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

Lindsey L. Cohen,1 PHD, Amanda Feinstein,1 MA, Akihiko Masuda,1 PHD, and Kevin E. Vowles,2 PHD

1Department of Psychology, Georgia State University, and 2Department of Psychology, University of New Mexico

All correspondence concerning this article should be addressed to Lindsey L. Cohen, PHD, Department of Psychology, Georgia State University, Atlanta, GA 30302-5010, USA. E-mail: llcohen@gsu.edu

Received March 15, 2013; revisions received and accepted July 25, 2013

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 used to augment data interpretation.

Key words  research design and methods; single-case designs; statistical applications.

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 provide 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 two 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 historical information; conversely, systematic or objective assessment, prospective evaluations, or observations following intervention are 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 (Kratochwill et al., 2013, Sidman, 1960).

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 to provide a comparison. (c) There is a baseline assessment phase before intervention. (d) Data on the outcome variable are collected continuously and repeated 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) 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) 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 aforementioned 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 (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 (Stinson, McGrath, & Yamada, 2003)—are available to guide in the development, and to allow for evaluation of single-case designs (for a review of seven different appraisal tools, see Wendt & Miller, 2012). Horner et al. (2005) developed a list of considerations for single-case studies, which we adapted for pediatric psychology research (Table I). 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 healthcare 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 Web site between 2000 and 2013 using the keywords “single-case” or “single-subject” revealed only seven studies meeting the definition of single-case design as described previously: Applegate, Kelley, Applegate, Jayasinghe, and Venters (2003); Bernard, Cohen, and Moffett (2009); Burke, Kuhn, and Peterson (2004); Cushing, Jensen, and Steele (2011); Hains, Davies, Parton, and Silverman (2001); Powers et al. (2006); Sil, Dahlquist, and Burns (2013); and Spaulding, Devine, Duncan, Wilson, and Hogan (2012). We hope that this article 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 (Risley, 1970). This experimental and applied distinction parallels that of statistical and clinical significance, which is a valued distinction to pediatric psychologists (Drota,
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 II). That said, at this point in time, recommendations emphasize that visual analysis be the central method of data evaluation (Barlow et al., 2009; Gast, 2010; Kazdin, 2011; Kratochwill et 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 produce 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 (Hayes et al., 1999; Kazdin, 2011; Table III), which include an evaluation of changes in six 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, on 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 on the introduction of an intervention. For example, after starting a sleep hygiene intervention, an adolescent’s erratic sleep pattern (4 hr one night, 12 hr the next) might change to a consolidated pattern of 8 hr 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 (Bowen & Neuhauser, 2013; Diaz & Neuhauser, 2005; Neuhauser, Provost, & Bergman, 2011; Tennant, Mohammed, Colman, & Martin, 2007; Thor et al., 2007). (e) Latency of change after the introduction of a phase (e.g., intervention) refers the immediacy of the change in the data. For example, changes in behavior after 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) Finally, 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 demonstrate similar patterns during baseline and subsequent return to baseline phases. Studies suggest that inter-rater agreement for visual analysis can be high when researchers are trained to criteria (Kahng et al., 2010); however, some data do not support this assertion (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, and Todman (2012); Edgington and Onghena (2007); Franklin, Allison, and Gorman (1997); and Satake,
|----------|--------------------------|--------------------------------------------------------|---------------------------------------------------------------------------------|----------------------------------------------------------|---------------------------------------------------------|
2. Minimal score overlap  
3. Change in trend  
4. Adequate length (≥3)  
5. Stable data (Franklin et al., 1997; Parsons & 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 | ABA—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 five studies  
2. The studies must be conducted by at least three different research teams at three different geographical locations  
3. The combined number of experiments (e.g., single-case design examples) across the studies totals at least 20 | 1. Three 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 |  

Jagaroo, and Maxwell (2008)—for single-case studies, but there is no consensus on best practices regarding statistical methodology for analyzing single-case studies (Table II). 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 (Campbell, 2004; Manolov & Solanas, 2012; Nourbakhash & Ottenbacher, 1994) and when the same statistical technique is used with different assumptions or different aspects (e.g., mean, trend) of the same data (Manolov & Solanas, 2009). Further, there are no agreed-on metrics for judging statistical results (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). Finally, general criticisms regarding statistics in psychology writ large can be levied against statistical techniques for single-case design (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 (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 (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 (Crosbie, 1999; Huitema, 1986). Finally, data suggest that quantitative methods can improve agreement among raters using visual analysis (Fisch, 2001; Hojen & Ottenbacher, 1988). Rather than provide only cursory information on a sampling of tests, we have listed a number of different statistical approaches in Table IV 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 (Sackett et al., 1996), 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 (Horner, Swaminathan, Sugai, & Smolkowski, 2012; Lane & Carter, 2013; Maggin & Chaloules, 2013; Manolov, Solanas, Sierra, & Evans, 2011; Schlosser, Lee, & Wench, 2008; Shadish, Rindskopf, & Hedges, 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 article. See a recent special journal issue edited by Maggin and Chafouleas (2013) for more information on this topic.

**Example Single-Case Analysis**

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. 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, Jenkins, & Reinsel, 1994). Some authors (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 preassign the treatment (e.g., number of sessions), which is required for randomization tests (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). Fourth, 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 complement visual analysis in single-case research.

The CDC method provides assistance in improving inter-rater 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 nonoverlapping 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 preprogrammed 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 to determine that a significant change has occurred from one to the subsequent phase. Investigations using the CDC method suggest that it assists visual analysis and balances Type I and II error (Fisher et al., 2003; Stewart, Carr, Brandt, & McHenry, 2007).

The CDC method can be applied to ABAB or multiple-baseline designs (Swoboda 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 eight data points in the treatment phase (B), it is necessary to have seven of these points fall below both criterion lines to conclude that there is a reliable treatment effect. In our example, seven 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., seven data points fell above the mean line and eight data points above the trend line) and the second treatment phase again led 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 (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.,

Figure 1. Example ABAB single-case design results.
RCTs; Jacobson & Christensen, 1996; Westen & Bradley, 2005; Westen, Novotny, & Thompson-Brenner, 2004) have 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 (Horner et al., 2005; Wendt & Miller, 2012; Table I). 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 are 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 on 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 article, we have reviewed the methodology of analyzing single-case studies and highlighted one of many of the statistical approaches to complement 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 article; 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.

Conflict of interest: None declared

References


(pp. 133–157). Hillsdale, NJ: Lawrence Erlbaum
Associates.

tests (4th ed.). Boca Raton, FL: Taylor & Francis
Group.

Fisch, G. S. (2001). Evaluating data from behavior analysis:
Visual inspection or statistical models.


Visual aids and structured criteria for improving
visual inspection and interpretation of single-case de-
signs. Journal of Applied Behavior Analysis, 36,
387–406.

Franklin, R. D., Gorman, B. S., Beasley, T. M., &
analysis. In R. D. Franklin, D. B. Allison, & B.
S. Gorman (Eds.), Design and analysis of single-case


chotherapy and how to evaluate it. New York, NY:
 Holt, Rinehan and Winston.

Guyatt, G., Haynes, B., Jaeschke, R., Meade, M. O.,
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 man-
ual for evidence-based clinical practice (2nd ed

Hains, A. A., Davies, W. H., Parton, E., &
intervention for distressed adolescents with
type I diabetes. Journal of Pediatric Psychology, 26,
61–66.


(1999). The scientist practitioner: Research and ac-
countability in the age of managed care (2nd ed.).

Hojen, M. A., & Ottenbacher, K. J. (1988). Empirical in-
vestigation of visual-inspection versus trend-line anal-
ysis of single-subject data. Physical Therapy, 68,
983–988.

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 sys-
tematic analysis and use of single-case research.

subject designs: Some misunderstandings. In
A. S. Bellack & M. Hersen (Series Ed.) and A. Poling
& R. W. Fugua (Vol. Eds.), Applied clinical psychology:
Research methods in applied behavior analysis: Issues

effectiveness of psychotherapy: How well can clinical
trials do the job? American Psychologist, 51,
1031–1039.

Bayesian analysis: A practitioner’s guide.
Measurement and Education in Counseling and

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.

for clinical and applied settings (2nd ed.). New York,
NY: Oxford University Press.

Kratochwill, T. R., Hitchcock, J. H., Horner, R. H.,
Levin, J. R., Odom, S. L., Rindskopf, D. M., &
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 re-
search design standards. Remedial and Special
Education, 34, 26–38.

special issue: Issues and advances in the meta-analy-
sis of single-case research. Remedial and Special

the special series: Issues and advances of synthe-
sizing single-case research. Remedial and Special
Education, 34, 3–8.

Regression-based techniques for statistical decision
making in single-case designs. Psicothema, 22,
1026–1032.

Manolov, R., & Solanas, A. (2009). Problems of the ran-
domization test for AB designs. Psicoligica, 30,
137–154.


