Journal of School Psychology 52 (2014) 243–248

Contents lists available at ScienceDirect

Journal of School Psychology

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

Commentary

It has been said that, "There are three degrees of falsehoods: Lies, damn lies, and statistics"[☆]

Wayne W. Fisher[a,*], Dorothea C. Lerman[b]

[a] University of Nebraska Medical Center, Monroe-Meyer Institute, USA
[b] University of Houston-Clear Lake, USA

ARTICLE INFO

Article history:
Received 22 January 2014
Accepted 22 January 2014
Available online 22 February 2014

Keywords:
Single-case designs
Applied behavior analysis
Experimental analysis of behavior

ABSTRACT

Applied behavior analysis (ABA) researchers have historically eschewed population-based, inferential statistics, preferring to conduct and analyze repeated observations of each participant's responding under carefully controlled and manipulated experimental conditions using single-case designs (SCDs). In addition, early attempts to adapt traditional statistical procedures for use with SCDs often involved trade-offs between experimental and statistical control that most ABA researchers have found undesirable. The statistical methods recommended for use with SCDs in the current special issue represent a welcome departure from such prior suggestions in that the current authors are proposing that SCD researchers add statistical methods to their current practices in ways that do not alter traditional SCD practices. Further refinement and use of such methods would (a) facilitate the inclusion of research using SCDs in meta-analyses and (b) aid in the development and planning of grant-funded research using SCD methods. Collaboration between SCD researchers and statisticians, particularly on research that demonstrates the benefit of these methods, may help promote their acceptance and use in ABA.

1. Historical perspectives on quantitative analyses in behavior analysis

A common belief among applied behavior analysis (ABA) researchers is that statistics are often used in ways that obfuscate or even disguise the true nature of our observations. This belief is certainly not specific to ABA researchers, as suggested from the title of this commentary, a quote that has variously been attributed to Mark Twain; the twentieth-century British Prime Minister Benjamin Disraeli; the first Earl of Balfour, Arthur James Balfour; and others. However, statistics are merely tools, and like most tools, they are morally neutral; it is how one uses such tools that determines whether they are good or evil, helpful or harmful. Thus, despite the longstanding misgivings about inferential statistics that have been promulgated in ABA, we have attempted to review these articles from a functional standpoint. We have tried to determine whether and to what extent ABA researchers might advance their science and benefit from the use of the specific statistical procedures described in this special issue. We encourage other readers of this special issue to do likewise.

The experimental analysis of behavior (EAB) began with (and has largely continued) the pursuit of elucidating the general principles of behavior using both qualitative and quantitative analyses. Early experimental investigations systematically manipulated both qualitative schedule parameters (variable interval [VI] v. variable ratio [VR]) and quantitative schedule parameters (VI 30″ v. VI 60″) while precisely and quantitatively measuring the dependent variable of interest (e.g., measuring

☆ This research was supported by Grant Number AR100184 from the Autism Research Program, which is a component of the Congressionally Directed Medical Research Programs within the Department of Defense.
* Corresponding author at: Wayne W. Fisher, Center for Autism Spectrum Disorders, 985450 Nebraska Medical Center, Omaha, NE 68198-5450, USA.
E-mail address: WFisher@UNMC.Edu (W.W. Fisher).
ACTION EDITOR: Randy Floyd.

0022-4405/$ – see front matter
http://dx.doi.org/10.1016/j.jsp.2014.01.001

choice responding in terms of the relative rates of responding between two concurrently available response options; Herrnstein, 1961).

Over time, mathematical models have increasingly been used in EAB to describe and quantify general behavioral processes and also to refine and test competing theoretical accounts of behavioral phenomena (for reviews, see Nevin, 1984; Mazur, 2006). Examples of major advances in EAB facilitated by the application of mathematical models include the use of (a) the generalized matching law to predict and explain how individuals allocate their time and responses in proportion to the relative rates or amounts of reinforcement produced by each response option (Baum, 1974); (b) behavioral economics principles (e.g., demand curves) to better understand how the price of one reinforcer affects the consumption of similar (or substitutable) reinforcers (Hursh, 1980); (c) delay-discounting models to predict how the subjective value of a reinforcer decreases with increasing delays to its delivery, which helps explain a wide variety of impulse–control and addictive disorders (Reynolds, 2006); (d) behavioral momentum theory to model and predict the extent to which the rate of a response will persist when it contacts a disrupter (e.g., Nevin, 1992), which has the potential to predict (and potentially lead to ways of preventing) many forms of treatment relapse (Podlesnik & Shahan, 2009); and (e) quantitative models of discriminated operant behavior that may predict and account not only for the effects of antecedent stimuli but also for the effects of attending and forgetting (e.g., Nevin, Davison, Odum, & Shahan, 2007).

Paralleling this increase in the use of mathematical models of behavior in EAB has been an increase in the use of traditional statistics to quantify response variability, statistical significance, and effect sizes (e.g., the generalized matching law often accounts for about 90% of the response variance observed in concurrent-choice arrangements using a generalized linear model). Although research in ABA also has seen an increase in the application of mathematical models and statistical analyses, ABA has lagged well behind EAB in the use of such methods. For example, we conducted a search of PDF versions of the 19 research articles published in the Fall 2012 edition of the *Journal of Applied Behavior Analysis* (*JABA*) and the 19 research articles published in the September and November editions of the *Journal of the Experimental Analysis of Behavior* (*JEAB*) using the following string characters representing statistical terms: “standard dev”, “variance”, “significan”, “ANOVA”, “GLM”, “p <”, “p =”, “p >”, and “z-score.” A second investigator independently searched 35% of the PDF files, and the two searches produced identical results on this subset of files (i.e., 100% agreement). The searches in *JABA* identified just one article (5.3%) that included one or more of these statistical-term character strings, whereas the searches in *JEAB* identified at least one such string in 17 articles (89.5%). This difference is obviously both clinically and statistically significant ($Z = 5.2$; $p < .001$). But does this difference imply that ABA researchers should markedly increase their application of mathematical models of behavioral processes, traditional statistical analyses, or both?

A fundamental distinction between the mathematical formulations used to model general behavioral processes in EAB and traditional statistical models is that the former models consist primarily of mathematical parameters that are closely tied to the behavioral processes of interest (e.g., B_1, r_1, B_2, and r_2 representing the response and reinforcement rates for Responses 1 and 2, respectively, in the generalized matching law), whereas the latter models are largely devoid of theoretical or conceptual relevance (e.g., they can be applied equally well to data sets that are designed to test for a functional relation between a behavioral intervention and its target response or to changes over time in attitudes about marriage equality). Each type of model has its advantages. Perhaps the most important advantage of the mathematical models typically used in EAB is that they can lead to specific, testable hypotheses about behavioral processes. Applied behavior analysts would probably benefit from the increased use of such mathematical models in the design and testing of behavioral interventions (for an illustration, see Nevin & Shahan, 2011).

2. Potential uses for statistical analyses in ABA

Shadish (2014–this issue) identified a number of potential advantages of applying traditional statistical methods to single-case designs (SCDs), including (a) the ability to model and determine the extent to which observed changes in behavioral patterns may have occurred by chance, (b) the provision of information about how outcomes from SCDs fit with various statistical distributions (e.g., normal, binomial, and Poisson), (c) testing the accuracy of visual inspection (rates of Type I and II errors) against data sets drawn from distributions with defined characteristics using Monte Carlo methods (cf. Fisher, Kelley, & Lomas, 2003); (d) potentially more accurate and less biased summarization of findings across participants and studies, and (e) the formalization and more precise quantification of SCD data characteristics (e.g., changes in level, trend, and stability) used to determine the presence or absence of a functional relation between the independent and dependent variables. Of these, the summarization of findings across studies (e.g., meta-analyses) may be the most consequential for ABA researchers.

That is, the development of a common statistical metric for estimating effect sizes for both group-comparison and SCDs may increase the likelihood that research involving SCDs will be included in meta-analyses designed to determine whether the effectiveness of a particular behavioral intervention is supported by empirical research. A prime example of how this can impact the field of ABA involves meta-analyses on the effectiveness of early intensive behavioral interventions (EIBI) for young children with autism. Warren et al. (2011) reviewed the extant literature on EIBI but included only articles using group-comparison methods and only those involving at least 10 participants. Based on this review, they concluded that the strength of evidence for this intervention was low. By contrast, the National Autism Center (2009) reviewed both group-comparison studies and SCDs using well-defined criteria and concluded that the effectiveness of EIBI was “established,” which represented the highest level on their strength-of-evidence scale. The stark differences between the findings of these two meta-analysis reports on the same intervention highlight the importance of including SCDs in summative reviews of the effectiveness of behavioral interventions.