Single-Case Research Design in Pediatric Psychology: Considerations Regarding Data Analysis

Cohen, L. L., Feinstein, A., Masuda, A., & Vowles, K. E. (2014). Single-Case Research Design in Pediatric Psychology: Considerations Regarding Data Analysis. *Journal of Pediatric Psychology*, *39*(2), 124-137. https://doi.org/10.1093/jpepsy/jst065

Published in:
Journal of Pediatric Psychology

Document Version:
Peer reviewed version

Please cite as: Cohen, L. L., Feinstein, A., Masuda, A., & Vowles, K. E. (2014). Single-case research design in pediatric psychology: Considerations regarding data analysis. *Journal of Pediatric Psychology, 28,* 124-137.

Single-Case Research Design in Pediatric Psychology: Considerations Regarding Data Analysis

Lindsey L. Cohen

Amanda Feinstein

Akihiko Masuda

Georgia State University

Kevin E. Vowles

University of New Mexico

All correspondence should be addressed to:
Lindsey L. Cohen, Ph.D.
Department of Psychology
Georgia State University
Atlanta, GA 30302-5010
Phone: 404-413-6263
Fax: 404-413-6207
E-mail: llcohen@gsu.edu

Abstract

Objective: Single-case research allows for an examination of behavior and can demonstrate the functional relation between intervention and outcome in pediatric psychology. This review highlights key assumptions, methodological and design considerations, and options for data analysis. *Methods:* Single-case methodology and guidelines are reviewed with an in-depth focus on visual and statistical analyses. *Results:* Guidelines allow for the careful evaluation of design quality and visual analysis. A number of statistical techniques have been introduced to supplement visual analysis, but to date there is no consensus on their recommended use in single-case research design. *Conclusions:* Single-case methodology is invaluable for advancing pediatric psychology science and practice, and guidelines have been introduced to enhance the consistency, validity, and reliability of these studies. Experts generally agree that visual inspection is the optimal method of analysis in single-case design; however, statistical approaches are becoming increasingly evaluated and employed to augment data interpretation.

Keywords: research design and methods, single-case designs, statistical applications

Single-Case Research Design in Pediatric Psychology: Considerations Regarding Data Analysis

The definition and mission of the Society of Pediatric Psychology (APA Division 54: Society of Pediatric Psychology, 2013), which is consistent with the vision of the American Psychological Association (APA, 2013), is to apply the scientific study of human behavior to improve the lives of the individual patient. Paralleling this focus on the individual, the Evidence-based Medicine Working Group of the American Medical Association places single-case research at the apex of the hierarchy of evidence that should guide evidence-based prevention and treatment practices (Guyatt et al., 2008). Thus, it is ironic if not a “double-standard” (Valsiner, 1986, p. 1) that the psychology research literature is dominated by group-aggregate data, which provides the predominate evidence-base to inform our work with individual patients (American Psychological Association Presidential Task Force on Evidence-Based Practice, 2006; Sackett, Rosenberg, Gray, Haynes, & Richardson, 1996). This is especially poignant given that the history of psychology is punctuated by landmark findings largely grounded in case studies or single-case methodology. Consider Fechner, Broca, Boring, Wundt, Pavlov, Ebbinghaus, Freud, Watson, Piaget, and Skinner who largely conducted their science via the careful examination of individuals.

Some reasons for this disconnect might be tied to misconceptions about the internal and external validity of single-case research design (also known as “single-subject design”, “N=1 research”, “time-series designs”, and “intrasubject replication designs”). Regarding internal validity, it is important to distinguish case studies from single-case research; these 2 lines of inquiry are oft confused, even in methods textbooks (Dermer & Hoch, 1999). Case studies typically share the following qualities (Kazdin, 2011; Yin, 2012): (a) An in-depth description of a single unit, which might be an individual person, a group of people, an institution, a country, or another unitary "case." (b) The data are typically descriptive and detailed rather than quantitative. (c) The unique context and characteristics of the case are a primary focus of the study. (d) The study is focused on qualitatively describing the relations between the current state and prior historical information; conversely, systematic or objective assessment, prospective evaluations, or observations following intervention is rare. Case studies allow the study of unique or rare cases or situations, contribute to theory or therapy development by fueling inventive or innovative research questions, provide vivid examples to highlight arguments, and have heuristic value to the research enterprise. Thus, they might be particularly useful in pediatric psychology to highlight unique medical conditions with psychological factors, describe novel psychosocial treatments, or inform hypotheses for subsequent systematic research (Drotar, 2011).

Similar to case studies, single-case research focuses the analysis on the individual unit of interest (patient, family, school, city, country, etc.), although the similarities largely end there. Rather than primarily aiming to richly describe a case, the goal of single-case research is to demonstrate functional relations among variables of interest – most frequently that is to determine whether a causal relation exists between a researcher-controlled independent variable and participant behavior. Single-case research emphasizes the repeated measurement within an individual across time, and ideally across different conditions. Thus, single-case research typically has an applied focus with rigorous scientific standards and high internal validity (e.g., Sidman, 1960, Kratochwill et al., 2013).

Single-case research should include the following elements (Kazdin, 2011; Kratochwill et al., 2010; Rapoff & Stark, 2008): (a) The focus of analysis is on the individual "case", (b) The case serves as its own control in order to provide a comparison. (c) There is a baseline assessment phase prior to intervention. (d) Data on the outcome variable are collected continuously and repeatedly within and across different levels (phases) of the independent variable. This allows an examination and comparison of the patterns and stability of the data before and during an intervention. Ideally, subsequent phases are added based on stability of data; however, this is not always the process. (e) In order to adequately describe and predict the behavior, it is recommended that the data are stable (i.e., lack of trend and excessive variability) within a given phase. If a trend or slope is present, it might be permissible if it is in the opposite direction of what will be predicted for the pattern of a subsequent phase. (f) In order to provide additional evidence of internal validity as well as external validity, results are replicated across cases, conditions, settings, or other variables.

There are several single-case designs that meet the above criteria, such as reversal (ABAB), multiple-baseline, and changing-criterion. Specifics of these and other single-case designs are readily available in the literature and will not be covered here (e.g., Barlow, Nock, & Hersen, 2009; Hayes, Barlow, & Nelson-Gray, 1999; Kazdin, 2011; Perone & Hursh, 2012). General quality indicators have been developed and a number of appraisal tools – akin to the CONSORT guidelines for randomized controlled trials (e.g., Stinson, McGrath, & Yamada, 2003) – are available to guide in the development, and to allow for evaluation of single-case designs (for a review of 7 different appraisal tools, see Wendt & Miller, 2012). Horner (2005) developed a list of considerations for single-case studies, which we adapted for pediatric psychology research (Table 1). If a single-case research design study adheres to these specified standards, then internal validity will be maximized. Despite the emphasis on establishing and demonstrating tight experimental control of the variables in the study, there is flexibility in single-case research. For example, it is permissible to add additional interventions or change the treatment via adding new phases during the study and monitor changes on the outcome variable. This process is not dissimilar to how actual psychotherapy is conducted; albeit with tighter control and assessment when part of a single-case design experiment. Thus, single-case designs are tailor-made for applied evidence-based pediatric psychologists.

The question of external validity in single-case research is more complicated, in part because this is not merely a methodological issue, but a conceptual and pragmatic one (Hayes, 1991). One important question is the goal of the external validity. If a researcher aims to generalize the results of a study to a large population (e.g., all adolescents with chronic pain), the aggregate score of a group of individuals (e.g., subgroup of adolescents with chronic pain) representative of that population would be more relevant than the responses of an individual. However, when attempting to apply the results of research to a specific individual, the optimal design is less clear. In other words, how do researchers go about answering Gordon Paul's (1967, p. 111) iconic question, "What treatment, by whom, is most effective for this individual with that specific problem, and under which set of circumstances?" In this vein, the results of replicated single-case studies of similar individuals in similar situations should be more useful than the average response of a group of individuals (Sidman, 1960). Thus, single-case research results might be especially relevant to a practitioner taking an evidence-based practice perspective.

Single-case designs are particularly useful in pediatric psychology when studying rare conditions and large samples are difficult to obtain; results also have direct applicability to health care professionals. In addition, the sharing of ideology between single-case designs and evidence- based practice is especially relevant to pediatric psychologists working with medical professionals familiar with the evidence-based medicine framework. Despite this, single-case research is rare in pediatric psychology. A search of the articles in the *Journal of Pediatric*

Psychology on the journal website between 2000-2013 using the keywords ‘single-case’ or ‘single-subject’ revealed only 7 studies meeting the definition of single-case design as described above: Applegate, Kelley, Applegate, Jayasinghe, & Venters (2003); Bernard, Cohen, & Moffett (2009); Burke, Kuhn, & Peterson (2004); Cushing, Jensen, & Steele (2011); Hains, Davies, Parton, & Silverman (2001); Powers et al. (2006); Sil, Dahlquist, & Burns (2013); Spaulding, Devine, Duncan, Wilson, & Hogan (2011). We hope that this paper helps to stimulate research in pediatric psychology using single-case methodology.

Analyses in Single-Case Research

In single-case research, evaluation of results primarily focuses on whether the change in the outcome variable is caused by the experimenter-controlled independent variable (intervention) and is reliable and not due to chance. In addition to this evaluation of *experimental* analysis, results should be evaluated on an *applied* criterion; examination of the importance and meaningfulness of the changes (Risely, 1970). This experimental and applied distinction parallels that of statistical and clinical significance, which is a valued distinction to pediatric psychologists (Drotar, 2010; Rapoff, 2010). Both foci of evaluation will be highlighted in the subsequent sections.

The analysis of single-case research is a long-standing controversial topic with passionate arguments about whether visual analysis, statistical techniques, or a combination should be used when examining the results of single-case research. A summary of single-case design standards was recently published (Smith, 2012), and, consistent with the historical state of the field, there are mixed opinions regarding the optimal approach to analyzing results of single-case research (Table 2). That said, at this point in time, recommendations emphasize that visual analysis be the central method of data evaluation (e.g., Barlow et al., 2009; Gast, 2010; Kazdin, 2011;

Kratochwill at al., 2010, 2013) and statistical analysis be used to augment – not replace – visual analysis.

Visual analysis. As the name implies, visual analysis or visual inspection refers to determining the outcome of a single-case study via viewing the raw data, typically in graphical format. At first blush, the notion of relying on visual inspection might appear overly subjective; however, it is important to consider the rationale as well as criteria that have been developed for visual analysis. Further, Wilkinson and the Task Force on Statistical Inference (1999) argue that visual inspection of graphs should be the first step regardless of design and before proceeding to any statistical analyses as it allows better understanding of the distribution of data, provides an opportunity for identification of potential influence of outliers, and can inform interpretation.

From a theoretical perspective, single-case researchers are encouraged to conduct studies with sufficiently stringent control and potent interventions that produces results that are obvious to the naked eye (Sidman, 1960). Inherent in this perspective is that there is a relatively lower likelihood of Type I than Type II error in single-case research (Baer, 1977), and that results would have both experimental and applied significance. Thus, the fact that visual inspection does not sufficiently detect small effects – which might not cross the applied/clinical significance threshold – is consistent with the rationale of the approach and might be seen as an advantage of this analytic method.

Criteria for visual inspection have been developed (e.g., Hayes, Barlow, & Nelson-Gray, 1999; Kazdin, 2011; Table 3), which include an evaluation of changes in 6 domains: means, levels, trends, variability, latency, and consistency. (a) Evaluation of the means across phases refers to inspecting changes in the average rate of responding on the dependent variable. (b) Separate from any mean changes, examining changes in the level refers to any discontinuity in the dependent variable from the end of one phase to the beginning of the next phase. Any changes evident in level might suggest an immediate response to a treatment being introduced, such as when a child's thumb-sucking behavior initially and immediately stops when a distasteful substance is applied to the fingernail. (c) The trend in data is the slope within a phase. For example, upon introducing an intervention, the dependent variable might systematically increase or decrease. For example, medical adherence behavior might gradually improve over time after introducing a sticker-chart reward system. (d) It is possible for the mean, level, and trend to stay unchanged but the variability in the dependent variable to indicate response upon the introduction of an intervention. For example, after starting a sleep hygiene intervention, an adolescent's erratic sleep pattern (4 hours one night, 12 hours the next) might change to a consolidated pattern of 8 hours each night. A unique but related procedure for analyzing variability in healthcare at the single unit of analysis is via statistical process control; this framework parallels the single-case design perspective but is particularly focused on variability in health behavior and introduces additional concepts and procedures (e.g., Bowen & Neuhauser, 2013; Diaz & Neuhauser, 2005; Neuhauser, Provost, & Bergman, 2011; Tennant, Mohammed, Colman, & Martin, 2007; Thor et al., 2007). (e) Latency of change following the introduction of a phase (e.g., intervention) refers the immediacy of the change in the data. For example, changes in behavior following slow-acting medications might be predicted to be gradual, but the introduction of a potent punishment (e.g., electric shock) might result in an immediate change in behavior. (f) Lastly, data can be inspected to determine if there is consistency in the pattern of data across similar phases. For example, it might be expected that similar trends and rates of behavior will be found each time the intervention is introduced and that data is demonstrates similar patterns during baseline and subsequent return to baseline phases. Studies suggest that interrater agreement for visual analysis can be high when researchers are trained to criteria (e.g., Kahng et al., 2010); however, some data do not support this assertion (e.g., Ottenbacher, 1993).

Statistical Analysis. As noted, the general consensus in the field gives precedence to visual analysis. Even with this emphasis, a number of statistical approaches have been developed to augment visual analysis. Some caveats should be highlighted before proceeding. There are many statistical approaches – see books on this topic by Dugard, File, & Todman, 2012; Edgington & Onhena, 2007; Franklin, Allison, & Gorman, 1997; and Satake, Jagaroo, & Maxwell 2008 – for single-case studies but there is no consensus on best practices regarding statistical methodology for analyzing single-case studies (Table 2). Philosophically, researchers argue that statistics might reveal statistically significant but clinically insignificant findings (Baer, 1977), and that statistics fail to consider the multiple facets (e.g., changes across different aspects of the study [mean, level, trend, etc.], patient or setting characteristics, unique changes in the course of the study) in idiographic single-case research. The issue is complicated by the fact that statistics in single-case studies can be used for a number of purposes. For example, statistics might examine variability within a phase (e.g., statistical process control techniques), changes across phases (time-series analysis), or the slope or trend of the data (e.g., split-middle technique). Widely different results can be found when different statistical tests are applied to the same data (e.g., Campbell, 2004; Manolov et al., 2011; Nourbakhsh & Ottenbacher, 1994) and when the same statistical technique is used with different assumptions or different aspects (e.g., mean, trend) of the same data (e.g., Manolov & Solanas, 2009). Further, there are no agreed-upon metrics for judging statistical results (e.g., Parker & Hagan-Burke, 2007). Some aspects of single-case studies violate assumptions of a number of statistical tests. For example, serial dependence and autocorrelation are common given that observations are temporally ordered and not independent (*note:* some statistical approaches control for autocorrelation and serial dependence, and autocorrelation and serial dependence also influence visual analysis). Lastly, general criticisms regarding statistics in psychology writ large can be levied against statistical techniques for single-case design (e.g., Perone, 1999). In general, although counter-intuitive, currently there are fewer guidelines and standards, greater subjectivity (in selection and use), and more variable outcomes possible with statistical than visual analysis of single-case studies (e.g., Parker et al., 2005).

Although we have listed significant concerns regarding the use of statistics for analyzing single-case studies, it is a growing area of study and debate for a number of reasons. First, research with humans in real-world settings introduces extraneous variables that challenge the ability to obtain stable data. In addition, interventions with human participants might be of weaker magnitude leading to less visually apparent differences between phases. Consider the likelihood of detecting intervention effects in the behavior of a rat in a highly controlled laboratory environment to changes in behavior of a child during a medical appointment. Thus, small but important effects might be missed with visual inspection but identified via statistical methods. Second, there can be error in determining stability and changes in data; researchers can have diverging perspectives in drawing conclusions based on visual inspection (e.g., DeProspero & Cohen, 1979; Normand & Bailey, 2006; Ottenbacher, 1993) and serial dependence can influence outcomes (Matyas & Greenwood, 1990). Third, there are practical reasons (e.g., federal grant reviewers might be more familiar with quantitative analyses) to incorporate statistical techniques into single-case design research (e.g., Crosbie, 1999; Huitema, 1986). Lastly, data suggest that quantitative methods can improve agreement among raters using visual analysis (e.g., Fisch, 2001; Hojem & Ottenbacher, 1988). Rather than provide only cursory information on a sampling of tests, we have listed a number of different statistical approaches in Table 4 with references to studies in which they have been applied.

The application of statistics to calculate an effect size for single-case research deserves additional discussion. Given the emphasis on summarizing results across studies (e.g., meta-analyses, systematic reviews) to inform evidence-based practice (e.g., Sackett et al., 2006), attention has focused on an appropriate metric of single-case treatment effect (i.e., effect size). To date, no consensus has been reached regarding the optimal technique for this task (e.g., Horner, Swaminathan, Sugai, & Smolkowski, 2012; Lane & Carter, 2013; Maggin & Chafouleas, 2013b; Manolov, Solanas, Sierra, & Evans, 2011; Schlosser, Lee, & Wench, 2008; Shadish, Rindskopf, & Hedges, 2008), although many have been proposed (e.g., a review in 2008 identified 40 indices of effect size for single-case studies; Swaminthan et al., 2008). Part of the challenge comes from the difficulty in applying a nomothetic paradigm (e.g., meta-analysis) to an idiographic one (i.e., single-case research design). Several nonparametric methods for estimating single-case effect size have been proposed with the percentage of nonoverlapping data being one of the first to be introduced (Campbell, 2013; Scruggs & Mastropieri, 2013; Scruggs, Mastropieri, & Casto, 1987); although there are improvements in this area (see Campbell, 2013; Parker, Vannest, & Davis, 2011). Nonparametric methods do not adequately detect trend. Numerous parametric tests have been evaluated too, and at least one panel recommends regression-based estimates (Kratochwill et al., 2010). An in-depth discussion of computing effect sizes for single-case research is beyond the scope of this paper. See a recent special journal issue edited by Maggin and Chafouleas (2013a) for more information on this topic.

Example Single-Case Analysis. In order to illustrate how statistical methods might augment visual analysis, we selected one approach to highlight – the Conservative Dual Criteria method (CDC; Fisher, Kelley, & Lomas, 2003). This technique was selected for several reasons. First, pediatric psychology single-case research might target behavior (e.g., pill swallowing, seizure activity) that can be difficult to monitor over a long period of time. Thus, it might be challenging to amass the necessary data points (e.g., 50 or more) for some techniques (e.g., interrupted time-series analysis; Box et al., 1994). Some authors (e.g., Kazdin, 2011) have recommended interrupted time-series analyses when the data permit. Second, single-case data often violate the assumptions of normality, which is necessary for conventional *t* and *F* tests. Third, in pediatric psychology research it might be impractical or unethical to randomly pre-assign the treatment (e.g., number of sessions), which is required for randomization tests (e.g., Dugard et al., 2012). In fact, it has been recommended that study phases change based on stability of data in real-time (Rapoff & Stark, 2008). Third, the CDC technique shares qualities with other approaches, and thus might facilitate learning of additional techniques. Fourth, this approach has a commonsense and accessible quality, which allows for comprehension by a wider audience as well as critical appraisal. We are not arguing that the CDC method is superior to any other statistical technique; we simply selected it to illustrate how a researcher might incorporate techniques to compliment visual analysis in single-case research.

The CDC method provides assistance in improving interrater agreement in visually detecting changes in level and trend within and across phases. Fisher et al. (2003) developed the CDC method based on refinements of the split-middle technique (White & Haring, 1980) and percentage of nonoverlaping data method (Scruggs et al., 1987). The CDC method involves using the data from one phase (e.g., A) to compute trend (least squares linear regression using slope and intercept) and level (mean) lines, plus 0.25 standard deviations further in the direction of the predicted treatment effect. These lines are then superimposed on the subsequent phase (e.g., B) (*note:* an MS Excel spreadsheet pre-programmed to compute these formulas and produce the criterion lines can be obtained from Swoboda, Kratochwill, & Levin, 2010). Based on binomial probability and typical probabilities associated with hypothesis testing, Fisher et al. (2003) developed a table (p. 399) to determine how many data points above (or below) both criterion lines are necessary in order to determine that a significant change has occurred from one to the subsequent phase. Investigations using the CDC method suggests that it assists visual analysis and balances Type I and II error (Fisher et al., 2003; Steward, Carr, Brandt, & McHenry, 2007).

The CDC method can be applied to ABAB or multiple-baseline designs (Swobodha et al., 2010). Our example case will use an ABAB design to evaluate coping skills for an adolescent with pediatric abdominal pain. In this hypothetical experiment, pain is rated on a 0-10 scale during a pain stimulus task at baseline and intervention (i.e., coping skills). Based solely on visual analysis, determining results of our invented data might be challenging (Figure 1). The CDC method provides criterion lines (Figure 2) and decision rules. Specifically, according to Fisher et al. (2003), given that there are 8 data points in the treatment phase (B), it is necessary to have 7 of these points fall below both criterion lines to conclude that there is a reliable treatment effect. In our example, 7 pain scores in the experimental (B) phase fell below the trend and mean lines. Thus, it can be concluded that the coping skills resulted in lower pain ratings than during baseline. Similarly, using the CDC procedure, the return to baseline phase effectively resulted in a significant increase in pain scores from the first treatment phase (i.e., 7 data points fell above the mean line and 8 data points above the trend line) and the second treatment phase again lead to significant reductions in pain scores (i.e., all points fell below both criterion lines). In addition to providing these decision rules and confidence in determining differences between phases, the

CDC method at least highlights if not reveals the trend present within the phases of this example study. Changes within (trend) and across (mean) phases might be particularly relevant depending on the clinical area of study.

Conclusions

Single-case research relies on a rigorous methodology that can produce results optimally relevant to evidence-based practice (e.g., American Psychological Association, 2002; Guyatt et al., 2008) and with great potential for bridging the scientist-practitioner gap (Drotar, 2010; Morgan & Morgan, 2001). Despite the lineage of essential findings from single-case research that provide the foundation for applied psychology, there is also a paradoxical historical inclination in the field toward group design methodology and quantitative statistics. The evidence-based movement coupled with growing skepticism about the clinical applicability of results from large group design studies (e.g., RCTs; Jacobson & Christensen, 1996; Westen & Bradley, 2005; Westen, Novotny, & Thompson-Brenner, 2004) has reinvigorated interest in single-case design research.

Fortunately, there are specified criteria in place to guide pediatric psychologists in the design and conduct of rigorous single-case studies (e.g., Horner et al., 2005; Wendt & Miller, 2012; Table 1). Visual analyses have a long history and have reasonably well-defined standards for judgment, which are reliable when clear effects are present. Visual analysis also allows for the incorporation of unique patient and setting characteristics as well as nuances of the study, which is vital in idiographic research. In contrast, there is heated debate and generally a lack of consensus regarding which – if any – statistical analyses to consider and how to include these in analyses of single-case studies. There are some circumstances in which statistics might be appropriate: (a) When baseline data is unstable or trending in the direction of the intervention, statistics might reveal effects difficult to decipher with visual analysis. (b) When the treatment effect is small, but important, statistics might be useful. For example, a new intervention might only produce weak effects, which might be improved upon in subsequent research. (c) Statistics might be advisable when there is considerable variability in behavior, possibly due to environmental influences, which are difficult or impossible to control (e.g., busy medical clinic). (d) Statistics are recommended when computing effect sizes to aggregate findings across studies. In this paper, we have reviewed the methodology of analyzing single-case studies and highlighted one of many of the statistical approaches to compliment visual analysis. We have provided some guidelines and recommendations for designing and implementing rigorous single-case studies, how to systematically conduct visual analysis, and when statistics might be considered. Regarding statistical approaches, the current state of the field places the responsibility on the researcher or clinician to determine if statistics should be used and which statistic might be most appropriate.

We recognize that we might have raised more questions than answers in this paper; a practice that is not uncommon in research. Although unsettling, we do not believe that there are many easy answers or definitive directions in our scientific enterprise, which is grounded in philosophical skepticism. That said, some designs and analytic approaches are more practical, ethical, or adept than others at answering some empirical questions. As sagely put by Sackett and Wennberg (1997), “It’s time to stop squabbling over the ‘best’ methods” (p. 1636); the empirical question should dictate the most appropriate method and analysis. In pediatric psychology, many of our research and clinical questions focus on individual patients with rare conditions or situations, and we often use novel or adapted treatment approaches in unique settings. Thus, single-case design is particularly suitable to the applied pediatric psychologist. Fortunately, there are established criteria to guide the design and conduct of rigorous single-case studies. Although there is no consensus on recommended statistical tests for single-case studies, there are a number of techniques available to augment visual analysis. In closing, single-case designs are particularly useful and practical in evidence-based pediatric psychology work, and they are optimally suited to inform and quantify our practice with individual pediatric patients.

References

American Psychological Association. (2013). About APA. Retrieved from http://www.apa.org/about/index.aspx

American Psychological Association. (2002). Criteria for evaluating treatment guidelines. *American Psychologist, 57*, 1052-1059.

American Psychological Association Presidential Task Force on Evidence-Based Practice (2006). Evidence-Based Practice in Psychology. *American Psychologist, 61,* 271-285.

APA Division 54: Society of Pediatric Psychology. (2013). Who we are. Retrieved from http://www.apadivisions.org/division-54/about/index.aspx

Applegate, H., Kelley, M. L., Applegate, B. W., Jayasinghe, I. K., & Venters, C. L. (2003). Pediatric residents' discussions of and interventions for children's behavioral and emotional problems. *Journal of Pediatric Psychology, 28,* 315-321.

Baer, D. M. (1977). Perhaps it would be better not to know everything. *Journal of Applied Behavior Analysis, 10,* 162-172.

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.

Barlow, D. H., Nock, M. K., & Hersen, M. (2009). *Single case experimental designs: Strategies for studying behavior change, 3rd Edition.* Boston: Pearson Education, Inc.

Bernard, R. S., Cohen, L. L., & Moffett, K. (2009). A token economy for exercise adherence in pediatric cystic fibrosis: A single-subject analysis. *Journal of Pediatric Psychology, 34,* 354-365.

Brossart, D. F., Meythaler, J. M., Parker, R. I., McNamara, J., & Elliott, T. R. (2008). Advanced regression methods for single-case designs: Studying propranolol in the treatment for agitation associated with traumatic brain injury. *Journal of Rehabilitation Psychology, 53,* 357-369.

Burke, R. V., Kuhn, B. R., & Peterson, J. L. (2004). A "storybook" ending to children's bedtime problems: The use of a rewarding social story to reduce bedtime resistance and frequent night waking *Journal of Pediatric Psychology, 29,* 389-396.

Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (1994). Time series analysis: Forecasting and control (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.

Bowen, M. E., & Neuhauser, D. (2013). Understanding and managing variation: Three different perspectives. *Implementation Science, 8*, s1-s2.

Campbell, J. M. (2004). Statistical comparison of four effect sizes for single-subject designs. *Behavior Modification, 28*, 234-246.

Campbell, J. M. (2013). Commentary on PND at 25. *Remedial and Special Education, 34*, 20-25.

Crosbie, J. (1993). Interrupted time-series analysis with brief single-subject data. *Journal of Consulting and Clinical Psychology, 61*, 966-974.

Crosbie, J. (1999). Statistical inference in behavior analysis: Useful friend. *The Behavior Analyst, 22,* 105-108.

Cushing, C. C., Jensen, C. D., & Steele, R. G. (2011). An evaluation of a personal electronic device to enhance self-monitoring adherence in a pediatric weight management program using a multiple baseline design. J*ournal of Pediatric Psychology, 36*, 301–307.

DeProspero, A., & Cohen, S. (1979). Inconsistent visual analysis of intrasubject data. *Journal of Applied Behavior Analysis, 12,* 573-579.

Dermer, M. L., & Hoch, T. A. (1999). Improving descriptions of the single-subject experiments in research texts written for undergraduates. *The Psychological Record, 49,* 49-66.

Diaz, M., & Neuhauser, D. (2005). Pasteur and parachutes: When statistical process control is better than a randomized controlled trial. *Quality & Safety in Health Care, 14,* 140-143.

Drotar, D. (2010). Editorial: Integrating research and practice: The role of the Journal of Pediatric Psychology. *Journal of Pediatric Psychology, 35,* 111-113.

Drotar, D. (2011). Editorial: Guidance for submitting and reviewing case reports and series in the Journal of Pediatric Psychology. *Journal of Pediatric Psychology, 36,* 951-958.

Dugard, P., File, P., & Todman, J. (2012). *Single-case and small-n designs: A practical guide to randomization tests, 2nd Edition.* New York: Taylor & Francis Group, LLC.

Edgington, E. S. (1982). Nonparametric tests for single-subject case designs. In T. R. Kratochwill, & J. R. Lewin (Eds.), *Single-case research design and analysis: New directions for psychology and education* (pp. 133-157). Hillsdal, NJ: Lawrence Erlbaum Associates.

Fisch, G. S. (1998). Visual inspection of data revisited: Do the eyes still have it? *The Behavior Analyst, 21*, 111-123.

Fisch, G. S. (2001). Evaluating data from behavior analysis: Visual inspection or statistical models. *Behavioural Processes, 54*, 137-154.

Fisher, W. W., Kelley, M. E., & Lomas, J. E. (2003). Visual aids and structured criteria for improving visual inspection and interpretation of single-case designs. *Journal of Applied Behavior Analysis, 36,* 387-406.

Franklin, R. D., Gorman, B. S., Beasley, T. M., & Allison, D. B. (1997). Graphical display and visual analysis. In R. D. Franklin, D. B. Allison, & B. S Gorman (Eds.), *Design and analysis of single-case research* (pp. 119–158). Mahwah, NJ: Erlbaum.

Gast, D. L. (2010). *Single subject research methodology in behavioral sciences.* New York:

Routledge.

Gottman, J. M., & Leiblum, S. R. (1974). *How to do psychotherapy and how to evaluate it.* New York, NY: Holt, Rinehan and Winston.

Guyatt, G., Haynes, B., Jaeschke, R., Meade, M. O., Wilson, M., Montori, V., & Richardson, S. (2008). The philosophy of evidence-based medicine. In G. Guyatt, D. Rennie, M. O. Meade, & D. J. Cook (Eds.). *User's guides to the medical literature: A manual for evidence-based clinical practice, 2nd Edition* (pp. 9-16). New York: McGraw Hill.

Hayes, S. C. (1991). The limits of technological talk. *Journal of Applied Behavior Analysis, 24*, 417-420.

Hayes, S. C., Barlow, D. H., & Nelson-Gray, R. O. (1999). *The scientist practitioner: Research and accountability in the age of managed care, 2nd Edition.* Needham Heights, MA: Allyn & Bacon.

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*, 165-179.

Horner, R. H., Swaminathan, H., Sugai, G., & Smolkowski, K. (2012). Considerations for the systematic analysis and use of single-case research. *Education & Treatment of Children, 35*, 269-290.

Hojen, M. A., & Ottenbacher, K. J. (1988). Empirical investigation of visual-inspection versus trend-line analysis of single-subject data. *Physical Therapy, 68*, 983-988.

Huitema, B. E. (1986). Statistical analysis and single subject designs: Some misunderstandings. In A. S. Bellack & M. Hersen (Series Ed.) & A. Poling & R. W. Fugua (Vol. Eds), *Applied clinical psychology: Research methods in applied behavior analysis: Issues and* *advances* (pp. 209-232). New York: Plenum.

Jacobson, N. S., & Christensen, A. (1996). Studying the effectiveness of psychotherapy: How well can clinical trials do the job? *American Psychologist, 51,* 1031–1039.

Jones, W. P. (2003). Single-case time series with Bayesian analysis: A practitioner's guide. *Measurement and Education in Counseling and Development, 36,* 28-39.

Kahng, S., Chung, K, Gutshall, K., Pitts, S. C., Kao, J., & Girolami, K. (2010). Consistent visual analyses of intrasubject data. *Journal of Applied Behavior Analysis, 43,* 35-45.

Kazdin, A. E. (2011). *Single-case research designs: Methods for clinical and applied settings, 2nd Edition.* New York: Oxford University Press.

Kratochwill, T. R., Hitchcock, J. H., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2010). Single-case deigns technical documentation. Retrieved from What Works Clearinghouse website: http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf

Kratochwill, T. R., Hitchcock, J. H., 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.

Lane, K. L., & Carter, E. W. (2013). Reflection on the special issue: Issues and advances in the meta-analysis of single-case research. *Remedial and Special Education, 34,* 59-61.

Maggin, D. M., & Chafouleas, S. M. (Eds.). (2013a). Issues and advances of synthesizing single-case research [Special issue]. *Remedial and Special Education, 34*.

Maggin, D. M., & Chafouleas, S. M. (2013b). Introduction to the special series: Issues and advances of synthesizing single-case research. *Remedial and Special Education, 34,* 3-8.

Manolov, R., Arnau, J., Solanas, A., & Bono, R. (2010). Regression-based techniques for statistical decision making in single-case designs. *Psicothema*, 22, 1026-1032.

Manolov, R., & Solanas, A. (2009). Problems of the randomization test for AB designs. *Psicologica, 30*, 137-154.

Manolov, R., Solanas, A., Sierra, V., & Evans, J. J. (2011). Choosing among techniques for quantifying single-case intervention effectiveness. *Behavior Therapy, 42,* 533-545.

Matyas, T. A., & Greenwood, K. M. (1990). Visual analysis of single-case time series: Effects of variability, serial dependence, and magnitude of intervention effects. *Journal of Applied Behavior Analysis, 23,* 341–351.

Maughan, D. R., Christiansen, E., Jenson, W. R., Olympia, D., & Clark, E. (2005). Behavioral parent training as a treatment for externalizing disorders and disruptive behavior disorders: A meta-analysis. *School Psychology Review, 34,* 267-286.

McCall, R. B., & Appelbaum, M. I. (1973). Bias in the analysis of repeated-measures designs: Some alternative approaches. *Child Development, 44,* 401-415.

McKnight, S., McKean, J. W., & Huitema, B. E. (2000). A double boot strapping method to analyze linear models with autoregressive error terms. *Psychological Methods, 5,* 87-101.

Morgan, D. L., & Morgan, R. K. (2001). Single-participant research design: Bringing science to managed care. *American Psychologist, 56,* 119–127.

Nelson, T. D., Aylward, B. S., & Rausch, R. R. (2011). Dynamic P-technique for modeling patterns of data: Applications to pediatric psychology Research. *Journal of Pediatric Psychology, 36,* 959-968.

Neuhauser, D., Provost, L., & Bergman, B. (2011). The meaning of variation in healthcare managers, clinical and health-services researchers, and individual patients. *Quality & Safety, 20,* i36-i40.

Normand, M. P., & Bailey, J. S. (2006). The effects of celebration lines on visual data analysis. *Behavior Modification, 30,* 295-314.

Nourbakhash, M. R., & Ottenbacher, K. J. (1994). The statistical analysis of single-subject data: A comparative examination. *Physical Therapy, 74*, 768-776.

Ottenbacher, K. J. (1993). Interrater agreement of visual analysis in single-subject decisions: Quantitative review and analysis. *American Journal of Mental Retardation, 98,* 135-142. Effect Sizes in Single Case Research: How Large is Large?

Parker, R. I., Brossart, D. F., Vannest, K. J., Long, J. R., De-Alba, R. G., Baugh, F. G., Sullivan, J. R. (2005). Effect sizes in single case research: How large is large? *School Psychology Review, 34,* 116-132.

Parker. R. L., & Hagan-Burke, S. (2007). Useful effect size interpretations for single case research. *Behavior Therapy, 38,* 95-105.

Parker, R. L., Vannest, K. J., & Davis, J. L. (2011). Effect size in single-case research: A review of nine nonoverlap techniques. *Behavior Modification, 35*, 303-322.

Parsonson, B. S., & Baer, D. M. (1978). The analysis and presentation of graphic data. In T. R. Kratochwill (Ed.), *Single subject research* (pp. 101–166). New York, NY: Academic Press.

Paul, G. L. (1967). Strategy of outcome research in psychotherapy. *Journal of Consulting Psychology, 31*, 109-118.

Perone, M. (1999). Statistical inference in behavior analysis: Experimental control is better. *The Behavior Analyst, 22,* 109-116.

Perone, M., & Hursh, D. E. (2012). Single-case experimental designs. In G. J. Madden (Ed.), *APA handbook of behavior analysis.* Washington, DC: American Psychological

Association.

Powers, S. W., Piazza-Waggoner, C., Jones, J. S., Ferguson, K., Daines, C., & Acton, J. D. (2006). Examining clinical trial results with single-subject analysis: An example involving behavioral and nutrition treatment for young children with cystic fibrosis. *Journal of Pediatric Psychology, 31*, 574-581.

Rapoff, M. A. (2010). Editorial: Assessing and enhancing clinical significance/social validity of intervention research in pediatric psychology. *Journal of Pediatric Psychology, 35,* 114-119.

Rapoff, M., & Stark, L. (2008). Editorial: Journal of Pediatric Psychology statement of purpose: Section on single-subject studies. *Journal of Pediatric Psychology, 33,* 16-21.

Revusky, S. H. (1967). Some statistical treatments compatible with individual organism methodology. *Journal of the Experimental Analysis of Behavior, 10,* 319-330.

Risley, T. R. (1970). Behavior modification: An experimental therapeutic endeavor. In L. A. Hamerlynch, P. O. Davidson, & L. E. Acker (Eds.), *Behavior modification and ideal mental health services.* Calgary, Alberta: University of Alberta Press.

Sackett, D. L., Rosenberg, W. M. C., Gray, J. A. M., Haynes, R. B., & Richardson, W. S. (1996). Evidence based medicine: What it is and what it isn't. *British Medical Journal, 312,* 71-72.

Sackett, D. L., & Wennberg, J. E. (1997). Choosing the best research design for each question: It's time to stop squabbling over the "best" methods. *British Medical Journal, 315,* 1636.

Satake, E., Jagaroo, V. & Maxwell, D. L. (2008). *Handbook of statistical methods: Single subject design.* San Diego, CA: Plural Publishing.

Schlosser, R. W., Lee, D. L., & Wendt, O. (2008). Application of the percentage of nonoverlapping data (PND) in systematic reviews and meta-analyses: A systematic review of reporting characteristics. *Evidence-based Communication Assessment and Intervention, 2,* 163-187.

Scotti, J. R., Evans, I. M., Meyer, L. H., & Walker, P. (1991). A meta-analysis of intervention research with problem behavior: Treatment validity and standards of practice. *American Journal on Mental Retardation, 96,* 233-256.

Scruggs, T. E., & Mastropieri, M. A. (2013). PND at 25: Past, present, and future trends in summarizing single-subject research. *Remedial and Special Education*, *34,* 9-19.

Scruggs, T. E., Mastropieri, M. A., & Casto, G. (1987). The quantitative synthesis of single-subject research: Methodology and validation. *Remedial and Special Education, 8,* 24-33.

Shadish, W. R., Rindskopf, D. M., & Hedges, L. V. (2008). The state of the science in the meta-analysis of single-case experimental designs. *Evidence-based Communication Assessment and Intervention, 3,* 188-196.

Sidman, M. (1960). *Tactics of scientific research.* New York: Basic Books.

Sil, S., Dahlquist, L. M., & Burns, A. J. (2013). Videogame distraction reduces behavioral distress in a preschool-aged child undergoing repeated burn dressing changes: A single-subject design. *Journal of Pediatric Psychology, 38,* 330-341.

Sierra, V., Solanas, A., & Quera, V. (2005). Randomization tests for systematic single-case designs are not always appropriate. *Journal of Experimental Education, 73,* 140-160.

Skillings, J, H., Mack, G. A. (1981). On the use of a Friedman-type statistic in balanced and unbalanced block designs. *Technometrics, 23,* 171–177.

Smith, J. D. (2012). Single-case experimental designs: A systematic review of published research and current standards. *Psychological Methods, 17,* 510-550.

Spaulding, S. A., Devine, K. A., Duncan, C. L., Wilson, N. W., & Hogan, M. B. (2012). Electronic monitoring and feedback to improve adherence in pediatric asthma. *Journal of Pediatric Psychology, 37*, 64-74.

Stewart, K. K., Carr, J. E., Brandt, C. W., & McHenry, M. M. (2007). An evaluation of the conservative dual-criterion method for teaching university students to visually inspection AB-design graphs. *Journal of Applied Behavior Analysis, 40,* 713-718.

Stinson, J. N., McGrath, P. J., & Yamada, J. T. (2003). Clinical trials in the Journal of Pediatric Psychology. *Journal of Pediatric Psychology, 28,* 159-167.

Stone, A. A., & Shiffman, S. (2002). Capturing momentary, self-report data: A proposal for reporting guidelines. *Annals of Behavioral Medicine, 24,* 236–243.

Swaminathan, H., Horner, R. H., Sugai, G., Smolkowski, K., Hedges, L., & Spaulding, S. A. (2008). Application of generalized least squares regression to measure effect size in single-case research: A technical report. Institute of Education Sciences, U.S. Department of Education.

Swoboda, C. M., Kratochwill, T. R., & Levin, J. R. (2010). Conservative dual-Criterion method for single-case research: A guide for visual analysis of AB, ABAB, and multiple-baseline designs (WCER Working Paper No. 2010-13). Retrieved from University of Wisconsin–Madison, Wisconsin Center for Education Research website: http://www.wcer.wisc.edu/publications/workingPapers/papers.php

Tate, R. L., McDonald, S., Perdices, M., Togher, L., Schultz, R., & Savage, S. (2008). Rating the methodological quality of single-subject designs and N-of-1 trials: Introducing the Single-Case Experimental Design (SCED) Scale. *Neuropsychological Rehabilitation, 18,* 385–401.

Tennant, R., Mohammed, M. A., Colman, J. J., & Martin, U. (2007). Monitoring patients using control charts: A systematic review. *International Journal for Quality in Health care, 19,* 187-194.

Thor, J., Lundberg, J., Ask, J., Olsson, J., Carli, C., Härenstam, K. P., & Brommels, M. (2007). Application of statistical process control in healthcare improvement: Systematic review. *Quality & Safety in Health Care, 18,* 387-399.

Tryon, W. W. (1982). A simplified time-series analysis for evaluating treatment interventions. *Journal of Applied Behavior Analysis, 15,* 423-429.

Valsiner, J. (1986). Where is the individual subject in scientific psychology? In J. Valsiner (Ed.), *The individual subject and scientific psychology* (pp. 1-16). New York: Plenum.

Wendt, O., & Miller, B. (2012). Quality appraisal of single-subject experimental designs: An overview and comparison of different appraisal tools. *Education and Treatment of Children, 35,* 235-268.

Westen, D., & Bradley, R. (2005). Empirically supported complexity. *Current Directions in Psychological Science, 14,* 266–271.

Westen, D., Novotny, C. M., & Thompson-Brenner, H. K. (2004). The empirical status of empirically supported psychotherapies: Assumptions, findings, and reporting controlled clinical trials. *Psychological Bulletin, 130,* 631–663.

White, O. R., & Haring, N. G. (1980). *Exceptional teaching.* Columbus Ohio: Charles E. Merrill Publishing Company.

Wilkinson, L., & The Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist, 54,* 594–604.

Yin, R. K. (2012). *Applications of Case Study Research.* Thousand Oaks, CA: Sage.

Table 1.

Quality Indicators for Single-Case Research

Participants and Setting

- Participants are described with sufficient detail (e.g., age, gender, race, ethnicity, medical condition) such that others can select individuals with similar characteristics.
- The process of participant selection and inclusion and exclusion criteria is described with sufficient detail to allow replication.
- Critical features of the setting (e.g., inpatient, outpatient, urban, rural) are described with sufficient detail to allow replication.

Dependent Variable

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

Independent Variable

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

Baseline

- The majority of single-case research studies will include a baseline phase that provides repeated measurement of a dependent variable and establishes a pattern of responding that can be used to predict the pattern of future performance if introduction or manipulation of the independent variable did not occur.

Experimental Control/internal Validity

- The design provides at least 3 demonstrations of experimental effect at 3 different points in time.
- The design controls for common threats to internal validity (e.g., permits elimination of rival hypotheses).
- The results document a pattern that demonstrates experimental control.

External Validity

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

Applied/Clinical/Social Validity

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

Note. Adapted from “The use of single-subject research to identify evidence-based practice in special education,” by R. H. Horner, E. G. Carr, J. Halle, G. McGee, S. Odom, and M. Wolery, 1005, *Exceptional Children*, p. 174. Copyright 2005 by the Council for Exceptional Children.

Table 2.

Analysis Standards and Guidelines for Single-Case Research

Variable	What Works Clearinghouse	APA Division 12 Task Force on Psychological Interventions	APA Division 16 Task Force on Evidence-Based Interventions in School Psychology	The Single-Case Experimental Design Scale (Tate et al., 2008)	Ecological momentary assessment (Stone & Shiffman, 2002)
1. Visual analysis	4-step, 6-variable procedure (based on Parsonson & Baer, 1978)	Acceptable (no specific guidelines or procedures offered)	1. Change in level 2. Minimal score overlap 3. Change in trend 4. Adequate length (≥3) 5. Stable data (Franklin, Gorman, Beasley, & Allison, 1997; Parsonson & Baer, 1992)	Not acceptable ("use statistical analyses or describe effect sizes"; p. 389)	
2. Statistical analysis procedures	Estimating effect sizes: nonparametric and parametric approaches, multilevel modeling, and regression (recommended)	Preferred when the number of data points warrants statistical procedures (no specific guidelines or procedures offered)	Rely on the guidelines presented by Wilkinson and the Task Force on Statistical Inference (1999)	Specific statistical methods are not specified, only their presence or absence is of interest in completing the scale	1. Aggregated or disaggregated approach 2. Model used in analyses 3. Details of procedures (e.g., autocorrelation approach, random effect levels)

3. Demonstrating an effect	1. Documented consistency of level, trend, and variability within each phase 2. Documented immediacy of the effect, the proportion of overlap, the consistency of the data across phases 3. Identify for whom the intervention is and is not effective, if available 4. Examine external factors and anomalies 5. Follow-up of original study participants and multiple intervals with same outcome measures	ABAB—stable baseline established during first A period, data must show improvement during the first B period, reversal or leveling of improvement during the second A period, and resumed improvement in the second B period (no other guidelines offered)	1. 0.05 alpha levels 2. Nonsignificant or negative outcomes noted 3. Type of effect size, type of data on which effect size is based, effect size statistic 4. Clinical/educational significance (e.g., social comparison)	Replication occurs across subjects, therapists, or settings
4. Replication	1. Minimum of 5 studies 2. The studies must be conducted by at least 3 different research teams at 3 different geographical locations 3. The combined number of experiments (i.e., single-case design examples) across the studies totals at least 20	1. 3 replications of ≥3 subjects each 2. Replications conducted by ≥2 independent research groups	1. Same intervention (treatment protocol and duration) 2. Same target problem and sample 3. Independent evaluation	

Note. Adapted from "Single-case experimental designs: A systematic review of published research and current standards," by J. D. Smith, 2012, *Psychological Methods*, pp. 525-526.

Table 3

Visual Analysis Criteria to Evaluate Data in Single-Case Research

1. Changes in Means—Visually detectable differences in the average rate of the data on the outcome measure as phases are changed (e.g., Baseline to Intervention).
2. Changes in Level—Differences in the pattern of the data from the very end of one phase (e.g., Baseline) to the very beginning of the next phase (e.g., Intervention).
3. Changes in Trend or Slope—Differences in the trend lines (e.g., upward or downward movement) on the outcome measure across phases.
4. Changes in Variability—Differences in the variability (range or standard deviation) of the data across phases.
5. Changes in Latency—The elapsed time between the beginning of a phase (e.g., Intervention) and the subsequent change in the data.
6. Consistency in Patterns—Similarities in patterns of data from phases with the same conditions (e.g., Baseline and Return to Baseline).

Note: These criteria should be assessed both individually and collectively.

Table 4

Example Statistics Used in Single-Case Research

Statistic	References
Binomial Sign test	White & Haring, 1980
C Statistic	Jones, 2003; Tryon, 1982; Satake, Jagaroo, & Maxwell, 2008
Double bootstrap method	McKnight, McKean, & Huitema, 2000
Dynamic P-Technique	Nelson, Aylward, & Rausch, 2011
t-test, Analysis of Variance	Satake, Jagaroo, & Maxwell, 2008
Friedman two-way analysis by ran ks	Fisch, 2001
Interrupted Time-Series Analyses	Crosbie, 1993; Maughan, Christiansen, Jenson, Olympia, & Clark, 2005
Kendall's τ	Fisch, 2001
Kruskal-Wallis test	Fisch, 2001
Mann-Whitney–Wilcoxon	Edgington, 1982
Percentage of zero data	Scotti, Evans, Meyer, & Walker, 1991
Randomization Tests	Dugard, File, & Todman, 2012; Sierra, Solanas, & Quera, 2005
Regression	Brossart, Meythaler, Paker, McNamara, & Elliott 2008; Manolov, Arnau, Solanas, & Bono, 2010
Repeated-Measures Analysis of Variance	Fisch, 2001; McCall & Appelbaum, 1973
Revusky's R_n	Revusky, 1967
Skillings-Mack	Skillings & Mack, 1981
Split-middle Method of Trend Estimation	White & Haring, 1980
Two-standard Deviation Band Method	Gottman & Leiblum, 1974
Wilcoxon signed rank test	Fisch, 2001

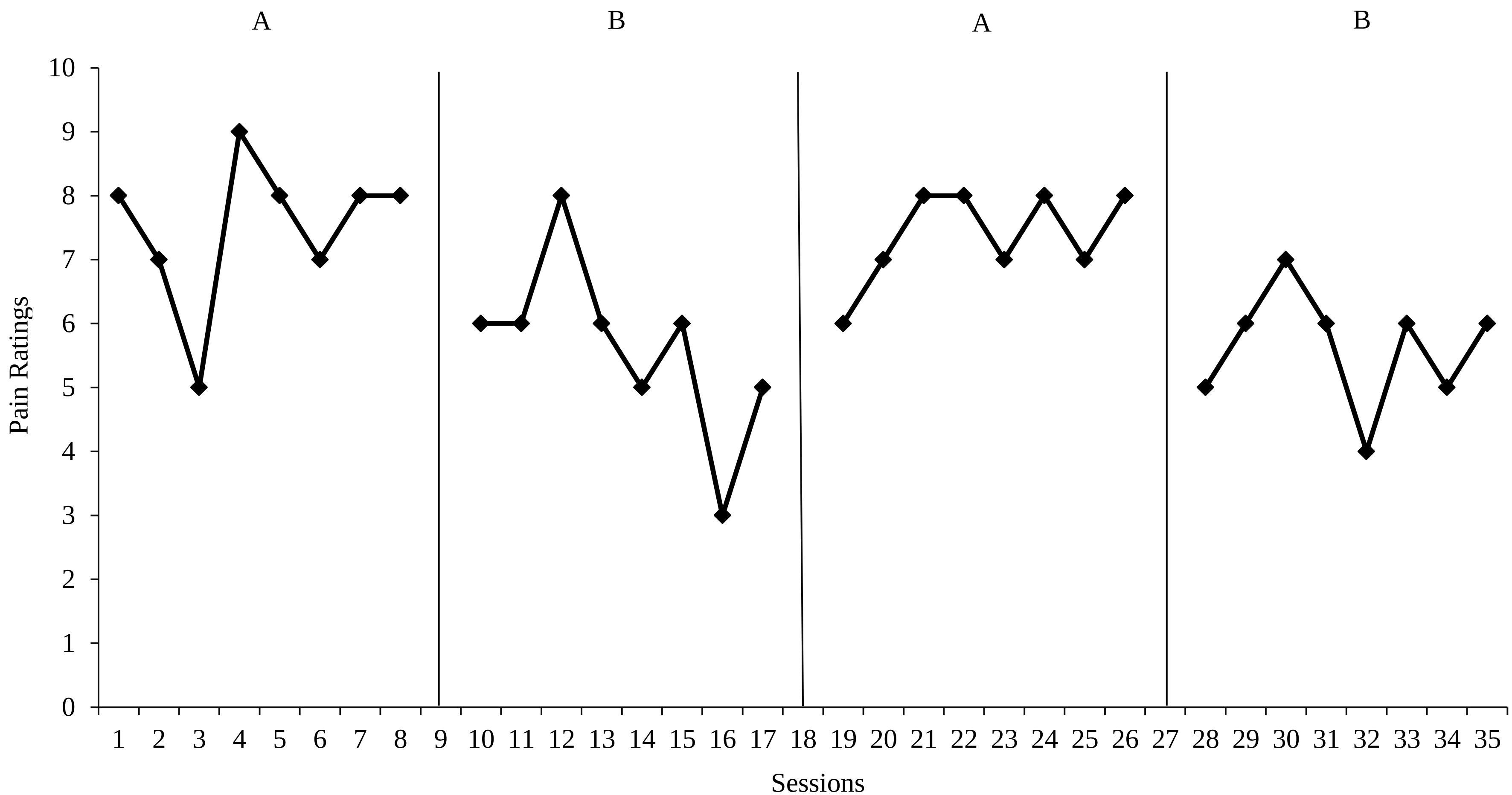

Figure 1. Example ABAB single-case design results.

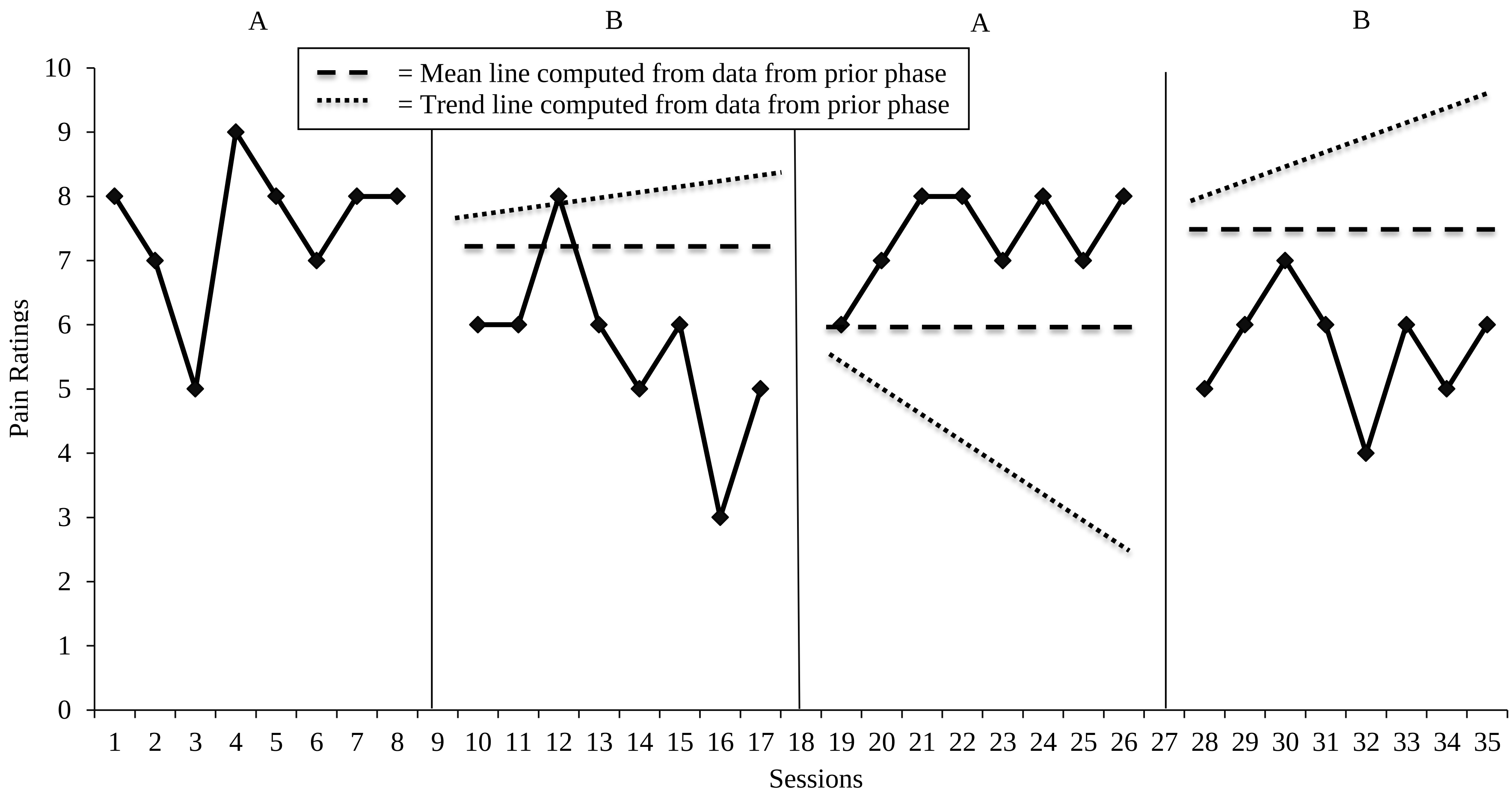

Figure 2. Example ABAB single-case design results with CDC mean and trend lines added.