisituational intervention for problem behavior in the community. *Journal of Positive Behavior Interventions*, *1*, 5-25.
- Chambless, D., & Hollon, S. (1998) Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology*, *66*, 7-18.
- Chambless, D., & Ollendick, T. (2001). Empirically supported psychological interventions: Controversies and evidence. *Annual Review of Psychology*, *52*, 685-716.
- Christenson, S., Carlson, C., & Valdez, C. (2002). Evidence-based interventions in school psychology: Opportunities, challenges, and cautions. *School Psychology Quarterly* *17*, 466-474.
- Cooper, L. J., Wacker, D. P., Sasso, G. M., Reimers, T. M., & Donn, L. K. (1990). Using parents as therapists to evaluate appropriate behavior of their children: Application to a tertiary diagnostic clinic. *Journal of Applied Behavior Analysis*, *23*, 285-296.
- Didden, R., Duker, P. C., & Korzilius, H. (1997). Meta-analytic study on treatment effectiveness for problem behaviors with individuals who have mental retardation. *American Journal on Mental Retardation*, *101*, 387-399.
- Dunlap, G., & Kern, L. (1997). The relevance of behavior analysis to special education. In J. L. Paul, M. Churton, H. Roselli-Kostoryz, W. Morse, K. Marfo, C. Lavelly, & D. Thomas (Eds.), *Foundations of special education: Basic knowledge informing research and practice in special education* (pp. 279-290). Pacific Grove, CA: Brooks/Cole.
- Durand, V. M., & Carr, E. G. (1987). Social influences on "self-stimulatory" behavior: Analysis and treatment application. *Journal of Applied Behavior Analysis*, *20*, 119-132.
- Durand, V. M., & Carr, E. G. (1991). Functional communication training to reduce challenging behavior: Maintenance and application in new settings. *Journal of Applied Behavior Analysis*, *24*, 251-264.
- Durand, V. M., & Rost, N. (in press). Selection and attrition in challenging behavior research. *Journal of Applied Behavior Analysis*.
- Faith, M. S., Allison, D. B., & Gorman, B. S. (1996). Meta-analysis of single-case research. In R. D. Franklin, D. B. Allison, & B. S. Gorman (Eds.), *Design and analysis of single-case research* (pp. 256-277). Mahwah, NJ: Erlbaum.
- Flay, B. R. (1986). Efficacy and effectiveness trials (and other phases of research) in the development of health promotion programs. *Preventive Medicine*, *15*, 451-474.
- Gettinger, M. (1993). Effects of invented spelling and direct instruction on spelling performance of second-grade boys. *Journal of Applied Behavior Analysis*, *26*, 281-291.
- Greenwood, C., Tapia, Y., Abbott, M., & Walton, C. (2003). A building-based case study of evidence-based literacy practices: Implementation, reading behavior, and growth in reading fluency, K-4. *The Journal of Special Education* *37*, 95-110.
- Gresham, F. M., Gansel, K. A., & Kurtz, P. F. (1993). Treatment integrity in applied behavior analysis with

- children. *Journal of Applied Behavior Analysis*, 26, 257-263.
- Hagopian, L. P., Fisher, W. W., Sullivan, M. T., Acquistio, J., & LeBlanc, L. A. (1998). Effectiveness of functional communication training with and without extinction and punishment: A summary of 21 inpatient cases. *Journal of Applied Behavior Analysis*, 31, 211-235.
- Hall, V. R., Axelrod, S., Tyler, L., Grief, E., Jones, F. C., & Robertson, R. (1972). Modification of behavior problems in the home with a parent as observer and experimenter. *Journal of Applied Behavior Analysis*, 5, 53-64.
- Hersen, M., & Barlow, D. H. (1976). *Single-case experimental designs: Strategies for studying behavior change*. New York: Pergamon.
- Hershberger, S. L., Wallace, D. D., Green, S. B., & Marquis, J. G. (1999). Meta-analysis of single-case designs. In R. H. Hoyle (Ed.), *Statistical strategies for small sample research* (pp. 109-132). Newbury Park, CA: Sage.
- Jayaratne, S., & Levy, R. L. (1979). *Empirical clinical practice*. New York: Columbia University.
- Kazdin, A. E. (1982). *Single-case research designs: Methods for clinical and applied settings*. New York: Oxford University Press.
- Kazdin, A. E. (1998). *Research design in clinical psychology* (3rd ed.). Boston: Allyn & Bacon.
- Kennedy, C. H. (in press). *Single case designs for educational research*. Boston: Allyn & Bacon.
- Koegel, L. K., & Koegel, R. L. (1986). The effects of interspersed maintenance tasks on academic performance in a severe childhood stroke victim. *Journal of Applied Behavior Analysis*, 19, 425-430.
- Koegel, R. L., & Koegel, L. K. (1990). Extended reductions in stereotypic behavior of students with autism through a self-management treatment package. *Journal of Applied Behavior Analysis*, 23, 119-127.
- Kratochwill, T., & Levin, J. R. (1992). *Single-case research design and analysis: New directions for psychology and education*. Hillsdale, NJ: Erlbaum.
- Kratochwill, T., & Stoiber, K., (2002). Evidence-based interventions in school psychology: Conceptual foundations for the procedural and coding manual of Division 16 and Society for the Study of School Psychology Task Force. *School Psychology Quarterly*, 17, 341-389.
- Mace, F. C., Hock, M. L., Lalli, J. S., West, B. J., Belfiore, P., Pinter, E. et al. (1988). Behavioral momentum in the treatment of non-compliance. *Journal of Applied Behavior Analysis*, 21, 123-141.
- Marquis, J. G., Horner, R. H., Carr, E. G., Turnbull, A. P., Thompson, M., Behrens, G. A. et al. (2000). A meta-analysis of positive behavior support. In R. M. Gerston & E. P. Schiller (Eds.), *Contemporary special education research: Syntheses of the knowledge base on critical instructional issues* (pp. 137-178). Mahwah, NJ: Erlbaum.
- Martella, R., Nelson, J. R., & Marchand-Martella, N. (1999). *Research methods: Learning to become a critical research consumer*. Boston: Allyn & Bacon.
- McReynolds, L. V., & Kearns, K. P. (1983). *Single-subject experimental designs in communicative disorders*. Baltimore: University Park Press.
- Mildon, R. L., Moore, D. W., & Dixon, R. S. (2004). Combining noncontingent escape and functional communication training as a treatment for negatively reinforced disruptive behavior. *Journal of Positive Behavior Interventions*, 6, 92-102.
- Miller, K., Gunter, P. L., Venn, M., Hummel, J., & Wiley, L. (2003). Effects of curricular and materials modifications on academic performance and task engagement of three students with emotional or behavioral disorders. *Behavior Disorders*, 28, 130-149.
- Moore, J. W., Edwards, R. P., Sterling-Turner, H. E., Riley, J., DuBard, M., & McGeorge, A. (2002). Teacher acquisition of functional analysis methodology. *Journal of Applied Behavior Analysis*, 35, 73-77.
- Nelson, R., Roberts, M., Mathur, S., & Rutherford, R. (1999). Has public policy exceeded our knowledge base? A review of the functional behavioral assessment literature. *Behavior Disorders*, 24, 169-179.
- Odom, S., & Strain, P. S. (2002). Evidence-based practice in early intervention/early childhood special education: Single-subject design research. *Journal of Early Intervention*, 25, 151-160.
- Parsonson, B., & Baer, D. (1978). The analysis and presentation of graphic data. In T. Kratochwill (Ed.), *Single-subject research: Strategies for evaluating change* (pp. 105-165). New York: Academic Press.
- Parsonson, B., & Baer, D. (1992). Visual analysis of data, and current research into the stimuli controlling it. In T. Kratochwill & J. Levin (Eds.), *Single-case research design and analysis: New directions for psychology and education* (pp. 15-40). Hillsdale, NJ: Erlbaum.
- Richard, S. B., Taylor, R., Ramasamy, R., & Richards, R. Y. (1999). *Single-subject research: Applications in educational and clinical settings*. Belmont, CA: Wadsworth.
- Rohena, E., Jitendra, A., & Browder, D. M., (2002). Comparison of the effects of Spanish and English constant time delay instruction on sight word reading by

Hispanic learners with mental retardation. *Journal of Special Education*, 36, 169-184.

Shavelson, R., & Towne, L. (2002). *Scientific research in education*. Washington, DC: National Academy Press.

Shernoff, E., Kratochwill, T., & Stoiber, K. (2002). Evidence-based interventions in school psychology: An illustration of task force coding criteria using single-participant research designs. *School Psychology Quarterly*, 17, 390-422.

Sidman, M. (1960). *Tactics of scientific research: Evaluating experimental data in psychology*. New York: Basic Books.

Tawney, J. W., & Gast, D. L. (1984). *Single-subject research in special education*. Columbus, OH: Merrill.

Todman, J., & Dugard, P. (2001). *Single-case and small-n experimental designs: A practical guide to randomization tests*. Mahwah, NJ: Erlbaum.

Wacker, D. P., Steege, M. W., Northup, J., Sasso, G., Berg, W., Reimers, T. et al. (1990). A component analysis of functional communication training across three topographies of severe behavior problems. *Journal of Applied Behavior Analysis*, 23, 417-429.

Weisz, J. R., & Hawley, K. M. (2002). *Procedural and coding manual for identification of beneficial treatments*. Washington, DC: American Psychological Association, Society for Clinical Psychology Division 12 Committee on Science and Practice.

Whitehurst, G. J. (2003). *Evidence-based education* [PowerPoint presentation]. Retrieved April 8, 2004, from <http://www.ed.gov/offices/OERI/presentations/evidencebase.ppt>

Wolery, M., & Dunlap, G. (2001). Reporting on studies using single-subject experimental methods. *Journal of Early Intervention*, 24, 85-89.

Wolery, M., & Ezell, H. K. (1993). Subject descriptions and single-subject research. *Journal of Learning Disabilities*, 26, 642-647.

Wolf, M. M. (1978). Social validity: The case for subjective measurement or how applied behavior analysis is finding its heart. *Journal of Applied Behavior Analysis*, 11, 203-214.

ABOUT THE AUTHORS

ROBERT H. HORNER (CEC OR Federation), Professor, Educational and Community Supports, University of Oregon, Eugene. **EDWARD G. CARR** (CEC #71), Professor, Department of Psychology, State University of New York at Stony

Brook. **JAMES HALLE** (CEC #51), Professor, Department of Special Education, University of Illinois, Champaign. **GAIL MCGEE** (CEC #685), Professor, Emory Autism Resource Center, Emory University, School of Medicine, Atlanta, Georgia. **SAMUEL ODOM** (CEC #407), Professor, School of Education, Indiana University, Bloomington. **MARK WOLERY** (CEC #98), Professor, Department of Special Education, Vanderbilt University, Nashville, Tennessee.

Address all correspondence to Robert H. Horner, Educational and Community Supports, 1235 University of Oregon, Eugene, OR 97403-1235; (541) 346-2462 (e-mail: robh@uoregon.edu)

Manuscript received December 2003; manuscript accepted April 2004.

INDEX OF ADVERTISERS

Council for Exceptional Children, cover 2, p 129, cover 4

DePaul University, p 134

East Carolina University, p 180

New York University, p 180

Rider University, p 134

Wadsworth, cover 3