



Anuario de Psicología

ISSN: 0066-5126

anuario-psicologia@ub.edu

Universitat de Barcelona

España

Manolov, Rumen

Reporting single-case design studies: Advice in relation to the designs' methodological
and analytical peculiarities

Anuario de Psicología, vol. 47, núm. 1, 2017, pp. 45-55

Universitat de Barcelona

Barcelona, España

Available in: <http://www.redalyc.org/articulo.oa?id=97053838005>

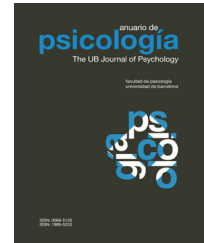
- How to cite
- Complete issue
- More information about this article
- Journal's homepage in redalyc.org

redalyc.org

Scientific Information System

Network of Scientific Journals from Latin America, the Caribbean, Spain and Portugal

Non-profit academic project, developed under the open access initiative



THEMATIC REVIEW

Reporting single-case design studies: Advice in relation to the designs' methodological and analytical peculiarities



Rumen Manolov

Department of Social Psychology and Quantitative Psychology, Faculty of Psychology, University of Barcelona, Spain

Received 12 February 2017; accepted 11 May 2017

Available online 9 June 2017

KEYWORDS

Single-case designs;
Reporting;
Data analysis;
Quality standards

Abstract The current text provides advice on the content of an article reporting a single-case design research. The advice is drawn from several sources, such as the *Single-case research in behavioral sciences reporting guidelines*, developed by an international panel of experts, scholarly articles on reporting, methodological quality scales, and the author's professional experience. The indications provided on the Introduction, Discussion, and Abstract are very general and applicable to many instances of applied psychological research across domains. In contrast, more space is dedicated to the Method and Results sections, on the basis of the peculiarities of single-case designs methodology and the complications in terms of data analysis. Specifically, regarding the Method, several aspects strengthening (or allowing the assessment of) the internal validity are underlined, as well as information relevant for evaluating the possibility to generalize the results. Regarding the Results, the focus is put on justifying the analytical approach followed. The author considers that, even if a study does not meet methodological quality standards, it should include sufficiently explicit reporting that makes possible assessing its methodological quality. The importance of reporting all data gathered, including unexpected and undesired results, is also highlighted. Finally, a checklist is provided as a summary of the reporting tips.

© 2017 Universitat de Barcelona. Published by Elsevier España, S.L.U. All rights reserved.

PALABRAS CLAVE

Diseños de caso
único;
Informes;
Análisis de datos;
Estándares de calidad

Informes en diseños de caso único: consejos en base a las peculiaridades metodológicas y estadísticas de los diseños

Resumen El texto proporciona consejo sobre el contenido necesario para aquellos artículos que informan sobre estudios que utilizan diseños de caso único. El consejo se basa en diferentes fuentes, como *Single-case research in behavioral sciences reporting guidelines*,

E-mail address: rmanolov13@ub.edu

<http://dx.doi.org/10.1016/j.anpsic.2017.05.004>

0066-5126/© 2017 Universitat de Barcelona. Published by Elsevier España, S.L.U. All rights reserved.

recomendaciones desarrolladas por un panel internacional de expertos, artículos científicos sobre informes, escalas de calidad metodológica y la experiencia profesional del autor. Las indicaciones proporcionadas sobre la Introducción, la Discusión y el Resumen son muy generales y aplicables a muchos ejemplos de investigación psicológica aplicada en diferentes ámbitos. En cambio, se dedica más espacio a las secciones Método y Resultados, en relación con las peculiaridades de la metodología de los diseños de caso único y las complicaciones en cuanto al análisis de datos. Específicamente, en cuanto al Método, se destacan aspectos que fortalecen (o permiten la evaluación de) la validez interna, además de la información relevante para valorar la posibilidad de generalizar los resultados. En cuanto a los Resultados, se focaliza la justificación del enfoque analítico seguido. El autor considera que, incluso si un estudio no cumple con los estándares de calidad metodológica, el informe debería ser lo suficientemente explícito para favorecer la valoración de la calidad metodológica. Se subraya la importancia de reportar todos los resultados obtenidos, incluidos los inesperados o indeseados. Finalmente, se proporciona una lista de verificación como resumen de los consejos.

© 2017 Universitat de Barcelona. Publicado por Elsevier España, S.L.U. Todos los derechos reservados.

In the current text we assume that the reader is already familiar with the main features of single-case designs (SCD, as described in depth in [Barlow, Nock, & Hersen, 2009](#); [Gast & Ledford, 2014](#); [Kennedy, 2005](#); [Kratochwill & Levin, 2014](#); [Vannest, Davis, & Parker, 2013](#); see also [Bono & Arnau, 2014](#) for a textbook in Spanish) and that s/he is an applied researcher considering the use of SCD or already with experience in the field. Thus, we assume that the reader is mainly interested in the key aspects that need to be reflected in the report describing a SCD study.

Recommendations about reporting

Reporting resources

When a research is performed and its results are to be shared publicly, it is important that the report reflects in a transparent way the process followed in order to make possible: (a) the assessment of the study's internal and external validity that each reader can perform independently, and (b) replicating the study, if considered necessary. The *Single-case research in behavioral sciences reporting guidelines* (SCRIBE; [Tate et al., 2016](#)) should be document of reference, because it is the result of a collaboration of an international panel of experts via a Delphi study. However, the SCRIBE is intended to refer to "minimum reporting standards" ([Tate et al., 2016, p. 11](#)), whereas we here point at some points that have been suggested, in several scholarly articles, for inclusion in a report. In what follows, we refer to different pieces of the article, paying special attention to the Method and Results sections. The Method section is crucial for evaluating the quality of the evidence provided by the study ([Tate, Perdices, McDonald, Togher, & Rosenkoetter, 2014](#)), whereas the Results section may entail certain complications due to the variety of analytical approaches and lack of consensus in the SCED context. A summary of the pieces of advice is presented in [Appendix A](#) in the form of a checklist.

Introduction, conclusion and abstract

For these three parts of the text, the rules usually followed for any kind of empirical psychological research are applicable to SCD. We recommend [Sternberg \(2003\)](#), for a textbook on the topic.

Introduction

This section should include a specification of the problem of interest and how it has been studied previously, plus what does the evidence published in peer-reviewed literature suggests about each of the approaches for dealing with the problem. It is also necessary to provide a rationale for choosing one of the existing approaches or for proposing a new one. Part of the theoretical framework are the definitions of the relevant terms, introducing any necessary abbreviations ([Sternberg, 2003](#)), but not too many, unless they are very common in the field (e.g., MBD, ATD are common abbreviations for designs and PND is a widely known abbreviation of a nonoverlap index for quantifying the difference between conditions). At the end of the Introduction aims are clearly specified, as well as any formal hypothesis, if available. If the structure of the article is complex or unusual, it is important to present its organization at the end of the Introduction. Finally, regarding the bibliographic basis, the authors should ensure that the references used are relevant, sufficient, and (at least some of them) recent.

Discussion

In this section it is necessary to relate the results to the aims and to compare these results with previous findings ([Wolery, Dunlap, & Ledford, 2011](#)). In case there are formal hypothesis postulated in the Introduction, the authors have to distinguish expected results from unexpected ones and explicitly state which explanations of the results are

made ad hoc or post hoc (Sternberg, 2003). Furthermore, a discussion of threats to internal validity and the possibility to generalize the results is called for. In relation to this point, limitations (foreseen by the design and unforeseen) should be explicitly stated, as well as specific lines for future research. Finally, theoretical and/or practical implications should be derived, whenever possible.

Abstract

It should be written after all other text is completed, in order for the author to be able to clearly understand the main aspects of the text and the main contribution of the study. Such understanding is basic for being able to transmit it to the reader. The abstract should contain information about the research question, population, design, intervention, target behavior, results, and conclusions (Tate et al., 2016).

Additional information

The author's note has to contain information about funding and the role of funders (Tate et al., 2016). It is possible to include an Appendix with an instrument used that is not easily accessed or with additional results (Sternberg, 2003).

Method – design and intervention

Many different single-case designs exist, but the most frequently used ones, according to several reviews (Hammond & Gast, 2010; Shadish & Sullivan, 2011; Smith, 2012) are multiple-baseline design (MBD), ABAB, and alternating treatments designs (ATD), whereas less frequently used designs include changing criterion, simultaneous (parallel) treatments designs, and multitreatment (e.g., ABCB) designs. All these descriptors of the design are useful and usually not ambiguous. However, what can be less informative is the term of the general type of design into which a specific design is classified. For instance, the MBD has been called a simultaneous replication design (Onghena & Edgington, 2005), a time-lagged design (Hammond & Gast, 2010), and an individual intervention design (Wolery et al., 2011). The ATD has been called an alternation design (Onghena & Edgington, 2005), a comparison design (Hammond & Gast, 2010) and a comparing interventions design (Wolery et al., 2011). Finally, the ABAB design is usually called "withdrawal" or "reversal design", although Hammond and Gast (2010) use the term "reversal" only when the design involves "applying the independent variable to one target behavior in baseline and another target behavior during intervention." (p. 189). Besides the name, it is also recommended to provide the rationale for using the specific design chosen (Wolery et al., 2011) and to make explicit the number of within-study attempts to demonstrate the treatment effect, which is a critical feature for establishing experimental control across quality standards (Maggin, Briesch, Chafouleas, Ferguson, & Clark, 2014). In case the analytical procedures stem from time-series analysis, it has to be specified whether measurements were obtained at uniform intervals (Smith, 2012). Moreover,

Skolasky (2016) recommends reporting whether there were wash-out periods before introducing or after withdrawing an intervention. Finally, it has to be stated whether the sequence of phases and their shifts were determined a priori or were data-driven (Tate et al., 2016).

Method – participants and setting: favor the assessment of external validity

The basic feature which all reports include is the number of participants, given that replication is crucial for both internal and external validity (Smith, 2012; Tate et al., 2013). We here refer to the number of participants that started the study, given that it is also necessary to describe when and why participants left the study or the intervention, in case this happened (Tate et al., 2016).

Additionally, when writing a report, the authors should consider that their study can be included in future meta-analyses, which are useful for establishing the evidence basis of interventions (Jenson, Clark, Kircher, & Kristjansson, 2007). This has consequences on the features of the participants that need to be described. For instance, one of the respected tools for assessing the methodological quality of systematic reviews, AMSTAR (Shea et al., 2007), includes items that prompt including relevant characteristics of the participants such as age, race, sex, relevant socioeconomic data, disease status, duration, severity, or other diseases. Romeiser-Logan, Slaughter, and Hickman (2017) stress the importance of reporting participant characteristics, so that each practitioner can decide to what extent the findings are related to their current client. Similarly, the Risk of Bias in N-of-1 Trials scale (RoBiNT; Tate et al., 2013, 2015) also highlights the same participant characteristics, although it refers to the etiology of the problem and its severity (which can be related to status and duration). Another aspect relevant is an assessment of the factors that maintain the problem behavior during the baseline (i.e., in absence of intervention). Regarding the access and admission of participants, it is important to state the inclusion and exclusion criteria (Wolery et al., 2011) and whether and how informed consent was obtained (Tate et al., 2016).

Besides participants' characteristics, the RoBiNT scale (Tate et al., 2013) includes an item requiring the description of the setting, both general (e.g., hospital, school, research laboratory) and specific environment (e.g., characteristics of the room, materials used, people present). In case such information is provided, the article will not only be assigned a higher score in methodological quality scales, but it will also favor assessing to what settings, participants, and target behaviors the results can be generalized.

Method – instrument: the dependent variable described

In SCD, measurements are frequently obtained by direct observation and the observational procedures followed (e.g., event coding, partial interval recording with a given interval length) need to be described, because they are relevant both in terms of interobserver agreement (IOA; Rapp, Carroll, Stangeland, Swanson, & Higgins, 2011) and in

terms of the quantifications of effect (Pustejovsky, 2015). The behavior to be observed has to be operatively defined providing examples and counterexamples of human actions that will or will not be counted as instances of the target behavior. IOA has to be assessed for at least 20% of the observational sessions and the exact quantification depends on the observational procedure followed (Hott, Limberg, Ohrt, & Schmit, 2015). IOA has to be reported both as an average and a range and the authors should be aware of the minimal standards: 80% percentage agreement and 60% Cohen's kappa (Horner et al., 2005).

If self-report measures (e.g., questionnaires, inventories, scales) are used, it is important to provide references to the manuals of the instruments, along with information about the sub-scales included (ideally provide examples of items; e.g., Bailey & Wells, 2014), and their psychometric properties (e.g., internal consistency, test-retest reliability). If an instrument is used in a population different from the one in which it was initially validated, it is relevant to mention whether any formal validation has taken place for the target population and/or to specify whether the original instrument has been translated and back translated (e.g., Callesen, Jensen, & Wells, 2014). Moreover, reporting cut-off scores representing normal vs. pathological functioning is also important in order to assess the potential practical significance of the intervention effect (e.g., Fitt & Rees, 2012). Finally, it is important to mention whether any additional measures are taken to explore the generalization of the effects of the interventions to behaviors and contexts that were not object of the treatment (Tate et al., 2015).

When a daily diary (e.g., Wells, 1990) or performance in tasks in which there is an objective correct answer (e.g., Tunnard & Wilson, 2014) is used, a description of the target behaviors is also required. For data gathered via technological devices it is important to describe those (Tate et al., 2016) and to report whether training of the participant in the use of the device took place (Smith, 2012) and whether the device failed at any point (Hott et al., 2015).

Method – intervention: the independent variable described

Procedural fidelity should be subjected to an independent assessment (Hott et al., 2015), just like the recording of the dependent variable. The importance of procedural fidelity is based on the need to implement research-supported interventions in typical environments as they were intended to be used (Ledford & Gast, 2014) and, thus, an ethical aspect is involved: not to offer a sub-optimal service to the person in need. Moreover, delivering the intervention accurately helps establishing the causal relation with the changes in the target behavior (Ledford & Wolery, 2013). Ledford and Gast (2014) recommend reporting procedural fidelity separately for each step, (e.g., percentage of behaviors correctly performed as assessed using observation and a checklist; Tate et al., 2015), for each participant and for each condition (baseline and intervention), distinguishing behaviors that have remain the same across conditions and behaviors that have to change due to being part of the intervention. Any differences across participants in terms of procedural fidelity could be relevant for explaining

differential response to and effectiveness of the intervention. Moreover, detecting steps which are not implemented as expected can be useful for identifying procedural components that are not practical for application by teachers, parents, etc. (Ledford & Wolery, 2013). Additional aspects that need to be reported are unforeseen changes in the participants and in the setting, not related to the treatment such as interruptions not due to the participant and change of medication (Hott et al., 2015).

Regarding treatment fidelity or adherence as a specific part of procedural fidelity, we recommend specifying whether a manual has been followed when applying the intervention, what the content of the different sessions was (e.g., McNicol, Salmon, Young, & Fisher, 2013), or in case the creator of the intervention also took part in the study (e.g., Callesen et al., 2014). Given that SCDs are flexible enough to allow for tailored interventions and changes in the conditions in response to the continuous measurements of the target behavior, such modifications need to be made explicit.

Method – procedure: favor the assessment of internal validity

The assessment the scientific quality of the studies in the context of a meta-analysis (see AMSTAR, Shea et al., 2007) would be made easier in case the authors explicitly state whether any randomization or counterbalancing of the order of the conditions has taken place, whether the researchers, participants, assessors (especially for overt behavior), and data analysts were blind to the aims, hypothesis, and specific conditions when performing their task.

In many SCD, it is possible and methodologically desirable to introduce randomization in the design in order to strengthen its internal validity (Kratzschwill & Levin, 2010). In terms of reporting, it is not sufficient to state that a design includes randomization, given that there are many different ways in which this randomization can be performed. For instance, focusing on MBD, there are several options for using randomization (Levin, Ferron, & Gafurov, 2016): (a) each participant can be assigned at random to one of the baselines (i.e., case randomization); (b) the starting point of the intervention can be determined at random, from a pre-defined set of options ensuring that the intervention start points do not overlap across baselines; and (c) case randomization and start point randomization can be combined. As another example of randomization, this time for ATDs, in one study (Schneider, Coddington, & Tryon, 2013) conditions were assigned to measurement times by drawn them from a hat without replacement in, whereas in another study (Yakubova & Bouck, 2014) this was achieved by flipping a coin. In the randomization scheme in Schneider et al. (2013) each condition is present necessarily the same number of times and there is no more than one consecutive administration of the same condition. In contrast, according to the randomization scheme followed in Yakubova and Bouck (2014), the number of consecutive administrations of the same condition was restricted to two and it was possible to obtain an unequal number of measurement occasions per condition.

Another aspect that has to be made clear in case there is a series of AB designs in the study is whether data are

collected from the individuals concurrently with a planned staggered introduction of the treatment (as in a MBD) or the participants are studied consecutively, according to the moment in which they agree to participate in the study. The control for history as a threat to internal validity is much stronger for the MBD than for the non-concurrent AB replications (see Tate et al., 2015).

Method – data analysis: justification

The lack of consensus on the most appropriate data analytical procedures (Kratochwill et al., 2010) entails the need to justify the analytical decisions made (Tate et al., 2013) in relation to the hypothesis (Smith, 2012) or to the appropriateness of the analytical procedures for the data at hand. For instance, O'Neill and Findlay (2014) provide the following justification: "The start date of the intervention was randomly assigned within a window of two weeks by allocating the start date to one of three envelope-concealed consecutive Mondays. This indicated two potential statistical approaches, randomization statistics [...] or non-overlap of all pairs [...]. The NAP statistic was chosen to compare baseline and intervention. [...] It has evidenced ability to discriminate among typical single case research results and has been correlated with established indices of magnitude of effect including Cohen's d " (p. 371).

Actually, some analytical techniques are complex enough to require their own reporting guidelines. For instance, Ferron et al. (2008) suggest that when using multilevel models, it is necessary to report the number of units per level (measurements, participants, studies), whether any of the predictors has been centered, the process followed for defining the model and its formulaic descriptions, the estimation methods, algorithms, and software program used for obtaining the results, the method for estimating the degrees of freedom, the results of hypothesis tests for comparing models, point estimates and confidence intervals for the key parameters representing the differences between conditions.

It is also important to make explicit any relevant decisions made in the process of data analysis. For instance, Zelinsky and Shadish (2016) describe the decisions made when applying the BC-SMD for meta-analytical purposes studies using different designs: "[b]ecause the SPSS macro required pairs of baseline and treatment phases, we excluded any extra nonpaired baseline or maintenance phases at the end of studies (e.g., excluding the last A-phase from an ABA design). Finally, if the case started with a treatment phase, we paired that treatment phase with the final baseline phase from the end of that case." (p. 5). As another example, Parker, Vannest, Davis, and Sauber (2011) describe the field test that they performed on a nonoverlap index as "[f]or complex, multiphase designs, only the initial A and B phases were included" (p. 293).

Results: raw data and all data

Given the variety of possible analyses, it is both common and expected (Tate et al., 2013) to report raw data in either tabular or graphical (readable) format so that further analyses can be performed on them (e.g., for demonstrating the

different conclusions that may be reached by using different analytical options) and to enhance interpretation of effect sizes (Pek & Flora, 2017). Moreover, the availability of raw data also favors future meta-analyses. The importance of re-analyses can be seen in the research on software tools for retrieving (i.e., digitizing) single-case data from plots (e.g., Drevon, Fursa, & Malcolm, 2016; Moeyaert, Maggin, & Verkuilen, 2016).

It is important to report all data and results obtained and not only the ones that support the hypotheses, favor the intervention, or are more salient (i.e., p values below nominal alpha and large effect sizes). The practice of omitting certain results has been detected in single-case research (Shadish, Zelinsky, Vevea, & Kratochwill, 2016) and it is a form of publication bias, which does not refer only to meta-analyses. We recommend an ethical attitude, based on the idea that SCD research should get published if the methodology followed is correct regardless of the results obtained (MacCoun & Perlmutter, 2015) and also considering that SCD research should contribute to identifying which interventions are effective for whom and also which are not, in order to avoid suffering and wasting time and money. As an example of good practice, Fitt and Rees (2012) describe the characteristics of a participant who eventually did not complete the intervention. Finally, Skolasky (2016) suggest reporting whether there are missing data.

Results: software output

An initial task is to get to know the existing options for data analysis. This task includes two steps. First, it is necessary to get familiar with the basis, strengths, and limitations of the analytical options, as well learning from examples of their application.¹ Second, it is important to know in which kind of software the analytical developments have been implemented. Although certain quantification can be obtained by hand calculation, the use of software eliminates the possibility of human error, in case the software code underlying the implementation is assumed to be flawless. A good starting point is the list of software tools available at <https://osf.io/t6ws6/> (which is a continuously updated and expanded version of the list provided in Manolov & Moeyaert, 2016). This list includes freely accessible web-based applications, packages and code in R (R Core Team, 2016), code for SAS (<http://www.sas.com/>), analyses based on Microsoft Excel (<https://products.office.com/en-us/excel>) and on IBM SPSS (<http://www.ibm.com/analytics/us/en/technology/spss/>). For some of the analytical procedures, there have been tutorials created for guiding their use. For instance, for the between-cases standardized mean difference (BC-SMD; Shadish, Hedges, & Pustejovsky, 2014) there is a tutorial about the SPSS macro (Marso & Shadish, 2015) and about the Shiny web application (Valentine, Tanner-Smith, & Pustejovsky, 2016). Additionally, for many of the software tools based on R there is also a tutorial (Manolov, Moeyaert, & Evans, 2016).

¹ See Appendix B for references to potentially useful texts.

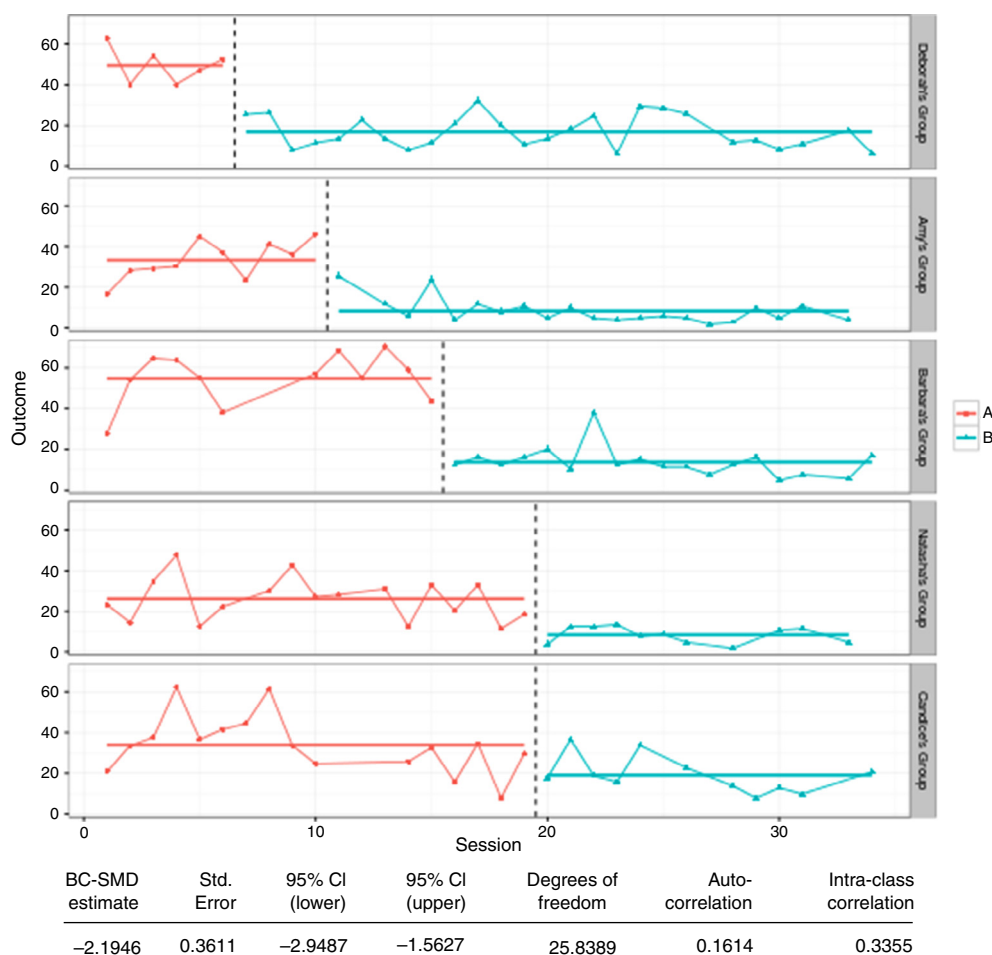


Figure 1 Graphical and numerical output obtained directly from <https://jepusto.shinyapps.io/scdhlm/> using the Rodriguez multiple-baseline data included in the website.

A second task is to decide how to report the information provided by the software. A positive aspect of the existing software is that certain tools offer a combination of visual and numerical information. For instance, the application of the BC-SMD via the web site <https://jepusto.shinyapps.io/scdhlm/> to one of the default data sets included (namely, the Rodriguez MBD data), leads to the output that we have combined in Fig. 1. As another example, see the implementation of piecewise regression (Center, Skiba, & Casey, 1985–1986) via <https://manolov.shinyapps.io/Regression/> represented in Fig. 2. Additional examples are included in Manolov and Moeyaert (2016), where the figures represent direct output of the R code used.

Combining visual and quantitative information is desirable for five reasons. First, visual analysis alone has been criticized for lacking formal decision rules (Ottenbacher, 1990) and the evidence suggests that the agreement between visual analysts is insufficient (Ninci, Vannest, Willson, & Zhang, 2015). Second, visual analysis itself is usually accompanied by graphical aids (Miller, 1985) or by quantitative summaries of different aspects of the data such as level, trend, overlap, and variability (Lane & Gast, 2014). Third, there is considerable consensus on the need to analyze single-case data by both visual and statistical analyses

(e.g., Fisch, 2001; Franklin, Gorman, Beasley, & Allison, 1996; Harrington & Velicer, 2015; Manolov, Gast, Perdices, & Evans, 2014). Even SCD quality standards usually include items on both visual and statistical analysis (Heyvaert, Wendt, Van den Noortgate, & Onghena, 2015). Fourth, presenting visual and numerical information jointly introduces objectivity to the visually-based decisions and makes possible validating the numerical summaries (Parker, Cryer, & Byrns, 2006) in relation to any salient data features (e.g., trend, outliers, variability). For instance, multilevel models have been used to augment visual analysis by providing quantifications and statistical significance (Davis et al., 2013), whereas visual analysis has been suggested for choosing the optimal multilevel model for the data (Baek, Petit-Bois, Van den Noortgate, Beretvas, & Ferron, 2016). Finally, for a masked visual analysis (Ferron & Jones, 2006), the graphical and the numerical expression of the results are inherently related (see, for instance, Lloyd, Finley, & Weaver, 2015).

Results: reporting of intervention effect

Kelley and Preacher (2012) emphasize the importance of reporting not only the point estimates of the (standardized or raw) effect size measures, but also their confidence

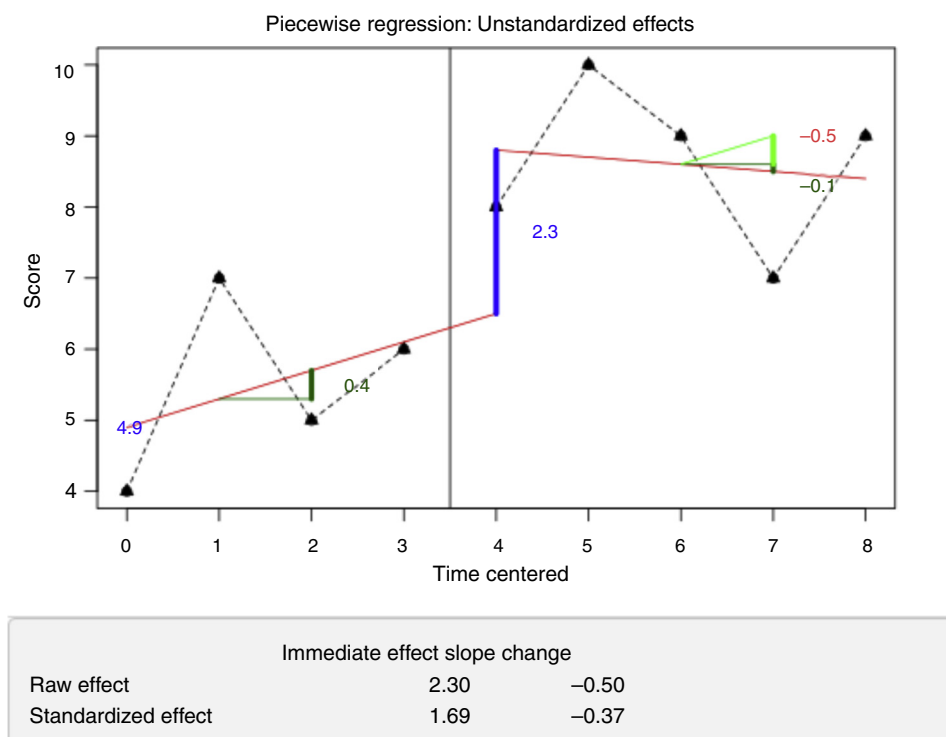


Figure 2 Graphical and numerical output obtained directly from <https://manolov.shinyapps.io/Regression/> using Piecewise regression and the data included as an illustration in the website.

interval, as has been generally recommended in psychology (Wilkinson & The Task Force on Statistical Inference, 1999). However, confidence intervals are available only for some indices such as the BC-SMD (Shadish et al., 2014) and for some nonoverlap indices (see Parker & Vannest, 2009, pp. 361–362). Additionally, it is relevant to specify the exact index used – for instance, there are several measures for quantifying data overlap (Parker, Vannest, & Davis, 2011) and several ways to compute a standardized mean difference (Beretvas & Chung, 2008; Shadish et al., 2014).

In terms of effect size interpretation, a review of benchmarks is provided by Kotrlik, Williams, and Jabor (2011), but it is not clear that such benchmarks are applicable to single-case data (Parker et al., 2005). For instance, Harrington and Velicer (2015) suggest alternative benchmarks for interpreting standardized mean difference values arising from single-case research. In relation to this aspect, Manolov, Jamieson, Evans, and Sierra (2016) offer a review of different ways in which benchmarks can be established to help interpreting the numerical outcomes.

If statistical significance is reported, it has to be clearly specified how a p value was obtained, because its interpretation is not necessarily equivalent. For instance, the p value of the Nonoverlap of all pairs (Parker & Vannest, 2009) stems from its correspondence to analytical procedures assuming random sampling and independent data, whereas the p value obtained in simulation modeling analysis (Borckardt & Nash, 2014) is based on Monte Carlo sampling or bootstrap, specifically taking the estimated autocorrelation into account and assuming normally distributed data. Still another option are randomization tests (Heyvaert & Onghena, 2014) in which the p value is based on the random assignment procedure

that is part of the design and does not entail any distributional assumptions about the data. In that sense, Skolasky (2016) underscores the importance after stating the assumptions and the effect of not meeting them.

The outcome of a SCD study is not necessarily only quantitative. Regarding systematic visual analysis following the What Works Clearinghouse *Standards* (Kratochwill et al., 2010), indications have been provided about the specific questions that need to be answered (Maggin, Briesch, & Chafouleas, 2013) and about the visual aids and quantifications regarding level, trend, variability, overlap, immediacy of effect, and consistency of data patterns that can make the assessment more objective (Horner, Swaminathan, Sugai, & Smolkowski, 2012; Lane & Gast, 2014). Moreover, as per Pek and Flora (2017), it is necessary to discuss whether the interpretation is made in terms of the operative definition (e.g., questionnaire scores) or in terms of the construct that supposedly underlies them.

Finally, regarding social validation, it has to be specified whether the intervention is feasible and socially important for clients; and whether they are satisfied with it (Hott et al., 2015): for instance, Fitt and Rees (2012) comment on the participants' feedback on the therapy. Additionally, it has to be considered whether the implementation of the intervention is practical and cost effective (Horner et al., 2005).

General personal tips on writing

The main aspect in reporting is elaborating a text that honestly, clearly, and concisely states what has been done,

why, and with what outcome. Such clear and concise text makes it easier for readers and reviewers to understand and assess the contribution of a study and it also makes replications possible. In that sense, a text such as the current one is potentially useful for authors, reviewers, and journal editors (Tate et al., 2014). On the one hand, a badly written text can make a good study (i.e., scoring high in a quality standard) unpublishable. On the other hand, being aware of the aspects that need to be reported may prompt researchers to perform more methodologically sound studies. In that sense, we should aim to perform better studies and not only to write better texts. This is why in the current article, we stressed the importance of taking not only reporting guidelines into account, but also methodological quality indicators. Specifically, we have mainly followed the RoBiNT scale (Tate et al., 2013), for which information is provided about its development and psychometric properties and it is also accompanied by an expanded manual (Tate et al., 2015). Moreover, this tool includes all methodological criteria reviewed by Maggin et al. (2014) for documenting an experimental effect and establishing generality. An additional relevant review of quality standards is performed by Smith (2012).

The best way to start writing is to start with the structure of the article (Luiselli, 2010), using headings and sub-headings. Afterwards, one gets motivated by one's own progress by writing the easiest content first. It is also recommended to describe the steps of the study, as they take place instead of relying on memory later on.

Writing is improved with experience – by reading scientific literature, writing reports, and answering to co-authors' and reviewers' concerns. The text written does not have to be perfect from the beginning; it can be edited continuously. Making the text's message clearer usually involves avoiding excessive extension and complication (Sternberg, 2003). Additionally, causing a positive impression is easier when refining, constructing and expanding on previous research rather than trying to demonstrate its uselessness and wrongness as a way of highlighting one's own contribution. Finally, getting published becomes a less difficult task, when submitting to journals interested in the content (Luiselli, 2010) and when referring to previous studies from the same journal, as its readers are likely to be more familiar with those.

Finally, note that the recommendations provided should not be considered as definitive or the only ones possible. They are not the product of a consensus of group of experts, but rather a synthesis of advice provided by SCD applied researchers and methodologists in published documents, as well as being based on the experience of the author of the current text.

Appendix A. Checklist of aspects to include in the report

Abstract

- Research question
- Population
- Design
- Intervention
- Target behavior

- Results
- Conclusions

Introduction

- Domain of the research question and formal statement of the research question
- Theoretical and methodological approaches to the domain
- Previous evidence
- Justification of the need for the study
- Justification of the theoretical approach and intervention selected
- Aim and, if applicable, Hypothesis
- Explanation of the organization of the following sections, if complex or uncommon

Method

- Design:
 - descriptive name
 - number of attempts to demonstrate an effect,
 - determination of phase sequence and moments of change in phase – a priori or data-driven,
- Participants:
 - inclusion and exclusion criteria,
 - number of individuals who begin the study,
 - number of individuals who complete the study,
 - demographical characteristics (age, gender, ethnicity, relevant socio-economic data),
 - main features of the problem behavior (diagnostic, severity, duration, etiology, factors that maintain it, any previous or current medication taken for dealing with it)
- Setting: description of the general and the specific context
- Instrument: measurement of the main target behavior and secondary measures: according to what is applicable to the specific study:
 - Specifications about observational coding schemes and interobserver agreement
 - Psychometric properties of self-report measures (incl. presence of cut-off scores for distinguishing normal from pathological functioning)
 - Technological devices: type of information obtained, need for training
 - Diaries: tasks and indications for the participants elaborating them
- Intervention:
 - manual and/or steps followed;
 - result of the assessment of treatment adherence
- Procedure:
 - presence of blinding/masking of experimenter, observer, participant, data analyst,
 - presence of randomization and how random assignment was performed
 - presence of counterbalancing,
 - result of the assessment of procedural fidelity for each step, each condition, and each participant
 - unexpected events
- Data analysis: Justification of the approach followed, in relation to:

- type of effect expected (e.g., immediate and sustained vs. progressive effect)
- characteristics of the data pattern (e.g., trend, variability)
- design features (e.g., randomization)

Results

- Raw data in graphical or tabular format
- Step-by-step specification of how visual analysis was performed (should include any visual aids actually used by the data analyst in the process)
- Quantifications:
 - effect size measures with confidence intervals, if possible; *p* values, if desired by the researcher
 - clear identification statistical indices and tests used to obtain them
 - explicit mention of the rules followed for interpreting an effect as small vs. large.
- Assessment of social validity:
 - importance of the change for everyday life functioning,
 - opinion of the client and significant others,
 - applicability of the intervention in everyday contexts
- Additional figures and tables, whenever necessary

Discussion

- Interpretation of the results in relation to aims (and hypothesis), previous research, and theoretical approaches to the problem
- Limitations to internal and external validity
- Proposals for future research
- Theoretical and/or practical implications of the results

Appendix B. List of texts on single-case data analysis

One of the key aspects when reporting the data analytical strategy used is to justify the choice made in a way that would convince the reviewers of the manuscript. We encourage the interest reader to get acquainted with the different analytical options available, by consulting some of the following journal special issues.

Special Issues on single-case data analysis and meta-analysis:

- Evidence-Based Communication Assessment and Intervention Vol. 2, Issue 3 in 2008
- Journal of Behavioral Education Vol. 21, Issue 3 in 2012
- Journal of School Psychology Vol. 52, Issue 2, in 2014
- Developmental Neurorehabilitation: planned for 2017

Special Issues on single-case methodology and data analysis:

- Remedial and Special Education Vol. 34, Issue 1, in 2013
- Journal of Applied Sport Psychology Vol. 25, Issue 1, in 2013
- Neuropsychological Rehabilitation Vol. 23, Issues 3–4, in 2014

- Journal of Contextual Behavioral Science Vol. 3, Issues 1–2, in 2014
- Journal of Counseling and Development Vol. 93, Issue 4, in 2015
- Aphasiology Vol. 29, Issue 5 in 2015
- Remedial and Special Education (“Issues and Advances in the Systematic Review of Single-Case Research: An Update and Exemplars”): planned for 2017

Book summarizing data analytical options:

- Kratochwill, T. R., & Levin, J. R. (2014). *Single-case intervention research. Methodological and statistical advances*. Washington, DC: American Psychological Association.

Articles summarizing data analytical options:

- Perdices, M., & Tate, R. L. (2009). Single-subject designs as a tool for evidence-based clinical practice: Are they unrecognized and undervalued? *Neuropsychological Rehabilitation*, 19, 904–927.
- Gage, N. A., & Lewis, T. J. (2013). Analysis of effect for single-case design research. *Journal of Applied Sport Psychology*, 25, 46–60.
- Manolov, R., & Moeyaert, M. (2017). Recommendations for choosing single-case data analytical techniques. *Behavior Therapy*, 48, 97–114.

References

- Baek, E. K., Petit-Bois, M., Van den Noortgate, W., Beretvas, S. N., & Ferron, J. M. (2016). Using visual analysis to evaluate and refine multilevel models of single-case studies. *The Journal of Special Education*, 50, 18–26.
- Bailey, R., & Wells, A. (2014). Metacognitive therapy in the treatment of hypochondriasis: A systematic case series. *Cognitive Therapy and Research*, 38, 541–550.
- Barlow, D., Nock, M., & Hersen, M. (2009). *Single case experimental designs: Strategies for studying behavior change* (3rd ed.). Boston, MA: Allyn and Bacon.
- Beretvas, S. N., & Chung, H. (2008). A review of meta-analyses of single-subject experimental designs: Methodological issues and practice. *Evidence-Based Communication Assessment and Intervention*, 2, 129–141.
- Bono, R., & Arnau, J. (2014). *Diseños de caso único en ciencias de la salud*. Madrid: Síntesis.
- Borckardt, J., & Nash, M. (2014). Simulation modelling analysis for small sets of single-subject data collected over time. *Neuropsychological Rehabilitation*, 24, 492–506.
- Callesen, P., Jensen, A. B., & Wells, A. (2014). Metacognitive therapy in recurrent depression: A case replication series in Denmark. *Scandinavian Journal of Psychology*, 55, 60–64.
- Center, B. A., Skiba, R. J., & Casey, A. (1985–1986). A methodology for the quantitative synthesis of intra-subject design research. *The Journal of Special Education*, 19, 387–400.
- Davis, D. H., Gagné, P., Fredrick, L. D., Alberto, P. A., Waugh, R. E., & Haardörfer, R. (2013). Augmenting visual analysis in single-case research with hierarchical linear modeling. *Behavior Modification*, 37, 62–89.
- Drevon, D., Fursa, S. R., & Malcolm, A. L. (2016). Inter-coder reliability and validity of WebPlotDigitizer in extracting graphed data. *Behavior Modification*, <http://dx.doi.org/10.1177/0145445516673998> (Advance online publication)

- Ferron, J. M., Hogarty, K. Y., Dedrick, R. F., Hess, M. R., Niles, J. D., & Kromrey, J. D. (2008). Reporting results from multilevel analyses. In A. A. O'Connell, & D. B. McCoach (Eds.), *Multilevel modeling of educational data* (pp. 391–426). Greenwich, CT: Information Age Publishing.
- Ferron, J. M., & Jones, P. K. (2006). Tests for the visual analysis of response-guided multiple-baseline data. *The Journal of Experimental Education*, 75, 66–81.
- Fisch, G. S. (2001). Evaluating data from behavioral analysis: Visual inspection or statistical models? *Behavioural Processes*, 54, 137–154.
- Fitt, S., & Rees, C. (2012). Metacognitive therapy for obsessive compulsive disorder by videoconference: A preliminary study. *Behaviour Change*, 29, 213–229.
- Franklin, R. D., Gorman, B. S., Beasley, T. M., & Allison, D. B. (1996). 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: Lawrence Erlbaum.
- Gast, D. L., & Ledford, J. R. (2014). *Single case research methodology: Applications in special education and behavioral sciences* (2nd ed.). New York, NY: Routledge.
- Hammond, D., & Gast, D. L. (2010). Descriptive analysis of single subject research designs: 1983–2007. *Education and Training in Autism and Developmental Disabilities*, 45, 187–202.
- Harrington, M., & Velicer, W. F. (2015). Comparing visual and statistical analysis in single-case studies using published studies. *Multivariate Behavioral Research*, 50, 162–183.
- Heyvaert, M., & Onghena, P. (2014). Analysis of single-case data: Randomisation tests for measures of effect size. *Neuropsychological Rehabilitation*, 24, 507–527.
- Heyvaert, M., Wendt, O., Van den Noortgate, W., & Onghena, P. (2015). Randomization and data-analysis items in quality standards for single-case experimental studies. *The Journal of Special Education*, 49, 146–156.
- 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 and Treatment of Children*, 35, 269–290.
- Hott, B. L., Limberg, D., Ohrt, J. H., & Schmit, M. K. (2015). Reporting results of single-case studies. *Journal of Counseling & Development*, 93, 412–417.
- Jenson, W. R., Clark, E., Kircher, J. C., & Kristjansson, S. D. (2007). Statistical reform: Evidence-based practice, meta-analyses, and single subject designs. *Psychology in the Schools*, 44, 483–493.
- Kelley, K., & Preacher, K. J. (2012). On effect size. *Psychological Methods*, 17, 137–152.
- Kennedy, C. H. (2005). *Single-case designs for educational research*. Boston, MA: Pearson.
- Kotrlik, J. W., Williams, H. A., & Jabor, M. K. (2011). Reporting and interpreting effect size in quantitative agricultural education research. *Journal of Agricultural Education*, 52, 132–142.
- Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2010). *Single-case designs technical documentation*. Retrieved from What Works Clearinghouse website: http://ies.ed.gov/ncee/wwc/Docs/ReferenceResources/wwc_scd.pdf
- Kratochwill, T. R., & Levin, J. R. (2010). Enhancing the scientific credibility of single-case intervention research: Randomization to the rescue. *Psychological Methods*, 15, 124–144.
- Kratochwill, T. R., & Levin, J. R. (2014). *Single-case intervention research. Methodological and statistical advances*. Washington, DC: American Psychological Association.
- Lane, J. D., & Gast, D. L. (2014). Visual analysis in single case experimental design studies: Brief review and guidelines. *Neuropsychological Rehabilitation*, 24, 445–463.
- Ledford, J. R., & Gast, D. L. (2014). Measuring procedural fidelity in behavioural research. *Neuropsychological Rehabilitation*, 24, 332–348.
- Ledford, J. R., & Wolery, M. (2013). Procedural fidelity: An analysis of measurement and reporting practices. *Journal of Early Intervention*, 35, 173–193.
- Levin, J. R., Ferron, J. M., & Gafurov, B. S. (2016). Comparison of randomization-test procedures for single-case multiple-baseline designs. *Developmental Neurorehabilitation*, <http://dx.doi.org/10.1080/17518423.2016.1197708> (Advance online publication)
- Lloyd, B. P., Finley, C. I., & Weaver, E. S. (2015). Experimental analysis of stereotypy with applications of nonparametric statistical tests for alternating treatments designs. *Developmental Neurorehabilitation*, <http://dx.doi.org/10.3109/17518423.2015.1091043> (Advance online publication)
- Luiselli, J. K. (2010). Writing for publication: A performance enhancement guide for the human services professional. *Behavior Modification*, 34, 459–473.
- MacCoun, R., & Perlmutter, S. (2015). Blind analysis: Hide results to seek the truth. *Nature*, 526, 187–189.
- Maggin, D. M., Briesch, A. M., & Chafouleas, S. M. (2013). An application of the What Works Clearinghouse standards for evaluating single-subject research: Synthesis of the self-management literature base. *Remedial and Special Education*, 34, 44–58.
- Maggin, D. M., Briesch, A. M., Chafouleas, S. M., Ferguson, T. D., & Clark, C. (2014). A comparison of rubrics for identifying empirically supported practices with single-case research. *Journal of Behavioral Education*, 23, 287–311.
- Manolov, R., Gast, D. L., Perdices, M., & Evans, J. J. (2014). Single-case experimental designs: Reflections on conduct and analysis. *Neuropsychological Rehabilitation*, 24, 634–660.
- Manolov, R., Jamieson, M., Evans, J. J., & Sierra, V. (2016). A discussion of alternatives for establishing empirical benchmarks for interpreting single-case effect sizes. *Psicológica*, 37, 209–234.
- Manolov, R., & Moeyaert, M. (2016). How can single-case data be analyzed? Software resources, tutorial, and reflections on analysis. *Behavior Modification*, <http://dx.doi.org/10.1177/0145445516664307> (Advance online publication)
- Manolov, R., Moeyaert, M., & Evans, J. J. (2016). *Single-case data analysis: Software resources for applied researchers*. Retrieved from https://www.researchgate.net/publication/289098041_Single-case_data_analysis_Software_resources_for_applied_researchers
- Marso, D., & Shadish, W. R. (2015). *Software for meta-analysis of single-case design: DHPS macro*. Retrieved from <http://faculty.ucmerced.edu/wshadish/software/software-meta-analysis-single-case-design>
- Miller, M. J. (1985). Analyzing client change graphically. *Journal of Counseling and Development*, 63, 491–494.
- Moeyaert, M., Maggin, D., & Verkuilen, J. (2016). Reliability, validity, and usability of data extraction programs for single-case research designs. *Behavior Modification*, 40, 874–900.
- McNicol, K., Salmon, P., Young, B., & Fisher, P. (2013). Alleviating emotional distress in a young adult survivor of adolescent cancer: A case study illustrating a new application of metacognitive therapy. *Clinical Case Studies*, 12, 22–38.
- Ninci, J., Vannest, K. J., Willson, V., & Zhang, N. (2015). Interrater agreement between visual analysts of single-case data: A meta-analysis. *Behavior Modification*, 39, 510–541.
- Onghena, P., & Edgington, E. S. (2005). Customization of pain treatments: Single-case design and analysis. *Clinical Journal of Pain*, 21, 56–68.

- Ottensbacher, K. J. (1990). When is a picture worth a thousand p values? A comparison of visual and quantitative methods to analyze single subject data. *Journal of Special Education*, 23, 436–449.
- O'Neill, B., & Findlay, G. (2014). Single case experimental designs in neurobehavioural rehabilitation: Preliminary findings on biofeedback in the treatment of challenging behaviour. *Neuropsychological Rehabilitation*, 24, 365–381.
- Parker, R. I., Brossart, D. F., Vannest, K. J., Long, J. R., Garcia De-Alba, R., 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. I., Cryer, J., & Byrns, G. (2006). Controlling baseline trend in single-case research. *School Psychology Quarterly*, 21, 418–443.
- Parker, R. I., & Vannest, K. J. (2009). An improved effect size for single-case research: Nonoverlap of all pairs. *Behavior Therapy*, 40, 357–367.
- Parker, R. I., Vannest, K. J., & Davis, J. L. (2011). Effect size in single-case research: A review of nine nonoverlap techniques. *Behavior Modification*, 35, 303–322.
- Parker, R. I., Vannest, K. J., Davis, J. L., & Sauber, S. B. (2011). Combining nonoverlap and trend for single-case research: Tau-U. *Behavior Therapy*, 42, 284–299.
- Pek, J., & Flora, D. B. (2017, March 9). Reporting effect sizes in original psychological research: A discussion and tutorial. *Psychological Methods*. Advance online publication. doi: 10.1037/met0000126.
- Perdices, M., & Tate, R. L. (2009). Single-subject designs as a tool for evidence-based clinical practice: Are they unrecognized and undervalued? *Neuropsychological Rehabilitation*, 19, 904–927.
- Pustejovsky, J. E. (2015). Measurement-comparable effect sizes for single-case studies of free-operant behavior. *Psychological Methods*, 20, 342–359.
- R Core Team. (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>
- Rapp, J. T., Carroll, R. A., Stangeland, L., Swanson, G., & Higgins, W. J. (2011). A comparison of reliability measures for continuous and discontinuous recording methods: Inflated agreement scores with partial interval recording and momentary time sampling for duration events. *Behavior Modification*, 35, 389–402.
- Romeiser-Logan, L., Slaughter, R., & Hickman, R. (2017 February 22). Single-subject research designs in pediatric rehabilitation: A valuable step towards knowledge translation. *Developmental Medicine & Child Neurology*, <http://dx.doi.org/10.1111/dmcn.13405>. Advance online publication.
- Schneider, A. B., Coddington, R. S., & Tryon, G. S. (2013). Comparing and combining accommodation and remediation interventions to improve the written-language performance of children with Asperger syndrome. *Focus on Autism and Other Developmental Disabilities*, 28, 101–114.
- Shadish, W. R., Hedges, L. V., & Pustejovsky, J. E. (2014). Analysis and meta-analysis of single-case designs with a standardized mean difference statistic: A primer and applications. *Journal of School Psychology*, 52, 123–147.
- Shadish, W. R., & Sullivan, K. J. (2011). Characteristics of single-case designs used to assess intervention effects in 2008. *Behavior Research Methods*, 43, 971–980.
- Shadish, W. R., Zelinsky, N. A. M., Vevea, J. L., & Kratochwill, T. R. (2016). A survey of publication practices of single-case design researchers when treatments have small or large effects. *Journal of Applied Behavior Analysis*, 49, 1–18.
- Shea, B. J., Grimshaw, J. M., Wells, G. A., Boers, M., Andersson, N., Hamel, C., . . . & Bouter, L. M. (2007). Development of AMSTAR: A measurement tool to assess the methodological quality of systematic reviews. *BMC Medical Research Methodology*, 7, 1. <http://dx.doi.org/10.1186/1471-2288-7-10>
- Skolasky, R. L., Jr. (2016). Considerations in writing about single-case experimental design studies. *Cognitive and Behavioral Neurology*, 29, 169–173.
- Smith, J. D. (2012). Single-case experimental designs: A systematic review of published research and current standards. *Psychological Methods*, 17, 510–550.
- Sternberg, R. J. (2003). *The psychologist's companion: A guide to scientific writing for students and researchers* (4th ed.). Cambridge, UK: Cambridge University Press.
- Tate, R. L., Perdices, M., McDonald, S., Togher, L., & Rosenkoetter, U. (2014). The conduct and report of single-case research: Strategies to improve the quality of the neurorehabilitation literature. *Neuropsychological Rehabilitation*, 24, 315–331.
- Tate, R. L., Perdices, M., Rosenkoetter, U., McDonald, S., Togher, L., & Vohra, S. (2016). The Single-Case Reporting Guideline In Behavioural Interventions (SCRIBE). Explanation and elaboration. *Archives of Scientific Psychology*, 4, 10–31.
- Tate, R. L., Perdices, M., Rosenkoetter, U., Wakima, D., Godbee, K., Togher, L., & McDonald, S. (2013). Revision of a method quality rating scale for single-case experimental designs and n-of-1 trials: The 15-item Risk of Bias in N-of-1 Trials (RoBiNT) Scale. *Neuropsychological Rehabilitation*, 23, 619–638.
- Tate, R. L., Rosenkoetter, U., Wakim, D., Sigmundsdottir, L., Doubleday, J., Togher, L., McDonald, S., & Perdices, M. (2015). *The risk-of-bias in N-of-1 trials (RoBiNT) scale: An expanded manual for the critical appraisal of single-case reports*. St Leonards, NSW: John Walsh Centre for Rehabilitation Research.
- Tunnard, C., & Wilson, B. (2014). Comparison of neuropsychological rehabilitation techniques for unilateral neglect: An ABACADAEAF single-case experimental design. *Neuropsychological Rehabilitation*, 24, 382–399.
- Valentine, J. C., Tanner-Smith, E. E., & Pustejovsky, J. E. (2016). *Between-case standardized mean difference effect sizes for single-case designs: A primer and tutorial using the scdhlw web application*. Oslo, Norway: The Campbell Collaboration. <http://dx.doi.org/10.4073/cmpn.2016.3>. Retrieved from https://campbellcollaboration.org/media/k2/attachments/effect_sizes_single_case_designs.pdf
- Vannest, K. J., Davis, D. L., & Parker, R. I. (2013). *Single case research in schools: Practical guidelines for school-based professionals*. New York, NY: Routledge.
- Wells, A. (1990). Panic disorder in association with relaxation induced anxiety: An attentional training approach to treatment. *Behavior Therapy*, 21, 273–280.
- Wilkinson, L., & The Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, 54, 694–704.
- Wolery, M., Dunlap, G., & Ledford, J. R. (2011). Single-case experimental methods: Suggestions for reporting. *Journal of Early Intervention*, 33, 103–109.
- Yakubova, G., & Bouck, E. C. (2014). Not all created equally: Exploring calculator use by students with mild intellectual disability. *Education and Training in Autism and Developmental Disabilities*, 49, 111–126.
- Zelinsky, N. A. M., & Shadish, W. R. (2016). A demonstration of how to do a meta-analysis that combines single-case designs with between-groups experiments: The effects of choice making on challenging behaviors performed by people with disabilities. *Developmental Neurorehabilitation*, <http://dx.doi.org/10.3109/17518423.2015.1100690> (Advance online publication)