

Abstract Title Page
Not included in page count.

Title: Synthesis of single-case experimental data: A comparison of alternative multilevel approaches

Authors and Affiliations:

Ferron, John¹; Van den Noortgate, Wim²; Beretvas, Tasha³; Moeyaert, Mariola²; Ugille, Maaïke²; Petit-Bois, Merlande¹; Baek, Eun Kyeng¹

¹ University of South Florida

² Katholieke Universiteit Leuven, Belgium

³ University of Texas at Austin

Acknowledgment:

This study is performed with a grant of the Institute of Education Sciences (IES Grant number R305D110024)

Abstract Body

Limit 4 pages single-spaced.

Background / Context:

Description of prior research and its intellectual context.

Single-case or single-subject experimental designs (SSED) are used to evaluate the effect of one or more treatments on a single case. The basic interrupted time series design has a baseline phase consisting of a series of observations prior to treatment introduction, and a treatment phase consisting of a series of observations following treatment introduction.

Although SSED studies are growing in popularity, the results are in theory case-specific. To enhance generalizability, researchers can replicate the design across cases, either across studies, or within a primary study such as through the application of a multiple-baseline design or a replicated ABAB design. By synthesizing SSED studies' results, we can investigate the overall effect of an intervention, explore the generalizability of this effect, and look for factors that moderate the effect.

One systematic and statistical approach for combining single-case data within and across studies is multilevel modeling (Nugent, 1996; Shadish & Rindskopf, 2007; Van den Noortgate & Onghena, 2003a, 2003b, 2008). Use of the multilevel modeling framework provides an appealing option because it can be used to provide estimates of individual treatment effects and how these effects change over time, estimates of the average treatment effect over subjects and how this effect changes over time, estimates of the variability in treatment effects, and estimates of the effects of moderators on the treatment effect and on the pattern of a treatment's effects over time. In addition, the models are flexible enough to handle (a) the nesting of observations and of outcomes within cases and the nesting of cases within studies, (b) a variety of forms for the growth trajectory within each phase of the design (e.g., linear, curvilinear), (c) alternative phase structures (e.g., AB, ABA, ABAB), (d) dependent error structures for the growth trajectories (e.g., first order autoregressive, Toeplitz), (e) heterogeneous variances (within cases, across cases, or across studies), (f) different types of outcomes (e.g., continuous, count), and (g) standardized or unstandardized raw data or effect size measures.

Purpose / Objective / Research Question / Focus of Study:

Description of the focus of the research.

The purpose of the study is to investigate the applicability of the multilevel approach for analyzing the data from a study by Lambert, Cartledge, Heward, and Yo (2006), which is being used as a common data set in this symposium to explore the applicability of a range of different approaches to analyzing single-case data. Within the multilevel modeling approach a variety of methodological questions will typically arise. For this data set, one could ask: Should trends be included? Should the first AB pair of the replicated ABAB design be modeled differently than the second AB pair, and how can we code the phase indicators and time variables to give the model parameters a meaningful interpretation? Should the outcome be treated as continuous or as a proportion? Should the variance be modeled to differ across phases? Should autocorrelation be modeled? Should differences between classrooms be modeled? Which method should be used to standardize the effect estimate? Like other single-case data sets, this one is small in size

and thus contains limited information on which to base the methodological choices. In addition, the small size of the data set limits the number of potential complexities that can be modeled. As a consequence, a series of methodological choices (or simplifying assumptions) have to be made. Part of our objective in analyzing this data set was to estimate the effect size under a variety of plausible methodological choices, and then to examine the sensitivity of the effect estimate to the choices made. With the obtained study results, we want to inform applied researchers about possibilities and limitations of the use of the multilevel model for combining SSED data. At the same time, the results will give indications about conditions for which the model or the estimation procedures should be further developed.

Setting:

Description of the research location.

(May not be applicable for Methods submissions)

Not Applicable

Population / Participants / Subjects:

Description of the participants in the study: who, how many, key features, or characteristics.

(May not be applicable for Methods submissions)

Not Applicable

Intervention / Program / Practice:

Description of the intervention, program, or practice, including details of administration and duration.

(May not be applicable for Methods submissions)

Not Applicable

Significance / Novelty of study:

Description of what is missing in previous work and the contribution the study makes.

Although the multilevel modeling approach and its flexibility are appealing, there is much about SSED data and the functioning of multilevel modeling with this type of data that is not fully understood. Much of the previous work has focused on the statistical functioning of the approach with simulated data sets (Ferron, Bell, Hess, Rendina-Gobioff, & Hibbard, 2009; Ferron, Farmer, & Owens, 2010; Moeyaert, Ugille, Ferron, Beretvas, & Van den Noortgate, in press). Results suggest that average effect estimates are unbiased, that interval estimates of these average effects are accurate, as are the interval estimates of individual effects, as long as the Satterthwaite or Kenward-Roger approaches are used, but that variance components tend to be biased. Although the results are in many ways promising, these studies are limited by the congruency between the models used to generate the data and the analyses conducted. In this study of a real data set, the generating model is not known and thus the possibility of misspecification is a concern. By estimating the effect across multiple plausible multilevel models, and a variety of reasonable methodological choices, we will be able to provide some limited, but practical, information to applied researchers about how much the effect estimates may depend on these methodological decisions.

Statistical, Measurement, or Econometric Model:

Description of the proposed new methods or novel applications of existing methods.

The models that will be investigated in this study are extensions of the multilevel model of Van den Noortgate and Onghena (2003a). In the basic model, the observed scores for case j are regressed on a dummy variable (D_{ij}) indicating the treatment condition ($0 = A, 1 = B$):

$$Y_{ij} = \beta_{0j} + \beta_{1j}D_{ij} + e_{ij} \quad (1)$$

The equation shows that the expected score in the baseline phase equals β_{0j} , while it is $\beta_{0j} + \beta_{1j}$ in the treatment phase, and thus β_{1j} can then be interpreted as the intervention effect for the j^{th} participant. The errors e_{ij} are assumed to be distributed normally with the covariance matrix $\sigma^2\mathbf{I}$. At the second level of the model, the variation over cases is described using two equations:

$$\begin{aligned} \beta_{0j} &= \theta_{00} + u_{0j} \\ \beta_{1j} &= \theta_{10} + u_{1j} \end{aligned} \quad \text{with } \mathbf{u} \sim N(0, \Omega_u) \quad (2)$$

The first equation indicates that the baseline performance for participant j equals an overall baseline performance, plus a random deviation from this mean; the subsequent equation indicates that the treatment effect for participant j equals an overall average treatment effect, plus a random deviation from this mean.

This basic model is then extended a number of ways. Instead of assuming Y is continuous with an error term that is normally distributed, Y is treated as a count out of n trials ($n = 10$ for this study), where $Y/10$ can be modeled using a binomial distribution and a logit link function. Other extensions include the loosening of the level-1 error structure to assume a first-order autoregressive structure, as opposed to $\sigma^2\mathbf{I}$, the addition of a classroom effect to the level 2 equations, and the addition of effects to the level-1 equation to obtain separate treatment effect estimates for the first and second AB pairing and to add trends within each phase. In addition, standardized regression coefficients from analyses of each individual time series can be used as standardized effect estimates and be combined in a multi-level meta-analytic model (Ugille, Moeyaert, Beretvas, Ferron, & Van den Noortgate, in press; Van den Noortgate & Onghena, 2003b, 2008). Furthermore, these standardized effect size estimates can be adjusted to also include between participant variability by multiplying the standardized effect by:

$$\sqrt{\frac{\hat{\sigma}_{\text{within}}^2}{\hat{\sigma}_{\text{within}}^2 + \hat{\sigma}_{\text{between}}^2}} \quad (3)$$

where the within- and between-participant variance estimates are obtained from the multilevel model.

Usefulness / Applicability of Method:

Demonstration of the usefulness of the proposed methods using hypothetical or real data.

The data used for the re-analysis was originally gathered in two urban fourth grade classrooms (Lambert, Cartledge, Heward, & Yo, 2006). The participants in the original study were nine fourth grade students who were nominated by their teachers as the most disruptive and least attentive. Two conditions were compared in the original study using a replicated ABAB design.

The A condition was a single-student responding condition in which the teacher called on one student that raised his or her hand, whereas the B condition allowed each student to respond to all teacher questions by writing answers on a laminated board, or response card.

Using Equations 1 and 2 to model the number of intervals with disruptive behavior using restricted maximum likelihood estimation and the Kenward-Roger approach to estimating degrees of freedom and standard errors, the estimated average treatment effect was -5.45, and when the effect was estimated by modeling the number of intervals with disruptive behavior out of 10 using a binomial distribution with a logit link function and transforming the parameters from the logit scale to the raw score scale the effect estimate was -5.45. When autocorrelation was modeled among the level-1 errors, the effect estimate from the models treating Y as continuous with a normal distribution and from modeling Y/10 with a binomial distribution were -5.37 and -5.36, respectively. Additional models along with standardized estimates and more complete results will be included in the presentation, as will the individual effect estimates for the nine participants.

Research Design:

Description of the research design (e.g., qualitative case study, quasi-experimental design, secondary analysis, analytic essay, randomized field trial).

(May not be applicable for Methods submissions)

Not Applicable

Data Collection and Analysis:

Description of the methods for collecting and analyzing data.

(May not be applicable for Methods submissions)

Not Applicable

Findings / Results:

Description of the main findings with specific details.

(May not be applicable for Methods submissions)

Not Applicable

Conclusions:

Description of conclusions, recommendations, and limitations based on findings.

Multilevel models are flexible enough to handle many of the complexities that arise in analyzing single-case data, and thus we expect them to be of general interest to single-case researchers. Research, such as this, needs to continue to document the use of multilevel models with single-case data, the methodological decisions that arise in such applications, and the sensitivity of the effect estimates to these decisions. Furthermore, simulation research on multilevel models for single-case data needs to extend the examination of the statistical functioning of these models to a wider array of single-case design and data conditions.