

MULTIPLE BASELINE DESIGNS: THE USE OF A SINGLE-CASE EXPERIMENTAL DESIGN IN LITERACY RESEARCH

RICHAEL BARGER-ANDERSON
Slippery Rock University of Pennsylvania

JOSEPH W. DOMARACKI
Indiana University of Pennsylvania

NEDRA KEARNEY-VAKULICK
California University of Pennsylvania

RICHARD M. KUBINA JR.
The Pennsylvania State University - University Park

Single-case experimental designs are becoming more popular and acceptable ways to conduct classroom-based research. Single-case designs can also be beneficial when conducting literacy research. There are a variety of single-case research designs that can accommodate a range of instructional strategies and research questions. This paper provides examples and reasons for using the multiple baseline with literacy research. The effectiveness of multiple baseline designs are also examined in relation to the following experimental design issues: control, replication, causal relationships, internal and external validity and ethical concerns.

Single-case experimental research designs are becoming more popular and acceptable ways to conduct classroom-based research (Birnbrauer, Peterson, Solnick, 1974; Gay & Airasian, 2000; Neuman & McCormick, 1995; Richards, Taylor, Ramasamy, & Richards, 1999). Since the mid sixties, single-case research designs have continued to increase in popularity (Gay, 1987). There is a need to increase awareness about the qualities and attributes of single-case research designs because many misconceptions still exist (Dermer & Hoch, 1999).

Single-case, or single-subject, experi-

mental research design can personalize the data collection process because data is collected on each subject, or participant, and is individually analyzed. The term single-subject is not used because there is only one participant; rather, it refers to the procedure for data collection and the focus of the study as opposed to the number of participants. (Neuman & McCormick, 1995). Additionally, experimental control is established with each participant (Cooper, Heron, & Heward, 1987; Johnston & Penningpacker, 1993).

Wolery and Gast (2000) suggest teachers do not always have access to a large

number of participants to use group designs therefore single-case designs offer an excellent opportunity for teachers to conduct research in their classroom. Gay and Airasian (2000) agree that single-subject designs may be beneficial for use in the classroom due to the limited size of accessible population. For example, in special education the numbers of students in self-contained or resource rooms are generally small making single-case designs useful. "Approximately one third of all data-based interventions conducted on students with learning disabilities use single-subject designs" (Swanson, Hoskyn, Sachse-Lee, & O'Shaughnessy, as cited in Swanson & Sachse-Lee, 2000, p. 144).

A meta-analysis of eighty-five single-subject research design studies involving specific instructional skill domains had reading as a focus (Swanson & Sachse-Lee, 2000). In fact, "most of the intervention studies conducted were in the domain of reading" (p. 149). When examining the meta-analysis Swanson and Sachse-Lee report that a multiple baseline design was used in 62% of the studies. Other findings were that 34% of the studies occurred in a resource classroom and 99% of the materials used were existing materials in the classroom. The meta-analysis indicates that single-case research designs help teachers and researchers examine variables that effect student learning.

There are several types of designs to choose from when utilizing single-case research. The most familiar types include the withdrawal design, reversal design, multiple-baseline designs, and the alternating treatments design (Gay, 1987; Gay

& Airasian, 2000; Neuman & McCormick, 1995; Richards et al., 1999). Of these options the multiple baseline design offers a considerable degree of flexibility in classroom research (Swanson & Sachse-Lee, 2000).

Multiple Baseline Designs

Multiple baseline designs provide a means for collecting multiple sets of data in a single-case experimental design (Neuman & McCormick, 1995). Kucera and Axelrod (1995) state multiple baseline designs are "particularly well-suited to literacy research" (p. 47). The authors also maintain that multiple baseline designs can help examine new techniques and strategies that are used and found to be effective in the area of teaching reading. The multiple baseline design is the design of choice when it is not possible for subjects to return to original baseline (Hersen & Barlow, 1984; Gay 1987; Gay & Airasian, 2000; McReynolds & Kearns, 1983). For example, once a student has learned a new strategy for decoding words is not desirable, and in many cases possible, to have the student unlearn the new skill.

There are three types of multiple baseline designs. These include the multiple baseline across behaviors, subjects or participants, and settings (Cooper, Heron, & Heward, 1987; Gay, 1987; Gay & Airasian, 2000; Hersen & Barlow, 1984; Kazdin & Kopel, 1975; McReynold & Kearns, 1983; Neuman & McCormick, 1995). Alberto and Troutman (2003) provide the following examples for potential behaviors of interest to classroom teacher; a multiple

baseline across behaviors can study talking out and out of seat behavior. A multiple baseline across subjects could target spelling accuracy for multiple students. And a multiple baseline across settings might examine a student's inappropriate behavior of swearing in recess and the cafeteria. In literacy research, any almost any question involving a dependent variable and independent variable can be asked provided the answer can be obtained by a few participants. For the purposes of illustrating a technical application, the multiple baseline across participants design will be used as example.

The multiple baseline across participants addresses the impact of the treatment of the independent variable on the dependent variable, the same behavior, for different participants. Once a baseline has been established, the treatment or independent variable is applied to one of the participants. During this time, baseline is maintained for the other participants. Once improvement is seen for the first participant, the treatment is started with the second subject, and so on. The reasoning behind this design is that if one participant shows improvement when treatment is started it is probable that improvement is due to the treatment. If gains were reflected in the other participants' behaviors even though they were in baseline, a conclusion could not be made that the independent variable was the most probable reason for the observed changes.

Baseline data is retrieved for a minimum of one behavior across multiple subjects (Kazdin, 1973). It is generally cautioned, however, that prolonged baselines for students who need intervention

on a more immediate basis may not be educationally sound (Neuman & McCormick, 1995; Schloss & Smith, 1998). For example, a student who engages in physical aggression should not be allowed to continue for the sake of a scientific study. While physical aggression can be studied experimentally, using a multiple baseline design would not be the best choice for the student and those around him.

Figure 1 shows an example of a multiple baseline design used with three elementary students. The data points represent words read correctly per minute on passages selected by the teacher. During the baseline phase, the students do not come in contact with the independent variable. To do so would be to compromise experimental control. When the experimenter implements the independent variable of "repeated readings," the first student, Tom, shows an increase in his reading rate. Inspecting the other two students still in baseline, an increase in reading rate is not observed. When Sue enters the independent variable phase her reading rate also increases with the repeated readings intervention. The third student, Eric, is not showing an increase during his baseline performance. Only when Eric receives the independent variable does his reading rate accelerate. The multiple baseline shows a possible functional relationship with the introduction of the independent variable repeated readings and the words read correctly per minute for the three 2nd grade students.

In the ever-continuing quest for knowledge surrounding effective literacy interventions, teachers and researchers observe behavior, gather data, analyze and

interpret the information and come to a conclusion regarding their research (Leedy & Ellis-Ormrod, 2001). As an additional research option, the multiple baseline is a single-case design well suited for experiments requiring small numbers of participants. The following section outlines the multiple baseline design and shows how it meets the standards for the following experimental benchmarks; internal and external validity, control, replication, causal relationships and weaknesses and ethical concerns.

Experimental Standards

Internal and External Validity. As with group designs, internal and external validity are important issues to be considered when using single-case designs. When changes occur that are attributable to the effects of the independent variable, the study has internal validity (Poling & Grossett, 1986). Internal validity of a single-case design is considered acceptable if an intervention is reliably associated with higher response levels while also revealing sound experimental control (Neuman & McCormick, 1995). In the case of a multiple baseline design, the changes occurring in each phase arise from the systematic application of the independent variable, not some extraneous variable.

External validity refers to the degree to which the results from a study can be generalized to other groups or settings (Gay & Airasian, 2000). Regardless of the type of experimental design used, single-case or group design, a number of factors encompass external validity (Poling & Grossett, 1986). Neuman and McCormick (1995) suggest that the best ways to attend to the

issues of external validity include: 1) providing a rich and detailed description of the setting and the intervention, 2) detailing the measures, and 3) generalizing the results to a particular theory.

Control. Gay and Airasian (2000) explain how control is a primary characteristic of experimental studies: "Direct manipulation by the researcher of at least one independent variable is the one single characteristic that differentiates experimental research from other types of research. Control refers to the researcher's efforts to remove the influence of any extraneous variable (other than the independent variable itself) that might affect scores on the dependent variable" (p. 370). Dermer and Hoch (1999) propose that the single-case researcher views control as a way to discard variability caused by anything other than the treatment or independent variable. Control should continue until the researcher can visually determine an effect of the treatment. Single-case research uses control procedures rather than control groups (Good, 2000). Each student, then, serves as her or his own control (Gay & Airasian, 2000; Wolery & Gast, 2000).

Control can be difficult to achieve in a research design, particularly when dealing with people (Gay & Airasian, 2000). Control is even more challenging to achieve in the multiple baseline design when different interventions are offered sequentially. Counterbalancing is a possible way to control for order effect when using two or more interventions in multiple baseline design research. "The counterbalancing of treatments permits a comparison of the effectiveness of the two instructional methods" (Domaracki, 1987, p. 57).

Counterbalancing is completed in a study to address the concern of order effects. The first subject receives a treatment, a second subject receives a different treatment, then once the subjects have shown improvements, the treatments are reversed. Counterbalancing is an attempt to control for order effects (McReynolds & Kearns, 1983).

Replication. Replication is important to all areas of science in two ways. First, to establish the reliability of previous findings and second, to determine the generality of the findings under differing conditions (Hersen & Barlow, 1976). Hersen and Barlow also point out that many researchers welcome replicating their single-subject design study or one of another researcher. Specifically because each replication not only helps to generalize results but also helps to validate the findings. In terms of generalizability, the more replications that are completed while retrieving similar results, the more convincing it is in generalizing the results to other subjects (Birnbrauer et al., 1974; Gay, 1987).

Replication of multiple baseline and other single-case designs depends on accurate, detailed, and rich descriptions of the study (Neuman & McCormick, 1995). "Readers need to know about the selection, assignment, and number of subjects who will participate in the experiment" (Creswell, 1994, p. 126). Without providing this information, replication in future studies will not be possible. Single-case research designs must include sufficient detail so that a study can be replicated (Birnbrauer et al., 1974).

The number of replications in the study and the treatment effectiveness demon-

strates control (Hersen & Barlow, 1976; Parsonson & Baer, 1978; McReynolds & Kearns, 1983). It has been recommended that a study should be replicated at least two to four times (Kazdin & Kopel, 1975; Wolf & Risley, 1971). Two replications are only suggestive but a set of three to four replications is more convincing in terms of experimental control and treatment effectiveness (Wolf & Risley, 1971). In terms of baseline collection, Hersen and Barlow (1976) recommend three to four baselines in a study.

Causal Relationships. Experimental research is the only form of research that can establish cause-effect relationships (Creswell, 1994; Gay & Airasian, 2000). Researchers must manipulate at least one independent variable while also controlling other relevant variables. The resulting effect on the dependent variable allows the experimenter to draw a conclusion of a relationship between the variables. Manipulation of the independent variable is the key difference between experimental research and other types of research (Gay & Airasian, 2000). As such, single-case designs like the multiple baseline present the opportunity to conduct experimental research. In the words of Good (2000): "Single-subject experimental design is a rigorous technique for evaluation that can typically stand alone in assessing program effectiveness" (p. 35).

Cause-and-effect relationships in single-case experimental research can be increased through treatments and replications (Birnbrauer et al., 1974). The purpose of the single-case design, even while using the multiple baseline approach, is to determine levels of causation for each individual

participant involved in the study (Dermer & Hoch, 1999). Functional relationships are probable if the student's performance is altered only in response to the systematic application of the independent variable (Neuman & McCormick, 1995). Changes that occur in a study are more likely to be indicative in proving the effect of the treatment when baselines are independent from the treatment (Kazdin & Kopel, 1975).

Weaknesses/Ethical Concerns. When scrutinizing multiple baseline designs, ethical concerns are very seldom an issue generating much discussion (McReynolds & Kearns, 1983; Neuman & McCormick, 1995). One reason is because multiple baseline designs do not withdrawal treatment. In other single-subject experimental research designs (e.g., withdrawal design), treatment is implemented followed by a return to baseline through the removal of the independent variable. For instance, once a student learns how to apply a specific comprehension strategy, the student does not have to unlearn it for the sake of demonstrating a functional relationship.

One area of concern however, is the prolonged nature of baselines. Extended baselines are not always educationally sound for students who need intervention (Neuman & McCormick, 1995). In special education, a student may be behind his same aged peers by a year or more. Some single-case experimental designs allow teachers and researchers the opportunity to immediately implement a treatment (e.g., alternating treatments design) rather than have the student wait in baseline for a protracted period of time.

A further potential weakness of the multiple baseline design is establishing

cause-and-effect relationships. Researchers must be aware of causal variables to make statements regarding generality (Dermer & Hoch, 1999). Multiple baseline designs contain potential weaknesses when it comes to proving the effect of the study (Kazdin, 1973). The subsequent phases used to verify that a treatment effect is occurring are done by inference. Inference is demonstrated by the following example; a student whose behavior changes after the independent variable is introduced is an A-B condition (i.e., A is the baseline and B the intervention). The second student also has an A-B condition albeit staggered from the first student. Any subsequent students in later implementation still undergo the A-B condition and the inference is made from all of the data that the effect is due to the intervention.

Another difficult task in the use of the multiple baseline design is choosing dependent variables that are independent of each other. If the dependent variables are not independent of one another, it is not possible to determine the effects of the independent variable. The possibility for the target behaviors to co-vary makes the results of the cause questionable (Hersen & Barlow, 1976; Kazdin, 1973; Neuman & McCormick, 1995). Using a multiple baseline across behaviors with narrative and persuasive writing may result in two behaviors that do not co-vary enough. The effects of the independent variable may not be clearly demonstrated with two similar behaviors.

An additional concern lies in the fact that the participant cannot return to baseline. The multiple baseline design is appropriate to utilize when is not possible

or tolerable to return the behavior to baseline (Neuman & McCormick, 1995). However, there are mixed reviews in the literature as to the need for this. For instance, Baer, Wolf, and Risely (1968) have long maintained that a return to baseline is not needed to prove cause-and-effect relationships. On the other hand, there is no way possible to be sure the treatment effects would be the same if the treatment phase came before the baseline phase (Gay & Airasian, 2000; Birnbrauer et al., 1974).

Many replications of any single-case research design are needed to prove cause-and-effect relationships. Further, no experiment, regardless of the method, should create a general assumption of a cause-and-effect relationship after only one completion of the study (Neuman & McCormick, 1995). Wolery and Gast (2000) and Birnbrauer et al. (1974) suggest the best way of proving cause-and-effect relationships is through direct and systematic replication of studies.

Conclusion

Single-case designs have been used in psychology and psychiatry. However, they can be very beneficial in educational settings (Gay, 1987; Scholss & Smith, 1998). An instructor of a classroom can easily implement single-case research designs. Single-case experimental research designs offer an avenue to more closely examine components of literacy research (Good, 2000). This approach to research can be of use especially when students involved in the study need remedial assistance (Neuman & McCormick, 1995).

There are many types of single-case research designs that may be employed

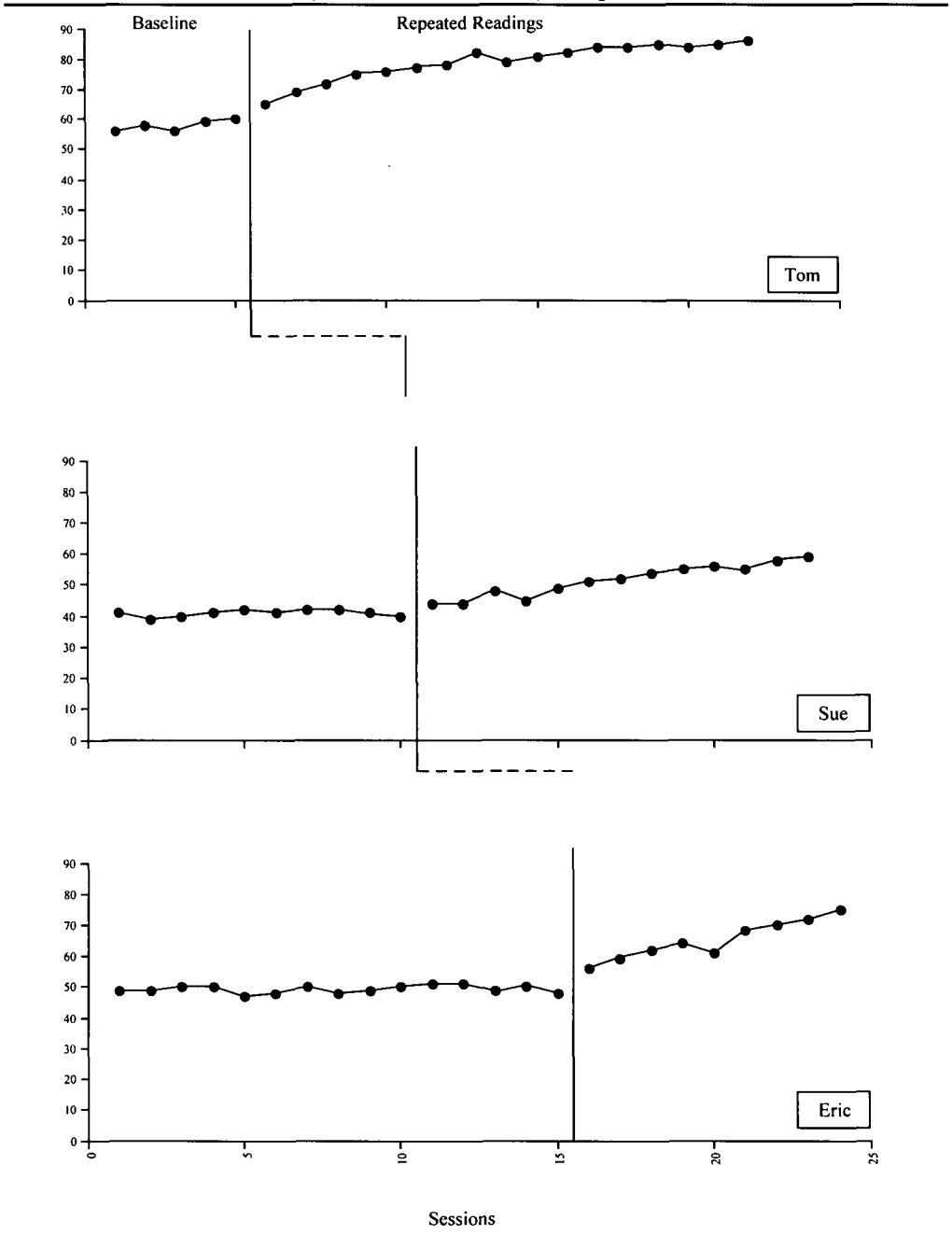
(Birnbrauer et al., 1974; Gay, 1987; Gay & Airasian, 2000; Neuman & McCormick, 1995) which can accommodate a variety of instructional strategies (Swanson & Sachse-Lee, 2000). The multiple baseline design can accommodate research variables ranging from decoding and comprehension to spelling and expressive writing. The experimental standards relied upon to produce sound studies (i.e., internal and external validity, control, replication, causal relationships and weaknesses and ethical concerns) and make subsequent generations are present in the multiple baseline design. And with our society continuing to demand competent readers and high degrees of literacy, teachers and researchers will need a number of research tools for their disposal. The multiple baseline is one such single-case design that can aid experimental studies of literacy.

References

- Alberto P. A., & Troutman A. C. (2003). *Applied behavior analysis for teachers* (6th ed.). Upper Saddle River, NJ: Merrill.
- Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis, 1*, 91-97.
- Birnbrauer, J. S., Peterson, C. R., & Solnick, J. V. (1974). Design and interpretation of studies of single-subjects. *American Journal of Mental Deficiency, 79*(2), 191-203.
- Creswell, J. W. (1994). *Research design qualitative and quantitative approaches*. Thousand Oaks, CA: SAGE.
- Cooper, J.O., Heron, T.E., & Heward, W. L. (1987). *Applied behavior analysis*. Columbus: Merrill.

- Dermer, M. L., & Hoch, T. A. (1999). Improving descriptions of single-subject experiments in research texts written for undergraduates. *Psychological Record*, 49(1), 49-66.
- Domaracki, J. W. (1987). *General case simulation instruction and naturalistic instruction: A descriptive and comparative analysis*. Unpublished doctoral dissertation, University of Pittsburgh, PA.
- Gay, L. R. (1987). *Educational research competencies for analysis and application*. New York: Macmillan.
- Gay, L. R., & Airasian, P. (2000). *Educational research competencies for analysis and application* (6th ed.). Upper Saddle River, NJ: Prentice Hall.
- Good, J. M. (2000). Evaluating developmental education programs by measuring literacy growth. *Journal of Developmental Education*, 24(1), 30-36.
- Hersen, M., & Barlow, D. H. (1976). *Single-case Experimental Designs Strategies for Studying Behavior Change*. Elmsford, NY: Pergamon Press Ltd.
- Johnston, J. M., & Pennypacker, H. S. (1993). *Strategies and tactics of behavioral research*. Hillsdale, NJ: Lawrence Erlbaum.
- Kazdin, A. E. (1973). Methodological and assessment considerations in evaluating reinforcement programs in applied settings. *Journal of Applied Behavior Analysis*, 6(3), 517-531.
- Kazdin, A. E., & Kopel, S. A. (1975). On resolving ambiguities of the multiple-baseline design: problems and recommendations. *Behavior Therapy*, 6, 601-608.
- Kucera, J., & Alexrod, S. (1995). Multiple-baseline designs. In S. B. Neuman and S. McCormick (Eds.), *Single-subject experimental research: Application for literacy*. Newark, DE: International Reading Association.
- Leedy, P. D., & Ellis-Ormrod, J. (2001). *Practical research: Planning and design* (7th ed.). Upper Saddle River, NJ: Merrill.
- McReynolds, L. V., & Kearns, K. R. (1983). *Single-subject experimental designs in communicative disorders*. Baltimore: University Park Press.
- Neuman, S. B., & McCormick, S. (1995). *Single-subject experimental research: Application for literacy*. Newark, DE: International Reading Association.
- Parsonson, B. S., & Baer, D. M. (1978). The analysis and presentation of graphic data. In T. R. Kratochwill (Ed.), *Single-subject research* (pp. 101-165). New York: Academic Press.
- Poling, A., & Grossett, G. (1986). Basic research designs in applied behavior analysis. In A. Poling and R. W. Fuqua (Eds.), *Research methods in applied behavior analysis: Issues and advances* (pp. 7-27). New York: Plenum Press.
- Richards, S. B., Taylor, S. B., Ramasamy, R., & Richards, R. Y. (1999). *Single-subject research: Applications in educational and clinical settings*. San Diego, CA: Singular.
- Schloss, P. J., & Smith, M. A. (1998). *Applied behavior analysis in the classroom* (2nd ed.). Needham Heights, MA: Allyn & Bacon.
- Swanson, H. L., & Sachse-Lee, C. (2000). A meta-analysis of single-subject design intervention research for students with learning disabilities. *Journal of Learning Disabilities*, 33(2), 114-166.
- Wolery, M., & Gast, D. (2000). Classroom research for young children with disabilities: Assumptions that guided the conduct of research. *Topics in Early Childhood Special Education*, 20(1), 49-55.
- Wolf, M. M., & Risely, T. R. (1971). Reinforcement: Applied research. In R. Glaser (Ed.), *The nature of reinforcement* (pp. 310-325). New York: Academic Press.

Figure 1
 A multiple baseline across subjects design showing the effects of the independent variable repeated readings on the dependent variable amount of correct words per minute for various passages.



Copyright of Reading Improvement is the property of Project Innovation, Inc. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.