

THEORY/REVIEW MANUSCRIPT

Alignment of Single-Case Design (SCD) Research With Individuals Who Are Deaf or Hard of Hearing With the What Works Clearinghouse Standards for SCD Research

Erica Wendel, Stephanie W. Cawthon*, Jin Jin Ge, S. Natasha Beretvas

The University of Texas at Austin

*Correspondence should be sent to Stephanie W. Cawthon, Department of Educational Psychology, The University of Texas at Austin, 1912 Speedway Avenue MC D5800, Austin, TX 78712 (e-mail: stephanie.cawthon@austin.utexas.edu).

Abstract

The authors assessed the quality of single-case design (SCD) studies that assess the impact of interventions on outcomes for individuals who are deaf or hard-of-hearing (DHH). More specifically, the What Works Clearinghouse (WWC) standards for SCD research were used to assess design quality and strength of evidence of peer-reviewed studies available in the peer-reviewed, published literature. The analysis yielded four studies that met the WWC standards for design quality, of which two demonstrated moderate to strong evidence for efficacy of the studied intervention. Results of this review are discussed in light of the benefits and challenges to applying the WWC design standards to research with DHH individuals and other diverse, low-incidence populations.

There are many challenges in establishing a robust foundation of interventions and evidence-based practices for individuals who are deaf or hard-of-hearing (DHH) (Luckner, 2006; Luckner, Sebald, Cooney, Young, & Muir, 2005).¹ Some of these challenges in the field stem from the demographic reality of assessing diverse DHH student populations across a broad range of educational settings that vary in the extent to which interventions are tailored to fit the needs of DHH students. Other challenges are a result of few resources available to conduct said studies, particularly needed replications of prior findings. Replications across research teams is also challenging because, in many cases, there are divergent perspectives on the theoretical underpinnings of the field. From a measurement perspective, the low-incidence and heterogeneity of the DHH population restricts the range of options for experimental research designs with sufficient numbers of students in control and treatment conditions (with sufficient covariates), thus limiting the breadth and scope of many educational research methodologies that can be feasibly implemented and

lead to strong causal inferences. This, in turn, reduces the depth of the research literature from which causal inferences about effective interventions for DHH students might be made.

Evidence-based practices in many areas in education are informed by results from large-scale experimental studies such as randomized control trials (Thompson, Diamond, McWilliam, P. Snyder, & S. Snyder, 2005). Alternative methods to traditional group-based experimental designs include descriptive approaches, such as case studies, and other empirical approaches, such as single-case designs (SCDs) (Horner, Carr, Halle, Odom, & Wolery, 2005). SCDs include repeated observations of participants across multiple phases of the study, from before the intervention through follow-up, using a participant's individual changes in behaviors as the point of reference for interpreting degree and cause of change. In recent years there has been increased recognition of the importance of the SCD for estimating the effectiveness of interventions for low-incidence populations (Kratochwill et al., 2013; Shadish & Sullivan, 2011),

including individuals who are DHH (e.g., Antia & Kreimeyer, 2003; Beal-Alvarez, Lederberg, & Easterbrooks, 2012).

In the past, a lack of explicit protocol related to experimental design and data interpretation has been an inherent challenge with evaluating the quality of SCD research and the strength of resultant findings. More specifically, the use of visual analysis as the primary means to interpreting interventions effects is a major criticism of SCDs, (Kazdin, 2011), because visually analyzed results are vulnerable to researcher bias, subjectivity, and inconsistency. Regardless of these drawbacks, the potential for SCDs to contribute to the advancement of research among low-incidence groups has been widely acknowledged. Research utilizing SCDs is a methodologically sound and functionally appropriate way to construct experimental research studies with individuals who are DHH. Most notably, SCDs are often compatible with research on topics in deaf studies and deaf education, such as language development, literacy, and communication strategies. These studies often demonstrate observable outcomes at the individual level, allowing for analysis of a specific individual's response to treatment and a meaningful interpretation of treatment outcomes that is responsive to population diversity (Bergeron, Lederberg, Easterbrooks, Miller, & Connor, 2009; Long, 1996; Odom et al., 2005; Schirmer & McGough, 2005).

Although there have been a number of SCD studies with DHH participants, a systematic review of the scope, quality, or methodological rigor of these publications has yet to be conducted. Given the paucity of intervention research that includes DHH individuals, and the compatibility of SCD approaches with measuring intervention outcomes among low-incidence groups (Horner et al., 2005), an exploration of the quality and scope of SCD studies with participants who are DHH is warranted. The purpose of this article is to evaluate the alignment of research using SCD across DHH literature with the guidelines provided by the What Works Clearinghouse (WWC), a project sponsored by the U.S. Department of Education to evaluate the strength of evidence for interventions studied within education research (<http://ies.ed.gov/ncee/wwc/>). (Specific WWC guidelines for SCDs are discussed in the Method section.) In this article we first provide a more detailed discussion of the implications of research design decisions in studies that include DHH participants, specifically looking at group-based designs and SCDs. We offer this discussion of group-based designs as a point of comparison because of the prevalence of this design approach in intervention research. We then describe the WWC criteria by which we evaluated each included study and the types of indicators that varied between SCDs with DHH participants. This article concludes with a discussion on the state of SCD research within the field (in accordance with the WWC standards) and future directions for research.

Group-Based Versus SCDs

In this section, we offer a comparison of a widely used method of demonstrating causal relationships within intervention research (i.e., group-based experimental research designs) and SCDs. We provide this comparison in order to illustrate the appropriateness of SCDs to demonstrate participant outcomes on an individual level, as group-based designs provide a suitable contrast to illustrate this concept. Please note that the following discussion does not encompass the full range of methodologies that can be used to draw causal inferences; for a comprehensive review of quantitative research designs that can be used to demonstrate causal relationships, please refer to Shadish, Cook, and Campbell (2002).

Group-Based Designs

In order to establish that a treatment or practice is effective, researchers need to demonstrate that a causal relationship exists between an intervention and student outcome variables (Tankersley, Harjusola-Webb, & Landrum, 2013). A causal relationship is established when participant outcomes are demonstrated to be the product of an intervention while external factors are simultaneously ruled out as the source of change (Barger-Anderson, Domaracki, Kearney-Vakulick, & Kubina, 2004). Although group-based research, specifically RCTs, has been historically regarded as the "gold standard" for determining causal relationships between intervention and outcome variables (Borckardt et al., 2008; Parker, Davidson, & Banda, 2007), there are significant limitations to using group designs with low-incidence populations. One limitation is the challenge of having enough DHH students in treatment and control conditions to have sufficient power to detect an effect of the intervention, if it exists. For example, there may not be enough DHH participants with similar demographic backgrounds to be able to "match" students in a matched-pair random assignment. As a further example, in order to gain the needed samples, researchers often face significant nesting issues (e.g., all students who are DHH in a district receive services in the same classroom or school location) or fidelity issues (e.g., treatments must be administered across multiple sites, introducing variance from different administrators, interpreters, and other providers of the intervention (Luckner, 2006; Luckner et al., 2005).

Even when sufficient numbers of DHH participants can be assigned to treatment and control conditions of an experimental study, there are still significant obstacles to drawing appropriate inferences of nuances of the DHH population that are challenging to account for in a group-based research design. Individuals who are deaf are a highly heterogeneous population with characteristics that often intersect between language, disability, communication, and cultural identity (Baker-Shenk & Kyle, 1990; De Clerck, 2010; Najarian, 2008; Wheeler-Scruggs, 2003). For research with DHH participants, it is particularly important to consider possible confounds related to constructs such as language and communication modalities or disability status (e.g., learning disability), as these could have a considerable effect on intervention outcomes across a highly heterogeneous group of individuals (Antia & Kreimeyer, 2003; Bullis & Anderson, 1986). Even when sufficient numbers of participants might be available for a group design, researchers may still opt for a SCD approach because group designs are limited in the extent to which a variety of differing demographic dimensions of participants can be represented the same time as investigating the specific effects of the intervention.

Because of the heterogeneity of DHH population, the degree of success for interventions and resultant practices are often dependent upon the match between the strategy or practice and the individual student characteristics. Group-based data, particularly those presented without an investigation of interaction effects, often mask important moderators related to participant demographic variables. As a result, researchers using the group design approach will have a limited understanding of specific treatment needs and intervention responsiveness between DHH subgroups. For example, individuals who are native deaf and those who are late-deafened are likely to respond differently to a training intervention for increasing functional communication skills, given the varying degrees of language exposure and divergent communication styles between these groups. This variability in both etiology and characteristics have implications for the

generalizability of intervention research across DHH subpopulations (e.g., individuals who are deaf with additional disabilities), in that aggregated outcome data are not ideal for interpreting the differences in how individuals may differentially respond to the intervention or strategy.

Single-Case Designs

Due to the focus on individual level participant outcomes (rather than aggregated group outcomes), SCDs allow for a more detailed examination of differences across subpopulations within the broader context of individuals who are DHH. Take, for example, a group design approach to looking at the effect of an intervention with students from a variety of home language contexts. A typical grouping might be to divide participants into three groups: families with at least one parent who is fluent in sign language, families with parents who learn sign while the child is growing up, and families with parents who do not use sign at all. This would give three groups to use in analysis as either an independent variable (IV) or a covariate. In an SCD approach, instead of forcing a group delineation, it would be possible to meaningfully utilize multiple variables that form a robust picture of the participants home language context. Each individual case would have a profile that could include a greater number of relevant variables such as the sign language proficiency of both parents (where applicable), siblings' use of sign and hearing status, access to language models outside of the home, early intervention resources, and so forth. The individual participant's full profile would be displayed alongside the baseline and response to interventions. In comparison with a group design, the SCD approach provides a greater context for interpreting the data related to outcome data.

Although there are multiple variations of the SCD, each type shares some common characteristics. The foundational requirement of SCD is related to the staggered application of the intervention across participants, and the use of repeated measures of outcome variables across cases. This means that whenever treatment is applied to one case (e.g., a participant, setting, or behavior) and not to another, there exists a comparison between the treatment and no-treatment group (Kazdin, 2011). For a visual depiction of staggered treatment application, see

Figure 1. Staggering the introduction of an intervention across cases allows for more stringent analysis of outcomes among different participants, behaviors, or settings. More specifically, gathering multiple data measurements allows researchers to closely examine trends in performance during the baseline and intervention conditions (referred to as "phases"), and allows researchers to rule out competing explanations for changes in the dependent variable between baseline and intervention phases (Kelly, 1998).

During the baseline phase, or the "A phase," data is collected on a participant's performance prior to the introduction of the intervention. The purpose of the baseline phase is to provide a basis of comparison between the treatment and no-treatment conditions (Kazdin, 2011). Researchers can determine the likelihood that an individual's performance will remain stable in the absence of an intervention by examining data patterns from baseline measurements. Conclusions can be drawn about the efficacy of an intervention when there are changes to the data pattern during the intervention phase, or "B phase," relative to baseline measurements. Additionally, data collected from the baseline phase allows researchers to observe an individual's performance or behavior in the absence of intervention. This type of descriptive analysis is useful for determining the extent of the problem and to confirm that the intervention and intended outcomes are well suited for the participant. One limitation to measuring participant outcomes over time is known as the threat of maturation, which refers to the natural changes that can occur in participants over time. Changes due to maturation threaten internal validity if they could have produced the outcome that would otherwise be attributed to the intervention (Shadish et al., 2002). Threats of maturation can be minimized by ensuring that participant groups are similar across variables that are susceptible to changes across time (i.e., age, geographic location) (Murray, 1998).

A commonly used variation of the SCD in deaf studies and deaf education research is the multiple-baseline design. Common to all types of multiple-baseline designs, treatment effects are demonstrated by systematically introducing an intervention at different points in time across different baselines (i.e., participants, settings, or behaviors), and causal relationships between two or more variables are determined by

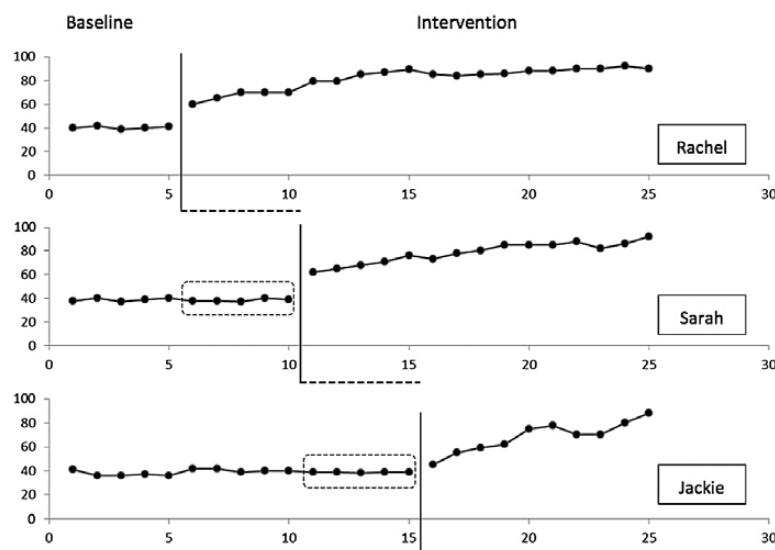


Figure 1. Example diagram of a single-case, multiple-baseline design across participants.

comparing data between intervention and baseline phases. The strength of documented intervention effects using a multiple-baseline design is dependent on the nonconcurrent application of the IV across cases.

The multiple-baseline design is often used with interventions that are intended to generate outcomes that are not reversible, such as literacy or communication interventions (Barger-Anderson et al., 2004). The multiple-baseline design allows investigators to determine both the immediate and generalized effects of a treatment across multiple settings, participants, or behaviors. SCD is a suitable methodology across a number of different research scenarios, such as communication, behavioral, or social skills interventions. When investigating participant outcomes related to communication skills, for example, it is important to assess the persistence and generalization of those skills across time and different settings (e.g., work, school, and home). SCD calls for continuous or periodic monitoring of skill generalization during intervention and maintenance phases, which allows investigators to determine how much of a particular treatment is needed to produce skill generalization across settings. Furthermore, because interventions for individuals who are DHH are often conducted in one type of setting, such as a single classroom or school, assessing the generality of outcomes outside of the direct training environment is necessary to determine the efficacy of the intervention.

In the multiple-baseline design, intervention data can be collected continuously throughout the experiment, or periodically, after a stable baseline is reached (Horner & Baer, 1978). Continuous (as opposed to periodic) data collection with the multiple-baseline design is a desirable approach because it provides a more complete representation of the effects of an intervention, especially with regard to the generality of outcomes. A significant advantage to the multiple-baseline design is that periodic data collection can be used strategically to demonstrate both the immediate and persisting effects of an intervention (Kaiser, 2014). A major drawback to using continuous data collection is that it is often very time-consuming and costly, especially when there are multiple behaviors or settings of interest or if data is collected across multiple contexts.

A variation on the multiple-baseline design that uses periodic data collection is known as the multiple-probe design (Horner & Baer, 1978; Kazdin, 2011; Richards, Taylor, & Ramasamy, 2014). Periodic data collection is especially efficient for assessing the generalizability of outcomes, or the persistence of effects after the intervention has ended, as is usually desired in behavioral, communication, or literacy interventions. Because many interventions that are developed for the DHH population are utilized within these domains, the multiple-baseline design can be considered an exceptionally compatible approach for experimental research in deaf studies and deaf education.

Figure 1 illustrates the use of a multiple-baseline design across participants for a hypothetical intervention study targeting expressive communication in DHH participants. The graph demonstrates the effects of a cued speech intervention (IV) on the percent of correctly expressed vocabulary words (dependent variable) for each participant. Consider the data for the first participant represented in this figure, Rachel. The five data points collected during the baseline phase reflect her performance in the absence of the intervention. This data provides a basis of comparison for her performance after the introduction of the intervention. Notice the intervention phase for the next participant, Sarah, began at session 10. Sarah's baseline data for sessions six through 10 serves as a control condition for the initial

sessions of Rachel's intervention phase. Another control condition occurs for Jackie's case, during sessions 11 through 15. Because there was a change in each participant's performance after the introduction of the intervention, but no changes to the baseline data during each participant's control condition, we can more confidently attribute the observed changes to the intervention.

The Current Study

Due to the limited number of identified evidence-based practices that have been linked to positive outcomes in DHH individuals (Easterbrooks, 2010), it is especially important to conduct and evaluate empirical research with consistent and rigorous standards of methodological quality. The purpose of the present analysis was to answer the following questions: (a) To what extent is published SCD research with DHH individuals aligned to the standards established by the WWC? (b) What are the specific methodological strengths and weaknesses of the studies within the SCD literature with individuals who are DHH?

Method

Article Search Strategy

The initial article search was conducted using the publications databases ERIC, PsycINFO, PsycARTICLES, Academic Search Complete, and Google Scholar. Manual searches were also used within the Journal for Deaf Studies and Deaf Education. The following key words were included in the search criteria: Deaf*, hard-of-hearing, hearing impaired, hearing loss, single-case, single-subject, intervention. These search terms do not assume to cover the full range of characteristics, identities, and terminology that are associated with individuals who are DHH; they are instead vocabulary and key words that are commonly used in the research literature. The search yielded 768 articles, dissertations, book chapters, and other literature that was potentially suitable for our analysis based on article abstracts. Publication titles and abstracts were screened to determine which studies should be considered for inclusion. After the initial screening, 97 studies were identified for further review.

Each publication was screened for inclusion eligibility with following criteria: (a) participants were identified as DHH (e.g., deaf, hearing impaired, with a hearing loss, late-deafened, hard-of-hearing) or DHH with an additional disability; (b) the research design was a variation of SCD that demonstrated experimental control and replication as outlined by the WWC (i.e., multiple-baseline or multiple-probe, reversal, alternating treatments, changing criterion); (c) used individual level unit of analysis (i.e., not at a classroom level); (d) the study had to present the data graphically to allow for visual analysis; and (e) published in a peer-reviewed journal. The decision to only include peer-reviewed research was because these articles are most likely to be used in summaries of evidence in support of specific interventions. Of the 97 publications that were initially retrieved, 12 were eligible for inclusion in this analysis. A diagram of the literature search process is provided in Figure 2 and a summary of included articles is provided in Table 1.

The WWC Standards for SCD Research

The specific criteria used to assess each study are based on the most recent pilot version of WWC single-case

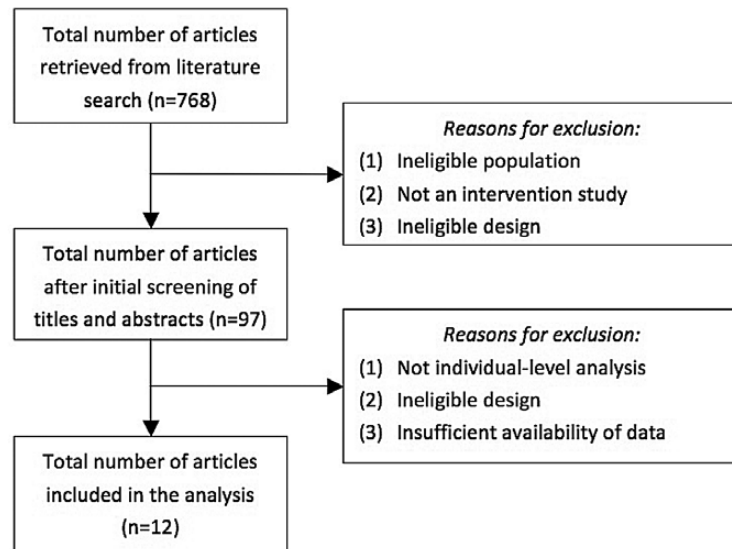


Figure 2. Single-case design article inclusion flowchart.

intervention research design standards, which are documented in the *Procedures and Standards Handbook*, version 3.0 (Kratohwill et al., 2010/2014). WWC was established in 2002 by the Institute for Education Sciences to provide evaluations of the quality of studies that test the effects of educational and psychological interventions. Although the WWC initially focused on the analysis of group-based methodologies, the *Procedures and Standards Handbook* now proposes guidelines for designing and evaluating a research study using SCD. The guidelines for evaluating the SCD research include protocol for determining the strength of the research design to ensure sufficient internal validity, and the strength of the evidence through visual analysis of data (Maggin, Chafouleas, Goddard, & Johnson, 2011). This protocol will be summarized in the following sections,² followed by a description of the quality and analysis of evidence conducted with the present sample.

Assessment of design quality

The first step in quality assessment was an evaluation of the studies' design properties, which yielded quality rankings of either (a) Meets WWC Pilot SCD Standards without Reservations, (b) Meets WWC Pilot SCD Standards with Reservations, or (c) Does Not Meet WWC Pilot SCD Standards. This process is summarized in Figure 3. The four criteria for WWC design quality (described below) were assessed on a dichotomous scale (yes or no). In order to meet design standards without reservations, all of the following criteria must be met:

1. Independent variable: The investigators must systematically manipulate the IVs. In other words, the researchers must determine how and when the IV is introduced.
2. Inter-observer agreement: Each dependent variable must be measured systematically over time by no less than two investigators, using an accepted measure of inter-observer agreement (IOA) to determine agreement for each baseline across each dependent variable. Documentation of IOA must be reported for each phase, for no less than 20% of the points within each phase. Agreement must be no less than 0.80 if measured by percentage agreement, and 0.60 if using Cohen's kappa.
3. Attempts to demonstrate effects over time: A study must include at least three attempts to demonstrate the effects

of the IV at different points in time.³ Intervention data that overlap across cases do not represent an attempt to demonstrate an effect. Further clarification on this criteria include the following:

- Some commonly used SCDs do not meet this standard, including AB, ABA, and BAB designs.
 - Multiple-baseline designs must have at least three baseline phases, and at least three intervention phases (three A and three B phases).
 - For alternating and simultaneous designs, five demonstrations of an effect are required.
4. Data points per phase: In order to constitute an attempt to demonstrate an effect, studies need to meet criteria with regard to the number of data points per phase. There is some flexibility in the application of these standards, depending on the unique properties of the research population and intervention. The principal investigator must provide strong justification for applying alternate criteria. Further clarification includes the following:
 - For reversal/withdrawal designs (e.g., ABAB) to meet design standards without reservations, there must be at least two cases with four phases per case, with a minimum of five data points per phase. In order to meet design standards with reservations, phases must have at least three data points per phase. Phases with fewer than three data points are not sufficient to demonstrate an effect.
 - For multiple-baseline and multiple-probe designs, there must be a minimum of six phases with at least five data points per phase to meet design standards without reservations. To meet standards with reservations, there must be a minimum of six phases with at least three data points per phase.
 - For alternating treatment designs, there must be a minimum of five data points per condition and two data points per phase at most to meet design standards without reservations. To meet design standards with reservations, there must be four data points per condition and two data points per phase. Phases with more than two data points cannot be used to demonstrate an effect because the design calls

Table 1. Analyzed studies

Study authors	n	Age	Hearing status			Additional disability	Intervention	Outcomes
			Degree of loss	Amplification				
S1 Allgood, Heller, Easterbrooks, and Fredrick (2009)	5	17–20	Moderate, profound	NR		Mild ID, moderate ID	Instructional: Picture dictionaries	Functional communication in vocational setting
S2 Heller, Allgood, Ware, Arnold, and Castelle (1996)	4	18–19	Moderate, profound	NR		Mild ID, moderate ID, visual impairment	Instructional: Dual communication boards	Functional communication in vocational setting
S3 Beal-Alvarez et al. (2012)	1	5	Moderate	HA		NR	Instructional: Emergent literacy curriculum combined with visual phonics administered individually	Grapheme–phoneme correspondence acquisition
S4 Beal-Alvarez et al. (2012)	3	4–5	Moderate, severe	HA		NR	Instructional: Emergent literacy curriculum combined with visual phonics administered in a small group	Grapheme–phoneme correspondence acquisition
S5 Bergeron et al. (2009)	5	3–7	Moderate, severe	CI, HA		NR	Instructional: literacy curriculum from children’s early intervention (CEI) program	Knowledge of correspondence between alphabetic phonemes and graphemes
S6 Bergeron et al. (2009)	5	3–4	NR	CI		NR	Instructional: Component of foundations for literacy program	Knowledge of correspondence between alphabetic phonemes and graphemes
S7 Cannon, Fredrick, and Easterbrooks (2009)	4	10–12	NR		NR	NR	Instructional: DVD expository books	Vocabulary words
S8 Cohen, Allgood, Heller, and Castelle (2001)	3	17–20	Moderate, severe, profound	NR		Borderline ID, mild ID	Instructional: Picture dictionaries	Functional communication in vocational setting
S9 Miller, Lederberg, and Easterbrooks (2012)	5	3–5	Moderate, severe	CI, HA		None	Instructional: Component of foundations for literacy program	Phonemic and phonological awareness
S10 Mueller and Hurtig (2010)	4	2–4	Mild, moderate, severe, profound	CI, HA, unaided		None	Instructional: Reading program from The Iowa Signing E-book	Time on-task during reading session; acquired sign vocabulary
S11 Neef and Iwata (1985)	2	23–26	NR	NR		None	Instructional: Sequential cued speech training	Acquisition of lip reading skills; expressive articulation responses
S12 Van Hasselt, Hersen, Egan, Mckelvey, and Sisson (1989)	2	21	Moderate, severe	NR		Blindness, cerebral palsy, severe ID, seizure disorder, cardiac disease	Behavioral reinforcement training	Increase of on-task behaviors; appropriate social interaction; decrease of self-injury, disruption, and stereotypy

Note. CI, cochlear implant; HA, hearing aid; NR, not reported.

for fast alternations between phases. Designs that attempt to compare multiple interventions are rated individually.

Studies that used a multiple-probe design were assessed with additional criteria, outlined below. Multiple-probe studies that

did not meet all of the following criteria WWC received a rating of Does Not Meet WWC Pilot SCD Standards.

1. Initial baseline sessions must overlap vertically. The design must include three consecutive probe points for each

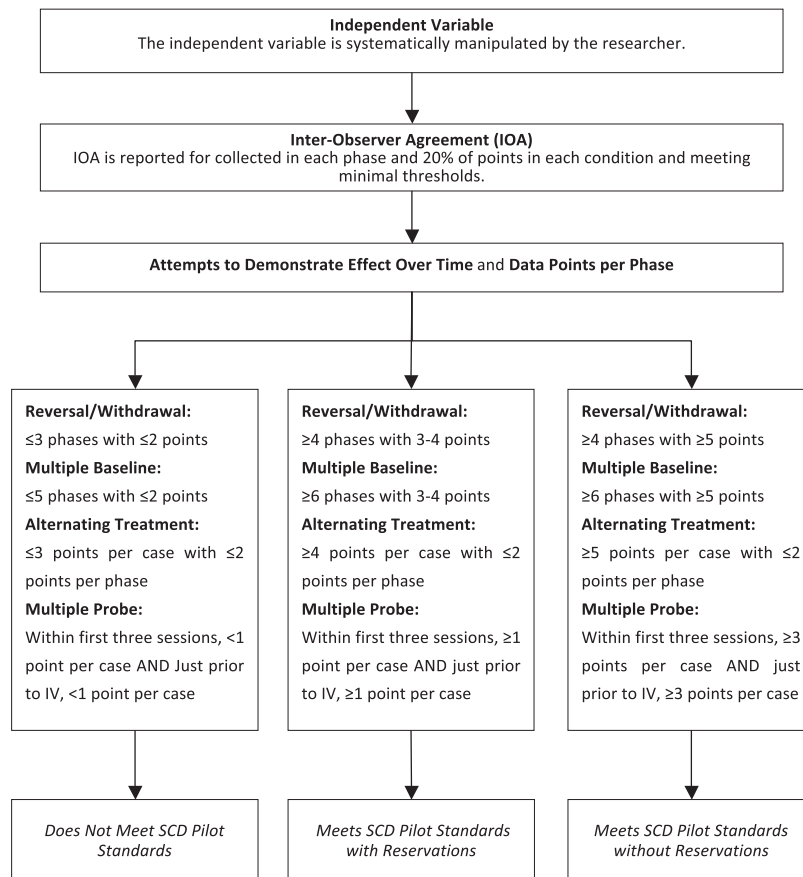


Figure 3. WWC quality rating criteria for SCD studies. SCD, single-case design; WWC, What Works Clearinghouse.

case within the first three baseline sessions. Studies with at least one probe point for each case within three initial baseline sessions Meet Pilot Standards with Reservations.

2. Probe data must be available in the sessions immediately prior to the introduction of the intervention. Within the three sessions just prior to the introduction of the intervention, there must be three consecutive probe points for each case. At least one probe point within the three sessions before the intervention phase is required to Meet Pilot Standards with Reservations.
3. For each case that is not currently receiving the intervention, there must be a probe data point during a session where another case either (a) first receives the intervention, or (b) reaches the predetermined intervention criterion level. This point must be consistent in level and trend with the case's previous baseline points.

Steps for visual analysis

The next step was to apply the WWC protocol for visual analysis of data⁴ to studies that meet design standards with or without reservations. Recall that the rationale for using visual analysis of data in SCD studies is that observed changes in the dependent variable can be associated with manipulation of the IV, which is based on the observation of predicted and replicated data across cases. Visual analysis is used to determine whether the data demonstrated at least three indications of an effect at different points throughout the study.

Quality of effect demonstrations will be assessed to determine whether the study provides strong evidence, moderate

evidence, or no evidence of a causal relationship. An intervention effect is demonstrated by documenting six experimental properties. The first three properties should be measured within phases, while the last three refer to properties of the data patterns across phases:

1. Level: The mean score of the outcome measures within a phase.
2. Trend: The slope of the best fitting line for the data within a phase.
3. Variability: The degree of overall scatter around the line of best fit. Generally speaking, increased data variability detracts from the strength of the relationship between variables.
4. Immediacy of the effect: Refers to the change in level between the last three data points in one phase, and the first three data points of the next. An immediate change in an effect between phases provides a stronger indication of an intervention effect.
5. Overlap: The proportion of data from one phase that overlaps with data from the previous phase. In general, a smaller proportion of overlapping data points leads to a stronger demonstration of the intervention effect.
6. Consistency of data in similar phases: Involves looking at the data from all phases within the same condition (e.g., baseline vs. intervention) and investigating the consistency in the data patterns from phases with the same conditions. Increased consistency is an indication of a causal relationship between the independent and dependent variables.

If a SCD study provided three demonstrations of an intervention effect, as determined through visual analysis of the above properties by at least two trained observers, then the study was considered to demonstrate strong evidence of a causal relationship. If a study does not provide three demonstrations of an effect at different points in time, then the study is rated as no evidence. If a study provides three demonstrations of an effect, plus one demonstration of a noneffect, then the study is considered to provide moderate evidence of an intervention effect. A noneffect is demonstrated with the following criteria: (a) data in the baseline phase are not stable enough to provide a meaningful comparison with data during the intervention phases; (b) high variability throughout phases, or inconsistent data trends across similar intervention phases; (c) a long latency between the introduction of an intervention and an observed effect; (d) overlap of outcome effects between baseline and intervention phases; or (e) inconsistency in outcomes between different cases (e.g., subjects, behaviors, or settings).

Results

The following section describes an analysis of design quality and strength of evidence of SCD intervention studies that include DHH participants utilizing the WWC standards. Table 2 summarizes the results of the design quality analysis for the included articles in this analysis.

Coder Reliability for Design Standards

The first three authors coded each article utilizing the WWC design quality standards. Each coder evaluated the design quality of each study across the individual criterion levels, which led to one of three comprehensive ratings in accordance with the WWC design standards: (a) Meets WWC Pilot SCD Standards without Reservations; (b) Meets WWC Pilot SCD Standards with Reservations; (c) Does not Meet WWC Pilot SCD Standards. After an initial training period, IOA in rating the sample of SCD studies was 100% across all criteria.

Design Quality Outcomes

The majority of studies within the sample ($n = 8$ out of 12) did not meet the standards for WWC design quality. Across these

studies, there were several weaknesses that related to application of SCD methodology, experimental validity, and overall design rigor. The weaknesses that compromised design quality fell under three categories, including (a) failure to meet IOA criteria; (b) insufficient demonstration of intervention effects; or (c) inadequate experimental control. According to the WWC framework, studies which fail to meet any one of the design quality criteria does not meet standards.⁵

IOA criteria

Among the studies that did not meet design standards, a common flaw was the failure to meet IOA criteria, with three studies not fulfilling these criteria. Failure to meet quality standards under IOA criteria was a result of insufficient percent agreement between observers (less than 0.80), or because the investigators did not measure or report IOA. For studies that reported and IOA above 0.80, it was often the case that the investigators did not report the number of sessions or cases for which an IOA measure was taken. Insufficient reporting of IOA resulted in questionable validity of the studies' data, which detracted from the strength of the reported effects.

Insufficient demonstration of intervention effects and inadequate experimental control

Three studies had several methodological flaws that compromised the integrity of the effect demonstration. An uncommon but significant design issue resulted from a failure to manipulate IVs under controlled conditions, which resulted in compromised experimental integrity. For example, one study did not administer the IV in an experimentally controlled condition (i.e., intervention was administered in the participants' homes and not supervised by a trained investigator or with checks for fidelity). Another design weakness was resulted from multiple intervention phases with fewer than three data points, and insufficient opportunities for comparison of outcome data across phases, baselines, and participants. Studies that did not meet standards under this category also failed to meet standards for adequate attempts to demonstrate an effect. However, the presence of a phase with fewer than three data points was not an automatic disqualifier for a study. If the remaining phases with three or more data points were sufficient to demonstrate

Table 2. Results of design quality analysis

Study	Manipulation of IV	Three attempts to demonstrate an effect	Three data points in each phase	IOA	Criteria for multiple probe	Design rating
S1	Yes	Yes	Yes	Yes	Yes w/reservations ¹	Meets Pilot Design Standards with Reservations
S2	Yes	Yes	Yes	Yes	Yes w/reservations ¹	Meets Pilot Design Standards with Reservations
S3	Yes	Yes	No ³	Yes	No ^{1,3,6}	Does not Meet Pilot Design Standards
S4	Yes	Yes	Yes	Yes	No ⁶	Does not Meet Pilot Design Standards
S5	Yes	Yes	Yes	No ²	No ⁵	Does not Meet Pilot Design Standards
S6	Yes	Yes	Yes	No ²	No ^{5,6}	Does not Meet Pilot Design Standards
S7	Yes	No ^{3,6}	No ³	Yes	No ^{1,6}	Does not Meet Pilot Design Standards
S8	Yes	Yes	Yes	Yes	Yes w/reservations ¹	Meets Pilot Design Standards
S9	Yes	No ⁶	No ³	Yes	No ^{1,5,6}	Does not Meet Pilot Design Standards
S10	No ⁴	Yes	No ³	No ²	Not applicable	Does not Meet Pilot Design Standards
S11	Yes	Yes	No ³	Yes	Not applicable	Does not Meet Pilot Design Standards
S12	Yes	Yes	Yes	Yes	Not applicable	Meets Pilot Design Standards

Note. 1 = Fewer than three probe points at the start of the baseline phase for each case, 2 = No IOA data gathered or reported, 3 = Baseline and/or experimental phase(s) with less than three total data points, 4 = Inadequate experimental control, 5 = Insufficient data points just prior to introduction of IV, 6 = Probe data missing or not vertically aligned according to multiple-probe standards. IOA, inter-observer agreement; IV, independent variable.

three effects, then the study was still eligible to meet standards in this domain.

Another issue under this category related to insufficient data, specifically in the baseline phases. Insufficient baseline data within and across baseline phases detracts from the internal validity of a multiple-baseline design, as adequate baseline data is necessary for effect demonstration across all variations of SCDs. Periodic data collection is not always an indicator of insufficient baseline data, as adequate baseline data is a feature of well-constructed multiple-probe designs, which was a commonly used variation of SCD within our sample. However, this flaw was most prevalent across studies using the multiple-probe design.

Other methodological issues

Another design problem within our sample resulted from insufficient number and placement of data points within studies that used the multiple-probe design. Because multiple-probe designs use periodic data collection, they are generally more cost efficient than designs that utilize continuous data collection. However, the cost for increased resource efficiency is the potential reduction of methodological rigor. Periodic data collection can detract from the strength of a demonstrated effect, especially if there is high variability in the data patterns. For example, if probe data is inconsistent or unexpected given the time sequencing of treatment application (i.e., if a behavior is observed that is unlikely to occur without prior training), then it would be more difficult to attribute the observed effects to the intervention, especially if the design was not executed with methodological rigor.

Results of visual analysis

Following the analysis of design quality, the studies that were determined to meet design standards with or without reservation were evaluated for strength of evidence using the guidelines for visual analysis outlined above. Out of the four studies that met design standards with or without reservations, all were determined to have moderate to strong evidence. (Moderate evidence: Van Hasselt et al., 1989, strong evidence: Allgood et al., 2009; Cohen et al., 2001; Heller et al., 1996.) The study that was rated as providing moderate evidence (Van Hasselt et al., 1989) had one documented noneffect, which resulted from a high level of variability in the data in one of the phases.

Discussion

In order to identify evidence-based practices for students who are DHH, researchers are called to produce more stringent study designs that are sensitive to the low-incidence and heterogeneity of the DHH population. The rising attention surrounding the SCD as a mechanism for drawing causal inferences about the effectiveness of interventions represents a promising shift in the field and includes a call to define criteria for structurally sound and methodologically rigorous SCD research.

The purpose of this article was to apply the WWC standards for SCD studies in a review of the existing body of DHH research in order to determine the present state of the evidence using this methodological approach. At present, there are an insufficient number of studies in any one domain that meet the standards of SCD design quality to constitute an evidence-based practice. We now discuss the challenges of applying the WWC design evaluation protocol to DHH research, as well as the limitations to the present analysis and implications for future research.

Limitations

There are several limitations to consider in the present analysis. First, although a thorough literature search was used to identify studies for inclusion in the present analysis, it is possible that other potentially relevant articles were unidentified, particularly unpublished dissertation studies that did not undergo external peer review. An additional limitation is due in part to the WWC standards being a new tool for evaluating SCD research. The standards were developed to bring increased consistency and objectivity to the evaluation of SCD research; however, the current standards do not take specific populations or applied contexts into account, an issue discussed in further detail below.

Development of the standards for SCD studies provides the first explicit definition of methodological quality in the history of SCD research (Kratichwill & Levin, 2014). The present analysis was conducted within an emerging area of inquiry, which on a field-wide level, is presently in a state of flux. Consequently, it is important to acknowledge the inherent limitations to using a newly developed set of criteria to assess existing work. Challenges to using the standards that apply specifically to DHH research include population neutrality (i.e., uniform application of guidelines across diverse populations) and unclear guidelines on areas for flexibility in their application. It would be impractical to create individualized standards for every population, yet it is also unreasonable to have all users of SCD designs adhere to a rigid, one-size-fits all set of standards for research design across multiple populations of interest and research questions. Although the present WWC standards for SCD purport to offer flexibility in the application of the standards for certain criteria, there are no explicit guidelines for when to exercise this flexibility when using the standards to evaluate a study for its adherence to them.

An additional challenge to applying the WWC standards is in the lack of evaluation criteria for reporting participant characteristics and demographic information. Although it is not the responsibility of WWC to know the specific details of relevant demographic information across all psychological and educational research, it would be helpful for the guidelines to include this category as a place where each domain reflects on the inclusion or exclusion of attributes that may be theoretically relevant to the intervention and inferences about its efficacy. In the case of DHH research, it is especially important to consider participant variables and research contexts that are shaped by a unique set of participant characteristics (such as age of cochlear implantation or presence of a disability) and circumstances surrounding the practical application of experimental conditions (such as interventions administered in individual vs. group settings). Operationalized descriptions of participant characteristics are necessary for not only DHH populations, but also for other heterogeneous populations within special education research. Because there were no specific guidelines within the standards for evaluating the quality of demographic reporting, this article lacks a systematic evaluation of such variables.

One final recommendation for future versions of the WWC standards for SCD research is the need for specific guidelines for the technical application of SCD principles (e.g., instructions for timing the introduction of an intervention across baselines in the multiple-baseline design). More specifically, the standards do not thoroughly address the need for appropriate sequencing of introduction of the intervention for multiple-baseline and multiple-probe designs. In the current version of the standards, the criteria states, "multiple-baseline and multiple-probe designs implicitly require some degree of concurrence in the

timing of their implementation across cases when the intervention is being introduced. Otherwise, these designs cannot be distinguished from a series of separate AB designs” (Kratochwill & Levin, 2014, p. E.4). This statement touches on an important component of SCD methodology but does not provide enough guidance to execute a multiple-baseline or multiple-probe design with solid internal validity. Therefore, it is important that future versions of the standards incorporate more rigorous and thorough guidelines in this area.

Implications for Future Research

The WWC Standards for SCD research provides a methodological framework that can potentially increase the consistency of rigorous SCD application, which in turn will hopefully promote identification of evidence-based practices and interventions for DHH and other low-incidence populations. Considering that the articles reviewed in the present study were conducted prior to the development of the WWC standards for SCD, it would be counterproductive to discount these studies based on imperfect methodology, particularly in a field where the evidence base for interventions and practices are still gaining capacity. One of the implications of this analysis is that scholars in deaf studies and deaf education may need to engage in a dialog as to where the SCD standards are applicable (and to what degree), and places where the unique context of the field may call for flexibility in how the WWC SCD standards are applied.

Within the field of applied research with individuals who are DHH, interventions are developed for a highly diverse group of individuals, for which group-based research methodologies are sometimes not appropriate. Consequently, the development of a consistent, documented methodology for evaluating SCD studies will likely lead to an increase in the identification and propagation of high-quality, evidence-based practices among traditionally underserved populations, including individuals who are DHH (Beal-Alvarez & Cannon, 2014; Maggin, Briesch, & Chafouleas, 2013). The present review serves as an indicator of the current state of single-case methodology applied to deafness research, and offers direction for future SCD research in the field. Ultimately, the continued use and refinement of the WWC method of quality and evidence analysis, tied with strong theoretical foundation and sufficient sampling across the broad range of individuals in this population, should lead to a stronger evidence base of interventions and practices within DHH research.

Conclusion

Although the present sample of SCD studies does not provide sufficient empirical evidence to identify an evidence-based practice in any one content area on their own merit, these articles undoubtedly contribute to the development of evidence-based interventions and practices designed for individuals who are DHH, summarized across all methodological approaches. To assess this body of SCD work without considering experimental context would be problematic and unduly critical. Furthermore, it is necessary for authors to report (and reviewers to consider) relevant covariates such as age of onset or language development in order for the significance of the study findings and applications to specific subpopulations to be adequately considered. SCDs hold the same responsibilities as other methodologies in clearly establishing the connections between the intervention and theoretical rationale for its efficacy with individuals who are DHH. We acknowledge that the studies in our

sample represent the foundation of an important branch of experimental research on issues related to interventions with individuals who are DHH, and that their findings still have significant impact on and application in the field.

Although there are limitations to using the WWC standards to evaluate the quality of SCD research in deaf studies and deaf education, the current standards do provide a solid framework for guiding future SCD research in the field. It is our hope that this article will support a thoughtful critique of SCD use in its current form, and their use in interpreting the effects of interventions for individuals who are DHH. We also encourage investigators to use the WWC standards as a platform for discussing the validity of research findings using SCDs and ways in our field might apply them to improve the strength of inferences about psychological and educational research in our field.

Funding

Research to Practice Division, Office of Special Education Programs and the U.S. Department of Education via Grant Award (H326D110003).

Notes

1. Many different terminologies for individuals who are DHH are used across the research literature. Studies that focus on the cultural, linguistic, and societal aspects of the Deaf community, including use of sign language, group membership, children of Deaf adults, and so forth, will identify participants as “Deaf,” with a capital D (Padden & Humphries, 1988). Either in conjunction with the above or in contrast, other studies refer to an individual’s “hearing loss” or as being “hearing-impaired,” either in terms of decibel loss, a range of loss such as “moderate” or “profound,” whether it is bilateral or in one ear, and in some cases, in terms of communication in their environment (e.g., has trouble hearing people on the telephone). Stemming from this audiological, medical perspective, “deaf” and “hard-of-hearing” are commonly used terms to describe study participants in a range of study contexts. Yet individual choices about communication modality and identities may vary not only between individuals, but also between different contexts and settings for an individual person (Harris, Holmes, & Mertens, 2009; Stanley, Ridley, Harris, & Manthorpe, 2011). In this paper we keep the terminology as used in the original research article, unless otherwise stated.
2. The WWC guidelines for evaluating SCD research are provided to illustrate the review process used by the authors of the present article and are not intended to be instructional. For a more comprehensive description of the article review process, readers are advised to refer to the complete WWC SCD standards handbook, available from http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf
3. Based on the discretion of the principal investigators, it is sometimes acceptable for a research design to demonstrate two replications, rather than three. Reviewers must provide an explanation as to why less than three demonstrations provide sufficient evidence of an intervention effect.
4. There are no universally accepted conventions for visual analysis. The steps for visual analysis described in the present article are from the WWC Handbook for evaluating SCD research and are provided in an abbreviated format. To access the full version of the steps for visual analysis by

the WWC, please refer to the complete Handbook: http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf

- The current standards allow the principal investigator to apply exceptions to the standards, based on the nature of the intervention, outcome variable(s), and population of interest. Exceptions to the standards must be justified and specified in the review protocol.

Conflicts of Interest

No conflicts of interest were reported.

References

- Allgood, M. H., Heller, K. W., Easterbrooks, S. R., & Fredrick, L. D. (2009). Use of picture dictionaries to promote functional communication in students with deafness and intellectual disabilities. *Communication Disorders Quarterly*, 31, 53–64. doi:10.1177/1525740108327078
- Antia, S., & Kreimeyer, K. (2003). Peer interaction of deaf and hard-of-hearing children. In M. Marschark & P. Spencer (Eds.), *Oxford handbook of deaf studies, language, and education* (pp. 164–176). New York, NY: Oxford University Press.
- Baker-Shenk, C., & Kyle, J. G. (1990). Research with deaf people: Issues and conflicts. *Disability, Handicap & Society*, 5, 65–75. doi:10.1080/02674649066780051
- Barger-Anderson, R., Domaracki, J. W., Kearney-Vakulich, N., & Kubina, R. M. (2004). Multiple baseline designs: The use of a single-case experimental design in literacy research. *Reading Improvement*, 41, 217–225.
- Beal-Alvarez, J. S., Lederberg, A. R., & Easterbrooks, S. R. (2012). Grapheme-phoneme acquisition of deaf preschoolers. *Journal of Deaf Studies and Deaf Education*, 17, 39–60. doi:10.1093/deafed/enr030
- Beal-Alvarez, J., & Cannon, J. E. (2014). Technology intervention research with deaf and hard of hearing learners: Levels of evidence. *American Annals of the Deaf*, 158, 486–505. doi:10.1353/aad.2014.0002
- Bergeron, J. P., Lederberg, A. R., Easterbrooks, S. R., Miller, E. M., & Connor, C. M. (2009). Building the alphabetic principle in young children who are deaf or hard of hearing. *The Volta Review*, 109, 87–119.
- Borckardt, J. J., Nash, M. R., Murphy, M. D., Moore, M., Shaw, D., & O'Neil, P. (2008). Clinical practice as natural laboratory for psychotherapy research: A guide to case-based time-series analysis. *The American Psychologist*, 63, 77–95. doi:10.1037/0003-066X.63.2.77
- Bullis, M., & Anderson, G. (1986). Single-subject research methodology: An underutilized tool in the field of deafness. *American Annals of the Deaf*, 131, 344–348. doi:10.1353/aad.2012.0966
- Cannon, J. E., Fredrick, L. D., & Easterbrooks, S. R. (2009). Vocabulary instruction through books read in American Sign Language for English language learners with hearing loss. *Communication Disorders Quarterly*, 31, 98–112. doi:10.1177/1525740109332832
- Cohen, E. T., Allgood, M. H., Heller, K. W., & Castelle, M. (2001). Use of picture dictionaries to promote written communication by students with hearing and cognitive impairments. *Augmentative and Alternative Communication*, 17, 245–254.
- De Clerck, G. A. M. (2010). Deaf epistemologies as a critique and alternative to the practice of science: An anthropological perspective. *American Annals of the Deaf*, 154, 435–446. doi:10.1353/aad.0.0121
- Easterbrooks, S. R. (2010). Evidenced-based curricula and practices that support development of reading skills. In M. Marschark & P. E. Spencer (Eds.), *The Oxford handbook of deaf studies, language, and education* (Vol. 2, pp. 111–126). New York, NY: Oxford University Press.
- Harris, R., Holmes, H. M., & Mertens, D. M. (2009). Research ethics in sign language communities. *Sign Language Studies*, 9, 104–131. doi:10.1353/sls.0.0011
- Heller, K. W., Allgood, M. H., Ware, S., Arnold, S. E., & Castelle, M. D. (1996). Initiating requests during community-based vocational training by students with mental retardation and sensory impairments. *Research in Developmental Disabilities*, 17, 173–184. doi:10.1016/0891-4222(95)00040-2
- Horner, R. D., & Baer, D. M. (1978). Multiple-probe technique: A variation of the multiple baseline. *Journal of Applied Behavior Analysis*, 11, 189–196. doi:10.1901/jaba.1978.11-189
- Horner, R. H., Carr, E. G., Halle, J., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children*, 71, 165–179. doi:10.1177/001440290507100203
- Kaiser, A. (2014). Using single-case research designs in programs of research. In T. R. Kratochwill & J. R. Levin (Eds.), *Single-case intervention research: Methodological and statistical advances (...)*. Washington, DC: American Psychological Association.
- Kazdin, A. E. (2011). *Single case research designs: Methods for clinical applied settings* (2nd Ed.). New York, NY: Oxford University Press.
- Kelly, L. (1998). Using silent motion pictures to teach complex syntax to adult deaf readers. *Journal of Deaf Studies and Deaf Education*, 3, 217–230.
- Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2010/2014). Single-case designs technical documentation, version 2.1, 3.0. Retrieved from What Works Clearinghouse website http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf
- Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2013). Single-case intervention research design standards. *Remedial and Special Education*, 34, 26–38. doi:10.1177/0741932512452794
- Kratochwill, T. R., & Levin, J. R. (2014). *Single-case intervention research: Methodological and statistical advances*. Washington, DC: American Psychological Association.
- Long, G. (1996). *Assessing workplace communication skills with traditionally underserved persons who are deaf*. Northern Illinois University Research and Training Center on Traditionally Underserved Persons who are Deaf.
- Luckner, J. L. (2006). Evidence-based practices with students who are deaf. *Communication Disorders Quarterly*, 28, 49–52. doi:10.1177/15257401060280010801
- Luckner, J. L., Sebold, A. M., Cooney, J., Young, J. 3rd, & Muir, S. G. (2005). An examination of the evidence-based literacy research in deaf education. *American Annals of the Deaf*, 150, 443–456. doi:10.1353/aad.2006.0008
- Maggin, D. M., Briesch, A. M., & Chafouleas, S. M. (2013). An application of the What Works Clearinghouse Standards for evaluating single-subject research: Synthesis of the self-management literature base. *Remedial and Special Education*, 34, 44–58. doi:10.1177/0741932511435176
- Maggin, D. M., Chafouleas, S. M., Goddard, K. M., & Johnson, A. H. (2011). A systematic evaluation of token economies as a classroom management tool for students with challenging behavior. *Journal of School Psychology*, 49, 529–554. doi:10.1016/j.jsp.2011.05.001

- Miller, E. M., Lederberg, A. R., & Easterbrooks, S. R. (2012). Phonological awareness: Explicit instruction for young deaf and hard-of-hearing children. *Journal of Deaf Studies and Deaf Education*, 18, 206–227. doi:10.1093/deafed/ens067
- Mueller, V., & Hurtig, R. (2010). Technology-enhanced shared reading with deaf and hard-of-hearing children: The role of a fluent signing narrator. *Journal of Deaf Studies and Deaf Education*, 15, 72–101. doi:10.1093/deafed/enp023
- Murray, D. M. (1998). *Design and analysis of group-randomized trials*. New York: Oxford University Press.
- Najarian, C. G. (2008). Deaf women: Educational experiences and self-identity. *Disability & Society*, 23, 117–128. doi:10.1080/09687590701841141
- Neef, N. A., & Iwata, B. A. (1985). The development of generative lipreading skills in deaf persons using cued speech training. *Analysis and Intervention in Developmental Disabilities*, 5, 289–305. doi:10.1016/0270-4684(85)90001-1
- Odom, S. L., Brantlinger, E., Russel, G., Horner, R. H., Thompson, B., & Harris, K. R. (2005). Research in special education: Scientific methods and evidence-based practices. *Exceptional Children*, 71, 137–148.
- Padden, C., & Humphries, T. (1988). *Deaf in America: Voices from a culture*. Cambridge, MA: Harvard University Press.
- Parker, A. T., Davidson, R., & Banda, D. R. (2007). Emerging evidence from single-subject research in the field of deaf-blindness. *Journal of Visual Impairment and Blindness*, 101, 690–700.
- Richards, S. B., Taylor, R. L., & Ramasamy, R. (2014). *Single subject research: Applications in educational and clinical settings* (2nd Ed.). Belmont, CA: Wadsworth.
- Schirmer, B. R., & McGough, S. M. (2005). Teaching reading to children who are deaf: Do the conclusions of the National Reading Panel apply? *Review of Educational Research*, 75, 83–117. doi:10.3102/00346543075001083
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. New York: Houghton Mifflin Company.
- Shadish, W. R., & Sullivan, K. J. (2011). Characteristics of single-case designs used to assess intervention effects in 2008. *Behavior Research Methods*, 43, 971–980. doi:10.3758/s13428-011-0111-y
- Stanley, N., Ridley, J., Harris, J., & Manthorpe, J. (2011). Disclosing disability in the context of professional regulation: A qualitative UK study. *Disability and Society*, 26, 19–32. doi:10.1080/09687599.2011.529663
- Tankersley, M., Harjusola-Webb, S., & Landrum, T. (2013). Using single-subject research to establish the evidence base of special education. *Intervention in School and Clinic*, 44, 2. doi:10.1177/1053451208321600
- Thompson, B., Diamond, K. E., McWilliam, R., Snyder, P. & Snyder, S. (2005). Evaluating the quality of evidence from correlational research for evidence-based practice. *Exceptional Children*, 71, 181–194. doi:10.1177/001440290507100204
- Van Hasselt, V. B., Hersen, M., Egan, B. S., Mckelvey, J. L., & Sisson, L. A. (1989). Increasing social interactions in deaf-blind severely handicapped young adults. *Behavior Modification*, 13, 257–272. doi:10.1177/01454455890132007
- Wheeler-Scruggs, K. (2003). Discerning characteristics and risk factors of people who are deaf and low-functioning. *Journal of Rehabilitation*, 69, 39–46.