Design Options for Home Visiting Evaluation

SINGLE CASE DESIGN BRIEF

Selecting Appropriate Single Case Designs for Evaluating MIECHV Funded Home Visiting Programs

March 2012

Overview

Single case design (SCD), often referred to as single subject design, is an evaluation method that can be used to rigorously test the success of an intervention or treatment on a particular case (i.e., a person, school, community) and to also provide evidence about the general effectiveness of an intervention using a relatively small sample size. Generally, SCDs use visual analysis of data to systematically compare participants' target behaviors before they receive an intervention to the same behaviors during and after the intervention has been introduced.[a] If the data illustrate that target behaviors change only after the intervention is initiated, this effect suggests that the intervention was responsible for the behavior change. In order to establish causal evidence that the intervention was in fact responsible for behavior change, replication of the effect is then attempted.

Although SCDs are often done with a single participant as the case, these designs can also be conducted on a group of individuals involved in a single experiment. In order for a group of individuals to be considered a case, their responses or behaviors must produce one score per measurement period[1] (e.g., the rate of temper tantrums performed by all 2-year old children in a classroom).

One of the most notable features of SCD research is that each case serves as its own control, therefore eliminating the need for a control group condition typically required of experimental designs.[2] Specifically, the baseline period serves as the control, which can then be compared to the period after and/or during the intervention phase that follows.

The purpose of this document is to support Maternal Infant and Early Childhood Home Visiting (MIECHV) Program grantees considering the use of SCD methods in their evaluations by providing general information on SCD methodology and presenting considerations that should be weighed when selecting a SCD method. The material presented in this document is intended to provide introductory information about SCD in relation to home visiting programs and is not a comprehensive review of the application of SCD to other types of interventions.

Basics of SCD[3]

Dependent variables (DVs) - In SCD research, dependent variables are the intended outcomes, or more specifically the observable target behaviors that in theory should be changed by the intervention. DVs should produce stable and consistent visual patterns at baseline and are continually observed during the onset and withdrawal of the intervention to assess for change. In effective SCDs, the DVs are measured repeatedly within and across

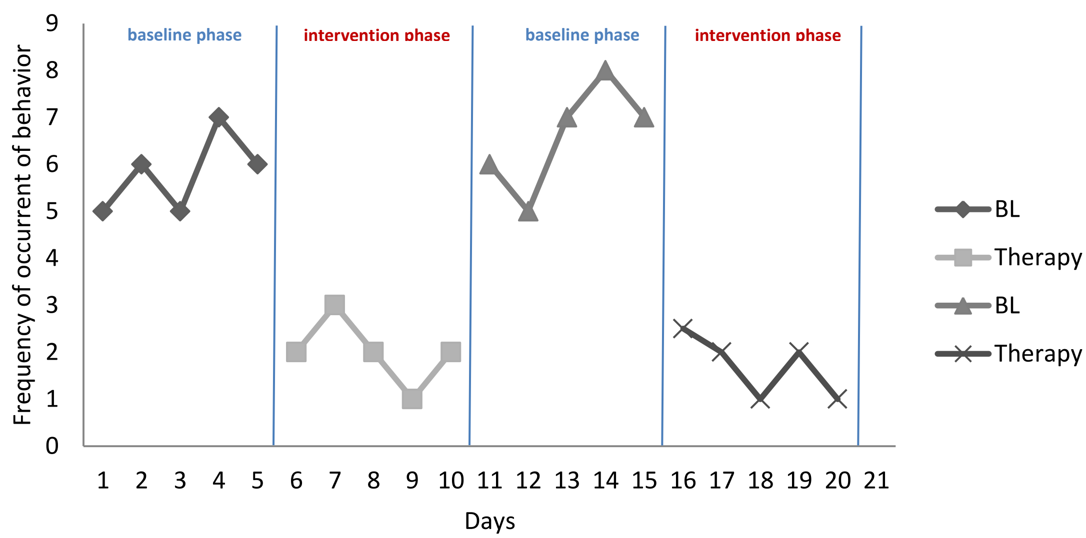
[a] In addition to examining behavioral change, it is also possible to examine intervention impacts on learning outcomes (e.g., academic performance).

conditions to compare response patterns before, during, and after the intervention is implemented.

Independent variables (IVs) - In SCD research, independent variables are the practice, treatment, or intervention being tested. The introduction of the IV should produce observable changes in the DV. Independent variables are actively manipulated and it must be decided in advance how and when the IVs will be introduced and/or removed.

Baseline period - In SCD research, the baseline period is similar to the 'treatment as usual condition' in group designs. During the baseline period, the DV is measured until a visibly consistent response pattern is present. As it is important to establish stability of a targeted behavior at baseline, the baseline period typically requires five or more data points in order to reasonably conclude that a change in performance is unlikely in the absence of introducing some sort of change in condition (e.g., introducing an intervention).

Example of a single case design (ABAB)



The above example illustrates a single case design methodology (ABAB) aimed at reducing anxiety demonstrated through nail-biting behavior in a 25-year old man by introducing relaxation therapy. In this example, the dependent variable (what we hope to change) is the nailing biting behavior the man exhibits and the independent variable (what we actively manipulate) is the relaxation therapy intervention. The baseline period, where we measure the dependent variable in the absence of the intervention, is illustrated by the first and third measurement phase.

Internal and External Validity – Internal and external validity are two important elements in all research designs. Internal validity refers to the degree to which one can be sure an intervention is responsible for changing a participant's behavior or responses. Threats to the internal validity of a single case design are essentially the same as with any other causal design. These threats can weaken any study design and render explanations other than a treatment effect plausible. By designing robust controls (e.g., documenting a researcher was able to manipulate the independent variable and could replicate the effect in a manner that supports prior expectation, gathering a sufficient number of data points to support analyses, etc.), the internal validity of a study can be strengthened. External validity, which is often weak in SCD research due to small sample sizes, refers to the generalizability of the treatment outcomes to new participants, settings, and outcome measures. A study is considered to have good external validity when the results can be generalized beyond the sample to other settings, circumstances, and populations. Since SCD research is usually conducted with a small number of participants and in highly contextualized settings, it is reasonable to assume these designs often have limited external validity. To address these concerns, it is often advisable to repeat the SCD study. According to the standards adopted by the Home Visiting Evidence of Effectiveness (HomVEE) review, three replications of an intervention effect within a single study are required to demonstrate sound experimental control.[4]

Causality – SCD research allows the researcher to draw a causal argument for the single case or group of cases. This can be distinguished from experimental and control group designs that make causal arguments at only a group level. Causality allows the researcher to infer that the treatment is responsible for the observed effect. For a strong causal argument, the researcher must be able to demonstrate that: (1) the independent and dependent variables are associated with one another, (2) one variable precedes the other, and (3) all other reasonable alternative explanations have been ruled out.[5]

Types of Single Case Designs

Rather than one single design model, SCD is a class of multiple designs. Although the basic SCD has many variations, there are two classes that warrant further description regarding their applicability to home visiting studies: (1) withdrawal/reversal designs and (2) multiple baseline designs. When considering which SCD design to use, attention must be paid to the nature of the research question(s) and objective(s). Specifically, less stable outcomes (e.g., smoking behaviors or alcohol use) may be more suited to a withdrawal/reversal design whereas learned outcomes (e.g., parental knowledge) tend to be more stable and may be better suited to a multiple baseline design. More information about each of these designs and their uses can be found below.

Withdrawal and Reversal Designs[6]

In withdrawal/reversal designs, changes in the DV are observable depending on the introduction and removal of the IV. In other words, in order for this design to provide evidence that the intervention was effective, it relies on a reversal of behavior during the treatment withdrawal period, in a manner that fits prior expectation of the researchers. Two variations on this design are presented here: the ABA design and the ABAB design.

ABA Design

In an ABA design, there are three measurement periods: baseline (A), during intervention (B), and following the removal of intervention (A). Specifically, the targeted condition is repeatedly measured during a baseline period before the intervention is introduced (A). Once a consistent pattern of baseline responses has been established, the intervention is introduced and the condition is again repeatedly measured (B). Next, treatment stops but the measurement of the condition continues for a period of time (A).

Example: A mother seeks a behavioral intervention for her child who is exhibiting disruptive behaviors during school. The child's disruptive behaviors are first measured during a baseline period (A). After a stable pattern of behavior is observed, the child receives the behavioral intervention while being re-assessed during the intervention period (B). After the child's behaviors are measured repeatedly during the intervention period, the intervention is stopped (removed) and the child's behaviors are measured once again (A). When the child's disruptive behaviors are measured during the treatment period, they are significantly reduced from where they measured previously at baseline. In order to conclude that the intervention was effective in reducing the behaviors, we would need to observe a decrease in disruptive behaviors during the treatment period, followed by an increase in disruptive behaviors to near baseline when the intervention was removed.

Considerations: This design is problematic for prevention/early intervention research, as the goal of these types of services is to have a more lasting effect. Additionally, with this design, it is difficult to make generalizations about the intervention effects because of difficulty controlling for outside variables. In view of the example above, it would be difficult to know if the intervention caused the reduction in disruptive behaviors or if outside factors were also responsible (e.g., the presence of an observer, the child's physical health that day, etc.). Additionally, because this design does not meet causal validity standards being followed by the HomVEE standards, it is only briefly mentioned here.[b]

<u>ABAB Design</u>

An ABAB design is similar to an ABA design, except after the post-treatment measurement period, the intervention is re-introduced (B) with the assumption that the treatment effect will return. In order for this design to provide evidence that the intervention was effective, it relies on a reversal of behavior during the treatment withdrawal period, in a manner that fits prior expectation of the researchers.

Example: An early intervention program wants to increase participation in program services (DV) by introducing incentives for participation (IV). Participation would be monitored during a baseline period before incentives are introduced (A) and again during a period of time after incentives are introduced (B). The researcher would then return to baseline conditions by removing incentives for participation and again monitoring participation for a period of time (A). Finally, incentives would be reinstated and participation would be continually tracked and measured (B). If incentives are tied to program participation, the withdrawal of the intervention would lead to a pattern of behavior similar to the pattern observed at baseline. Likewise, by reintroducing the intervention, a change in the pattern of behavior should be evident during the intervention phase.

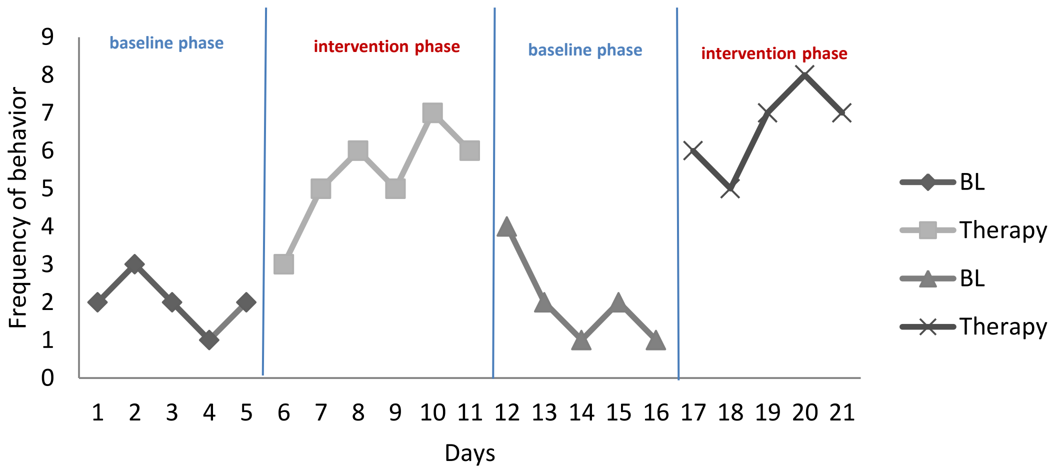
Visual analysis is an effective way to document the relationship between the DV and IV in an SCD, as illustrated by the ABAB design example in Figure 1 below. When visually examining the data, consistency within each of the phases should be observed. In Figure 1, it is clear that there is a stable pattern in behavior during the baseline phase, as evident by the low variability in behavior during this period. Once the intervention is introduced, the frequency

[b] ABA designs only allow two demonstrations of an effect and therefore do not meet causal validity standards being followed by the HomVEE review. On the other hand, the second intervention phase of the ABAB designs helps provide additional causal evidence that it is in fact the intervention that is responsible for the change in behavior. Note: HomVEE adopted the single case design standards set forth by the What Works Clearinghouse (Retrieved from What Works Clearinghouse website: http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf).

of the behavior increases, suggesting that the intervention may be effective.[c] In order to prove strong causality, these two phases would be repeated for further visual analysis.

In addition, it is important to note that there are no overlapping data points illustrating that the frequency of behavior observed during the intervention phases was not also observed during the baseline phases, and vice versa.

Figure 1. *Example of an ABAB Design*[7]



Considerations for all withdrawal/reversal designs: The intrinsic challenge with using withdrawal/reversal designs in home visiting evaluation is that, in order to prove an intervention effect, these designs are dependent on the pattern of responses returning to baseline conditions when the intervention is removed. It is important to know that reversal may not always occur if the treatment effect continues even when the treatment is removed. If this were the case, there would be visible overlapping data points across the intervention and baseline phases, which would likely be attributed to residual intervention effects. Therefore, if it is suspected that performance change and consequently reversal may not occur due to spillover from the intervention (i.e., learned skills which cannot be unlearned), it would be inappropriate to use a withdrawal/reversal design. As noted earlier, this strong effect is one of the goals of most home visiting interventions, so this must be taken into consideration when choosing this design.

Additionally, the ethics of treatment withdrawal need to be weighed against the long-term benefit of confirming that a particular treatment or intervention is responsible for some level of participant improvement. For example, if an intervention plan is designed to stop a child from engaging in repetitive self-injurious behaviors like head banging, an ABAB design may be an unethical choice due to risk of traumatic brain injury. In contrast, if the intervention plan is to increase a particular student's focus during class, short-term treatment removal may be

[c] When applied to the study of processes, the first pattern of variation, which reflects a stable process, is also known as "random" variation whereas the second pattern is often called "special cause" variation. See Donald J. Wheeler: *Understanding Variation*. 2000."

acceptable even if strong evidence for the treatment during the intervention phase can be established. As with all research, the best interest and safety of the participants should always be carefully balanced with the potential for greater knowledge learned.

Finally, some skills cannot be unlearned, therefore making it challenging to accomplish a complete reversal (e.g., if the intervention is teaching an individual how to ride a bike, a reversal would be difficult to achieve as it would be hard for an individual to unlearn those skills). In sum, treatment removal designs should be considered in the context of the long-term benefits, the nature of the intervention, and the outcome measure.

Multiple Baseline Design[8]

A multiple baseline design examines three or more people, behaviors, or settings to pinpoint the effect of an intervention. In this design, multiple baselines are established, but treatment is implemented one participant at a time by staggering the start of the intervention while continuing to assess baseline performance of participants who are not yet treated. The behaviors during the intervention period are then compared to those during the baseline stages. Multiple baseline designs can be used to demonstrate generality of intervention effects with:

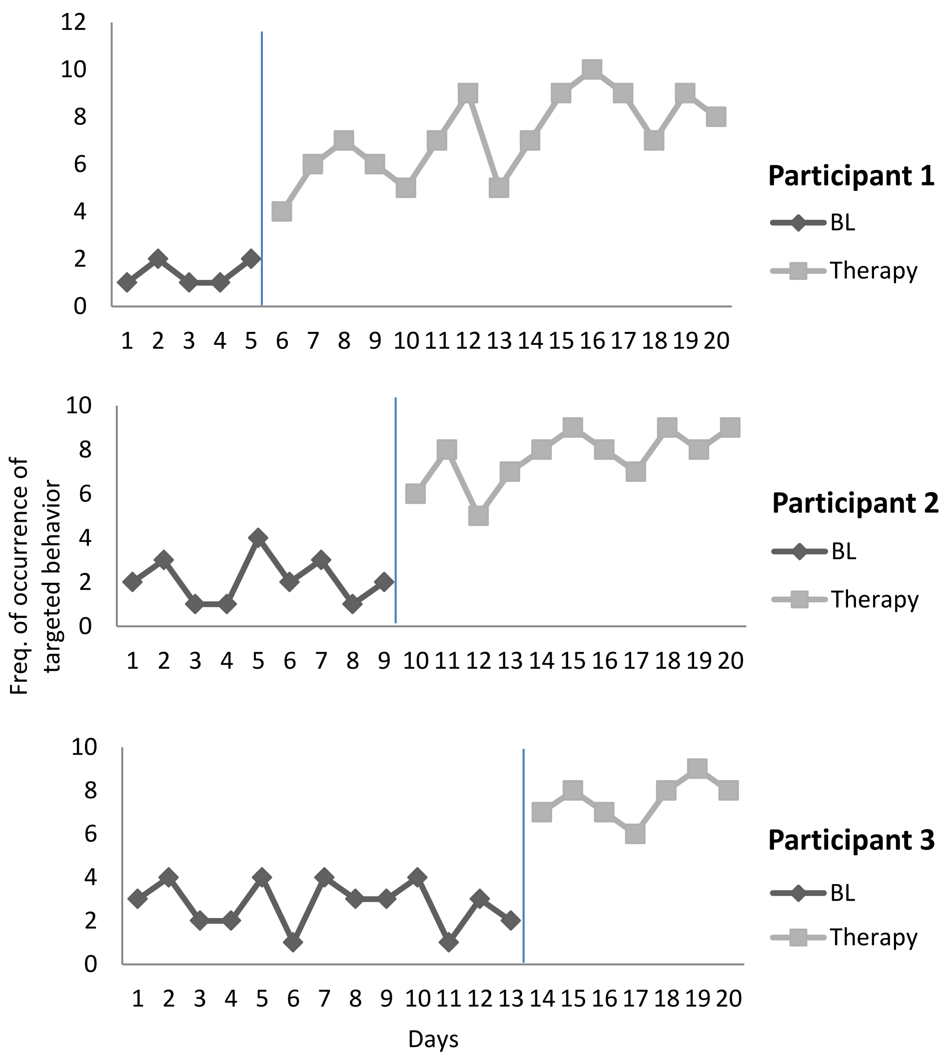
1) Additional participants (across individuals);
2) Additional problems for the same participant (across behaviors);
3) Same problem and same participant in different settings (across settings).

Example: Three women are referred to a program, each of whom reports suffering from overeating behaviors. Baseline data is gathered for each participant, but the length of the baseline period varies across the participants. Participant 1 begins treatment, and data collection continues for all three women while the second and third women continue to remain in the baseline period. The second woman does not begin treatment until there is a clear indication that the first woman who is receiving treatment experiences a reduction in her overeating behaviors. Similarly, the third woman does not begin treatment until the second woman experiences a reduction in her behaviors. Repeated measurement would continue for all three women throughout the process, but the second and third women would have a longer baseline period than the first.

If the treatment is the sole determinant of improvement, then behavior should change only when the treatment is implemented and not before. Therefore, participants who remain in the baseline phase should not show behavior change until they themselves have been treated. In the example above, if the treatment is directly related to overeating behavior reduction, the effect should be evident in the data from all three participants. Specifically, overeating behaviors in the second and third women would continue at the baseline level until the onset of treatment.

Figure 2 below illustrates the relationship between a DV and an IV in a multiple baseline design with 3 participants. In this example, Participant 1 received the intervention at week 6, Participant 2 received the intervention at week 10, and Participant 3 received the intervention at week 14. Although the baseline phase varied in length for each participant, a trend in behavior was present across participants, as evident by the stability of behavior during this period. Once the intervention was introduced, another pattern of behavior emerged at the onset of the intervention phase.

Figure 2. *Example of a multiple baseline design with 3 participants*



Considerations for multiple baseline designs: This design avoids the problem of failure to revert to baseline that can exist with withdrawal/reversal designs, but still has challenges. Though all participants in a multiple baseline design will receive treatment, some will wait longer than others to receive it. If it is assumed that the intervention will be effective, withholding treatment during the baseline period may be questionable. However, there may be other reasons that some participants would need to wait to receive services (e.g., wait list due to full caseloads), in which case the baseline period may be a more natural one.

Other considerations for Single Case Designs:

- Varying, increasing, or decreasing trends in behavior during baseline can lead to difficulty interpreting treatment effects.

- Interventions must be powerful (with large and prompt effects) in order for the effects to be visibly observed.[9]
- SCDs require a substantial amount of information on each participant before beginning the intervention. Detailed operational definitions of the participants, the setting, and the participant selection process are required before the treatment can begin.[10]
- Behaviors need to be repeatedly and rapidly measured. This is often difficult to achieve when using screeners or measures that aim to capture fairly stable aspects of a particular outcome (e.g., the Beck Depression Inventory and depression).
- Because of the various ethical and practical issues described in this document, SCDs are not an appropriate measurement design for many commonly targeted home visiting outcomes (e.g., birth outcomes, breastfeeding, subsequent births, child development, etc.)

- The DOHVE technical assistance team would like to thank John Hitchcock, PhD, for his contribution to the content of this brief.

- For more specific information about evidence-based practice standards for single case design methodology, please access the What Works Clearinghouse website link below:

  http://ies.ed.gov/ncee/WWC/documentsum.aspx?sid=229

- Additional DOHVE TA resources are available at:

  http://www.mdrc.org/dohve/dohve_resources.html

- For more information about using single case design methods to evaluate MIECHV-funded home visiting programs, please contact a DOHVE[4] TA Liaison at:

Kerry Ryan, MA
TA Liaison
James Bell Associates
1001 19th Street North, Suite 1500
Arlington, VA 22209
703-528-3230
Ryan@jbassoc.com

Jill Filene, MPH
DOHVE TA Project Director
James Bell Associates
1001 19th Street North, Suite 1500
Arlington, VA 22209
703-528-3230
Filene@jbassoc.com

[1] Horner, R.H., Carr, E.G., Halle, J., McGee, G., Odom, S., Wolery, M. (2005). The use of single subject research to identify evidence-based practice in special education. Exceptional Children. 71(2), 165-179.
[2] Spiegler, M.D., & Guevremont, D.C. (2003). *Contemporary Behavior Therapy.* 4th Edition. (pp. 45-74). Belmont, CA: Wadsworth/Thomson Learning.
[3] Horner, R.H., Carr, E.G., Halle, J., McGee, G., Odom, S., Wolery, M. (2005). The use of single subject research to identify evidence-based practice in special education. Exceptional Children. 71(2), 165-179.
[4] Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M. & Shadish, W. R. (2010). Single-case designs technical documentation. Retrieved from What Works Clearinghouse website: http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf.
[5] Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M. & Shadish, W. R. (2010). Single-case designs technical documentation. Retrieved from What Works Clearinghouse website: http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf.
[6] Spiegler, M.D., & Guevremont, D.C. (2003). *Contemporary Behavior Therapy.* 4th Edition. (pp. 45-74). Belmont, CA: Wadsworth/Thomson Learning.
[7] Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M. & Shadish, W. R. (2010). Single-case designs technical documentation. Retrieved from What Works Clearinghouse website: http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf.
[8] Spiegler, M.D., & Guevremont, D.C. (2003). *Contemporary Behavior Therapy.* 4th Edition. (pp. 45-74). Belmont, CA: Wadsworth/Thomson Learning.
[9] DOHVE TA Webinar: Program Evaluation For Promising Home Visiting Approaches. (2010). Retrieved from DOHVE Project Resources website: http://www.mdrc.org/powerpoint/dohve_webinar_4.pdf
[10] Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M. & Shadish, W. R. (2010). Single-case designs technical documentation. Retrieved from What Works Clearinghouse website: http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf.

[4] The purpose of the Design Options for Home Visiting Evaluation (DOHVE) is to provide research and evaluation support for the Maternal, Infant and Early Childhood Home Visiting (MIECHV) Program. The project is funded by the Administration for Children and Families in collaboration with the Health Resources and Services Administration.