The Journal of Educational Research & Policy Studies

The *Journal of Educational Research & Policy Studies* is a national peer-reviewed publication that seeks to provide an interdisciplinary forum for the consideration of meaningful educational research initiatives and policy analyses.

Volume (13), No. (1), 2013

The Journal of Educational Research & Policy Studies
February 2013 Vol. 1, pp. 18-31

Can Multilevel (HLM) Models of Change Over Time Adequately Handle Missing Data?

Anat Zaidman-Zait
Tel-Aviv University

Bruno D. Zumbo
University of British Columbia

Abstract

The focus of this paper was whether, and under what conditions, can multilevel (HLM) analysis of longitudinal data adequately handle missing data. Two studies were reported: the first was an investigation of real data and the second was a Monte Carlo simulation. For the first study, the focus was on the effect of types of missing data – missing completely at random (MCAR); missing not at random (MNAR); and two types of data missing at random (MAR) – on model parameter estimates and statistical results. The results of this study highlighted the importance of model specification. That is, it highlighted that when the model is correctly specified missing data, of all types, are correctly handled using HLM analyses. The second study focused on the role of level-l model mis-specification in the context of the four different types of missing data. Results indicated that when the level-1 model is correctly specified the parameter estimates are nearly always the same (and the same as the complete data case) irrespective of the degree or kind of missing data. However, this finding did not hold when the level-1 model is mis-specified. Overall, the findings highlight that claims in the research literature of the ability of HLM analyses to handle missing data should be qualified by the statement that this is only true when the correct model is specified.

Author Note

Anat Zaidman-Zait, Ph.D., is Associate Professor, Tel-Aviv University. During this project she was a Azrieli Foundation Post-Doctoral Research Fellow at Tel-Aviv University as well as a Michael Smith Foundation for Health Research (MSFHR), and Social Sciences and Humanities Research Council of Canada (SSHRC) post-doctoral fellow at UBC. Her research focuses jointly on Human Development, and Measurement, Evaluation and Research Methodology.

Bruno D. Zumbo, Ph.D., is Professor of Measurement, Evaluation and Research Methodology as well as member of the Department of Statistics and the Institute of Applied Mathematics at the University of British Columbia.

Address Correspondence to Professor Bruno D. Zumbo, The University of British Columbia, Department of ECPS, Vancouver, B.C. CANADA V6T 1Z4. E-mail: bruno.zumbo@ubc.ca http://educ.ubc.ca/faculty/zumbo/cv.htm

Introduction

Longitudinal studies are widely used in educational and policy research (e.g., Goldenberg, Gallimore, Reese, & Garnier, 2001; Jimerson, Egeland, Sroufe, & Carlson, 2000; Zumbo, 1999). Since the mid-1970s methodologists have developed a class of statistical models that enable researchers to study change – see Singer & Willett (2003) for a review. This class of statistical models goes by various names in the research literature*: individual growth models, random coefficient models, multilevel models, mixed models, or hierarchical linear models.* Longitudinal data can be viewed as multilevel data with repeated measurements nested within individuals. This structure involves a two-level model, with the series of repeated measures at the lowest level, and the individual persons at the next level. Therefore, in the first level of an analysis of change, *within-individual change* over time is examined. In the second level of the analysis *inter-individual differences in change* are examined.

Missing Data in Multilevel Models for Studying Change

Statistical analysis with missing data has always been a challenge. In general across a broad class of commonly used statistical techniques, missing data may result in biased parameters estimates, inflated Type I and Type II error rates during hypothesis testing, and hence a degradation of the performance of confidence intervals. In addition, missing data can reduce statistical power (Collins, Schafer, & Kam, 2001).

One of the advantages of multilevel analysis of longitudinal data is its ability to handle missing data (Bryk & Raudenbush, 1992; Snijders, 1996). Hox (2000, 2002) is among the few to clarify that this advantage refers to the ability to handle models with varying time points. Multilevel models do not assume equal numbers of observations, or even fixed time points, so respondents with missing data do not cause special problems (Hox, 2000). This advantage focuses on the response (dependent) variable in the analysis; however, if a measurement on an explanatory (independent) variable is missing the typical treatment is to remove the case completely.

The central purpose of this study is to explore effects of various missing data causes on the results of a multilevel analysis of longitudinal data. We report two studies. In this first study we investigated whether manipulating the missing data source with a real data example would change the parameter estimates in the multilevel model. We used real data as our base because we wanted to have as much generalizability to commonly found data as possible. The results of this first study raised a question for us as to what would happen in terms of missing data in varying settings of level-1 model mis-specification. Therefore the results of the first study informed the second study, a Monte Carlo simulation study wherein we manipulated the level-l model mis-specification.

Causes of Missingness

A significant theoretical advance in the statistical and mathematical study of missing data occurred with the formalization of, and focuses on, why the data is missing. Little and Rubin (1987) presented a statistical framework to understand the processes that generated the missing data with respect to the information provided about the unobserved data, rather than the data one has at hand. The focus now was on the processes that lead to the missing data. They presented three missing data processes: missing completely at random (MCAR), missing at random (MAR), and missing not at random (MNAR). See Collins, Schafer, and Kam (2001) as well as Wothke (2000) for detailed descriptions of the MCAR, MAR, and MNAR missing data processes.

Collins, Schafer, & Kam (2001) emphasized that the distinction among MCAR, MAR and MNAR is based on the relationships between variables of interest and variables explaining the missingness. Another important distinction is whether the cause of missingness is accessible or inaccessible (Graham & Hofer, 2000). An accessible cause occurs when the cause has been measured (often fortuitously) and it can be included in the statistical data model. Inaccessible cause of missingness is when the cause has not been measured and it is not available for analysis.

To help make the abstract missing data concepts more concrete, let us imagine an example (that will also be used again later in the paper) wherein one has collected repeated measures data about cognitive ability over four time points. Throughout this example the focus is on missing data for the cognitive variable, the dependent variable. MCAR implies that the reason for the missing cognitive ability score data is not related to the cognitive variable itself (e.g., only high scoring data is missing) and the missingness is also not due to any another variables, whether these other variables are accessible or inaccessible in your data at hand. Hence the missingness is, as the name suggests, completely at random. There are two types of MAR missingness that are of interest. MAR scenario-A implies that the reason for the missing cognitive ability score data is not related to the variable itself but is related, for example, to other data accessible in the sample (e.g., gender); however, in this scenario gender is not related to the cognitive ability variable. Unlike MAR scenario-A, in MAR scenario-B the other accessible data (e.g., gender) would be related to the cognitive ability variable. Therefore, in MAR scenario-B the missingness is explainable through another variable, in this case gender. MNAR would be a scenario in which the missingness on the cognitive is due to the cognitive variable itself – for example, the missing data (on cognitive ability) is only for those with low scores on cognitive ability.

Study 1 – Simulation with Real Data

Study one focused on investigating the effects of various missing data sources (i.e., causes) on the resulting parameter values of a multilevel growth model. This complete data was manipulated in a variety of ways described below to mimic the various causes of missing data, as described by Little and Rubin (1987), and Collins et al. (2001). This first study was not a "simulation" in the statistical sense of repeated sampling from a population but rather a simulation in the purely mathematical sense wherein one mimics a model with one set of data.

Method

Procedure

The data. Willet's (1988) classic data from a study of change in cognitive ability over time (Singer & Willet, 2003) were used. The repeated measure dependent variable was scores on the opposite naming cognitive task. Thirty-five individuals were measured on four occasions. Time was coded 0, 1, 2, and 3, hence the intercept estimates the true value of opposite naming skill at occasion 0 (initial status) and the slope estimates the rate of change in opposite naming skill across occasions.

Simulated Missing Data Mechanism

The Little and Rubin mathematical framework for missing data was operationalized in the four scenarios described above: MCAR, MNAR, and two types of MAR. In addition, these four missing data scenarios were investigated with 7 and 14 percent missing data—10 and 20 cases missing the dependent variable.

In the MCAR condition, missing values were randomly imposed on the Y variable, the dependent variable in the study, independent of any variable.

In the missing not at random (MNAR) condition, missing values were imposed on the Y variable depending on its own value. To achieve this, the distribution of Y was divided into quartiles. Next, missing values were randomly selected from all the cases in the lowest quartile. In this way, the missingness is dependent on the level of the cognitive variable—i.e., the lowest quartile.

Two different types of MAR were investigated. In the first MAR scenario, denoted (MAR) scenario-A, the missingness was not related to the cognitive ability itself but related to another variable, gender. Under this condition, missing values were imposed on the cognitive ability variable depending on the participant's gender. To achieve this, roughly half of the participants were (randomly) coded as girls, and the rest were coded as boys. Next, the data were sorted according to gender and the missing values were randomly selected only from the girls. Therefore, the missingness depended on gender, but gender was not related to the cognitive task.

In the second MAR scenario, scenario-B, the missingness was related to gender, however in this scenario, gender was related to the cognitive task itself. Under this scenario, the missingness was higher among girls who had low scores on the cognitive variable. To achieve this, the distribution of the cognitive variable was divided into quartiles only for the girls. Next, the data were sorted according to gender and according to the cognitive variable. Finally, the missing values were randomly selected from girls in the lowest 25% of the cognitive task distribution.

Results

Complete Data Analysis

The complete data were used as a comparative baseline condition. Exploratory analyses were conducted in order to describe how individuals in the data set change over time. Scatter diagrams were used to visualize the empirical growth plot (Singer & Willet, 2003). Like Singer and Willet (2003), inspection of the empirical growth plots lead us to conclude that the cognitive variables (i.e., the opposite naming skill) increases over time, and that the pattern of change appears to be linear for most of the sample. In addition, an exploratory ordinary least squares regression model was fit to each individual to summarize the growth trajectory. The resultant R-squared statistics reveal —that 68.6% of the R^2 statistics were above 0.90. In summary, a linear change trajectory seems to be very reasonable for many individuals in the sample.

An unconditional linear growth model. In the next step, as is commonly recommended, an unconditional linear growth model was fit to the data. The unconditional growth model written in the "levels" notation (Bryk & Raudenbush, 1992) is:

Level 1: $Y_{ij} = \pi_{0j} + \pi_{1j}(TIME_{ij}) + r_{ij}$

Level 2: $\pi_{0j} = \beta_{00} + u_{0j]}$

$\pi_{1j} = \beta_{10} + u_{ij}$

The combined model is:

$$Y_{ij} = \underbrace{[\beta_{00} + \beta_{10}TIME_{ij}]}_{\text{Fixed effects}} + \underbrace{[u_{0j} + u_{ij}(TIME_{ij}) + r_{ij}]}_{\text{Random effects}}$$

In the unconditional linear growth model, for the complete data, both fixed effects were statistically significant. The results are displayed in Table 1. The second column, from the left, is the results of the complete data – our baseline condition. As we can see in Table 1, both fixed effects were statistically significant and the average intercept score was 164.37 and the rate of change was 26.96 points per testing time.

A linear growth model with a person-level covariate. Next, a linear growth model with a person-level covariate was fit to the complete data. This model in the "levels" notation is:

Level 1: $Y_{ij} = \pi_{0j} + \pi_{1j}(TIME_{ij}) + r_{ij}$

level 2: $\pi_{0j} = \beta_{00} + \beta_{01}(CCOVAR_J) + u_{0j}$

$\pi_{1j} = \beta_{10} + \beta_{11}(CCOVAR_J) + u_{1J}$

The combined model is:

$$Y_{ij} = \underbrace{[\beta_{00} + \beta_{10}(TIME_{ij}) + \beta_{01}(CCOVAR) + \beta_{11}(CCOVAR)(TIME_{ij})]}_{\text{Fixed effects}} + \underbrace{[u_{0j} + {}_{1j}(TIME_{ij}) + r_{ij}]}_{\text{Random effects}}$$

In order to explain individual variation in opposite naming ability, Willett considered a person level covariate model, denoted CCOVAR (the grant mean centered covariate), which models whether variation in intercepts and slopes is related to a covariate (e.g., child SES). The results of this model fit to the complete data can be seen second column, from the left, in Table 2. Using child SES as the covariate, the results of the complete data show that the SES is not a statistically significant predictor of the intercept (i.e., the starting point of the trajectory of change) but is a statistically significant predictor of the slope (i.e., the rate of change).

Table 1

An Unconditional Linear Growth Model (LGM)

Data Model	Complete Data	MCAR	MNAR	MAR Scenario-A	MAR Scenario-B
Fixed part					
Intercept	164.37 (6.12)*	165.03 (6.09)*	170.62 (6.31)*	164.94 (6.56)*	175.57 (5.05)*
Time	26.96 (2.16)*	26.76 (2.18)*	24.66 (2.26)*	26.78 (2.30)*	23.17 (2.13)*
Random part					
σ^2_e	159.48 (26.96)	164.83 (31.71)	139.62 (27.48)	131.06 (25.88)	141.39 (27.38)
$\sigma^2_{intercept}$	1198.77 (318.38)	1125.61 (313.59)	1170.92 (334.88)	1322.55 (382.87)	654.81 (207.33)
σ^2_{time}	132.40 (40.21)	117.64 (39.33)	131.09 (43.97)	136.75 (44.57)	110.95 (38.34)
$\sigma^2_{int*time}$	-179.25 (88.96)	-134.39 (86.31)	-178.53 (96.35)	-213.98 (108.80)	-67.61 (69.04)

Note: The results are for the 14% missing data. The “*” denotes statistical significance at a .05 level.

Table 2

Linear Growth Model (LGM) with a Person-Level Covariate

Data Model	Complete Data	MCAR	MNAR	MAR Scenario-A	MAR Scenario-B
Fixed part					
Intercept	164.37 (6.21)*	164.57 (6.14)*	170.19 (6.41)*	164.60 (6.62)*	175.13 (5.11)*
Time	26.96 (1.99)*	26.94 (2.04)*	24.82 (2.13)*	27.03 (2.14)*	23.33 (1.98)*
cov	-.11 (.50)	-.17 (.50)	.03 (.50)	-.07 (.53)	.02 (.39)
Cov*time	.43 (.16)*	.43 (.17)*	.37 (.16)*	.41 (.17)*	.37 (.15)*
Random part					
σ^2_e	159.47 (26.96)	164.67 (31.66)	139.95 (27.58)	130.73 (25.75)	141.90 (27.50)
$\sigma^2_{intercept}$	1236.41 (332.40)	1146.98 (322.28)	1215.86 (351.68)	1349.98 (394.11)	677.30 (216.31)
σ^2_{time}	107.25 (34.68)	97.32 (34.42)	111.06 (39.94)	112.86 (39.59)	90.45 (34.25)
$\sigma^2_{int*time}$	-178.23 (85.43)	-132.34 (82.97)	-185.41 (95.29)	-210.94 (105.84)	-69.20 (67.07)

Note: The results are for the 14% missing data. The "*" denotes statistical significance at a .05 level.

Applying Missing Data Mechanisms

The results of the unconditional and conditional models with the different mechanisms of missing data at a rate of 14% are displayed in the last four columns of Tables 1 and 2, respectively. We did not report the results of the 7% rate of missingness because they are the same as the more extreme case of 14%. Under all conditions of missingness both fixed effects remained statistically significant. Moreover, the estimates of the intercept and time fixed effect coefficient, as well as the standard errors, demonstrated only minor changes all reasonably close to the complete data values. Note that under MNAR and MAR scenario-B mechanism the estimates were somewhat worse. With regard to the estimates of the covariance-variance matrix, some changes occurred; however, not in a consistent way that reflects the different missingness mechanisms.

In the linear growth model with a person-level covariate, the fixed effects are the same as in the unconditional model. The effect for time and the time-by-covariate interaction remained significant in all the analyses regardless of the missing mechanism and the rate of missingness (7% and 14%). Moreover, the estimated values are close to the population (completed data) values.

The variance covariance matrix, the last four rows of the tables, resulted in somewhat different variance estimates for the slopes and intercepts. According to the complete data analysis, and computing the appropriate intraclass correlation, the covariate accounts for 19% of the variation in growth rate. However, MNAR and MAR scenario-B mechanisms inflated the variance in growth rate explained by the covariate by additional 7% to 15%.

Conclusions

Overall, the results of Study 1 are a good demonstration of a multilevel model's ability to handle missing data as described by, for example, Bryk & Raudenbush (1992), Snijders (1996), and Hox (2000, 2002). This empirical illustration is a useful addition to these earlier texts' description. There is some persuasive power to be had by seeing an abstract statistical claim illustrated with real data, as we did in Study 1.

The findings, however, are not just illustrative for the multi-level modeling user but also touch at the core statistical issue of model mis-specification. That is, in the complete data in Study 1, for most individuals, the linear change model fits very well—their observed and fitted values nearly coincide (based on data visualization and R-squared statistics). The available data points that are left in the model are strong enough and provide an indication of the correct model, even without the full information due to missingness. Hence, missing data did not cause any changes in the model parameter estimates. When the model is correctly specified missing data will most likely not cause dramatic change in the parameters estimation even when the supposed cause of missingnesss is MNAR or MAR where the missingness is related indirectly (i.e., MAR Scenario-B). Missing not at random (MNAR) is usually considered the worst scenario of missing data that increases the risk for reaching incorrect conclusions. It is worth noting that the case of MAR condition B there is a risk for increased biases, similarly to a MNAR scenario. It is important to keep in mind that when the magnitudes of the effects in the multilevel model are so large the presence of missing data (even in the MNAR setting) would not influence the conclusions.

Study 2 – Monte Carlo Simulation

An interesting feature of the data used in Study 1 is that a linear growth curve fit the data well (nearly 69%) and showed an R-squared of greater than 0.90 for the level-1 unconditional linear growth curve; with the level-l model accounting for 83% of the variation. In statistics it is well-known that model mis-specification can accentuate biases in parameter estimates in certain contexts. For example, in the field of complex survey data it is well-known that model mis-specification can alter the effects of applying sample weights, wherein sample weights play less of a role when the model is correctly specified. Therefore, the purpose of the study 2 was to investigate the influence of different missingness scenarios when the level-1model is not correctly specified. For the purposes of this study it is sufficient to investigate the unconditional growth curve model because any effects on it will carry forward to the conditional model.

Method

Procedure

The dependent variables (i.e., the parameter estimates in the various conditions) in the simulation will be the same as those in Study 1, above. All of the independent variables in Study 1 were included in Study 2 (i.e., the various causes of missingness). In addition, we varied model mis-specification by simulating linear, quadratic, and cubic level-1 growth curves. It should be noted, however, that we again fit linear level-l models in all cases—and hence demonstrating a naïve model mis-specification error, as commonly found in practice. That is, this simulation mimics the empirical case wherein one fits a linear level-l model irrespective of whether the level-l data relationship is strictly linear in the population. Figure 1 depicts the degrees of mis-specification in the simulation. As one can see, given that the range of time is only four points, the cubic function is just a more extremely non-linear version of the quadratic.

To allow us to compare our results to the findings above, and to give us generalizability to real data situations, we based our simulation parameters on the Willett (1988) data described above. That is, we simulated the same sample size as in Study 1, used the same coding of time, and used the Willett (1998) data parameters to generate the level-1 curve for each simulated study participant. In essence, the Willett (1998) data served as the population structure for the simulation.

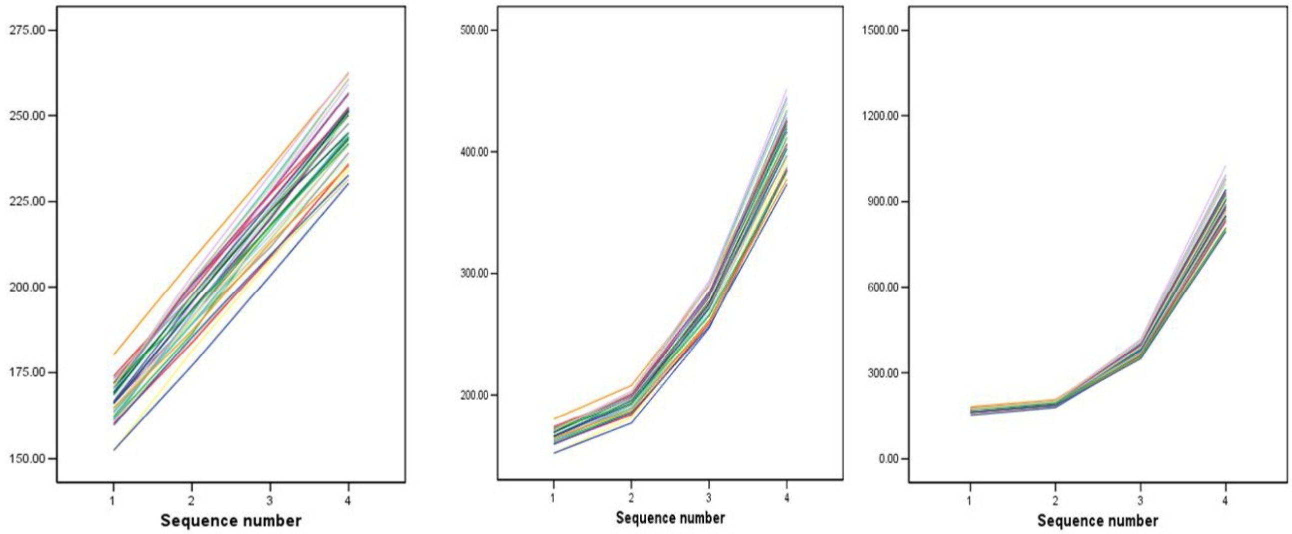

Figure 1. Plots of The Linear, Quadratic, And Cubic Level-1 Curves

For each simulation condition (i.e., for the complete data and for each of the four simulated missing data mechanisms described above) ten replications of the simulation experiment were conducted. We stopped at ten replications because the replication-to-replication variability was very small obviating the necessity for more replications at this point.

Results and Conclusions

Table 3 displays the results of the study. Table 3 is arranged in four sets of columns. The column on the far left of the table describes the data missingness mechanisms investigated in this simulation: no missingness (i.e., complete data), and MCAR, MNAR, MAR Scenario-A, and MAR Scenario-B for both 7 and 14 percent missing data—as used in Study 1 above. Starting from the far right of the table we list the average values (over the 10 replications) of the intercept and time parameters in the level-1 unconditional model. Starting from the far right of the table one can see the columns for the intercept and time parameters for the cubic, quadratic, and linear level-l growth curves. As a reminder, the unconditional linear model was fit for each replication therefore the linear condition represents a correctly specified level-l model whereas quadratic and cubic are mis-specified. It is important to note that one should interpret the values reported in Table 3 for each section (linear, quadratic, and cubic) separately and relative to the baseline condition, the complete data, reported in the first row of the data. Therefore, to interpret the effect of the missing data mechanisms on the correctly specified (linear) model one notes that the complete data intercept is 166.06 and that following down that column one could compare the intercept values for the eight missing data mechanisms to this complete data value.

Table 3

An Unconditional Linear Growth Model (LGM); Level-1 Data Simulated as Linear, Quadratic, and Cubic

		Linear		Quadratic		Cubic	
Data Missingness		*Intercept*	*Time*	*Intercept*	*Time*	*Intercept*	*Time*
Complete Data		166.06	27.03	52.50	237.99	52.47	238.03
MCAR	7%	166.07	27.03	50.47	239.54	52.75	238.05
	14%	166.05	27.03	55.21	233.10	52.56	237.43
MNAR	7%	165.93	27.08	66.97	216.28	66.84	216.34
	14%	165.97	27.08	85.09	189.11	85.04	189.17
MAR Scenario-A	7%	166.00	27.06	53.85	237.45	52.08	238.18
	14%	165.95	27.10	51.15	236.98	49.48	240.74
MAR Scenario-B	7%	165.96	27.10	61.74	224.13	62.05	223.74
	14%	165.97	27.07	72.35	209.95	72.35	210.00

Note: The averages reported in this table are across the ten replications in that corresponding cell. For the ‘linear’ columns all the statistical tests (intercept and time) are statistically significant. For the ‘quadratic’ and ‘cubic’ columns all the statistical tests were statistically non-significant except for the MAR Scenario-B cells wherein all the statistical effects (intercept and time) were statistically significant.

The results are a striking demonstration of the interaction of the cause of missing data with model mis-specification. Focusing on the columns of Table 3 containing the results for the linear (correctly specified) case, the eight values (for each of the intercept and time parameters) are strikingly similar irrespectively of the degree of missingness or the missingness mechanism. In fact, one finds the result that even in what is often considered the “worst case scenario” of MAR Scenario-B or MNAR the intercept and time parameters are nearly the same (to two decimal points) as the complete data. Focusing on statistical significance, it is expected from the simulation design, that the intercept and time parameters would be statistically significant. This was found to be true in all of the correctly specified linear cases.

When one focuses on the quadratic or cubic cases (i.e., when the model we are fitting is mis-specified) then findings from the correctly specified case do not hold up. That is, when compared to their respective complete data cases, the data missingness mechanisms have a distorting effect on the findings—especially, the MNAR and MAR Scenario-B cases. Focusing on the statistical significance, all of the significance tests of the parameters were, as expected by simulation design, statistically significant except for MAR Scenario-B wherein all cases the time effect was statistically non-significant.

In summary, when the level-1 model is correctly specified the parameter estimates are nearly always the same (and the same as the complete baseline data case) irrespective of the degree or kind of missing data. However, when the model is mis-specified the results are quite different than the reference cell of complete data. Furthermore, the results of the mis-specified models all triggered a warning in the software that there was a lack of convergence. We increased the number of iterations for convergence but this warning continued. It is interesting that this warning only occurred in the mis-specified cases; however, in general this warning is neither a necessary nor sufficient sign of model mis-specification.

General Discussion

The results reported herein remind us that Bryk & Raudenbush (1992), Snijders (1996), Hox (2000, 2002) and others are correct in stating that one of the advantages of multilevel (HLM) analysis of longitudinal data is its ability to handle missing data; however, the underlying condition of a correctly specified level-l model is not explicitly stated. Our results highlight that if the model is correctly specified then multilevel (HLM) analysis of longitudinal data is able to handle missing data. If the level-1 model is not correctly specified, multilevel models are not able to handle missing data well and the conclusions are distorted by the various missing data mechanisms. Our findings also highlight the importance of visual inspection of the level-1 curves, data inspection, via graphics. Of course, as we know in statistics more generally, this visual inspection does not guarantee that one will fit the correct model.

References

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical Linear Models: Applications and Data Analysis Methods.* Newbury Park: Sage.

Collins, L. M., Schafer, J. L., & Kam, C. (2001). A comparison of inclusive and restrictive strategies in modern missing data. *Psychological Methods, 6*, 330 - 351.

Goldenberg, C., Gallimore, R., Reese, L., & Garnier, H. (2001). Cause or Effect? A longitudinal study of immigrant Latino parents' aspirations and expectations and their children's school performance. *American Educational Research Association Journal*, *38*, 547-582.

Graham, J. W., & Hofer, S. M. (2000). Multiple imputation in multivariate research. In T. D. Little, K. U. Schnabel & J. Baumert (Eds.), *Modeling longitudinal and multilevel data: Practical issues, applied approaches, and specific examples*. (pp. 201-218). Mahwah, NJ: Lawrence Erlbaum Associates.

Hox, J. J. (2000). Multilevel analyses of grouped and longitudinal data. In T. D. Little, K. U. Schnabel & J. Baumert (Eds.), *Modeling longitudinal and multilevel data: Practical issues, applied approaches, and specific examples.* (pp. 15-32). Mahwah, NJ: Lawrence Erlbaum Associates.

Hox, J. J. (2002). *Multilevel analysis techniques and applications*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Jimerson, S., Egeland, B., Sroufe, L. A., & Carlson, E. (2000). A prospective longitudinal study of high school dropouts: Examining multiple predictors across development. *Journal of School Psychology*, *38*, 525-549.

Little, R. J. A., & Rubin, D. B. (1987). *Statistical analysis with missing data.* New York: Wiley.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. London: Oxford University Press.

Snijders, T. (1996). Analysis of longitudinal data using the hierarchical linear model. *Quality & Quantity, 30*, 405-426.

Willett, J. B. (1988). Questions and answers in the measurement of change. In E. Rothkpof (Ed.), *Review of research in education (1988-1989)* (pp. 345-422). Washington, DC: American Educational Research Association.

Wothke, W. (2000). Longitudinal and multigroup modeling with missing data. In T. D. Little, K. U. Schnabel & J. Baumert (Eds.), *Modeling longitudinal and multilevel data: Practical issues, applied approaches, and specific examples.* (pp. 219-240). Mahwah, NJ: Lawrence Erlbaum Associates.