

12-2017

Multilevel Modeling of Single-Case Data: An Introduction and Tutorial for the Applied Researcher

Emily Rodabaugh

University at Albany, State University of New York, erodabaugh@albany.edu

Mariola Moeyaert

University at Albany, State University of New York, mmoeyaert@albany.edu

Follow this and additional works at: <http://opencommons.uconn.edu/nera-2017>

Recommended Citation

Rodabaugh, Emily and Moeyaert, Mariola, "Multilevel Modeling of Single-Case Data: An Introduction and Tutorial for the Applied Researcher" (2017). *NERA Conference Proceedings 2017*. 8.
<http://opencommons.uconn.edu/nera-2017/8>

**Multilevel modeling of single-case data: An introduction and tutorial for the applied
researcher**

Emily Rodabaugh and Mariola Moeyaert

Department of Educational Psychology and Methodology, State University of New York, NY

Contact author

Correspondence concerning this article should be addressed to Emily Rodabaugh, Department Educational and Counseling Psychology, State University of New York, NY, 1400 Washington Ave., 12222, Albany, US. E-mail: erodabaugh@albany.edu, Phone: (518) 442-5055

Abstract

Two- and three-level hierarchical linear modeling is introduced and demonstrated as a meta-analysis technique to estimate treatment effects across cases and studies. Multilevel modeling techniques have been empirically validated and are promising (Moeyaert, Ugille, Ferron, Beretvas, & Van den Noortgate, 2013a, 2013b). However, detailed instruction is needed to bridge the gap between methodologists and applied researchers. An empirical illustration estimating the effectiveness of choice-making interventions on decreasing students with special needs' problem behaviors is discussed. Step-by-step formulations of equations and R code for applied researchers are provided in order for researchers to conduct multilevel analyses using their own SCED data.

Multilevel Modeling of Single-Case Data: An Introduction for the Applied Researcher

Single-Case Experimental Designs

Single-Case Experimental Design

Single-case experimental designs (i.e., SCEDs) are commonly used in educational research due to the researcher's ability to actively introduce and manipulate a case-specific intervention (Barlow, Nock, & Herson, 2009). The influence of the intervention on the outcome score for each individual case can be evaluated and estimated (i.e., treatment effect or intervention effect) (Kazdin, 2011; Kratochwill et al., 2010). The use of this design also allows for researchers to measure one or multiple outcome variables over time, including before, during, and after the intervention has been implemented (Kratochwill et al., 2010).

The What Works Clearinghouse created a set of design standards that can be used to assess the quality of SCED studies (see Kratochwill et al., 2010). Ideally SCEDs are characterized by a systematic manipulation of the independent variable, multiple assessors measuring the dependent variable throughout time, at least three attempts to indicate the existence of an intervention effect, and at least three data points collected within each phase (Kratochwill et al., 2010). A study that meets the aforementioned criteria can be considered a strong experimental SCED. There are multiple types of SCED studies, with the multiple baseline across participants design (MBD) being the most popular and strongest experimental SCED (See Figure 1 copied from Moeyaert, Ugille, Ferron, Beretvas, & Van den Noortgate, 2013b) (Shadish & Sullivan, 2011; Smith, 2012). Shadish and Sullivan (2011) indicated that at least 73.5% of SCED studies have three or more study participants. As a consequence, one SCED study can be characterized by a hierarchical structure: repeated measures are nested within cases (see Figure 1).

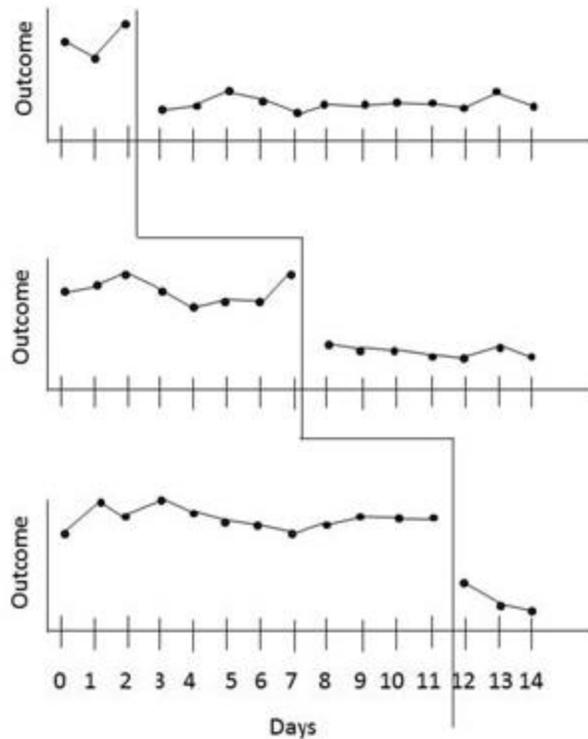


Figure 1. Graphical display of the multiple baseline design.

Recently there has been an exponential increase in the number published of SCED studies (See Figure 2, displaying the number of published items for the keywords “single-case” or “single-subject” or “interrupted time-series” between 1995 and 2016 using the Social Science Citation Index within the Web of Sciences). This leads to an increased number of SCED studies available for quantitative summary.

One technique that is promising and has been demonstrated to be valid for this purpose is three-level hierarchical linear modeling (Moeyaert et al., 2013a, 2013b). Three-level modeling takes the hierarchical structure of the SCED data into account as repeated measures are nested within cases and cases in turn are nested within studies. The quantitative synthesis has the potential to make significant contributions to policy, practice and theory.

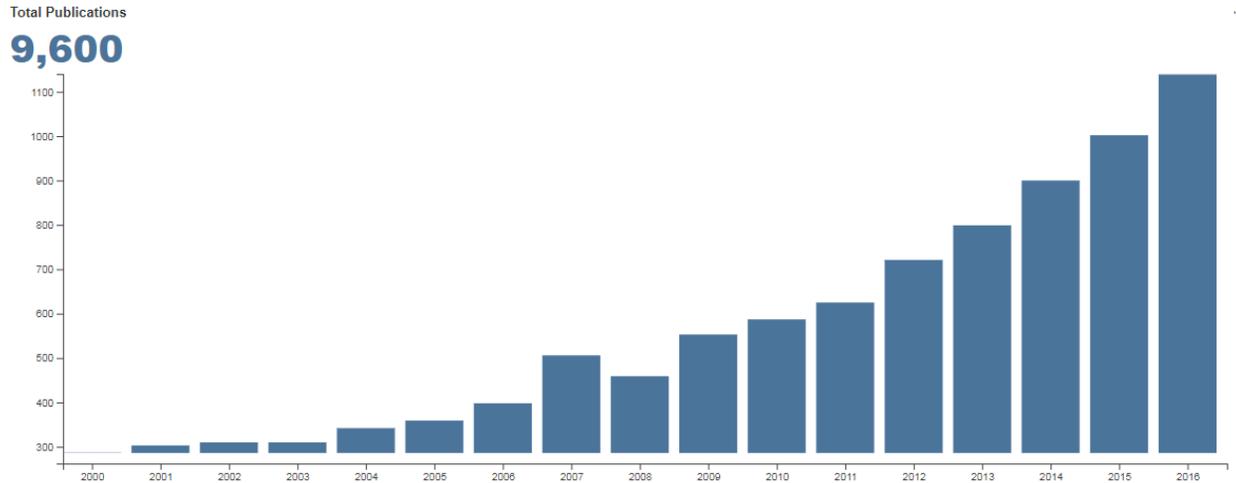


Figure 2. Graphical display showing the increase in the number of published items for the keywords “single-case” or “single-subject” or “interrupted time-series” between 1995 and 2016 using the Social Science Citation Index within the Web of Sciences.

Multilevel Modeling

In the field of SCEDs, researchers commonly use overlap indices, such as percent of data points exceeding the median (PEM; Ma, 2006), percent of nonoverlapping data (PND; Scruggs, Mastropieri, & Castro, 1987), percent of all nonoverlapping data (PAND; Parker, Hagan-Burke, & Vannest, 2007), and nonoverlap of all pairs (NAP; Parker & Vannest, 2009) to reflect the effectiveness of a treatment (Baek, Moeyaert, Petit-Bois, Beretvas, Van den Noortgate, & Ferron, 2014; Parker & Vannest, 2009). However, some issues arise when using these effect size metrics: (a) failure to account for replication, (b) procedural sensitivities, and (c) mischaracterization as an estimate of magnitude. This calls for the need to use more complex statistical analysis techniques to estimate treatment effects using SCED data.

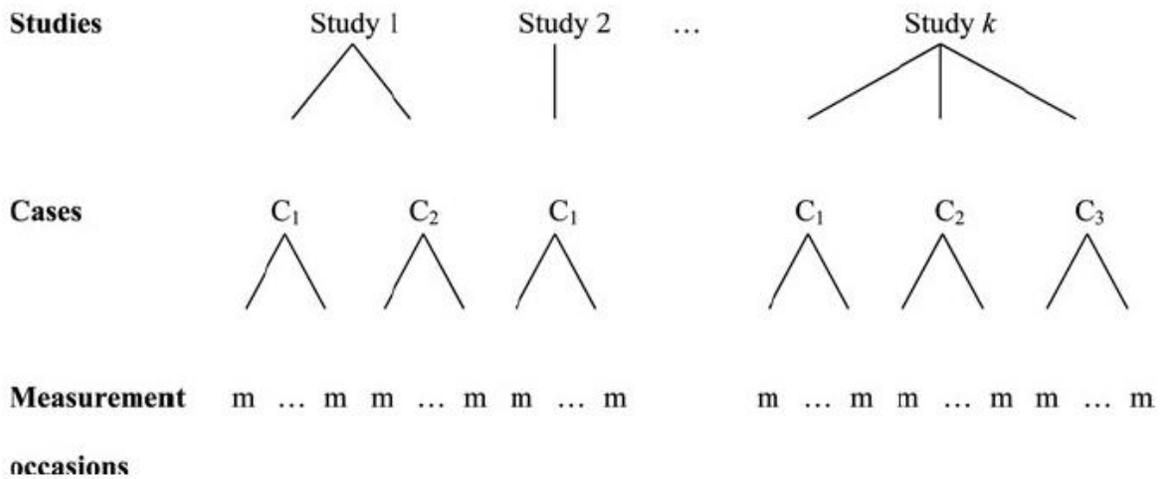


Figure 3. SCED three-level hierarchical structure.

Hierarchical linear models (HLM) are among the most promising analytical techniques for data analysis and summary of single-case designs (Kratochwill et al., 2010; Shadish, Kyse, & Rindskopf, 2013). Multilevel modeling provides many advantages such as allowing for the estimation of the overall treatment effects across SCED cases and studies in addition to individual treatment effects, the estimation of variance between cases and studies in these treatment effects, and adding moderators (Shadish et al., 2013). Figure 3 (copied from Moeyaert et al., 2013b) demonstrates the ability of multilevel modeling to take the hierarchical nature of SCED data from multiple studies into account. Multiple studies may be included in the multilevel model with each study (k) including at least one case (j) with multiple measurement occasions (i).

In the following sections, a step-by-step introduction to the HLM methodology is given followed by empirical demonstrations.

Methodological Framework

Single-Level Model

In order to summarize SCED data from one case, a simple linear regression can be used:

$$y_i = \beta_0 + \beta_1 Treatment_i + e_i \text{ with } e_i \sim N(0, \sigma_e^2) \quad (1)$$

This equation includes the outcome variable for measurement occasion i (y_i) regressed on a dummy coded variable ($Treatment_i$). The treatment variable is substituted with a zero when the measurement occasion is gathered prior to the start of the intervention (i.e., the baseline) and a value of one when the measurement occasion takes place during the treatment phase. As a

consequence, β_0 indicates the expected baseline level (i.e., $Treatment_i = 0$) and β_1 reflects the change in outcome score, y_i , due to the introduction of the treatment. Additional coefficients may be included to model the trend during the treatment phase (Center, Skiba, & Casey, 1985-1986; Moeyaert, Ugille, Ferron, Beretvas, and Van den Noortgate, 2014) which allows for the estimation of the change of the effectiveness of the treatment across time. The following equation may be used for this purpose:

$$y_i = \beta_0 + \beta_1 Treatment_i + \beta_2 Treatment_i Time_i + e_i \text{ with } e_i \sim N(0, \sigma_e^2) \quad (2)$$

In order to model the trend during the treatment phase, an interaction between the dummy coded variable, $Treatment_i$, and the time variable is created, namely $Treatment_i Time_i$. The time variable ($Time_i$) is centered on the first treatment measurement. For example, the intervention/treatment may be implemented at the fourth measurement. Therefore, at the fourth measurement, $Time_i = 0$. As a consequence, β_0 is the expected baseline level, β_1 is the treatment effect, and β_2 is the trend during the treatment. More information about coding the design matrix and the influence of centering time on the obtained treatment effect estimates can be found in Moeyaert et al. (2014).

Two-Level Model

When multiple cases are included within a study, researchers can summarize SCED data across cases. An extension of the regression introduced in Equation 2 can be used to summarize SCED data across subjects within one study:

$$y_{ij} = \beta_{0j} + \beta_{1j} Treatment_i + \beta_{2j} Treatment_i Time_{ij} + e_i \text{ with } e_i \sim N(0, \sigma_e^2) \quad (3)$$

In Equation 3, i stands for the measurement occasion ($i = 0, 1, \dots, I$), and j for the case ($j = 0, 1, \dots, J$). Similar regression coefficients as in Equation 2 are obtained with the only difference being the inclusion of the index j , referring to the case in which β_{0j} indicates the expected baseline level for case j , β_{1j} referring to the immediate treatment effect for case j , and β_{2j} referring to the time trend during the intervention phase.

The three regression coefficients of interest from Equation 3 can be allowed to vary at the second level of the two-level model to capture variability between cases:

$$\text{Level 2 of the two-level model: } \begin{cases} \beta_{0j} = \theta_{00} + u_{0j} \\ \beta_{1j} = \theta_{10} + u_{1j} \\ \beta_{2j} = \theta_{20} + u_{2j} \end{cases} \text{ with } \begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Sigma_u) \quad (4)$$

These equations indicate that the β coefficients from Equation 3 equal an overall study-specific performance, the θ coefficients, plus a random variation around the means.

As can be deduced from Equations 3 and 4, the residuals at the two levels are assumed to be multivariate normally distributed. If violations of normality assumptions occur, simulation studies have determined that fixed effects will remain unbiased (Moeyaert, Ugille, Ferron, Beretvas, & Van Den Noortgate, 2016). By combining Equations 3 and 4, the combined multilevel-model is as follows:

$$y_{ij} = \theta_{00} + u_{0j} + (\theta_{10} + u_{1j})Treatment_{ij} + (\theta_{20} + u_{2j})Time_{ij} * Treatment_{ij} + e_i \text{ with } e_i \sim N(0, \sigma_e^2) \quad (5)$$

The researchers are especially interested in the estimate of the immediate treatment effect (θ_{10}) and the trend during the treatment (θ_{20}) across cases.

The effect of a predictor variable such as type of intervention or gender can be added to the two-level model in order to explain between-case variance. For the purpose of the empirical illustration, an equation including the immediate treatment effect and a predictor effect was utilized. For instance, in Equation 6, a dummy coded variable ‘‘Predictor’’ can be added and as a consequence, β_{2j} indicates the effect of the predictor on the baseline level. This two-level equation is as follows:

$$y_{ij} = \beta_{0j} + \beta_{1j}Treatment_i + \beta_{2j}Predictor_i + e_i \text{ with } e_i \sim N(0, \sigma_e^2) \quad (6)$$

Three-Level Model

Using an even more extended regression model, researchers can quantify the immediacy of a treatment effect on the level and the slope across cases and across studies. The regression model of interest looks as follows:

$$y_{ijk} = \beta_{0jk} + \beta_{1jk}Treatment_{ijk} + \beta_{2jk}Treatment_iTime_{ijk} + e_{ijk} \text{ with } e_{ijk} \sim N(0, \sigma_e^2) \quad (7)$$

where i stands for the measurement occasion ($i = 0, 1, \dots, I$), j for the case ($j = 0, 1, \dots, J$) and k for the study ($k = 0, 1, \dots, K$).

The three regression coefficients of interest from Equation 7 can be allowed to vary at the second level and third level of the three-level model to capture, respectively, variability between cases and between studies in these regression coefficients as suggested by Moeyaert et al. (2013b):

$$\text{Level 2 of the three-level model: } \begin{cases} \beta_{0jk} = \theta_{00k} + u_{0jk} \\ \beta_{1jk} = \theta_{10k} + u_{1jk} \\ \beta_{2jk} = \theta_{20k} + u_{2jk} \end{cases} \text{ with } \begin{bmatrix} u_{0jk} \\ u_{1jk} \\ u_{2jk} \end{bmatrix} \sim N(0, \Sigma_u) \quad (8)$$

These equations indicate that the β coefficients from Equation 7 equal an overall study-specific performance, the θ coefficients, plus a random variation around the means. At the third level, the variation of the study-specific regression coefficients from the second level equations is described:

$$\text{Level 3 of the three-level model: } \begin{cases} \theta_{00k} = \gamma_{000} + v_{00k} \\ \theta_{10k} = \gamma_{100} + v_{10k} \\ \theta_{20k} = \gamma_{200} + v_{20k} \end{cases} \text{ with } \begin{bmatrix} v_{00k} \\ v_{10k} \\ v_{20k} \end{bmatrix} \sim N(0, \Sigma_v) \quad (9)$$

As can be deduced from Equations 7 to 9, the residuals at the three levels are assumed to be multivariate normally distributed. By combining equations 7 to 9, the combined multilevel-model is as follows:

$$y_{ijk} = \gamma_{000} + v_{00k} + u_{0jk} + (\gamma_{100} + v_{10k} + u_{1jk})Treatment_{ijk} + (\gamma_{200} + v_{20k} + u_{2jk})Time_{ijk} * Treatment_{ijk} + e_{ijk} \text{ with } e_{ijk} \sim N(0, \sigma_e^2)$$

The researchers are especially interested in the estimate of the immediate treatment effect (γ_{10}) and the trend during the treatment (γ_{20}) across cases and across studies.

Empirical Illustration

To provide researchers with an applied example of how to use multilevel modeling to analyze SCED data, R software was used to run both a two- and three-level model. The R code for each of the models is provided and can be used by the applied researchers to conduct multilevel modeling analyses on their own SCED data.

Two-Level Model

Romaniuk, Miltenberger, Conyers, Jenner, Jurgens, and Ringenberg (2002)'s study regarding the use of an activity choice intervention to decrease problems behaviors was used to run a two-level model. Five participants from the study were used. Each of the participants exhibited problem behaviors such as using violence, crying, leaving seat, using materials inappropriately, yelling, or throwing objects. The study met WWC criteria for Meeting Evidence Standards (Kratochwill et al., 2010). Specifically, the independent variable in the study is manipulated by the researchers, inter-assessor agreement was collected for more than 20% of the data with a percentage agreement greater than .80, and at least three demonstration of the effect of the intervention at three different points in time were included (Kratochwill et al, 2010).

An ABAB reversal design (see Figure 4, copied from Romaniuk et al., 2002) was used to collect data about the percentage of session time that problem behaviors occurred for the five participants in the study. This percentage acted as the dependent variable in the study. The intervention/treatment that the students received involved the opportunity to choose the task to

participate in. During the baseline phase, the students were not given a choice. The type of intervention students received was also included in the study. Students were either given an intervention focused on escape-maintained problem behaviors or attention-maintained behaviors.

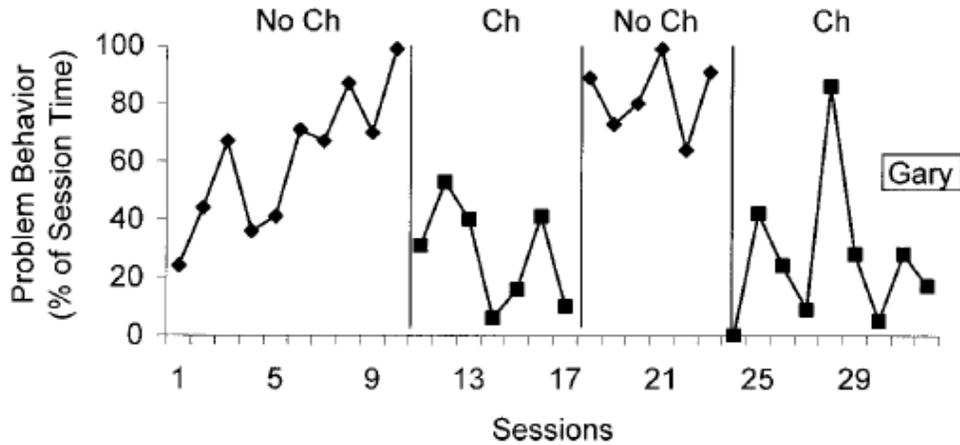


Figure 4. Graphical display of the ABAB reversal design for one case from the Romaniuk et al. (2002) study.

A two-level model can be used to assess the treatment effects of this intervention across all participants in the study. In addition, the type of intervention can be included in the two-level model as a predictor as shown in Equation 6. This dummy-coded predictor allows for the analysis of how participating in an intervention focused on problem behaviors maintained by escape (0) or problem behaviors maintained by attention (1) may affect overall percentage of session time spent exhibiting problem behaviors. For an example of the coding used for the study, see Figure 5.

The R code used to create the two-level model is provided below. The average immediate treatment effect is allowed to vary between cases. In addition, the predictor, intervention type, is included in the two-level model as a fixed effect. The R code can be found in the box below:

```
Model <- lme(Y~1+Treatment+Predictor, random=~Treatment|Case, data=Roman,
control=list(opt="optim"), na.action="na.omit")
Summary(Model)
VarCorr(Model)
Coef(Model)
Intervals(Model)
```

The first two lines of the above code create the two-level model to be analyzed in R with Y representing the outcome which was percent of session time displaying problem behaviors, the

immediate treatment effect (*Treatment*), and the intervention predictor (*Predictor*). The immediate treatment effect was allowed to vary at the second level, therefore it is included after the “random=” section of the code.

Study	Author	Case	D	T	T_cnr	DVY	Dt	Gender	Age	Diagnosis	Setting	TypeBeha	Functiona	TypeInter	TypeActiv	Preinterv	Design	TypeOutcome
9	Romaniuk	1	0	0	-9	25	0	1	7	5	1	1	1	0	0	0	0	0
9	Romaniuk	1	0	1	-8	73	0	1	7	5	1	1	1	0	0	0	0	0
9	Romaniuk	1	0	2	-7	30	0	1	7	5	1	1	1	0	0	0	0	0
9	Romaniuk	1	0	3	-6	39	0	1	7	5	1	1	1	0	0	0	0	0
9	Romaniuk	1	0	4	-5	26	0	1	7	5	1	1	1	0	0	0	0	0
9	Romaniuk	1	0	5	-4	71	0	1	7	5	1	1	1	0	0	0	0	0

Figure 5. Section of data used for analysis.

The code “Summary(Model)” can be used to provide the parameter estimates of the fixed and random effects. The parameter estimates of the two-level model appear in Table 1.

Variance components are generated using the R code “VarCorr(Model)”. The covariance between the baseline level and the average immediate treatment effect ($\sigma_{u_0u_1}$) can be calculated by multiplying the correlation between the baseline level and the treatment effect by the standard deviation of the baseline level and the standard deviation of the average immediate treatment effect.

The R code “coef(Model)” generates the estimated baseline level, average immediate treatment effect, and average predictor effect for each case in the study. This output may give further insight into the variability between cases in the study and whether or not the parameter estimates should be interpreted cautiously.

The statistical significance of the fixed and random effects can also be calculated using R. The code “intervals(Model)” provides users with the approximated 95% confidence intervals for the fixed and random effects in addition to the within-group standard error. Parameter estimates can be found in Table 1. The immediate treatment effect across participants [$\theta_{01} = -11.06, t(65) = -0.79, p > .05$] and the predictor effect [$\theta_{02} = 5.11, t(3) = 0.93, p > .05$] were not found to be significant compared to the baseline level which was significant [$\theta_{00} = 56.22, t(65) = 15.32, p < .05$]. Although there were not significant changes in problem behaviors from baseline to treatment, it can be noted that problem behaviors decreased when the treatment was implemented and that students who participated in attention-maintained interventions displayed more problem behaviors compared to those in the escape-maintained interventions.

Table 1

Parameter and standard error estimates resulting from estimation of two-level model

	Parameter	Parameter Estimates	SE	p
Two-level model	Fixed Coefficient			
Average baseline level	θ_{00}	56.22*	3.67	< .001
Average immediate treatment effect	θ_{01}	-11.06	13.95	.4309
Average effect of predictor (intervention type) during baseline	θ_{02}	5.11	5.50	.4212
	(Co)variance component			
Baseline level	$\sigma_{u_0}^2$	2.10*	/	< .05
Immediate treatment effect	$\sigma_{u_1}^2$	857.40*	/	< .05
Covariance of baseline and immediate treatment effect	$\sigma_{u_0u_1}$	-40.55	/	/
Residual variance	σ_e^2	369.50	/	> .05

Note. *p < .05

Three-Level Model

The two-level model can be expanded into a three-level model. A three-level hierarchical linear model was used to analyze the single-case experimental data presented in Shogren, Faggella-Luby, Bae, and Wehmeyer's (2004) meta-analysis of the effects on problem behaviors when student choice is utilized as a treatment. An applied illustration of the three-level model includes a total of 9 studies and 21 cases from the original study. A three-level model was used to determine average immediate treatment effect and average trend during treatment. The average immediate treatment effect was allowed to vary and was included as a random effect in addition to as a fixed effect. The average trend during treatment is included as a fixed factor.

The three-level model can be created in R using the code:

```
Model <- lme(Y~1+Treatment+TreatmentTime, random=~1+Treatment|Study/Case,
data=Shogren, control=list(opt="optim"), na.action="na.omit")
Summary(Model)
VarCorr(Model)
Coef(Model)
Intervals(Model)
```

The parameter estimates of the three-level model can be found in Table 2. The average immediate treatment effect across studies was found to be significant [$\gamma_{100} = -47.01$, $t(249) = -2.11$, $p < .05$] in addition to the baseline effect [$\gamma_{000} = 62.55$, $t(249) = 3.00$, $p < .05$]. The average trend during treatment was not found to be significant [$\gamma_{200} = -0.24$, $t(249) = -0.34$, $p > .05$]. These results indicate that when the treatment was implemented, there was a significant treatment effect in which problem behaviors decreased. However, although there was a negative average trend during treatment, it was not significant.

The same R code that was used in the two-level model can be applied to a three-level model. The first two lines of the code provide the model to be analyzed. This model includes the addition of the trend during treatment (*TreatmentTime*) as a fixed effect. In addition, the inclusion of the immediate treatment effect (*Treatment*) after the “random=” code allows the immediate treatment effect to vary across cases and studies by including “|Study/Case”.

The code “summary(Model)” provides the parameter estimates for the fixed and random effects. However, parameter estimates for fixed effects at the third-level are interpreted as average effects across all studies, and the between-case and between-study variance is included. Additionally, the covariance between the baseline level and immediate treatment effect can be calculated between cases ($\sigma_{u_0u_1}$) and between studies ($\sigma_{v_0v_1}$).

Table 2

Parameter and standard error estimates resulting from estimation of three-level model

	Parameter	Parameter Estimates	SE	p
Three-level model	Fixed Coefficient			
Average baseline level	γ_{000}	62.55*	20.83	.0029
Average immediate treatment effect	γ_{100}	-47.01*	22.30	.0360
Average trend during treatment	γ_{200}	-0.24	0.69	.7315
	Variance component			
Between-study (co)variance				
Baseline level	$\sigma_{v_0}^2$	939.08*	/	<.05
Immediate treatment effect	$\sigma_{v_1}^2$	1284.98*	/	<.05
Covariance between baseline level and immediate treatment effect	$\sigma_{v_0v_1}$	-1017.21	/	/
Between-case (co)variance				
Baseline level	$\sigma_{u_0}^2$	6334.32*	/	<.05
Immediate treatment effect	$\sigma_{u_1}^2$	6603.30*	/	<.05
Covariance between baseline level and immediate treatment effect	$\sigma_{u_0u_1}$	-6350.99	/	/
Residual Variance	σ_e^2	412.19	/	/

Note. *p < .05

Discussion and Implications

The purpose of this paper is to introduce the use of HLM to the novice researcher and to make multilevel modeling more accessible to applied researchers by providing an empirical demonstration including sample R code. Multilevel models provide useful analyses of SCED studies by allowing for the evaluation of overall effects and the variance between cases, the use of complex statistical techniques, and the opportunity to analyze multiple types of SCED data (Shadish et al., 2013). Extensions can be made to the two-level model in an attempt to explain increased portions of the between-case variance such as including additional predictors, modeling time and trend in treatment effect, and including autocorrelation and heterogeneous phase variance (Baek & Ferron, 2013).

The two-level model can also be extended to a three-level model. This model includes the same attributes as the two-level model in addition to modeling variance between studies. To

account for more of the variance between cases and studies, the three-level model can also be extended. This includes including more predictors and modeling autocorrelation and heterogeneous phase variance.

In addition, standardizing data becomes increasingly important when multiple studies are being analyzed. This allows the researcher to make more appropriate conclusions regarding the data. If the raw SCED data were not collected by the researcher conducting the multilevel model, it can easily be retrieved from the original study article. Software such as WebPlotDigitizer can be used to collect data from the graphs provided in the articles (Moeyaert, Maggin, & Verkuilen, 2016). According to Maggin et al. (2011), using raw data analyzation techniques such as visual analysis can become cumbersome due to the need to evaluate each graph in order to make conclusions regarding whether a treatment effect occurred. It is recommended that effect sizes be used to analyze SCED data (Maggin et al., 2011). The effect sizes are included in the model as the parameter estimates, signifying average effect sizes across cases or across studies.

When using unstandardized effect sizes, the scale of the original variables can greatly impact the outcome scores in addition to how variables are interpreted (Van den Noortgate & Onghena, 2008). For example, problem behavior may be measured on a percentage scale in one study versus on a count scale in another. To account for differences in scale, it is proposed by Van den Noortgate and Onghena (2008), effect sizes can be standardized by dividing individual outcome scores (y_{ijk}) by the estimated residual within-subject standard deviation ($\hat{\sigma}_{e_{jk}}$).

The educational implications of this study show the need for applied SCED researchers to consider using complex statistical techniques such as multilevel modeling to analyze the effectiveness of interventions. Multilevel modeling provides the applied researcher with an opportunity to analyze data in a hierarchical nature by providing information about the treatment effect averaged across cases and across multiple studies. These values can be used by policy makers and applied researchers in that they provide additional information about the effect of an intervention (Kratochwill & Levin, 2010).

References

- Baek, E., & Ferron, J. M. (2013). Multilevel models for multiple-baseline data: Modeling across participant variation in autocorrelation and residual variance. *Behavior Research Methods, 45*, 65-74. doi:10.3758/s13428-012-0231-z
- Baek, E. K., Moeyaert, M., Petit-Bois, M., Beretvas, S. N., Van den Noortgate, W., & Ferron, J. M. (2014). The use of multilevel analysis for integrating single-case experimental design results within a study and across studies. *Neuropsychological Rehabilitation, 24*(3-4), 590-606. doi:10.1080/09602011.2013.835740
- Barlow, D. H., Nock, M. K., & Hersen, M. (2009). *Single case experimental designs: Strategies for studying behavior change* (3rd ed.). Boston, MA: Allyn & Bacon.
- Center, B. A., Skiba, R. J., & Casey, A. (1985-1986). A methodology for the quantitative synthesis of intra-subject design research. *Journal of Special Education, 19*(4), 387-400. doi:10.1177/002246698501900404
- Kazdin, A.E. (2011). *Single-case research designs: Methods for clinical and applied settings* (2nd ed.). New York, NY: Oxford University Press.
- Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2010). *Single-case designs technical documentation*. Retrieved from What Works Clearinghouse website: http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf
- Kratochwill, T. R., & Levin, J. R. (2010). Enhancing the scientific credibility of single-case intervention research: Randomization to the rescue. *Psychological Methods, 15*, 124-144. doi:10.1037/a0017736
- Ma, H. H. (2009). The effectiveness of intervention on the behavior of individuals with autism a meta-analysis using percentage of data points exceeding the median of baseline phase (PEM). *Behavior Modification, 33*(3), 339-359. doi:10.1177/0145445509333173
- Maggin, D. M., Swaminathan, H., Rogers, H. J., O'keeffe, B. V., Sugai, G., & Horner, R. H. (2011). A generalized least squares regression approach for computing effect sizes in single-case research: Application examples. *Journal of School Psychology, 49*(3), 301-321. doi:10.1016/j.jsp.2011.03.004
- Moeyaert, M., Maggin, D.M., & Verkuilen, J. (2016). Reliability and validity of extracting data from image files in contexts of single-case experimental design studies. *Behavior Modification, 40*, 874-900. doi:10.1177/0145445516645763

- Moeyaert, M., Ugille, M., Ferron, J., Beretvas, S., & Van den Noortgate, W. (2013a). Modeling external events in the three-level analysis of multiple-baseline across-participants designs: A simulation study. *Behavior Research Methods*, *45*(2), 547-559. doi:10.3758/s13428-012-0274-1
- Moeyaert, M., Ugille, M., Ferron, J., Beretvas, S., & Van den Noortgate, W. (2013b). Three-level analysis of single-case experimental data: Empirical validation. *Journal of Experimental Education*, *82*, 1-21. doi:10.1080/00220973.2012.745470
- Moeyaert, M., Ugille, M., Ferron, J., Beretvas, S., Van den Noortgate, W. (2014). The influence of the design matrix on treatment effect estimates in the quantitative analyses of single-subject experimental design research. *Behavior Modification*, *38*, 665-704. doi:10.1177/0145445514535243
- Moeyaert, M., Ugille, M., Ferron, J., Beretvas, S., & Van Den Noortgate, W. (2016). The misspecification of the covariance matrix in the three-level modeling of single-cases. *Journal of Experimental Education*, *84*(3), 473-509. doi:10.1080/00220973.2015.1065216
- Parker, R. I., Hagan-Burke, S., & Vannest, K. J. (2007) Percent of all nonoverlapping data (PAND): An alternative to PND. *Journal of Special Education*, *40*(4), 194-204. doi:10.1177/00224669070400040101
- Parker, R. I., & Vannest, K. J. (2009). An improved effect size for single case research: Nonoverlap of All Pairs (NAP). *Behavior Therapy*, *40*, 357-367. doi:10.1016/j.beth.2008.10.006
- Romaniuk, C., Miltenberger, R., Conyers, C., Jenner, N., Jurgens, M., & Ringenberg, C. (2002). The influence of activity choice on problem behaviors maintained by escape versus attention. *Journal of Applied Behavior Analysis*, *35*(4), 349-362. doi:10.1901/jaba.2002.35-349
- Scruggs, T. E., Mastropieri, M. A., & Casto, G. (1987). The quantitative synthesis of single subject research: Methodology and validation. *Remedial and Special Education*, *8*(2), 24-33. doi:10.1177/074193258700800206
- Shadish, W. R., Kyse, E. N., & Rindskopf, D. M. (2013). Analyzing data from single-case designs using multilevel models: New applications and some agenda items for future research. *Psychological Methods*, *18*(3), 385-405. doi:10.1037/a0032964

- Shadish, W. R., & Sullivan, K. J. (2011). Characteristics of single-case designs used to assess intervention effects in 2008. *Behavior Research Methods*, 43(4), 971-980. doi:10.3758/s13428-011-0111-y
- Shogren, K. A., Faggella-Luby, M. N., Bae, S. J., & Wehmeyer, M. L. (2004). The effect of choice-making as an intervention for problem behavior: A meta-analysis. *Journal of Positive Behavior Interventions*, 4, 228-237. doi:10.1177/10983007040060040401
- Smith, J. D. (2012) Single-case experimental designs: A systematic review of published research and current standards. *Psychological methods*, (17)4, 510-550. doi:10.1037/a0029312
- Van den Noortgate, W., & Onghena, P. (2008). A multilevel meta-analysis of single-subject experimental design studies. *Evidence-Based Communication Assessment and Intervention*, 2(3), 142-151. doi:10.1080/17489530802505362