Available online at www.sciencedirect.com

ScienceDirect

Behavior Therapy 48 (2017) 97–114

Behavior Therapy

www.elsevier.com/locate/bt

Methodological Paper

Recommendations for Choosing Single-Case Data Analytical Techniques

Rumen Manolov
University of Barcelona

Mariola Moeyaert
State University of New York

The current paper responds to the need to provide guidance to applied single-case researchers regarding the possibilities of data analysis. The amount of available single-case data analytical techniques has been growing during recent years and a general overview, comparing the possibilities of these techniques, is missing. Such an overview is provided that refers to techniques that yield results in terms of a raw or standardized difference and procedures related to regression analysis, as well as nonoverlap and percentage change indices. The comparison is provided in terms of the type of quantification provided, data features taken into account, conditions in which the techniques are appropriate, possibilities for meta-analysis, and evidence available on their performance. Moreover, we provide a set of recommendations for choosing appropriate analysis techniques, pointing at specific situations (aims, types of data, researchers' resources) and the data analytical techniques that are most appropriate in these situations. The recommendations are contextualized using a variety of published single-case data sets in order to illustrate a range of realistic situations that researchers have faced and may face in their investigations.

Keywords: single-case designs; data analysis; recommendations

Address correspondence to Rumen Manolov, Department of Behavioural Sciences Methods, University of Barcelona, Departament de Metodologia de les Ciències del Comportament, Facultat de Psicologia, Universitat de Barcelona, Passeig de la Vall d'Hebron, 171, 08035-Barcelona, Spain; e-mail: rrumenov13@ub.edu.

During the last decade there has been a great proliferation of data analytical techniques for single-case experimental designs (SCEDs) and an intensified discussion on the topic. A bibliographic search performed on September 8, 2015, via the PsycINFO database for years 2005–2014 using "single-case" or "single-subject" and "analysis" as keywords to be found in the abstract suggested the following number of papers: 3 in 2005 and 2006, 7 in 2007, 6 in 2008, 7 in 2009, 5 in 2010, 10 in 2011, 13 in 2012, 15 in 2013, and 35 in 2014. The amount of works (including papers, Ph.D. dissertations, and book chapters) that propose, test, or discuss SCED data analysis illustrates the current relevance of the topic. Despite this increased attention to SCED analysis, a common requirement made by SCED article reviewers and journal editors has been to provide concrete recommendations regarding connecting specific conditions (e.g., design and data characteristics, and purpose of the study) with appropriate SCED analytical techniques. In contrast to data analysis, guidelines for conducting SCEDs are already available in the form of rubrics and standards for assessing the methodological quality of SCED studies (see Maggin, Briesch, Chafouleas, Ferguson, & Clark, 2014; Smith, 2012, for reviews). A similar broad overview regarding SCED data is lacking and this is why we provide it here.

The current SCED data analysis situation is well illustrated by Waddell, Nassar, and Gustafson's (2011) statement that "the problem of how to statistically analyze the data . . . is perhaps the most confusing, daunting, and disjointed element of this experimental method" (p. 161). These authors also state that the amount of analytical techniques and

formulae makes the issue even more confusing. In the current paper, we offer an overview and tentative recommendations based on the idea that there is no single analytical technique that is optimal for all situations (aims, data features, researchers' resources), but that one data analytic technique may be more appropriate in certain conditions compared with another.

In Search for Criteria and Recommendations

Solid and updated state-of-the-art summaries of SCED analysis are expected in special issues of peer-reviewed journals. However, there are a few problems with the assumption that special issues might provide sound recommendations. First, the choice of focus of the special issue may not be based on the appropriateness of the techniques, but rather on (a) a desire to provide the full spectrum of possibilities, (b) the guest editors knowing some of the techniques or some of the authors better than others, and (c) the need to cover different topics as compared with previous special issues. Second, there may not be an explicit effort to point to the most appropriate analytical technique(s)—as each research team presents the techniques it has been working on and the guest editors might not be willing to act as judges—due to (a) lack of knowledge, or (b) lack of journal space for a formal public discussion with the authors of the different papers.[1] Third, it is possible to find different foci and recommendations in different special issues. Accordingly, an informal review of all the SCED data analysis special issues that we know of shows that some of the special issues pay more attention to techniques related to regression analysis (Shadish, 2014), whereas others focus on randomization tests (Vilardaga, 2014) or nonoverlap indices (Burns, 2012; Maggin & Chafouleas, 2013). Another group of special issues covers a variety of techniques (Evans, Gast, Perdices, & Manolov, 2014; Shadish, Rinsdskopf, & Hedges, 2008). Finally, two papers dealing with data analysis from special issues on SCED methodology ought to be mentioned. One of them (Vannest & Ninci, 2015) is focused on nonoverlap indices, whereas the other one (Gage & Lewis, 2013) reviewed several techniques before stressing the lack of agreement among researchers, stating that "a preference for standard mean difference, non-overlap, or regression-based approaches is also without empirical support" (p. 55). Thus, the lack of clear consensus (Kratochwill et al., 2010; Smith, 2012; Tate et al., 2013) and indications suggest that the current paper is necessary, as we consider that more discussion is needed apart from more research (Gage & Lewis, 2013).

Wolery, Busick, Reichow, and Barton (2010) were the first to suggest a set of criteria for SCED analytical procedures: (a) focus on the replication logic of SCED; (b) use all the data of the study; (c) estimate the magnitude of the effects across replications; (d) take into account all the characteristics of the data: level, trend, and variability; (e) show high agreement with careful visual analysis; (f) do not violate the assumptions about the nature of the data, such as serial dependency; and (g) have some method of allowing analyses of moderator variables. On the other hand, Manolov, Gast, Perdices, and Evans (2014) suggest that (a) the technique chosen should reflect the aim of the analysis (statistical significance versus effect size in a common metric versus unstandardized effect size); (b) the output of the analysis should be easy to interpret (includes whether the quantification provided is meaningful and whether there are any interpretative benchmarks available); (c) the analysis should be easy to compute (includes hand calculation and software availability and user-friendliness); (d) the technique must take into account design requirements and data assumptions (includes randomization, absence of trend, absence of serial dependence); and (e) the technique should be supported by evidence of appropriate performance with typical SCED data (includes both simulation studies and field tests with real data).

Additionally, it is possible to trace criteria closely associated to specific procedures. For instance, Kratochwill et al. (2013) suggest that an effect size should be comparable to the ones obtained in group design studies (in reference to the *d* statistics by Hedges, Pustejvosky, & Shadish, 2012, 2013; hereinafter referred to as HPS *d* statistics). Another criterion is that the analytical technique should not rely on rarely possible random sampling to ensure the validity of inferential results (Dugard, 2014), with randomization tests being a procedure that meets this criterion.

We also consider the following additional criteria relevant: (a) in relation to the general recommendations for reporting results in psychology (Wilkinson & Task Force on Statistical Inference, 1999), it could be useful for the technique to offer the possibility of constructing a confidence interval around the effect size estimate; (b) regarding design structures that meet evidence standards (Kratochwill et al., 2010), it is necessary that the procedure be easily extensible to designs beyond AB (which is related to criteria b an c; Wolery et al., 2010); (c) considering that visual analysis is commonly the initial step and sometimes the only step in data analysis (Perdices & Tate, 2009;

[1] It would not be ethical to invite a research team to submit a paper, review it, accept it for publication, and then publicly criticize the technique proposed/described without providing them the opportunity to respond.