Abstract Title Page
Not included in page count.

Title: The Synthesis of Single-Subject Experimental Data: Extensions of the Basic Multilevel Model

Authors and Affiliations:

Van den Noortgate, Wim[1]; Moeyaert, Mariola[1]; Ugille, Maaike[1], Beretvas, Tasha[2]; Ferron, John[3]

[1] Katholieke Universiteit Leuven, Belgium
[2] University of Texas at Austin
[3] University of South Florida

Wim.vandennoortgate@ppw.kuleuven.be (contact)
Mariola.moeyaert@ppw.kuleuven.be
Maaike.Ugille@ppw.kuleuven.be
Tasha.Beretvas@mail.utexas.edu
Ferron@usf.edu

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.

Due to an increasing interest in the use of single-subject experimental designs (SSEDs), appropriate techniques are needed to analyze this type of data. Regression analysis (Center, Skiba, & Casey, 1985-1986), and ANOVAs (Gentile, Roden, & Klein, 1972) - amongst other models - have been proposed as ways to estimate and/or test the treatment effect per case. However, (single-case) researchers may not only be interested in the effect of a treatment for a specific case, but also in the effect for a broader range of cases. In order to examine the external validity, single-case researchers often replicate their single-subject experiment within a study, such as in the multiple-baseline design, or across studies. Multiple-baseline designs across participants are the most popular designs amongst the SSEDs (Shadish & Sullivan, 2011) because they are characterized by a high internal validity (i.e., the introduction of a treatment is staggered across time, which allows disentangling changes in outcome scores due to some external event from real treatment effects, Barlow & Herson, 1984; Barlow, Nock, & Hersen, 2009). These designs are characterized by a hierarchical structure since measurement occasions are nested within cases. Ignoring this structure can lead to misleading treatment effect estimates and their standard errors. Therefore the use of a two-level regression model, which is a simple extension of the linear regression model has been proposed (Van den Noortgate & Onghena, 2003) and validated using simulation studies (Ferron, Bell, Rendina-Gobioff, & Hibbard, 2009; Ferron, Farmer, & Owens, 2010). However, as many multiple-baseline design studies' results are published and because some of them have the same research purpose, there is a good reason to estimate the treatment effect not only within a study, but also across studies. This need motivates use of the three-level model suggested by Van den Noortgate and Onghena (2003a, 2003b, and 2008) and validated using intensive simulation studies (Owens & Ferron, 2010, Moeyaert, Ugille, Ferron, Beretvas, & Van den Noortgate, 2013a, 2013b; Ugille, Moeyaert, Beretvas, Ferron, & Van den Noortgate, 2012). Because previous research has only focused on the basic two and three-level model, making a lot of assumptions (e.g., linear trends, continuous outcomes, independent, identically and normally distributed errors at the different levels), a lot of methodological questions and challenges remain. Because of the flexibility of the multilevel model, we will extend the basic two- and three-level model which enables analyzing more complex and realistic SSED data.

Purpose / Objective / Research Question / Focus of Study:

Description of the focus of the research.

The purpose of this paper proposal is to present four studies (Beretvas, Hembry, Van den Noortgate, & Ferron, 2013; Bunuan, Hembry & Beretvas, 2013; Moeyaert, Ugille, Ferron, Beretvas, & Van den Noortgate, 2013c; Petit-Bois, Baek & Ferron, 2013) each studying empirically one particular extension of the basic two-level or basic three-level model that can be used to synthesize SSED data across cases (i.e., two-level model) and across cases and studies (three-level model). Each extension is discussed and examined through intensive simulation studies by looking at the bias and precision of the parameter estimates, and the validity of statistical inferences. The results indicate whether the extended two- and three-level models are recommended to summarize SSEDs and in which conditions. This is a timely issue, since the number of published SSEDs is increasing rapidly together with the need for adequate techniques to analyze and meta-analyze SSEDs answering the same underlying research question.

Significance / Novelty of study:

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

Although the basic two-level and three-level model have been validated and extensively studied, a lot of methodological issues remain. Therefore, in this paper, we would like to present and discuss four extensions of the basic two- and three-level model. Each extension is validated through simulation studies and is an original contribution to the field of multilevel modeling of raw SSED data. In a first part, we present two extensions of the basic two-level model. A first extension (i.e. study 1) involves the modeling of dependent (or autocorrelated) level-1 error structures and in a second extension (i.e., study 2), a model that can be used for non-linear trajectories is investigated and validated. In a second part, we discuss two extensions of the basic three-level model (i.e., study 3 and study 4 respectively). In study 3, we study standardization in situations where the within-case variance is heterogeneous and in a fourth study we investigate the covariance matrix specification at the second and third level of the multilevel model.

Statistical, Measurement, or Econometric Model:

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

Part 1: Two-Level Modeling

In a first part, we present two extensions of the basic two-level model. The basic two-level model is a simple extensions of a regression equation in which the observed scores, *i*, within case *j* are regressed on a dummy variable, D_{ij}, indicating the treatment phase (D_{ij} equals zero if measurement *i* within case *j* belongs to the baseline, one otherwise):

$$Y_{ij} = \beta_{0j} + \beta_{1j}D_{ij} + e_{ij} \qquad e_{ij} \sim N(0, \sigma_e^2) \tag{1}$$

β_{0j} indicates the expected baseline level and the coefficient β_{1j} can be interpreted as the effect of the treatment on the outcome. At the second level, the coefficients of Equation 1 are allowed to vary over cases:

$$\begin{aligned} \beta_{0j} &= \theta_{00} + u_{0j} \\ \beta_{1j} &= \theta_{10} + u_{1j} \end{aligned} \qquad \begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{u_0}^2 & \sigma_{u_0 u_1} \\ \sigma_{u_1 u_0} & \sigma_{u_1}^2 \end{bmatrix}\right) \tag{2}$$

Single-case researchers are especially interested in the estimate of θ_{10}, indicating the average estimated treatment effect across cases and in the estimate of $\sigma_{u_1}^2$, referring to the between-case variance of this estimated treatment effect.

Study 1: Modeling dependencies among the within-case errors. In a first extension, the within-case errors, presented by e_{ij} in Equation 1, were modeled to follow a first-order autoregressive process. The level-1 covariance structure, as shown in Equation 3 for four points in time, implies that errors that are closer together in time covary more strongly than errors that are separated further in time, a pattern that is often reasonable for behavioral data (Baek & Ferron, 2013).

$$\sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{bmatrix} \tag{3}$$

Similar to the basic two-level model, the level-1 regression coefficients are allowed to vary at the second level.

Study 2: Modeling non-linear trajectories. In a second extension, the focus of interest is the modeling of non-linear trends and therefore, the regression equation at the first level looks as follows:

$$Y_{ij} = \beta_{0j} + e_{A,ij}(1 - D_{ij}) \left\{ \frac{f - \beta_{0j}}{1 + \exp\left[-a(T_{ij} - b)\right]} + e_{B,ij} \right\} D_{ij}$$

During the treatment phase, the model follows a logistic curve where *f* is the upper right asymptote, *a* is the rapidity with which the trajectory accelerates toward the upper right asymptote, and *b* is the location of the steepest slope for the trajectory. The coefficient, β_{0j} is allowed to vary at the second level.

Part 2: Three-Level Modeling

In the second part of this study, we discuss two possible extension of the basic three-level model in which a time trend is modeled. The model that will be investigated in this study is a multilevel extension of the model of Center, et al. (1985-1986). More specifically, the observed scores for case *j* from study *k* are regressed on a time indicator, T_{ijk}, that is centered around the first observation of the intervention phase, a dummy variable, D_{ijk}, for the treatment phase, and an interaction term of these variables, $T_{ijk}D_{ijk}$:

$$Y_{ij} = \beta_{0jk} + \beta_{1jk}T_{ijk} + \beta_{2jk}D_{ijk} + \beta_{3jk}T_{ijk}D_{ijk} + e_{ijk} \quad (4)$$

β_{0jk} indicates the expected baseline level at the start of the treatment phase (when T_{ijk} = 0), β_{1jk} is the linear trend in the baseline scores. The coefficient β_{2jk} can then be interpreted as the immediate effect of the intervention on the outcome, whereas β_{3jk} gives an indication of the effect of the intervention on the trend. These four coefficients are allowed to vary at the second and at the third level.

Study 3: Standardization when heterogeneous within-case variance is observed. Previous research validated the standardizing method suggested by Van den Noortgate and Onghena (2008) to standardize SSED data in contexts where the within-case variance is homogeneous (Moeyaert et al., 2013b). The standardizing method involves dividing the raw SSED data by the estimated within-case standard deviation. However, it is more reasonable to assume heterogeneous within-case variance and then the question rises which residual within-case standard deviation to use to standardize (i.e., pooled within-case standard deviation, within-baseline standard deviation, etc.). In this third study, we suggest three standardizing methods and present the empirical validation.

Study 4: covariance specification at the second and the third level. In previous research, the covariance at the second and third level was not estimated. However, it is reasonable that the coefficients modeled at the second and third levels covary. Therefore, at the second and third level, we estimate covariance between each pair of regression coefficients (i.e., covariances among the errors within the level). In this study, we also examine the consequences of misspecification of the covariance matrix on the estimated average treatment effects across cases and across studies.

Research Design:

Description of the research design.

(May not be applicable for Methods submissions)

The extensions of the basic multilevel models were each evaluated by means of an extensive simulation study. More specifically, we simulated data for MBD studies with starting points of the treatment phase staggered across the series. Based on a study of Shadish and Sullivan (2011), a survey of multiple baseline studies by Ferron, Farmer, and Owens (2010), and meta-analyses of SSED data (including, e.g., Alen, Grietens, & Van den Noortgate, 2009; Denis, Van den Noortgate, & Maes, 2011; Kokina & Kern, 2010; Shogren, Fagella-Luby, Bae, & Wehmeyer, 2004; Swanson & Sachse-Lee, 2000; Wang, Cui, & Parrila, 2011), we varied a number of variables. The chosen variables depend on the specific study. For instance, in studies 3 and 4, more variables were varied in comparison to studies 1 and 2, such as the number of studies, the between-study variance and other study-specific variables.

Data Collection and Analysis:
Description of the methods for collecting and analyzing data.
(May not be applicable for Methods submissions)
Data are simulated and analyzed in SAS, using the restricted maximum likelihood estimation procedure implemented in the procedure MIXED for multilevel or mixed models and using MCMC estimation – using SAS's MCMC procedure - for the nonlinear model study.

Findings / Results:
Description of the main findings with specific details.
(May not be applicable for Methods submissions)
In the first study, treatment effect estimates were found to be unbiased (and the inferences accurate) when the level-1 error structure was assumed to be first-order autoregressive and the data were generated based on either a first-order autoregressive model, a moving average model, or an independence model. In the second study, the use of a logistic function for nonlinear treatment phase trajectories was investigated. This study supports use of the logistic function to describe MBD trajectories when the data exhibit upper asymptotes during the intervention phase. The study also supports use of the simpler change in level model and of a model with a quadratic function when the series are sufficiently long. The third study, in which several methods for standardizing data from MBDs across participants and settings are investigated in scenarios with heterogeneous phase variances, shows that one standardizing method clearly works better in terms of fixed effect parameter recovery. In the final study, possible covariances among the level-2 errors, and among the level-3 errors, were considered and again fixed effects were unbiased, but under some conditions problems arose in the fixed effect inferences.

Conclusions:
Description of conclusions, recommendations, and limitations based on findings.
In this proposal, we presented four studies that are part of a larger study on combining the results of SSED data (funded by IES Grant R305D110024). We conclude that the multilevel model is flexible enough to be extended and is found to be appropriate for combining SSEDs. However, there are still a lot of other extensions that are not discussed here that are being investigated and that still remain to be investigated such as other estimation procedures (i.e., bootstrapping and Bayesian estimation), count outcomes, and multiple dependent outcome variables. Some of these topics will be shortly mentioned during our presentation.

Appendices

Not included in page count.

Appendix A. References

References are to be in APA version 6 format.

Alen, E., Grietens, H., & Van den Noortgate, W. (2009). Meta-analysis of single-case studies: An illustration for the treatment of anxiety disorders. Unpublished manuscript.

Baek, E., & Ferron, J. (2013). Multilevel models for multiple-baseline data: Modeling across participant variation in autocorrelation and residual variance. Behavior Research Methods, *45*, 65-74.

Barlow, D.H., Nock, M., & Hersen, M. (Eds.). (2009). *Single case experimental designs: strategies for studying behavior change.* Boston: Pearson Education.

Barlow, D. H., & Hersen M., H. (1984). *Single case experimental designs: Strategies for studying behavior change.* (2nd ed. ed.). New York: Pergamon.

Beretvas, S.N. Hembry, I., Van den Noortgate, W., & Ferron, J. (2013). *Estimation of a nonlinear intervention phase trajectory for multiple baseline design data*. Manuscript submitted for publication.

Bunuan, R., Hembry, I. & Beretvas, S.N. (2013). *Meta-analysis of single-case design studies with multiple baselines across participants and settings*. Manuscript in preparation.

Denis, J., Van den Noortgate, W., & Maes, B. (2011b). Self-injurious behavior in people with profound intellectual disabilities: A meta-analysis of single-case studies. *Research in Developmental Disabilities*, *32*, 911-923.

Ferron, J. M., Bell, B. A., Hess, M. F., Rendina-Gobioff, G., & Hibbard, S. T. (2009). Making treatment effect inferences from multiple-baseline data: the utility of multilevel modeling approaches. *Behavior Research Methods, 41*, 372-384.

Ferron, J. M., Farmer, J. L., & Owens, C. M. (2010). Estimating individual treatment effects from multiple-baseline data: A Monte Carlo study of multilvel-modeling approaches. *Behavior Research Methods, 42*, 930-943.

Gentile, J., Hoden, A., & Klein, R. (1972). An analysis-of-variance model for the intrasubject replication design. *Journal of Applied Behavior Analysis, 5*, 193-198.

Heyvaert, M., Maes, B., Van den Noortgate, W., Kuppens, S., & Onghena, P. (2011). A multilevel meta-analysis of single-case and small-n research on interventions for reducing challenging behavior in persons with intellectual disabilities. Manuscript submitted for publication.

Kokina, A., & Kern, L. (2010). Social story interventions for students with autism spectrum disorders: a meta-analysis. *Journal of Autism and Developmental Disorders*, *40*, 812-826.

Owens, C. M., & Ferron, J. M. (2010). Synthesizing single-case studies: A Monte Carlo examination of a three-level meta-analytic model. *Behavior Research Methods*.

Petit-Bois, M., Baek, E. K., & Ferron, J. M. (2013, April). *The effect of error structure specification on the meta-analysis of single-case studies: A Monte Carlo study*. Poster presented at the American Educational Research Association Annual Meeting, San Francisco, CA.

Shadish, W. R., & Sullivan, K. J. (2011). Characteristics of single-case designs used to assess intervention effects in 2008. *Behavior Research Methods*. Retrieved from http://www.springerlink.com/content/f3r091184t30757j/fulltext.pdf

Shogren, K. A., Fagella-Luby, M. N., Bae, J. S., & Wehmeyer, M. L. (2004). The effect of choice-making as an intervention for problem behavior. *Journal of Positive Behavior Interventions*, *6*(4), 228-237.

Swanson, H. L., & Sachse-Lee, C. (2000). A meta-analysis of single-subject-design intervention research for students with LD. *Journal of Learning Disabilities*, *33*, 114-136.

Ugille, M., Moeyaert, M., Beretvas, N., Ferron, J., & Van den Noorgate, W. (2013). Multilevel meta-analysis of single-subject experimental designs: A simulation study. *Behavior Research Methods*, *44*, 1244-1254.

Moeyaert, M., Ugille, M., Ferron, J., Beretvas, S. N., & Van den Noortgate, W. (2013a). *Three-level analysis of single-case experimental data: Empirical validation*. The Journal of Experimental Education. doi: 10.1080/00220973.2012.745470

Moeyaert, M., Ugille, M., Ferron, J., Beretvas, S. N., & Van den Noortgate, W. (2013b). Three-level analysis of standardized single-case experimental data: Empirical validation. *Multivariate Behavior Research*, *48*, 719-748.

Moeyaert, M., Ugille, M., Ferron, J., Beretvas, S. N., & Van den Noortgate, W. (2013c). *The misspecification of the covariance matrix in the three-level modeling of single-cases*. Manuscript submitted for publication.

Van den Noortgate, W., & Onghena, P. (2003). Combining single-case experimental data using hierarchical linear models. *School Psychology Quarterly, 18,* 325-346.

Van den Noortgate, W., & Onghena, P. (2008). A multilevel meta-analysis of single-subject experimental design studies. *Evidence Based Communication Assessment and Intervention, 2,* 142-151.

Wang, S., Cui, Y., & Parrila, R. (2011). Examining the effectiveness of peer-mediated and video-modeling social skills interventions for children with autism spectrum disorders: a meta-analysis in single-case research using HLM. *Research in Autism Spectrum Disorders, 5*, 562-569.