Journal of School Psychology 52 (2014) 231–235

Contents lists available at ScienceDirect

Journal of School Psychology

journal homepage: www.elsevier.com/locate/jschpsyc

Commentary

Meta- and statistical analysis of single-case intervention research data: Quantitative gifts and a wish list

Thomas R. Kratochwill [a,*], Joel R. Levin [b]

[a] University of Wisconsin-Madison, USA
[b] University of Arizona, USA

ARTICLE INFO

Article history:
Received 29 January 2014
Accepted 29 January 2014
Available online 25 February 2014

Keywords:
Single-case research
Meta-analysis
What Works Clearinghouse Standards
Visual analysis
Effect size

ABSTRACT

In this commentary, we add to the spirit of the articles appearing in the special series devoted to meta- and statistical analysis of single-case intervention-design data. Following a brief discussion of historical factors leading to our initial involvement in statistical analysis of such data, we discuss: (a) the value added by including statistical-analysis recommendations in the What Works Clearinghouse *Standards* for single-case intervention designs; (b) the importance of visual analysis in single-case intervention research, along with the distinctive role that could be played by single-case effect-size measures; and (c) the elevated internal validity and statistical-conclusion validity afforded by the incorporation of various forms of randomization into basic single-case design structures. For the future, we envision more widespread application of quantitative analyses, as critical adjuncts to visual analysis, in both primary single-case intervention research studies and literature reviews in the behavioral, educational, and health sciences.

We applaud Editor Floyd's decision to invite a special issue of the *Journal of School Psychology* that is devoted to the quantitative analysis of single-case design (SCD) intervention research. The invitation was a timely one in that—as is well documented in this special issue—exciting new strategies for analyzing SCD-produced outcomes are emerging at a rapid rate. We also applaud special issue Editor Shadish's ability to assemble such a rich assortment of quantitative gifts to the field. The contributions are uniformly outstanding and have the potential to be enormously impactful for prospective SCD intervention researchers. Not to get lost in the special issue shuffle are (a) Shadish (2014–in this issue) masterful summary of the individual contributions and (2) his perceptive reflections on, and projections of, the additional wish list of meta- and statistical-analysis techniques, techniques that should be welcomed by SCD intervention researchers as having the potential to further advance the field. In our commentary, we primarily elaborate and expand on that wish list.

1. Introduction and reflections

For over 40 years, we have been involved in the discussion and debate surrounding the application of statistics in SCD intervention research. Our interest in the topic was crystallized in a 1973 advanced research seminar at the University of Wisconsin-Madison (with JRL as professor and TRK as graduate student), in which we raised concerns over the application of a conventional analysis of variance to then "single-subject" data from an ABAB design. Our response to an article by Gentile, Roden, and Klein (1972) in the *Journal of Applied Behavior Analysis* (Kratochwill et al., 1974), along with that of others in the applied behavior analysis field, initiated an ongoing debate about not only the appropriateness of certain statistical applications in SCD methodology but also whether any statistical (as opposed to visual/clinical) analysis should be conducted at all (see also Levin, Marascuilo, & Hubert, 1978, pp. 167–168, 193). Since that time, we have written on the single-case design-and-analysis topic through journal publications (e.g., Kratochwill & Levin, 1979) and edited books

* Corresponding author at: Department of Educational Psychology, 1025 W. Johnson St., University of Wisconsin, Madison, WI 53706, USA.
E-mail address: tomkat@education.wisc.edu (T.R. Kratochwill).
ACTION EDITOR: Randy Floyd.

0022-4405/$ – see front matter © 2014 Society for the Study of School Psychology. Published by Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.jsp.2014.01.003

(Kratochwill, 1978; Kratochwill & Levin, 1992, 2014). In addition, thanks to the sponsorship of the Institute of Education Sciences (IES), we (along with colleague Robert Horner) have offered three research-training institutes (2011–2013) targeting exclusively the design and analysis of SCD intervention research. The articles featured in this special issue of the *Journal of School Psychology* (hereafter referred to as the series) continue the dialog about the use of various statistical models in SCD intervention research, offer new options for data analysis, and extend the discussion of possibilities for calculating effect-size measures in both individual SCD studies and meta-analytic research syntheses.

As might be concluded from our introductory comments, there has been an exciting (and often excited) history related to the application of statistical methods in SCD intervention research. Aspects of this history are summarized in Shadish (2014–in this issue) lead article (see also Kratochwill & Levin, 1979; Kratochwill, Levin, Horner, & Swoboda, 2014; Swaminathan, Rogers, & Horner, 2014–in this issue). In our commentary, we focus on several issues that are either given minimal emphasis in the series articles or have not been considered in them. These issues include, among others, (a) the value added by including statistical-analysis recommendations in the What Works Clearinghouse *Standards* for SCD intervention designs; (b) the importance of visual analysis in SCD intervention research, combined with the vital contributions that could be made by accompanying statistical-analysis procedures and effect-size measures; and (c) the elevated internal validity and statistical-conclusion validity afforded by the incorporation of various forms of randomization and randomization tests into SCD investigations. We will proceed to discuss each of these issues in turn.

2. Standards for SCD statistical analysis

One of our most memorable encounters with statistical analysis of SCD intervention data was in helping to develop IES's What Works Clearinghouse (WWC) *Standards* for SCDs (Kratochwill et al., 2010). It became clear to the panel that we could reach consensus on and develop *design standards* for common classes of SCDs, as well as offer guidance on *evidence criteria* for SCD outcomes when visual analysis comprises the primary data-analysis tool. Our panel elected not to make recommendations regarding sensible or appropriate statistical analyses and effect-size measures, in part due to SCD researchers' traditional reliance on visual analysis, combined with a lack of agreement about what are currently regarded as the "best" statistical-analysis procedures and effect-size measures.[1]

So, what is left to be done in the way of prompting and promoting the use of quantitative analysis by SCD researchers? The appraisal protocols that have been developed for SCD literature reviews do not include guidelines for evaluating the researchers' statistical-analysis techniques—see reviews of the numerous appraisal rubrics for SCD research provided by Maggin, Briesch, Chafouleas, Ferguson, and Clark (2013), Smith (2012), and Wendt and Miller (2012). The present series and recent advances in several areas of SCD quantitative data analysis (see, for example, Kratochwill & Levin, 2014) are testimony to important developments that can now guide literature reviews of SCD research. The next step in this process is to formulate a panel for developing *statistical* evidence criteria for the analysis of SCD data. Such criteria could complement the guidelines advanced by Wilkinson and the Task Force on Statistical Inference (1999) for conventional group designs, in that they would take into account methods for both analyzing original SCD data and applying effect-size statistics for SCD research reviews. Given the increasing acceptance of SCD research in proposals submitted to IES, the National Science Foundation (NSF), and beyond, the formation of a panel of this kind should be a high—if not *the* highest—imminent priority.

3. Visual-analysis and effect-size issues

As we have noted throughout this commentary, one of the hallmarks of SCD research has been its reliance on visual analysis for making decisions about intervention effects—or in the preferred language of SCD researchers, decisions about whether there appears to be sufficient visual evidence for demonstrating a "functional relationship" between the intervention and the outcome measure(s). In that spirit, all of the just-referenced appraisal protocols rely on visual analysis in coding studies. Over the years, however, serious questions have been raised about SCD researchers' exclusive reliance on visual analysis, including those indicated by Shadish (2014–in this issue). Our review of research on visual inspection of SCD outcome data (Kratochwill et al., 2014) has precipitated a few major issues that need to be addressed.

First, there is considerable need for research on the *psychometric properties* of visual analysis—in particular, its reliability, validity, and accuracy. Despite over 50 years of work in this domain, there is a surprising lack of empirical support to bolster SCD researchers' confidence in these properties. In addition, much of the existing empirical evidence that does exist is based only on analog studies, rather than on studies representing SCDs that are implemented in actual intervention research contexts. Fortunately, visual-analysis researchers are beginning to recognize this gap and are assessing the psychometric properties of visual analysis in designs that better match those of research protocols published in contemporary scientific journals (e.g., Kahng et al., 2010).

Second, there is a need for the development and empirical testing of visual-analysis *training protocols*, including whether and how such training improves the reliability, validity, and accuracy of the method. In the past, for SCD studies in which visual-analysis training has been initially implemented, either the general procedures have not been described or the specific components of the process have not been specified.

[1] The current WWC *Standards* are similarly silent with respect to the benefits of SCD researchers incorporating various forms of randomization into their designs (see, for example, Ferron & Levin, 2014; Kratochwill & Levin, 2010). Advocating for design randomization schemes in the *Standards* would constitute an important contribution to elevating the internal validity and concomitant scientific credibility (Levin, 1994) of SCD intervention research.