

An Investigation of Power Analysis Approaches for Latent Growth Modeling

by

Bethany Lucía Van Vleet

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved June 2011 by the
Graduate Supervisory Committee:

Marilyn S. Thompson, Chair
Samuel B. Green
Craig K. Enders

ARIZONA STATE UNIVERSITY

August 2011

©2011 Bethany Lucía Van Vleet
All Rights Reserved

ABSTRACT

Designing studies that use latent growth modeling to investigate change over time calls for optimal approaches for conducting power analysis for a priori determination of required sample size. This investigation (1) studied the impacts of variations in specified parameters, design features, and model misspecification in simulation-based power analyses and (2) compared power estimates across three common power analysis techniques: the Monte Carlo method; the Satorra-Saris method; and the method developed by MacCallum, Browne, and Cai (MBC). Choice of sample size, effect size, and slope variance parameters markedly influenced power estimates; however, level-1 error variance and number of repeated measures (3 vs. 6) when study length was held constant had little impact on resulting power. Under some conditions, having a moderate versus small effect size or using a sample size of 800 versus 200 increased power by approximately .40, and a slope variance of 10 versus 20 increased power by up to .24. Decreasing error variance from 100 to 50, however, increased power by no more than .09 and increasing measurement occasions from 3 to 6 increased power by no more than .04. Misspecification in level-1 error structure had little influence on power, whereas misspecifying the form of the growth model as linear rather than quadratic dramatically reduced power for detecting differences in slopes. Additionally, power estimates based on the Monte Carlo and Satorra-Saris techniques never differed by more than .03, even with small sample sizes, whereas power estimates for the MBC technique appeared quite discrepant from the other two techniques. Results suggest the choice between using the Satorra-

Saris or Monte Carlo technique in a priori power analyses for slope differences in latent growth models is a matter of preference, although features such as missing data can only be considered within the Monte Carlo approach. Further, researchers conducting power analyses for slope differences in latent growth models should pay greatest attention to estimating slope difference, slope variance, and sample size. Arguments are also made for examining model-implied covariance matrices based on estimated parameters and graphic depictions of slope variance to help ensure parameter estimates are reasonable in a priori power analysis.

This project is dedicated to those loving and supportive people who have surrounded me and supported me throughout the years. Despite having watched my mom complete her dissertation as I grew up and saying “I will NEVER write a paper that long,” her example and encouragement were indispensable in completing my own dissertation! My dad’s constant encouragement and example of amazing work ethic helped me maintain motivation and drive. I thank both my parents and my husband who played an integral role in entertaining Ty while I “worked on homework” and I thank Ty, who helps keep life interesting and reminded me that I was more than my dissertation. Finally, I dedicate this work to my husband, who could not be more supportive of me and whose pride in my work was invaluable, and to Grayson, whose timing was incredibly motivational and appreciated!

ACKNOWLEDGMENTS

I would like to acknowledge the incredible help, support, and guidance I received from my chair, Dr. Marilyn Thompson! Her patience and assistance throughout the dissertation process was truly appreciated. I am also indebted to Drs. Green, Enders, and Iida who not only served on my committee, but provided expert input and guidance.

TABLE OF CONTENTS

	Page
LIST OF TABLES.....	viii
LIST OF FIGURES	x
CHAPTER	
1 INTRODUCTION AND BACKGROUND	1
Statistical Approaches for Modeling Growth	6
Multilevel Modeling	7
Latent Growth Modeling.....	9
Latent Growth Modeling of a Treatment Effect	12
Comparison of Growth Modeling Approaches.....	13
Power Analysis Approaches	15
Monte Carlo Simulation.....	16
Satorra-Saris.....	18
MacCallum-Browne-Cai (MBC)	20
Comparison of Power Analysis Approaches	24
Purpose of the Study	30
2 METHOD	32
Data Simulation Procedures and Conditions	33
Monte Carlo Parameter Values: Base Conditions	33
Slope Variance	34
Level-1 Error Variance	35
Slope Difference/Effect Size.....	36

CHAPTER	Page
Sample Size.....	38
Number of Repeated Measures.....	38
Monte Carlo Parameter Values: Error Variance Conditions	38
Monte Carlo Parameter Values: Misspecification	40
Power Analyses Across Techniques	41
Analysis Conditions.....	43
Analysis of Results	45
3 RESULTS	47
Type I Error.....	48
Monte Carlo Power Analyses: Base Conditions.....	49
Sample Size.....	52
Slope Difference/Effect Size.....	53
Number of Repeated Measures.....	53
Slope Variance	54
Level-1 Error Variance	56
Bias, Efficiency, and Coverage.....	57
Global Fit Indices.....	59
Level-1 Error Variance (Conditions with No Misspecification)	61
Misspecification in Error Variance and Form of the Growth Model.....	66
Comparison of Results across Power Analysis Approaches.....	73
4 DISCUSSION	81
Type I Error.....	81

CHAPTER	Page
Number of repeated measures.....	81
Slope variance.....	87
Level-1 Error Variance	88
Misspecification in Level-1 Error Structure and Form of the Growth Model	91
Power Analyses across Techniques	93
Limitations and Future Directions	97
Summary and Suggestions.....	101
REFERENCES	104
 APPENDIX	
A. SAMPLE MPLUS SYNTAX FOR MONTE CARLO POWER ANALYSIS FOR A LINEAR GROWTH CURVE MODEL WITH A DUMMY CODED COVARIATE	108
B. SAMPLE MPLUS SYNTAX FOR THE THREE STEPS IN THE SATORRA-SARIS POWER ANALYSIS FOR A LINEAR GROWTH CURVE MODEL	110
C. SAMPLE SAS SYNTAX TO DETERMINE POWER GIVEN A NONCENTRALITY PARAMETER DETERMINED USING THE SATORRA-SARIS OR MBC TECHNIQUE.....	113
D. SAMPLE SAS SYNTAX TO DETERMINE POWER USING THE MBC TECHNIQUE ACROSS A VARIETY OF RMSEA VALUES FOR A NULL HYPOTHESIS OF NO DIFFERENCE.....	115

LIST OF TABLES

Table	Page
1. Comparison of Selected Power Investigations for Latent Growth Modeling.....	28
2. Model-Implied Correlation Matrices Resulting from Selected Parameter Values in the Linear Growth Model	34
3. Simulation Design.....	44
4. Type I Error and Power when Level-1 Error Variance is Equal across Time	50
5. Absolute Bias (Population Value Minus Model Estimate) and Efficiency (in parentheses) when Level-1 Error Variance is Equal across Time but not Constrained to be Equal in the Model.....	58
6. Mean RMSEA and Standard Deviation (in Parentheses) when Level-1 Error Variance is Equal across Time.....	60
7. Type I Error and Power across Level-1 Error Variance Conditions.....	62
8. Absolute Bias (Population value Minus Model Estimate) and Efficiency (in parentheses) across Level-1 Error Variance Conditions.....	64
9. Mean RMSEA across Level-1 Error Variance Conditions.....	65
10. Average Chi-Square and Standard Deviation (in Parentheses) across Misspecified Level-1 Error Variance Conditions	68
11. Type I Error and Power under Correctly or Incorrectly Specified Form of the Growth Model	69

Table	Page
12. Absolute Bias (Population value Minus Model Estimate) and Efficiency (in parentheses) under Correctly or Incorrectly Specified Form of the Growth Model.....	71
13. Mean RMSEA and Standard Deviation (in Parentheses) under Correctly or Incorrectly Specified Form of the Growth Model.....	72
14. Average Chi-Square and Standard Deviation (in Parentheses) under Correctly or Incorrectly Specified Form of the Growth Model	72
15. Power Across Power Analysis Techniques.....	77

LIST OF FIGURES

Figure	Page
1. Path diagram for linear latent growth curve model with three repeated measures. The target parameter for which power was evaluated was γ_{11} . . .	12
2. Monte Carlo Power for 6 repeated measures and equal level-1 error variance across time (not constrained to be equal in the model).....	51
3. Observed Individual Growth Trajectories, Fit to Linear Growth, for 30 Randomly Selected cases in the Treatment Condition, for Slope Variance Values of 10 and 20.	56
4. Comparison of Power with and without Growth Form Misspecification.....	70
5. Comparison of Power using Different Power Analysis Techniques with 6 Repeated Measures	78
6. Comparison of Power using Different Power Analysis Techniques with 3 Repeated Measures	79
7. Comparison of Power using Different Power Analysis Techniques with 3 and 6 Repeated Measures	80

CHAPTER 1

INTRODUCTION AND BACKGROUND

The study of change and development is integral in the social and behavioral sciences, including education, family studies, psychology, and sociology. Whether researchers are interested in language development (e.g., Peterson & Dodsworth, 1991), video game violence and aggression (e.g., Möller & Krahé, 2008), children's social and academic performance (e.g., Brock, Nishida, Chiong, Grimm, & Rimm-Kaufman, 2008), or the influence of stress on health (e.g., Öhman, Bergdahl, Nyberg, & Nilsson, 2007), statistical techniques for handling data collected over multiple points in time are necessary for longitudinal research. The two most popular frameworks for modeling this type of data are the multilevel modeling (MLM) and SEM-based latent growth modeling (LGM) traditions (Bauer, 2003). Both of these techniques explicitly consider change and growth at the individual and group levels. SEM-based latent growth modeling offers great flexibility. SEM models, in general, are very accommodating to predictive paths, with any given variable able to act as a predictor, outcome, or both (Raudenbush, 2001; Singer & Willett, 2003). Examples of this flexibility can be seen in growth modeling contexts; growth can be treated as both an outcome and predictor, and multiple growth factors can be included in a single model.

However, the increasing popularity of latent growth models as a method of investigating change over time also calls for a better understanding of optimal approaches for conducting power analysis within a growth context. Power

analysis, and, in particular, a priori determination of required sample size, is an important step in designing longitudinal research studies (Muthén & Muthén, 2002; Zhang & Wang, 2009). Through power analysis, investigators can become equipped to understand the likelihood of appropriately rejecting a null hypothesis under various design and analysis conditions. Accordingly, this understanding can assist a researcher in the planning stages of a study (e.g., determining appropriate sample size) and is often a required step in applications for funding. Miles (2003) further argued that power analysis is “not just a statistical or methodological issue, but an ethical issue” (p. 7). Although a number of techniques have been developed to estimate power for certain statistical tests, as explained by Duncan, Duncan, and Li (2003), a number of issues relating to the measurement of power in growth modeling (including intervention effects) must be further examined.

Currently, researchers tend to use one of three methods for conducting power analysis in growth models based on a latent variable framework: the Monte Carlo simulation method (Muthén & Muthén, 2002; Partridge & Lerner, 2007); the Satorra-Saris method (Satorra & Saris, 1985; Muthén & Curran, 1997); or the RMSEA-based MacCallum-Browne-Cai (MBC) method (MacCallum, Browne, & Sugawara, 1996). Whereas the Monte Carlo technique relies on multiple replications generated from a specified population model to determine power for particular parameters, the Satorra-Saris and MBC techniques rely on the estimation of the noncentrality parameter to determine power. Once a noncentrality parameter is calculated, one simply needs to use a noncentral chi-

square distribution table, degrees of freedom, and alpha to determine power. However, the method of estimating the noncentrality parameter differs greatly for the MBC and Satorra-Saris methods. The Satorra-Saris method uses a two-step estimation procedure to estimate the noncentrality parameter by using the estimated mean vectors and covariances matrices derived from a properly specified model to estimate a second, misspecified model (Satorra & Saris, 1985). The resulting likelihood ratio chi-square value thus approximates the noncentrality parameter (Satorra & Saris, 1985). The MBC technique, on the other hand, makes use of overall model fit (based on RMSEA), model degrees of freedom, and sample size to calculate the noncentrality parameter (MacCallum, Browne, & Cai, 2006). In both the Satorra-Saris technique and the MBC technique, once the noncentrality parameter is estimated, the same procedures follow to determine estimated power.

This investigation first and foremost focused on the application of the Monte Carlo approach to power analysis, specifically the process of parameter value estimation, a critical step in the Monte Carlo process. By investigating power under conditions that employed different slope variance values, error variance values, slope values, sample sizes, numbers of repeated measures, and misspecification, results may suggest which parameters are most influential in power estimates, thereby aiding researchers in selecting their values in a Monte Carlo power analysis.

In addition to this focus on the Monte Carlo technique, this investigation also considered the other common power analysis techniques, explicitly

comparing the performance of the Monte Carlo simulation method (Muthén & Muthén, 2002), Satorra-Saris method (1985), and MacCallum-Browne-Cai method (2006) in assessing power under varying model conditions. Because of the accuracy and precision of the Monte Carlo technique (von Oertzen, 2010), it was regarded as the most accurate power estimate among those compared in this investigation; therefore, it was against this approach that the less labor intensive power analysis methods were compared.

Using simulated data, both parts of this investigation focused on the power to detect the difference in slope parameters between a treatment and control group for latent growth models that included a repeatedly-measured continuous outcome, an intercept factor, a slope (i.e., growth) factor, and a two-group dummy-coded covariate. As in Muthén and Muthén (2002), this covariate might be conceptualized as representing a treatment variable, indicating a control and treatment group, or a variable such as gender. For the purposes of this study, it will be referred to as a treatment variable.

Power is the probability of rejecting a false null hypothesis; therefore, in this study, the focus was on the power to detect the statistical significance of a particular parameter that was nonzero in the population. The Monte Carlo approach derives power estimates from the proportion of samples in which a particular parameter is significant, thus investigators using this technique most typically use the Wald test in judging significance. Users of the Satorra-Saris approach, on the other hand, typically employ a likelihood ratio chi-square test in estimating the noncentrality parameter used to determine power. Finally, the

MBC approach focuses on overall model fit and power to detect differences in fit between more- and less-constrained models as a whole using the RMSEA. Despite the typical methods of determining significance (e.g., the Wald test in the Monte Carlo technique), other options do exist. For example, Hertzog, von Oertzen, Ghisletta, and Lindenberger (2008) explain that a likelihood ratio chi-square test can be considered a better option than the Wald test, with two types of likelihood ratio tests available in latent growth modeling. In fact, Zhang and Wang (2009) used a Monte Carlo power analysis approach; however, rather than focusing on the significance of a particular value using the Wald test, they focused on the 2 times log-likelihood difference between nested models. Thus, it is important to recognize unique approaches in judging significance within these common power analysis techniques.

As in Muthén and Muthén (2002), the parameter of interest in this investigation was the regression coefficient that results when regressing the slope growth factor on the treatment variable. This coefficient is of great interest because the influence of a covariate on rate of growth for an outcome is a popular hypothesis investigated in latent growth modeling applications. The power to find this parameter significant (i.e., to detect a decrement in fit when this parameter was constrained vs. unconstrained, as in the MBC approach) was estimated under various conditions.

Data generation conditions varied, included slope variance, error variance, effect size for the difference in slopes, sample size, number of repeated measures, and misspecification. Additionally, when using the MBC power analysis

technique, choices of target RMSEA values and intervals were also varied. Outcomes examined were estimates of power from each of the three power analysis techniques, noncentrality parameters, and, for conditions in which the null hypothesis for the difference in slopes was true, Type I error rates.

This study yields recommendations for the selection of parameter values when using the Monte Carlo technique as well as guidance in selecting power analysis techniques under particular circumstances. Ultimately, it is hoped that findings from this study will make the use of a priori power analyses more accessible and effective for applied investigators.

The application of latent growth modeling coupled with the necessity of power analyses for study design and sample size determination demands explicit attention to the various power analysis techniques used in a latent growth context and the conditions that influence that power. Thus, this chapter comprises a description of approaches to modeling growth and a review of the literature on power analysis in latent growth modeling contexts.

Statistical Approaches for Modeling Growth

Longitudinal data analysis is a popular and important analysis technique. Even with thoughtful research design, if improper analytic methods are employed, rich longitudinal data can become less meaningful, neglecting the importance of individual change and variation (Bryk & Raudenbush, 1987). Therefore, it is important that researchers understand methods of analyzing longitudinal data. Two key frameworks for modeling growth include the multilevel modeling (MLM) and SEM-based latent growth modeling (LGM) traditions (Bauer, 2003).

While these two frameworks are essentially comparable in many circumstances, they grow out of distinct statistical traditions and offer their own advantages and disadvantages.

Multilevel modeling. Multilevel modeling (MLM; also referred to as random coefficient modeling or hierarchical linear modeling) grew in response to data analysis techniques that did not take the clustering of data into account (Chou, Bentler, & Pentz, 1998). Bryk and Raudenbush (1987) described previous methods of analyzing change as being fraught with “inadequacies in conceptualization, measurement, and design” (p. 147). By ignoring the hierarchical structure of data (e.g., students within schools), researchers were faced with biased and inaccurate results (e.g., smaller standard errors, inaccurate coefficients; Chou, Bentler, & Pentz, 1998), the tendency to overlook change on an individual level, and failure to recognize variation in growth rates and initial status among individuals (Bryk & Raudenbush, 1987). Multilevel modeling, therefore, made the important step of explicitly taking the hierarchical nature of data into account by modeling multiple levels of analysis (Singer & Willett, 2003). Because longitudinal data can be framed as having a hierarchical structure (measurement occasions within individuals), studies of individual change eventually found their place in multilevel modeling (Chou, Bentler, & Pentz, 1998), thus allowing one to investigate changes and variations both within individuals and across individuals, rather than ignoring change happening across levels (Bryk & Raudenbush, 1987).

The MLM approach to growth modeling conceptualizes growth at two levels or stages (Bryk & Raudenbush, 1987). Level 1 of the model focuses on individual growth parameters (intercept, π_{0i} , and growth rate, π_{1i}) and measurement error or level-1 residuals (ε_{ti}) predicting an outcome (Y_{ti} for a particular individual, i , at time t) (Singer & Willett, 2003).

$$\text{Level 1: } Y_{ti} = \pi_{0i} + \pi_{1i}(\text{Time}_{ti}) + \varepsilon_{ti} \quad (1)$$

In this manner, every individual in a given longitudinal dataset comes with his or her own level 1 equation that, while assuming the same overall form of growth, allows growth to differ in intercept and slope between individuals (Singer & Willett, 2003). Level 2, on the other hand, captures differences between individuals the influences of time-invariant predictors on growth parameters can be measured. At level 2, the individual growth parameters become the outcomes of their own regression equations and are predicted based on an average growth rate (γ_{10}), average intercept (γ_{00}), measured predictors, and deviations (residual or error) from the averages (ζ_{0i} and ζ_{1i}) (Singer & Willett, 2003).

$$\text{Level 2: } \pi_{0i} = \gamma_{00} + \zeta_{0i} \quad (2)$$

$$\pi_{1i} = \gamma_{10} + \zeta_{1i} \quad (3)$$

This method of modeling change at two levels allows a researcher to simultaneously investigate individual growth as well as group averages and individual variance around those averages (Singer & Willett, 2003). However, these two levels can be combined into a single equation or “composite model” by substituting the level 2 equations for the equivalent growth terms in the level 1 model (Singer & Willett, 2003, p. 80). A basic, linear growth model with no

predictors (unconditional growth model) can thus be represented by the following equations:

$$\text{Composite model: } Y_{ti} = \gamma_{00} + \gamma_{10}(\text{Time}_{ti}) + \zeta_{0i} + \zeta_{1i}(\text{Time}_{ti}) + \varepsilon_{ti} \quad (4)$$

Ultimately, these equations can be conceptualized as consisting of two key components: (1) fixed effects (γ_{00} and γ_{10}) and (2) random effects (ζ_{0i} , ζ_{1i} , and ε_{ti} ; Singer & Willett, 2003). The fixed effects correspond to the influence that predictors and average growth parameters have on individual change trajectories (Singer & Willett, 2003, p. 69). The random effects, on the other hand, capture the differences between expected and observed values of the outcome variable for each individual over time, and thus reflect the variance among individuals (Singer & Willett, 2003, p. 84).

Latent growth modeling. In contrast to the MLM framework, latent growth modeling is part of the structural equation modeling (SEM) tradition. SEM is a flexible analytic approach that is popular in various fields of research, including psychology, education, and biology (Tomarken & Waller, 2005). Structural equation models allow a researcher to conceptualize latent constructs based on measured indicators (i.e., measurement models), and then create relationships between latent constructs (i.e., structural models) (Tomarken & Waller, 2005). Thus, the measurement model takes observed scores, error, and true scores into account for each latent construct while the structural model allows for predictive relationships among various latent and measured variables (Singer & Willett, 2003). In fact, because of the flexibility of SEM, variables (measured or latent) need not be labeled only as “predictor” or “outcome,” but rather “one

predictor's outcome may be another outcome's predictor" (Singer & Willett, 2003, p. 269). It is precisely because of the flexibility of the SEM model that Meredith and Tisak (1990) were able to specify an SEM model that could model growth over time. Although, on the surface, it may appear that using an SEM framework to model repeated measures would result in problems with dependency among measures, by carefully specifying various aspects of the model, one can successfully create a model of change over time (Curran, 2003, p. 530).

In an LGM model, a growth model is conceptualized as a single-level model in which the repeated measures are influenced by time, in conjunction with two required latent factors: One factor that represents the intercept (similar to π_{0i} in equation 1) and one factor that represents growth (similar to π_{1i} in equation 1; Chou, Bentler, & Pentz, 1998). For an unconditional growth model, these factors and their related indicators can be expressed as measurement and structural models as follows:

$$\text{Measurement: } y_{ti} = v_t + \lambda_{0t}\eta_{0i} + \lambda_{1t}\eta_{1i} + \varepsilon_{ti} \quad (5)$$

$$\text{Structural (intercept): } \eta_{0i} = k_0 + \zeta_{0i} \quad (6)$$

$$\text{Structural (growth): } \eta_{1i} = k_1 + \zeta_{1i} \quad (7)$$

These equations are similar in form to equations 1, 2 and 3 from the MLM framework, with the MLM level 1 model represented as a measurement model and the MLM level 2 model represented as structural models (Raudenbush, 2001).

Here, an individual's score on a given outcome at a given time (y_{ti}) is predicted by the measurement model's intercept (v), underlying factors η_{0i}

(intercept) and η_{1i} (growth), and an error or residual term (Bauer, 2003; Chou, Bentler, & Pentz, 1998). Further, as in the MLM level 2 models, the latent factors are predicted based on the sum of a given value and an error or residual term. However, while the structural model equations appear quite similar to the MLM level-2 models, at this point, the LGM measurement model is not fully comparable to the MLM level 1 model (notice the presence of extra parameters). However, by appropriately constraining parameters, the MLM and LGM models become comparable. First, the intercept (ν) in the measurement model must be constrained to zero in order to pass the information carried by that parameter onto the latent variables in the form of k_0 (the intercept factor mean) and k_1 (the growth factor mean). Next, the value of λ_{0t} must be set to equal 1 and the values of λ_{1t} must be set equal to the appropriate measures relating to the passage of time (Bauer, 2003; Chou, Bentler, & Pentz, 1998; Singer & Willett, 2003). Following these constraints, one can more readily recognize the similarities between the MLM and the LGM models:

$$\text{LGM Measurement Model: } y_{ti} = \eta_{0i} + \eta_{1i}(\text{Time}) + \varepsilon_{ti} \quad (8)$$

$$\text{MLM Level 1 Model: } Y_{ti} = \pi_{0i} + \pi_{1i}(\text{Time}_{ti}) + \varepsilon_{ti}$$

Similarities are also apparent in considering how LGM captures MLM parameters through a path diagram for a linear latent growth curve model (see Figure 1). As with more general SEM models, additional predictors (time-variant, time-invariant, and even additional measurement models) can also be incorporated into the structural model(s). For example, Figure 1 includes a time-invariant dummy-coded predictor variable (representing treatment vs. control conditions).

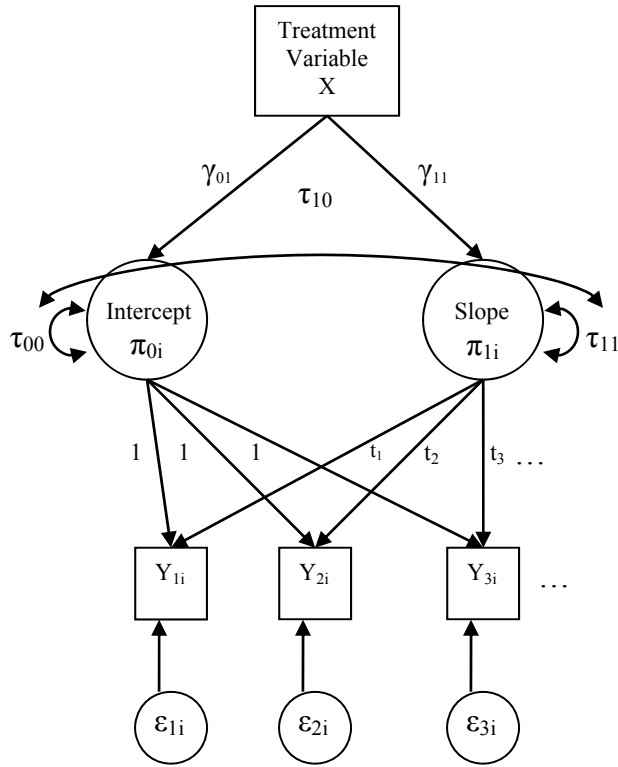


Figure 1. Path diagram for linear latent growth curve model with three repeated measures. The target parameter for which power was evaluated was γ_{11} .

Latent growth modeling of a treatment effect. When the focus of a power analysis is on the influence of a treatment on growth in a growth model, a number of options for how to model that treatment in a latent growth context are available. Muthén and Muthén (2002) opted to use a time-invariant binary variable to represent control versus treatment conditions. This technique involves the analysis of just one population, which offers simplicity to the practitioner. According to Muthén and Curran (1997), however, this technique is restrictive in that it does not allow for variations in other parameters between the treatment and control groups. Muthén and Curran (1997) used a two-group model, arguing that

only this approach offers “generality in the modeling” (p. 376). Therefore, rather than conceptualizing the sample as coming from a single population, where participants are coded as belonging to one of two treatment conditions, Muthén and Curran modeled two distinct populations (a treatment population and control population). By formulating the model in this way, one can then constrain the intercept factor and growth factor to be equal between these two groups and add a third growth factor for the treatment population. This third factor then represents the influence of the treatment on growth (in addition to any growth expected in the absence of the treatment) (Muthén & Curran, 1997). In practice, both the dummy coded technique (e.g., Muthén & Muthén, 2002; Fan, 2003) and multi-group technique (e.g., Muthén & Curran, 1997) are employed, with selection of a particular technique being influenced by analysis goals and contexts (e.g., a power analysis that does not manipulate differences in variance between the treatment and control group may lend itself to the simplicity of the dummy coded technique).

Comparison of Growth Modeling Approaches. In comparing the MLM and LGM approaches and how the MLM parameters were incorporated into an SEM framework, it is possible to see why Curran (2003) explained that “the boundaries between these two modeling strategies are becoming increasingly porous... We seem to be approaching a point in which the terms SEM and MLM better distinguish historical roots and commercial software rather than the underlying statistical models” (p. 565). Further, not only are these models mathematically comparable with sufficient constraints of the SEM model, but

applied studies have revealed that, with appropriate parameterizations, they do indeed yield comparable, if not identical, results (Bauer, 2003; Chou, Bentler, & Pentz, 1998; Curran, 2003; Tomarken & Waller, 2005). However, because of differences in strategy, in combination with statistical software idiosyncrasies, each framework does come with different advantages and disadvantages.

One drawback to latent growth modeling relates to measurement occasions and spacing between occasions. In MLM, by modeling time as a predictor in a “person-period” dataset where each measurement occasion is a row of data, it is quite simple to include individuals who vary in the number of measurement occasions and time between measurement intervals (one simply needs to add additional rows of data to the dataset; Raudenbush, 2001; Singer & Willett, 2003). However, in LGM, when time is treated as fixed loading values, it becomes more difficult to allow for varied measurement schedules and occasions. In the past, LGM was not considered optimal when observations were unbalanced on time (Farkas, 2008; Mroczek, 2007; Singer & Willett, 2003); however, advances in analysis software (such as Mplus 6) have made it easier to vary time intervals and measurement occurrences in LGM (Mehta & West, 2000; Mroczek, 2007). Thus, MLM is no longer the only framework that can accommodate longitudinal data unbalanced on time; although it could be argued that the process of analyzing data that is unbalanced on time is still simpler using an MLM approach and software program. Despite this potential complication with the latent growth approach, the flexibility of latent growth modeling makes LGM an attractive and prevalent analysis technique.

The latent growth modeling technique has become a widely used method for longitudinal data analysis primarily because of its flexibility. Mediation effects, for example, can be easily included in a latent growth model (Mroczek, 2007). Additionally, an LGM can extend beyond conceptualizing growth as only an outcome variable, with growth trajectory and intercept becoming potential predictors of other outcomes. Further, a given growth model can actually be predicted by another growth model (“growth on growth analysis”) (Singer & Willett, 2003, p. 274). Another cited advantage of the LGM framework is that error can be better estimated through the use of measurement models (Mroczek, 2007). Based on the advantages and popularity of the latent growth approach, it is this approach that is the focus of the present investigation. In fact, the flexibility of SEM is what allows for a number of power analysis techniques to be employed in analyzing the power of latent growth curves.

Power Analysis Approaches

A priori power analyses are an integral part of planning a study, allowing an investigator to understand how various conditions will influence the ability to find an effect of a particular parameter (or differences between models), and determine required sample sizes for doing so. While a number of power analysis techniques exist, the most popular techniques in the latent growth modeling context include the Monte Carlo simulation method (Muthén & Muthén, 2002), the Satorra-Saris method (Satorra & Saris, 1985), and the MacCallum-Browne-Cai (MBC; 2006) method.

Monte Carlo simulation. The Monte Carlo simulation approach makes use of multiple replications based on generated population data (Muthén & Muthén, 2002). Sample syntax for using the Monte Carlo approach in Mplus is available in Appendix A. In this technique, parameter values are theoretically estimated based on past research, pilot data, and/or theory. It is this step of selecting parameter values that often proves to be the most difficult part of the Monte Carlo power analysis process and it is not uncommon to lack pilot data or appropriate past research to assist in selecting these values. Therefore, one must consider broad findings relating to parameter values and, to an extent, rely on trial and error in finding parameter values that behave realistically. For example, Hertzog et al. (2008) report that slope variance is typically small to moderate compared to intercept variance. Based on this information, Hertzog et al. opted to set intercept variance at 100 with slope variance equal to either 50 (1:2 variance ratio) or 25 (1:4 variance ratio), noting that variance ratios are often smaller than 1:4. Hertzog et al. also set error variance to 100, 25, 10, or 1, treating error variance as homogenous across time. Muthén and Muthén (2002), on the other hand, set the intercept variance at .5 and slope variance at .1 (1:5 variance ratio) in a basic growth model. Similarly, Muthén and Curran (1997) also report that a 1:5 variance ratio is often seen in applied growth models. Once the slope variance, intercept variance, and residual variance values are set (based on past research, pilot data, or general parameter value recommendations, as above), it becomes much easier to explore possible values for remaining parameters.

Using these estimated parameter values, many samples can be generated and a growth model can be estimated for each sample. Using an appropriate analysis program, one can quickly see the proportion of samples that correctly rejected the null hypothesis for a particular parameter (i.e., produced a significant parameter estimate based on the parameter of interest, such as a treatment effect parameter). The proportion of samples that produced the significant parameter is the power estimate (Muthén & Muthén, 2002). When using Mplus for this technique of power estimation, output also includes information regarding bias (in standard error and parameter estimates) and coverage (how often the parameter value was contained in a 95% confidence interval across replications; Duncan & Duncan, 2004).

According to Muthén (2002), the Monte Carlo power analysis technique is often accurate with 500 replications, although increasing replications increases the precision and dependability of the estimates (with 10,000 replications being a very conservative, yet potentially time consuming, choice). Von Oertzen (2010) indicated that “with increasing repetitions, the result of the Monte Carlo simulation converges to the precise power of the SEM” (p. 260). It is precisely because Monte Carlo simulations can converge on the “exact power” (von Oertzen, 2010, p. 262) that this technique is typically used to determine the accuracy of other power analysis techniques (e.g., Muthén & Curran, 1997; Satorra & Saris, 1985). In addition to potential accuracy in determining power, another key advantage to this technique is that an investigator can easily alter the population parameters, missingness of data, normality of distributions, and sample

sizes in order to investigate how power is influenced by these factors. Overall, Abraham and Russell (2008) described this simulation approach as very flexible and Maxwell, Kelley, and Rausch (2008) proclaimed that, “a general principle of sample size planning appears to hold: Sample size can be planned for any research goal, on any statistical technique, in any situation with an a priori Monte Carlo simulation study” (p. 553).

Satorra-Saris. The Satorra-Saris method is another common method of analyzing power. This two-step technique focuses on estimating power for very specific parameters (effects) in the growth model (Duncan et al., 2003). Sample syntax for using the Satorra-Saris approach in Mplus is available in Appendix B. Note that although the Satorra-Saris technique can ultimately be completed in two steps, three steps are used in Mplus, with the second step included in order to double check that the model was properly set up and parameter estimates are correctly retrieved based on what was entered in step one.

Ultimately, two models are compared in this technique: a model that is assumed to be correctly specified and a second, more-constrained model that is nested within the correctly specified model and is assumed to be misspecified. The first step involves specifying the correct model and analyzing that model in order to obtain mean and covariance values. As with the Monte Carlo method, one may determine parameter values based on past research, pilot data, and/or theory. The second step involves using the estimated mean vectors and covariance matrices derived from step one to specify an incorrect model by fixing the effect of interest (e.g., the treatment effect, as in Muthén & Curran, 1997) to

zero. Based on the findings of Satorra and Saris (1985), the likelihood ratio chi-square value derived from step two approximates the noncentrality parameter, which can then be used to determine power based on degrees of freedom, alpha, and chi-square tables (Muthén & Curran, 1997). Appendix C presents SAS syntax that can be used to determine power based on the noncentrality parameter, degrees of freedom, alpha, and chi square. An alternative (yet comparable) approach, in which one need only specify a single model, is also available (Duncan, Duncan, Strycker, & Li, 2002). Rather than using the likelihood ratio chi-square test derived from two specified models, a researcher can rely on the Lagrange multiplier (LM) or Wald (W) test statistics that, like the likelihood ratio chi-square, also correspond to the noncentrality parameter. A researcher could, therefore, specify a model with the parameter of interest set to zero and focus on the LM statistic or specify a model that estimates the parameter of interest and focus on the W statistic. As with the likelihood ratio test discussed previously, one can use these values in conjunction with the degrees of freedom, alpha, and sample size to determine power using the chi-square distribution tables (Duncan, Duncan, Strycker, & Li, 2002).

Simulation studies suggest these Satorra-Saris based techniques for estimating power are accurate even when sample size is “quite small” (although a specific sample size is not reported) (Satorra & Saris, 1985, p. 89). Similarly, Muthén and Curran (1997) report that simulation studies indicate this technique is “sufficiently accurate for practical purposes at small sample sizes,” citing Curran’s (1994) study as an example, in which he found “very good results at

sample sizes of 100” (p. 382). However, Duncan and Duncan (2004) and Muthén (2002) explain that the Satorra-Saris technique is inappropriate when data is missing or nonnormal.

MacCallum-Browne-Cai (MBC). MacCallum, Browne, and Sugawara (1996) proposed a power analysis technique that aims to be easier for applied users of SEM to utilize than other popular techniques, such as those discussed above. MacCallum, Browne, and Cai (2006) further developed this technique to include comparisons of nested models. Rather than focusing on the power of a particular parameter or parameters, as done in the Satorra-Saris and Monte Carlo techniques, MBC assesses overall model fit, thus eliminating the need to estimate various model parameters. Although the MBC technique was not intended, and has not been recommended, for calculating power to detect a specific parameter, it is not unlikely that this technique has been employed by applied investigators in this way, which is why this technique, although different from the Monte Carlo and Satorra-Saris techniques, is considered here.

According to MacCallum, Browne, and Sugawara (1996), the MBC method shares identical assumptions and distributional approximations as the Satorra-Saris technique and also employs the same calculations to determine power once the noncentrality parameter is estimated (see Appendix C for SAS code that can convert the noncentrality parameter into a power estimate). Therefore, it is the processes of estimating the noncentrality parameter that sets the MBC technique apart. Overall, the MBC technique defines the noncentrality parameter, using model RMSEA values, as follows:

$$\lambda = n (d_A \varepsilon_A^2 - d_B \varepsilon_B^2) \quad (9)$$

where d_A and d_B are degrees of freedom for models A and B, ε_A and ε_B equal RMSEA values for models A and B, and n represents one minus sample size (MacCallum, Browne, & Cai, 2006). Through this calculation, it is possible to determine the power of detecting defined differences in model fit. Similar to the Satorra-Saris approach, it is also possible to investigate power related to difference in fit between nested models using the MBC technique (MacCallum, Browne, & Cai, 2006). In order to approximate the Satorra-Saris and Monte Carlo approaches of determining power to detect a particular parameter, one can simply determine the degrees of freedom that result from constraining the parameter of interest to zero (Model A) and use that information in conjunction with the degrees of freedom for the unconstrained model (Model B).

Typical hypothesis testing and power analysis techniques (e.g., Satorra-Saris method) in an SEM context tend to consider tests of exact fit, with a null hypothesis predicting no discrepancy between models where the distribution of the null hypothesis is a central chi square distribution. However, MacCallum et al. (2006) have argued that exact fit is not realistic and the likelihood of finding exact fit (accepting the null hypothesis) diminishes as sample sizes increase. As an alternative, MacCallum et al. have also proposed a method of testing a null hypothesis of small difference. This approach is similar to the process of testing a null hypothesis of exact fit; however, with a null hypothesis of small difference, an investigator can specify the difference in fit for the null hypothesis using

RMSEA values, which then makes both distributions (under the null and alternative hypotheses) noncentral chi square distributions.

According to MacCallum et al., the selection of RMSEA values to include in the MBC equation for the noncentrality parameter is an important decision that must be made carefully, particularly when “ N is moderate and d_A and d_B are also moderate (probably less than 50 or so, and greater than 5 or so, as rough guidelines), with the difference not being very large” (p. 30). MacCallum et al. first recommend that RMSEA values fall in the “midrange of the scale, roughly .03 to .10” (p. 31). Next, they recommended “choosing pairs of values with moderate to large differences (e.g., at least .01-.02)” (p. 31). Finally, they recommended exploring prior research relating to model fit, selecting smaller RMSEA values when a given model has been shown by past research to fit reasonably well. MacCallum et al. further explained that nested models differing greatly in the numbers of parameters likely merit RMSEA values that differ more than nested models differing by only a parameter or two. Despite these guidelines, MacCallum et al. ultimately recognized that it is not uncommon to have little guiding information in selecting these RMSEA values. Therefore, they suggested computing power or sample size using multiple pairs of RMSEA values. The basic suggestion was to compute power with pairs of RMSEA values that (1) “[represent] a small difference in the low range (e.g., = .05 and = .04),” (2) “[represent] a larger difference in the higher range (e.g., = .10 and = .07)” and then (3) explore a variety of pairs within “the recommended range” (p. 31). For this last step, MacCallum et al. suggest a SAS program that computes power or

required sample size by incrementally varying pairs of RMSEA values (see Appendix D). MacCallum, Browne, and Cai (2006) also recommend against using observed RMSEA values (from sample data) in determining power, arguing that “no new information is provided by observed power” (p. 32).

In exploring the MBC approach, one might consider the use of alternative measures of fit in estimating power. In fact, MacCallum, Browne, and Cai (2006) suggest that the GFI and AGFI initially appear to be appropriate for these power analysis calculations. However, after further explorations showed these indices yielded problematic results, MacCallum et al. concluded that RMSEA was the best fit index for this power analysis technique.

The MBC technique offers several advantages in comparison with other power analysis methods. According to MacCallum, Browne, and Sugawara (1996), one benefit is that this technique is not model specific and specification of an alternative model is not required. Therefore, the MBC technique is simple to apply and requires only RMSEA values and model degrees of freedom for power estimation and sample size determination. Additionally, this technique is flexible in that researchers are not restricted to testing exact fit. Rather, the hypotheses specified by designated RMSEA values can represent hypotheses of close or not close fit as well. Indeed, compared to the Satorra-Saris and Monte Carlo techniques, the MBC technique is simple and greatly decreases the amount of parameter estimation required to calculate power; however, in considering the guidelines for selecting appropriate RMSEA values, it is evident that this technique does not completely eliminate the need to make decisions that could

potentially influence resulting power calculations. Further, one may wonder if the benefit of eliminating the need for model specifications could result in decreased precision in power and required sample size estimates.

Comparison of power analysis approaches. Among the three described approaches to power analyses in SEM are similarities and differences that make each technique potentially useful. Assuming sufficient replications are used, the Monte Carlo approach provides a means of assessing the accuracy of other power analysis methods. For example, Satorra and Saris (1985) compared their power approximation technique to results from a Monte Carlo study (with 300 replications) to indicate the accuracy of their technique. Satorra and Saris reported that, for a significance level of .05 and sample size of 100 for a given model, the Monte Carlo technique estimated power of .390 whereas the Satorra-Saris technique estimated power of .407. Increasing sample size to 600 further narrowed this gap, with the Monte Carlo technique yielding a power of 1 and the Satorra-Saris technique resulting in power of .998. Muthén and Curran similarly used the Monte Carlo technique, based on 1,000 replications, to verify the accuracy of the Satorra-Saris technique for their two-group model of interest. Muthén and Curran found that “for a treatment effect size of 0.30 and a total sample of 200 divided equally among control and treatment group observations, the Satorra-Saris method obtained a power of 0.734 as compared with 0.755 from the simulation. An even better agreement was obtained at the higher total sample size of 500 with a treatment effect size of 0.20 where the Satorra-Saris method obtained a power of 0.783, whereas the simulation resulted in 0.780” (p. 383).

Despite some slight differences in power estimates resulting from the Monte Carlo and Satorra-Saris approaches, particularly for smaller sample sizes, the Satorra-Saris technique has enjoyed many years of popularity due to its relative simplicity in comparison with the Monte Carlo approach. However, the Satorra-Saris technique also has its drawbacks. First, in some conditions, the Satorra-Saris technique may result in inaccurate power estimates (Hertzog, von Oertzen, Ghisletta, and Lindenberger, 2008). For example, Hertzog, Lindenberger, Ghisletta, and von Oertzen (2006) had used the Satorra-Saris technique with success, indicating that their findings were comparable to findings of the Monte Carlo simulation technique. However, in their 2008 study, Hertzog et al. reported that the Satorra-Saris technique produced unacceptable results, finding errors in approximation with the Satorra-Saris technique compared to the Monte Carlo technique. These errors appeared to be related to the fact that Hertzog et al. were comparing two slightly misspecified, nested models. Because of these errors, Hertzog et al. (2008) ultimately used the Monte Carlo simulation technique to measure power.

Von Oertzen (2010) also suggested there can be problems with accuracy with the Satorra-Saris technique, explaining that with this technique “power can be computed very rapidly, but the accuracy of the result cannot be improved beyond a small approximation error (2% in this example)” (p. 262). Although Monte Carlo simulations may take longer to run, von Oertzen suggests that this technique will “eventually give the exact power with a very low standard deviation” (p. 262). In fact, von Oertzen, in his investigation of power

equivalence in SEM, compared the traditional Monte Carlo and Satorra-Saris techniques on accuracy and run time. His findings indeed support the accuracy of the Monte Carlo technique (1,000 iterations run in 3580 milliseconds with mean power of .9625, $SD = .0059$) but the speed of the Satorra-Saris technique (run in .6208 milliseconds with power of .9863). The Satorra-Saris technique, therefore, offers a faster method of estimating power that differs slightly from the more accurate Monte Carlo technique, with the Satorra-Saris power estimate of .9863 and the Monte Carlo power estimate of .9625.

MacCallum, Browne, and Cai (2006), on the other hand, outlined situations in which either the MBC or Satorra-Saris technique may be good options for estimating power, with their focus on these less labor intensive power estimate techniques (compared to the Monte Carlo approach). First, MacCallum et al. argued that one should make use of all available information (e.g., pilot data, prior research) when calculating power. Therefore, in situations in which a great deal of prior work has been done, providing a foundation for estimating all required parameters, and sufficient knowledge about the model and model parameters merits specific questions relating to particular parameters, the Satorra-Saris technique is an appropriate choice for power estimation. Similar reasoning would suggest that the Monte Carlo technique could also be appropriately employed in such instances. Additionally, more simple models that require fewer estimates of parameter values (such as latent growth curve models, according to MacCallum et al.) are also good candidates for the Satorra-Saris (or Monte Carlo) technique. However, MacCallum et al. (2006) posited that in those circumstances

in which prior knowledge and theory do not provide sufficient information to create reasonable parameter estimates, the MBC technique would be the appropriate choice.

Overall, the Monte Carlo technique allows for accuracy, whereas the Satorra-Saris technique is a simplification over the Monte Carlo technique that allows for faster, slightly less accurate power estimations, and the MBC technique is the simplest of the power analysis methods that requires no model specification (beyond degrees of freedom). Although each of these techniques has a place in power analysis, with MacCallum et al. even recommending power analysis techniques other than the MBC when possible, it is still quite feasible that investigators do not properly consider the strengths and weaknesses of each technique based on specific contexts. In fact, it is not surprising if some investigators simply default to the simpler MBC approach without considering more complex, yet potentially more precise and accurate, power analysis alternatives. It is for these reasons that these techniques should be explicitly compared under varying data and model contexts, thus allowing investigators to be more informed when selecting a power analysis approach.

In addition to considering how to determine power, it is important to consider what influences power to aid in a comparison of the approaches across various contexts in which power changes. A summary of investigations of power in latent growth modeling contexts is presented in Table 1.

Table 1

Comparison of Selected Power Investigations for Latent Growth Modeling

	Power Analysis Technique	Model Type	Focal Parameter	Manipulated Factors	Key Results
Muthén & Curran (1997)	Satorra-Saris	Linear, multiple-population	Treatment effect (on growth, intercept, and both)	Sample size (100 – 1,000), effect size (.2, .3, .4, .5), number of measurement occasions (3, 5), length of study (3-7 time points), balanced/unbalanced data across groups (proportion in treatment group from .1 to .9)	Balanced data does not necessarily maximize power when more variance is in treatment group Increase sample size = greater power, increase number of measurements = greater power, increase length of study = greater power, increase effect size = greater power Missing data = less power
Muthén & Muthén (2002)	Monte Carlo	Linear with 4 measurement occasions and a dummy coded covariate	Regression coefficient for the regression of the growth parameter on a dummy coded covariate	Effect size (.1, .2), data missingness (no missing vs. covariate=0, t1-t4 missing-12%, 18%, 27%, 50%; covariate=1, t1-t4 missing-12%, 38%, 50%, 73%)	
Fan (2003)	Monte Carlo	2-group, linear growth with 5 measurement occasions	Group differences in slope and intercept	Sample size (50, 100-1,000 in increments of 100), 5 patterns of group differences in growth trajectory - intercepts and slope (with effect sizes of 0, .2, .5, .8)	Pattern of group differences influences power
Hertzog, Lindenberger, Ghisletta, and von Oertzen (2006)	Satorra-Saris	Linear simultaneous growth	Covariance of slopes for two variables over time	Sample size (200, 500), effect size (.25, .5, .75), number of measurement occasions (3, 4, 5, 6, 10, 20), growth curve reliability (.5-.99 in .005 increments), slope variance (50 or 25 at t19)	Increase growth curve reliability = greater power; With effect size of .5, sample of 200, 3-6 measurement occasions, GCR needed to be .85 for power of .80; Best when GCR is above .90

Hertzog, von Oertzen, Ghisletta, and Lindenberger (2008)	Monte Carlo	Linear simultaneous growth	Slope variance	Sample size (100, 200, 500, 1,000), growth curve reliability (.5, .8, .91, .99), slope variance (50, 25, 0), intercept-slope correlation (-.5, -.25, 0, .25, .5), occasions of measurement (0/2/4, 0/2/4/6, 0/2/4/6/8, etc.), variance tests (generalized variance test, specific variance test, Wald test)	“Profound effect of GCR on power” (p. 557)
Zhang and Wang (2009)	Chi-square difference simulation technique	Linear model and non-linear model with 6 measurement occasions		Sample size (50-1,000), effect size (.1, .2, .3), number of measurement occasions (3-6), data missingness (no missing vs. attrition of 10%)	Focus was on the development of a SAS macros to calculate power

A great deal of research has been conducted to help inform investigators as to which factors can influence power, including some recommendations specific to latent growth modeling. For example, in the context of growth curve models, it has been reported that power increases as the number of measurement occasions increases (Hertzog, Lindenberger, Ghisletta, & von Oertzen, 2006; Zhang & Wang, 2009), as growth curve reliability increases (level-1 error decreases; Hertzog et al., 2006), and as correlations between repeated measures of an outcome become stronger (Murphy & Myors, 2004). Additionally, as in other types of models, power in growth models is also influenced by sample size, effect size, and missing data (Hertzog et al., 2008; Jung & Ahn, 2003; Zhang & Wang, 2009).

Purpose of the Study

This investigation had two aims related to power for detecting the difference in slopes between populations in a latent growth curve modeling context. First, the impacts of variations in particular specified parameters (slope variance, error variance, and slope difference), design features (sample size, number of repeated measures), and model misspecification in simulation-based power analyses were studied. Second, power estimates were compared across three common power analysis techniques: the Monte Carlo method (Muthén & Muthén, 2002); the Satorra-Saris method (Satorra & Saris, 1985); and the MacCallum-Browne-Cai (MBC) method (MacCallum, Browne, & Cai, 2006). By understanding how parameter values and model design features within the Monte Carlo technique influence power, it may be possible to provide applied investigators with guidance in parameter value selection, model specification, and study design. Additionally, by comparing powers and, for null conditions, Type I errors across different power analysis techniques, investigators can also make informed decisions regarding the selection of power analysis techniques.

The model of interest in this investigation included a repeatedly measured continuous outcome, with the growth trajectory on this outcome specified by an intercept factor and a slope factor. A dummy-coded covariate was also included to represent two treatment conditions (control vs. treatment; see Figure 1). This model is similar to the latent growth model explored in Muthén and Muthén (2002). The focal parameter was the coefficient for regression of the slope growth factor on the treatment variable.

Power for this model was estimated for manipulated conditions including slope variance, error variance, difference in slopes, sample size, number of repeated measures, and misspecification of the error structure and the form of the growth trajectory. Additionally, when using the MBC power analysis technique, choices of target RMSEA values and intervals were varied.

CHAPTER 2

METHOD

This investigation first focused on the process of parameter estimation in the Monte Carlo power analysis technique. By varying slope variance, error variance, and slope difference (effect size) under varying conditions, it was possible to consider how parameters may influence resulting power estimates. In addition to the focus on the Monte Carlo technique, all three popular methods of a priori power analysis in a latent growth modeling context (Monte Carlo simulation technique, Satorra-Saris technique, and MacCallum-Browne-Cai technique) were compared.

The focal parameter in each of these investigations was the regression coefficient that resulted when regressing the slope growth factor on a dummy coded treatment variable. Therefore, when using the Monte Carlo technique, the model was specified to include this significant covariate. When using the Satorra-Saris technique, on the other hand, two models were specified, including one that was specified correctly to include a regression coefficient based on the slope being regressed on the treatment variable and one that misspecified the model by setting the regression coefficient of interest to zero. Similarly, in using the MBC power technique, the model was treated as a nested model comparison based on the same degrees of freedom for the two models specified for the Satorra-Saris technique.

The varying parameter estimates and the three power analysis approaches were utilized under various simulated data generation conditions in order to

understand how these techniques compared across common scenarios known to influence power estimation and sample size determination.

Data Simulation Procedures and Conditions

In order to investigate power under various conditions and across power analysis techniques, multivariate normal data were simulated for both the Monte Carlo and Satorra-Saris techniques. The MBC technique did not require the use of simulated data as only model degrees of freedom and RMSEA values were required for that technique.

Monte Carlo parameter values: Base conditions. An initial investigation focused on parameters and conditions that influence power using a Monte Carlo power analysis. In specifying model parameters, intercept variance was set at 100 (as in Hertzog, Lindenberger, Ghisletta, & von Oertzen, 2006). The intercept-slope covariance was also held constant with a correlation value of .2. Initial status for both groups, the control group growth over time, and the intercept for the outcome variables were all set to zero. Additionally, the regression coefficient for the regression of the intercept growth factor on the dummy coded treatment variable and the correlations among level-1 error variances across time were set to zero. These parameter values, and those that follow, were selected based on past simulation studies as well as an investigation of the model-implied population correlation matrices (for the outcome variable) that resulted based on these parameter values. In terms of the model-implied correlation matrices, the goal was to select parameter values that resulted in relatively realistic correlation matrices in which correlations were not unusually

high or unlikely (although, it is recognized that some correlation values were likely higher than one might see in an applied setting). See Table 2 for the model-implied population correlation matrices that resulted from the selected parameter values.

Table 2

Model-Implied Correlation Matrices Resulting from Selected Parameter Values in the Linear Growth Model with an Intercept Variance of 100

		Slope Variance = 10						Slope Variance = 20					
		Y1	Y2	Y3	Y4	Y5	Y6	Y1	Y2	Y3	Y4	Y5	Y6
Error Variance = 50	Y1	1.00						1.00					
	Y2	0.66	1.00					0.65	1.00				
	Y3	0.63	0.72	1.00				0.59	0.75	1.00			
	Y4	0.58	0.71	0.78	1.00			0.53	0.73	0.83	1.00		
	Y5	0.54	0.69	0.78	0.83	1.00		0.48	0.70	0.83	0.88	1.00	
	Y6	0.50	0.66	0.77	0.84	0.87	1.00	0.43	0.68	0.82	0.89	0.92	1.00
Error Variance = 100	Y1	1.00						1.00					
	Y2	0.50	1.00					0.50	1.00				
	Y3	0.49	0.57	1.00				0.47	0.61	1.00			
	Y4	0.46	0.57	0.65	1.00			0.43	0.61	0.72	1.00		
	Y5	0.44	0.57	0.66	0.72	1.00		0.39	0.60	0.73	0.79	1.00	
	Y6	0.41	0.56	0.66	0.73	0.78	1.00	0.36	0.59	0.73	0.81	0.85	1.00

Slope variance. Because Hertzog et al. (2008) found that slope variance related to power, and in order to explore the effect of selected slope variance values on resulting power estimates in Monte Carlo analyses for the model of interest in this investigation, slope variance was set to be equal to either 10 or 20.

These values include the range of recommended variance ratios suggested by Hertzog et al. (2008), Muthén and Muthén (2002), and Muthén and Curran (1997).

Level-1 error variance. Level-1 error variance of the continuous outcomes were set to either 50 or 100. These values are similar to those used in Muthén and Muthén (2002) (who set intercept variance to .5, slope variance to .1, and error variance to .5) and Hertzog et al. (2008) (who set intercept variance to 100, slope variance to 25 or 50, and used error variance values of 1-100). Hertzog et al. (2008) conceptualized changes in error variance as changes in growth curve reliability (measured at wave one as intercept variance divided by total variance), resulting in growth curve reliability values of .5 to .99. With an intercept of 100 in the present investigation, an error variance of 50 corresponds to a growth curve reliability of .67 and an error variance of 100 corresponds to a growth curve reliability of .5.

Muthén and Muthén's parameter values (intercept variance of .5, slope variance of .1, 0 slope-intercept covariance, and error variance of .5) resulted in an *R* squared value for error variance of .5 at wave one and would result in an *R* square value of .86 at wave six. In most cases, similar *R* square values were obtained using the selected error variance values in the present investigation. Using an error variance of 50 and slope variance of 10 resulted in a wave one *R* square value of .67 and a wave six *R* square value of .89. An error variance of 50 and slope variance of 20 resulted in a wave one *R* square value of .67 and a wave six *R* square value of .93. An error variance of 100 and slope variance of 10

resulted in a wave one R square value of .5 and a wave six R square value of .81. Finally, an error variance of 100 and slope variance of 20 resulted in a wave one R square value of .5 and a wave six R square value of .87.

When using the Monte Carlo approach, error variance was specified to be homogenous across time in the population but was not constrained to be equal in the model.

Slope difference/effect size. In considering the parameter of interest (the regression of the slope on the treatment group covariate), it was important to determine a method of operationalizing the magnitude of the slope difference between the control and treatment conditions. Feingold (2009) explained that calculating effect size in repeated measures models has been “controversial because there are two possible denominators that can be used in the formula: (1) the standard deviation of the pretest-posttest change scores that reflect within-group variations in improvement over the course of the trial (Gibbons, Hedeker, & Davis, 1993; Mullen & Rosenthal, 1985; Rosenthal, 1991) or (2) the standard deviation of the raw scores (often based on the preset or baseline data) that estimate variations in the outcome measure of the population (Becker, 1988)” (pp. 3-4). Because an effect size based on variations in improvement is not available in latent growth modeling, one must utilize Feingold’s (2009) recommended calculation for an effect size based on standard deviation at the onset of the investigation:

$$\beta_{11}(\text{time})/SD_{\text{RAW}} \quad (11)$$

Here, β_{11} represents the difference in mean growth rates between the treatment and control groups and SD_{RAW} refers to the initial standard deviation of raw scores.

In using this equation to determine effect size for the slope parameter, however, difficulties arise because of the relationship between error variance, effect size, and slope difference. Using this equation, an increase in error variance necessarily results in an increase in slope difference if effect size is held constant (e.g., given this study's model parameters, for an effect size of .3, an error variance of 50 results in a slope difference of .7348 whereas an error variance of 100 results in a slope difference of .8485). This increase in slope difference results an increase in power as error increases when effect size held constant across error variance values. As seen in Hertzog et al. (2008), an increase in error variance has not been found to increase power.

Therefore, to understand the influence of increased error on power, slope difference was held constant in the present investigation rather than effect size. However, average effect size across error variance values was taken into account in the selection of slope differences. For example, rather than holding effect size equal to .30 across error variance values of 50 and 100 (resulting in slope differences of .7348 and .8485 respectively), a single slope difference value was used for both the error variance conditions of 50 and 100, approximating an effect size of .3. This value was determined by averaging the slope difference values for an effect size of .3 across error variance values (.7348 and .8485), resulting in a slope difference of .7917 (which equates to an effect size of .32 when error

variance is 50 and an effect size of .28 when error variance is 100). Using this technique of averaging slope difference values across error variances of 50 and 100 for three effect sizes (.2, .3, and .4), the following slope difference values were used in the present investigation: 0 (to test Type I error), .5278 (approximating an effect size of .2), .7917 (approximating an effect size of .3), and 1.0556 (approximating an effect size of .4). This range in effect size is similar to Muthén and Curran, who investigated small to moderate effect sizes of .2 to .5.

Sample size. As with most power analysis investigations, sample size was manipulated, with total sample sizes set to 200, 500, or 800 (similar to sample sizes and ranges used in Hertzog et al., 2006; Muthén & Curran, 1997; Zhang & Wang, 2009).

Number of repeated measures. In addition to the specified parameter and variance values above, the number of repeated measures was also varied to be equal to three or six time points, values similar to past investigations of how power was influenced by number of repeated measures (e.g., Zhang & Wang, 2009). The length of the study was held constant. For example, one could conceptualize the number of repeated measures as either three equally spaced repeated measures taken across six months or six equally spaced repeated measures taken across six months. In order to model this, the factor loadings for the growth factor were set to equal 0, 1, 2, 3, 4, and 5 or 0, 2.5, and 5.

Monte Carlo parameter values: Error variance conditions. Following the investigation of the base conditions using the Monte Carlo technique,

additional aspects of the growth model were explored, including error variance and its homogeneity or heterogeneity across repeated measures. In addition to investigating power when variance was equal but unconstrained in the model (EU), power was also considered when variance was equal in the population but constrained to be equal in the model (EC), unequal in the population and unconstrained in the model (UU), and unequal in the population but constrained to be equal in the model (UC), with the last error variance condition resulting in model misspecification. Error variance at wave one was set to be 50 all (EC, UU, and UC) conditions. Because variations in error variance across time can take on a number of forms with applied data (increasing over time, increasing AND decreasing over time, and increasing/decreasing by different increments), this investigation simply investigated a single, simplified form of unequal error variance over time: (1) error increasing in increments of 5 across the six time points (resulting in a variance of 50 at wave one and 75 at time six) and (2) error increasing in increments of 15 across the six time points (resulting in a variance of 50 at wave one and 125 at time six). While these values are necessarily arbitrary due to the wide range of error variance patterns in applied research, they do provide a foundation for investigating the potential influence of different error variance patterns and model specifications (and misspecifications) on power.

In this investigation of error variance, slope variance was held constant at 20. As in the base conditions, the full range of sample size (200, 500, 800) and slope difference (0, .5278, .7917, and 1.0556) were investigated.

Monte Carlo parameter values: Misspecification. Various forms of model misspecification and their influence on the probability of rejecting correctly the null hypothesis for the difference in slopes parameter were also investigated. Although power is investigated under circumstances where the model is assumed to be properly specified, proportions of replications that correctly reject the null hypothesis for the difference in slopes parameter under conditions of misspecification are referred to here as power. In investigating error variance conditions using the Monte Carlo technique, some model misspecification was introduced. Specifically, the misspecification occurred in the UC conditions in which error variance was unequal in the population but the model constrains the error variance to be equal. By increasing the increment of increase in error variance across time (5 versus 15), the amount of misspecification is also increased and resulting power could be investigated.

In addition to investigating model misfit relating to error variance, growth form was also investigated in terms of model misspecification using the Monte Carlo technique. Specifically, the form of growth was set to be quadratic in the population but growth was linear in the model. In this manner, it was possible to explore differences in power when a model was correctly specified as linear (when linear in the population) as opposed to a model that was incorrectly specified as linear (when quadratic in the population). Investigating quadratic growth required additional parameter value selection. According to Yu (2002) a common variance ratio between intercept variance, slope (linear) variance and quadratic growth variance is 1, .3, and .1. Wu (2008), however, employed a

smaller quadratic variance value (in terms of its ratio to intercept and slope variance) than suggested by Yu. In considering the values from each of these studies, a quadratic variance of 3 was selected (which falls between the proportional variance used by Yu and Wu). Additionally, the quadratic growth factor was set to $-.03$ (as done in many of the models in Wu, 2008).

In this investigation of misspecification, slope variance was held constant at 20. As in the base conditions, the full range of sample size (200, 500, 800) and slope difference (0, .5278, .7917, and 1.0556) were investigated. In investigating error variance misfit, error variance was set to either 50, increasing in increments of 5 or 50, increasing in increments of 15 across time in the population (and was constrained to be equal in the model). In investigating growth form misfit, error variance was set to be equal to 50 in the population, but was not constrained to be equal in the model.

Power analyses across techniques. In addition to investigating power using the Monte Carlo power analysis technique, power analyses were also conducted using the Satorra-Saris and MBC power analysis methods. In comparing these three power analysis techniques, slope variance was set to 20. As in the initial base conditions using the Monte Carlo technique, the full range of values was investigated for level-1 error variance (50 or 100, with equal variance across time in the population and unconstrained in the model), number of measurement occasions (3 or 6), slope difference (0, .5278, .7917, or 1.0556), and sample size (200, 500, or 800).

In analyzing power using the MBC approach, it was also necessary to define RMSEA values. Because a key aspect of using the MBC power analysis approach is selecting RMSEA values to test fit between nested models, it was important to explore how selected pairs of RMSEA values influenced power estimates. Power analyses typically employ a null hypothesis of no difference in fit between models (Kim, 2005); therefore, a null hypothesis of no difference was employed with the MBC technique to maintain comparability among power analysis techniques. Thus, any given power analysis using the MBC approach required the specification of two RMSEA values for models A and B (where model A is nested within model B), as well as the degrees of freedom for these models.

In this investigation, the values considered were either .01 (model A) and 0 (model B), .02 (model A) and 0 (model B), or .05 (model A) and .04 (model B). The two pairs with a model B RMSEA value of 0 were selected to mirror the Satorra-Saris technique, in which a correctly specified model (which one would expect to have a zero or near zero RMSEA) is compared to a model that is misspecified by setting a single parameter to 0 (MacCallum, Browne, and Cai recommend a difference in RMSEA values of .01 or .02 when the models differ by only one parameter). The .04-.05 pairing uses the RMSEA criteria suggested by MacCallum, Browne, and Cai, who suggest selected RMSEA values in the mid-range of the RMSEA scale (which assumes a model is not perfect, but may be near perfect). Additionally, this pairing (.04 and .05) was commonly used as an example in MacCallum, Browne, and Cai (2006).

Analysis Conditions

The Monte Carlo technique was conducted with Mplus (version 6.1) using 1,000 replications and maximum likelihood estimation. In total, six conditions were manipulated using the Monte Carlo technique: sample size, slope difference, slope variance, error variance, number of repeated measures, and misspecification conditions. These manipulated conditions resulted in a total of 96 combinations using the Monte Carlo technique. The investigation of error variance conditions using the Monte Carlo technique involved the manipulation of slope difference and sample size, resulting in 72 combinations. Finally, the investigation of growth form misspecification involved 12 conditions (as slope variance, error variance, quadratic growth, and quadratic growth variance were held constant).

The Satorra-Saris technique was also conducted using Mplus (version 6.1) and maximum likelihood methods. In using the Satorra Saris technique, four conditions were manipulated (with slope variance held at 20): sample size, slope difference, error variance (EU), and number of repeated measures. Misspecification could not be considered using the Satorra Saris technique. These manipulated conditions resulted in a total of 36 combinations using the Satorra Saris technique.

The MBC technique, on the other hand, uses only model degrees of freedom, desired alpha level, and specified RMSEA values to calculate power. Therefore, an Excel spreadsheet was used to complete these calculations. Because the MBC technique considers only model degrees of freedom and RMSEA values, it was not influenced factorially by the manipulated generation

conditions. Thus, it was only calculated for three versus six measurement occasions and across sample sizes. In considering RMSEA pairs in conjunction with sample size and the number of repeated measures, a total of 18 conditions were investigated using the MBC power analysis technique.

See Table 3 for a summary of manipulated conditions and constant parameter values across the power analysis techniques.

Table 3

Simulation Design

Monte Carlo: Base Conditions				
Intercept variance	100			
Initial status for both groups	0			
Intercept with slope correlation	.20			
Sample size	200	500	1000	
Slope difference between treatment and control	0	.5278	.7917	1.0556
Slope variance	10	20		
Number of repeated measures	3	6		
Level-1 Error variance (EU)	50	100		
Monte Carlo: Specification of Level-1 Error Variance				
Slope variance	20			
Number of repeated measures	6			
Sample size	200	500	1000	
Slope difference between treatment and control	0	.5278	.7917	1.0556
Error variance at time 1	50			
Error variance specification	EU	EC	UU, increments of 5 or 15	UC, increments of 5 or 15
Monte Carlo: Misspecification of growth form				
Slope variance	20			
Number of repeated measures	6			

Sample size	200	500	1000	
Slope/linear growth	0	.5278	.7917	1.0556
Error variance across time	50			
Quadratic variance	3			
Quadratic growth	-.03			
Comparison of Analysis Techniques				
Slope variance	20			
Sample size	200	500	1000	
Slope difference between treatment and control	.5278	.7917	1.0556	
Number of repeated measures	3	6		
Level-1 error variance (EU)	50	100		
MBC RMSEA values	0, .01	0, .02	.04, .05	

Analysis of Results

In using the Monte Carlo technique, power and Type I error was determined based on the % Sig Coeff. column in the Mplus output. This column indicates the proportion of replications that found the parameter of interest to be significant at the .05 level (two-tailed test, critical value = 1.96). Therefore, whereas the % Sig Coeff. column indicates power when the parameter is nonzero in the population, the same column provides the Type I error rate when the parameter is equal to zero in the population (Muthén & Muthén, 2002). Coverage (available in the Mplus output column labeled Cover.) refers to the proportion of replications that a focal parameter values fall within a 95% confidence interval of the true parameter value (Muthén & Muthén, 2002); this was also summarized for conditions in this study. Finally, the proportion of nonconverging samples were recorded for each analysis condition in order to monitor potential problems.

Power calculations with the Satorra-Saris technique, on the other hand, involved using the resulting noncentrality parameter related to the chi-square test. Similarly, power was calculated with the MBC technique by using the resulting noncentrality parameter based on pairs of RMSEA values and model degrees of freedom (for the nested models specified for the Satorra-Saris technique). Thus, the results compiled from these two techniques were the noncentrality parameter and power estimate.

Resulting power estimates from each of the three analysis techniques were compiled into tables presenting changes in power based on model attributes and power analysis approach. Power curve plots were also created for select conditions to illustrate changes in power across various conditions.

CHAPTER 3

RESULTS

In all Monte Carlo simulation conditions, 1,000 replications were generated to investigate the power to detect significance of the coefficient based on regression of the slope growth factor on a dummy-coded treatment variable. Under no circumstances did replications fail to converge. However, error messages for particular replications were observed in some conditions. Specifically, in some instances, Mplus indicated that “the residual covariance matrix (Theta) is not positive definite.” This could indicate a negative variance/residual variance for an observed variable, a correlation greater or equal to one between two observed variables, or a linear dependency among more than two observed variables.” This error message was received only for conditions with three repeated measures when sample size was either 200 or 500.

The error was most common (5% of the replications) for the condition in which slope variance was 20, error variance was 50, and sample size was 200. When sample size for this condition was increased to 500, this error was reported for only 0.5% of replications and no errors were reported for a sample size of 800. The error was seen in 4% of replications when slope variance was 10, error variance was 100, and sample size was 200. Again, increasing the sample size dramatically decreased this occurrence, with a sample of 500 resulting in only 0.4% of the replications with errors. The error was observed in approximately 2% of replications with a sample size of 200 when slope variance was 20 and error variance was 100 or when slope variance was 10 and error variance was 50.

However, this decreased to 0.2% when sample size was increased to 500.

Replications for which an error message was indicated were not included in the calculation of empirical powers or other summary statistics for those conditions.

Type I Error

When using the Monte Carlo approach, Type I error was investigated by setting the slope difference between the treatment and control groups to zero. Mplus simulation output reports the proportion of replications that indicated a significant slope difference, which was Type I error for these null conditions. Based on an alpha of .05, the empirical Type I error rate would be expected to be approximately .05. In no cases did Type I error exceed .059, with this value occurring with the smallest sample size (200), the smallest error variance (50), the largest slope variance (20), and three repeated measures. The smallest Type I error value was .037, with this value occurring with the largest sample size (800), the largest error variance (100), the largest slope variance (20), and six repeated measures. See Tables 4, 7, and 12 for Type I error rates across conditions.

The trend of decreasing Type I error rates as sample size increases suggests that Type I error tends to be slightly conservative with higher sample sizes. A similar trend is apparent in terms of number of repeated measures, with more repeated measures corresponding to lower Type I error rates. Despite slightly conservative Type I error rates, all of the Type I error rates do fulfill Bradley's (1978) liberal criterion (Type I error values fall between .025 and .075) for an alpha of .05. Further, all but one Type I error rate (.037) fulfill Bradley's moderate criterion (Type I error values fall between .04 and .06) for an alpha of

.05. This conservative value of .037, however, does fall well within Bradley's liberal criterion (Type I error values fall between .025 and .075).

Monte Carlo Power Analyses: Base Conditions

Table 4 indicates empirical powers for the base conditions using the Monte Carlo analysis; these are depicted graphically in Figure 2. In these conditions, error variances were generated to be equal across time, but were unconstrained in the analysis model.

Table 4

Type I and Power when Level-1 Error Variance is Equal across Time

Level-1 error variance	Number of measurement occasions	Slope difference	Slope variance = 10			Slope variance = 20		
			Sample size			Sample size		
			200	500	800	200	500	800
50	3 (df=2)	0	.056	.053	.043	.059	.048	.048
		.5278	.184	.350	.533	.143	.225	.337
		.7917	.326	.659	.844	.218	.436	.645
	1.0556	.527	.891	.976	.328	.675	.861	
	6 (df=20)	0	.051	.046	.040	.054	.047	.040
		.5278	.174	.340	.540	.112	.220	.343
.7917		.337	.677	.853	.205	.437	.631	
1.0556	.545	.913	.989	.346	.676	.853		
100	3 (df=2)	0	.056	.050	.051	.056	.051	.049
		.5278	.144	.273	.426	.115	.197	.291
		.7917	.260	.545	.750	.194	.387	.564
	1.0556	.430	.808	.937	.289	.609	.791	
	6 (df=20)	0	.056	.042	.043	.052	.051	.037
		.5278	.154	.284	.452	.104	.199	.307
.7917		.273	.595	.785	.188	.382	.583	
1.0556	.458	.839	.966	.305	.630	.809		

**Note:* Slope difference calculated as average slope difference across level-1 error variance values of 50 and 10, yielding effect sizes of approximately 0, 0.2 (slope difference of .5278), 0.3 (slope difference of .7917), and 0.4 (slope difference of 1.0556).

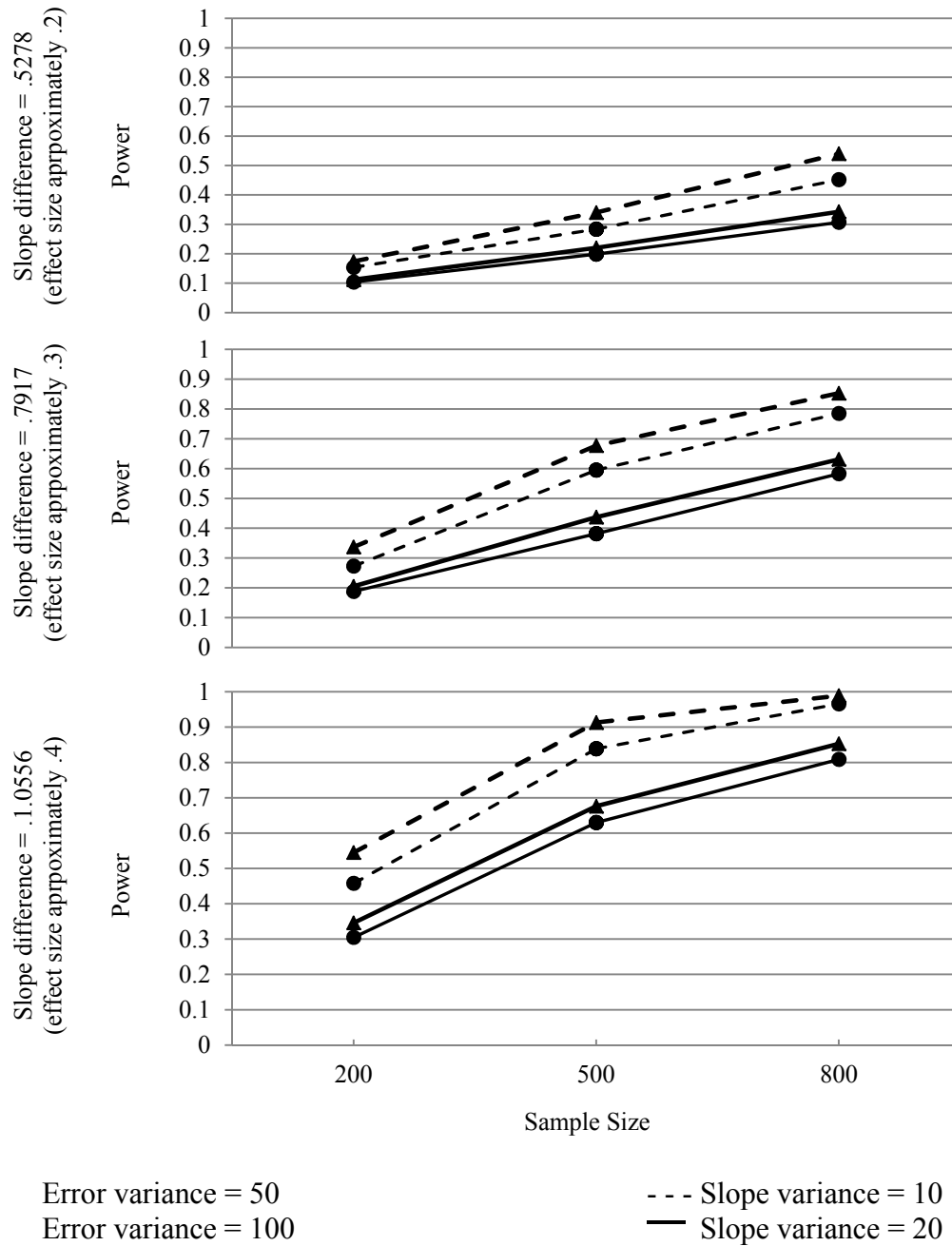


Figure 2. Monte Carlo power for 6 repeated measures and equal level-1 error variance across time (not constrained to be equal in the model).

Sample size. Sample size contributed to power for detecting the slope difference between treatment and control conditions, as expected. As indicated in Figure 2, power increased as sample size increased for all combinations of base conditions, with an approximate increase of .3 to .4 from the smallest to the largest sample size. The maximum power achieved with a sample size of 200 was .55, which occurred with a slope difference of 1.0556 (approximately an effect size of .4), level-1 error variance of 50, slope variance of 10, and 6 repeated measures. Thus, no conditions yielded acceptable power with a sample size of 200.

In contrast with a sample size of 200, sample sizes of 500 and 800 did yield acceptable power under certain conditions. With a sample size of 500, all conditions with a 1.0556 slope difference and slope variance of 10 exceeded a power of .80, ranging from .81 to .91, with the highest value occurring with an error variance of 50, slope variance of 10, and 6 repeated measures. No conditions with a sample size of 500 and slope variance of 20 reached a power of .80 (the maximum power was .68). With a sample size of 500, no other slope differences (.5278 or .7917) yielded power near .80 (the highest value among these slope differences was .68). With a sample size of 800, all but one condition with a 1.0556 slope differences exceeded .80, with values ranging from .81 to .99. The exception occurred when slope variance was 20, error variance was 100, and the model included three repeated measures, in which case power equaled .79. When slope variance was set to 10, power for a sample size of 800 and slope difference of .7917 (an effect size of approximately .2) ranged from .75 to .85.

Slope difference/effect size. As expected, the magnitude of the slope difference was also a major contributor to power to detect the difference in slope between populations. Figure 2 indicates the improvement in power as slope difference increased from .5278 (a small effect size of approximately .2) to 1.0556 (a moderate effect size of approximately .4). For example, with a small slope difference, the highest power achieved in the investigated conditions was approximately .50 (slope variance of 10, error variance of 100, sample size of 800). However, this same condition, with a moderate effect size, resulted in a power of over .90. In considering slope difference in conjunction with sample size, when the slope difference is small (.5278, an approximate effect size of .2), the differences between power estimates across conditions becomes greater as sample size increases. For example, with a small difference in slopes, whereas power ranges from about .10 to .18 (a spread of .08) when sample size was 200, a sample size of 800 yielded a range in values from .29 to .53 (a spread of .24). However, this trend is less apparent when slope differences are larger, with the range of power values not necessarily increasing as sample size increased. With a large slope difference of 1.0556, the range of power estimates for a sample size of 500 was wider than for a sample size of 800. Additionally, the power increase from sample sizes 200 to 500 was more substantial for larger slope differences than for smaller slope differences.

Number of repeated measures. Based on initial plots of power comparing models with three or six repeated measures, it was evident that this difference in number of measurement occasions did not substantially influence

power. Patterns of results for conditions with three and six repeated measures were very similar. The maximum difference in power between three and six repeated measures was .05, when error variance was 100, slope variance was 10, sample size was 500, and the slope difference was .7917). However, the majority of conditions differed by approximately .02 or less across number of repeated measures. Accordingly, further investigations of power considered only models with six repeated measures.

Slope variance. Figure 3 displays observed growth (fit to a linear trajectory) for 30 randomly selected cases from the treatment group for conditions with 10 versus 20 for slope variance (holding other conditions constant; sample size of 800, population intercept variance of 100, error variance of 50, slope difference of 1.0556). Slope variance influenced power such that larger slope variances resulted in decreased power. A slope variance of 10 resulted in increases of power ranging from .05 to .24 over conditions with a slope variance of 20. A moderate slope difference combined with sample sizes of 500 or 800 typically resulted in the most improvement in power when slope variance was 10 rather than 20. Because of the consistent difference and patterns of powers in comparing slope variances of 10 and 20 across conditions, many of the additional power investigations employed only a slope variance of 20, which is consistent with past literature (e.g., Muthén & Curran, 1997; Muthén & Muthén, 2002) suggesting specification of a 1:5 ratio for slope to intercept variance.

In addition to considering how slope variance influenced power, an additional investigation was conducted to examine power for conditions in which

the ratio of the slope difference to slope standard deviation was held constant while varying the components of this ratio, that is, the slope difference and slope variance. This ratio of slope difference to slope standard deviation was used by Muthén and Muthén (2002) to calculate effect size. In this initial investigation, level-1 error variance was 50, sample size was 500, and the model included 6 repeated measures. First, a slope difference to slope standard deviation ratio of .33 was considered. Power for this condition was calculated for the following pairs of values (slope difference, slope variance): 0.5278, 2.5; 1.0556, 10; 0.7917, 5.625; 2.1112, 40. Although the resulting power estimates differed across the slope variance and slope difference values, the range was relatively narrow (.71 to .96) considering the range of effect sizes represented (.22 to .86). A slope difference to slope standard deviation ratio of .24 was also considered using the following pairs of values (slope difference, slope variance): 0.5278, 5; 1.0556, 11.25; 0.7917, 20; 2.1112, 80. Powers for a ratio of .24 ranged from .54 to .71.

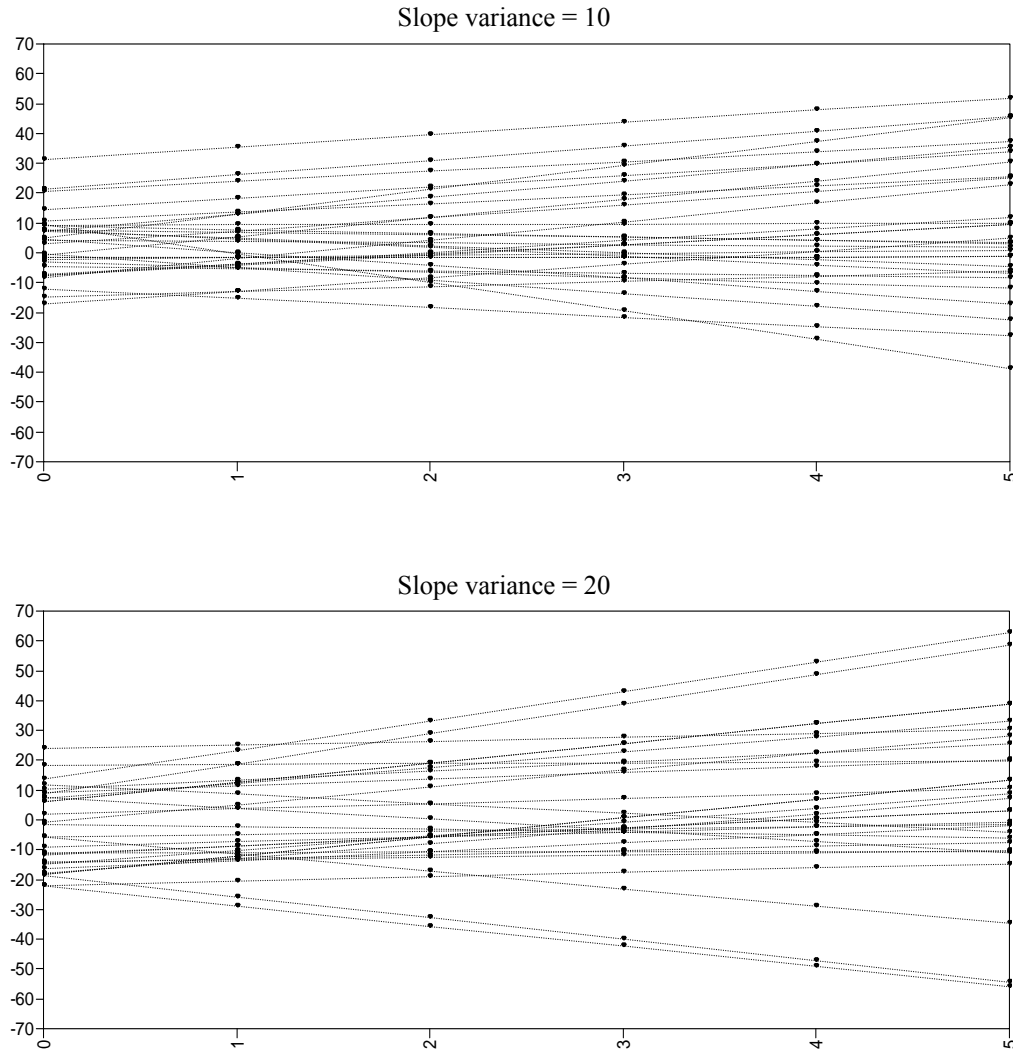


Figure 3. Observed individual growth trajectories, fit to linear growth, for 30 randomly selected cases in the treatment condition, for slope variance values of 10 and 20. Plots were for a sample size of 800 with the following population parameters held constant across slope variance conditions: intercept variance (100), error variance (50), and slope difference (1.0556).

Level-1 error variance. The magnitude of error variance also influenced power, with greater error variance slightly decreasing power. With an error variance of 50, power was from approximately .01 to .09 larger in comparison with conditions with an error variance of 100.

Bias, efficiency, and coverage. Absolute biases and efficiencies for the slope difference parameters are available in Table 5. Absolute bias was calculated as the difference between the estimated slope difference and the slope difference defined in the population. As seen in Table 5, absolute bias was typically very small, with the largest absolute bias (-0.03) occurring with a sample size of 200, slope variance of 20, error variance of 100, and 6 repeated measures. Relative bias, defined as the difference between the estimated parameter value and the population value divided by the population parameter value, was also computed. According to Muthén and Muthén (2002), investigations of sample size determination based on desired power must consider parameter values and conditions that yield a relative bias of no more than 10%. In the correctly specified models under investigation, relative bias reached no more than 4.89%. Muthén and Muthén also suggested that standard error bias for the parameter of interest (calculated as the standard error estimate across replications minus the population standard error, divided by the population standard error) not exceed 5%. In the base conditions, standard error bias ranged from 0 to 4%. Finally, Muthén and Muthén suggested parameter coverage should range between .91 and .98. This was also satisfied in this investigation, with coverage for the slope difference parameter, across all conditions, ranging from .94 to .96.

Table 5

Absolute Bias (Population Value Minus Model Estimate) and Efficiency (in parentheses) when Level-1 Error Variance is Equal across Time but not Constrained to be Equal in the Model

Level-1 error variance	Number of measurement occasions	Slope difference	Slope variance = 10			Slope variance = 20		
			Sample size			Sample size		
			200	500	800	200	500	800
50	3 (df=2)	0	.004 (.54)	-.0004 (.33)	-.003 (.27)	.008 (.72)	.0002 (.43)	-.002 (.35)
		.5278	.004 (.54)	-.0004 (.33)	-.003 (.27)	.008 (.72)	.0002 (.43)	-.002 (.35)
		.7917	.004 (.54)	-.0004 (.33)	-.003 (.27)	.008 (.72)	.0002 (.43)	-.002 (.35)
		1.0556	.004 (.54)	-.0004 (.33)	-.003 (.27)	.008 (.72)	.0002 (.43)	-.002 (.35)
	6 (df=20)	0	-.018 (.51)	-.012 (.32)	-.012 (.25)	-.021 (.67)	-.014 (.42)	-.015 (.34)
		.5278	-.018 (.51)	-.012 (.32)	-.012 (.25)	-.021 (.67)	-.014 (.42)	-.015 (.34)
		.7917	-.018 (.51)	-.012 (.32)	-.012 (.25)	-.021 (.67)	-.014 (.42)	-.015 (.34)
		1.0556	-.018 (.51)	-.012 (.32)	-.012 (.25)	-.021 (.67)	-.014 (.42)	-.015 (.34)
100	3 (df=2)	0	-.0001 (.61)	-.0003 (.38)	-.004 (.30)	.003 (.76)	.0002 (.47)	-.004 (.38)
		.5278	-.0001 (.61)	-.0003 (.38)	-.004 (.30)	.003 (.76)	.0002 (.47)	-.004 (.38)
		.7917	-.0001 (.61)	-.0003 (.38)	-.004 (.30)	.003 (.76)	.0002 (.47)	-.004 (.38)
		1.0556	-.0001 (.61)	-.0003 (.38)	-.004 (.30)	.003 (.76)	.0002 (.47)	-.004 (.38)
	6 (df=20)	0	-.021 (.56)	-.014 (.35)	-.014 (.28)	-.025 (.72)	-.016 (.45)	-.017 (.36)
		.5278	-.021 (.56)	-.014 (.35)	-.014 (.28)	-.025 (.72)	-.016 (.45)	-.017 (.36)
		.7917	-.021 (.56)	-.014 (.35)	-.014 (.28)	-.025 (.72)	-.016 (.45)	-.017 (.36)
		1.0556	-.021 (.56)	-.014 (.35)	-.014 (.28)	-.025 (.72)	-.016 (.45)	-.017 (.36)

*Note: Slope difference calculated as average slope difference across level-1 error variance values of 50 and 10, yielding effect sizes of approximating 0, 0.2 (slope difference of .5278), 0.3 (slope difference of .7917), and 0.4 (slope difference of 1.0556).

Global fit indices. In addition to collecting power estimates, mean chi-square and RMSEA (see Table 6) indices were also recorded for each condition. In all correctly specified models, the mean chi-square statistic was approximately equal to the model degrees of freedom, with the difference between the chi-square and degrees of freedom ranging from -.09 to .49. Average RMSEA values for correctly specified models with six repeated measures were .017 ($SD=.02$) for a sample size of 200, .01 ($SD=.01$) for a sample size of 500, and .008 ($SD=.01$) for a sample size of 800. In correctly specified models with three repeated measures, mean RMSEA values were .025 ($SD=.04$) for a sample size of 200, .014 to .015 ($SD=.02$) for a sample size of 500, and .01 ($SD=.02$) for a sample size of 800.

Table 6

Mean RMSEA and Standard Deviation (in Parentheses) when Level-1 Error Variance is Equal across Time

Level-1 error variance	Number of measurement occasions	Slope difference	Slope variance = 10			Slope variance = 20		
			Sample size			Sample size		
			200	500	800	200	500	800
50	3 (df=2)	0	.025 (.04)	.014 (.02)	.011 (.02)	.025 (.04)	.015 (.02)	.011 (.02)
		.5278	.025 (.04)	.014 (.02)	.011 (.02)	.025 (.04)	.015 (.02)	.011 (.02)
		.7917	.025 (.04)	.014 (.02)	.011 (.02)	.025 (.04)	.015 (.02)	.011 (.02)
		1.0556	.025 (.04)	.014 (.02)	.011 (.02)	.025 (.04)	.015 (.02)	.011 (.02)
	6 (df=20)	0	.017 (.02)	.010 (.01)	.008 (.01)	.017 (.02)	.010 (.01)	.008 (.01)
		.5278	.017 (.02)	.010 (.01)	.008 (.01)	.017 (.02)	.010 (.01)	.008 (.01)
		.7917	.017 (.02)	.010 (.01)	.008 (.01)	.017 (.02)	.010 (.01)	.008 (.01)
		1.0556	.017 (.02)	.010 (.01)	.008 (.01)	.017 (.02)	.010 (.01)	.008 (.01)
100	3 (df=2)	0	.025 (.04)	.014 (.02)	.011 (.02)	.025 (.04)	.014 (.02)	.011 (.02)
		.5278	.025 (.04)	.014 (.02)	.011 (.02)	.025 (.04)	.014 (.02)	.011 (.02)
		.7917	.025 (.04)	.014 (.02)	.011 (.02)	.025 (.04)	.014 (.02)	.011 (.02)
		1.0556	.025 (.04)	.014 (.02)	.011 (.02)	.025 (.04)	.014 (.02)	.011 (.02)
	6 (df=20)	0	.017 (.02)	.010 (.01)	.008 (.01)	.017 (.02)	.010 (.01)	.008 (.01)
		.5278	.017 (.02)	.010 (.01)	.008 (.01)	.017 (.02)	.010 (.01)	.008 (.01)
		.7917	.017 (.02)	.010 (.01)	.008 (.01)	.017 (.02)	.010 (.01)	.008 (.01)
		1.0556	.017 (.02)	.010 (.01)	.008 (.01)	.017 (.02)	.010 (.01)	.008 (.01)

60

**Note:* Slope difference calculated as average slope difference across level-1 error variance values of 50 and 10, yielding effect sizes of approximating 0, 0.2 (slope difference of .5278), 0.3 (slope difference of .7917), and 0.4 (slope difference of 1.0556).

Level-1 Error Variance (Conditions with No Misspecification)

In addition to using the Monte Carlo technique to investigate the influence of smaller versus larger error variances (50 vs. 100) when population variances were specified to be equal over time and were unconstrained in the analysis model, additional error variance patterns and model specifications with regard to error were also investigated (see Table 7). In examining the influence of constraining error variances to be equal in the model (when they were in fact equal in the population; EC) as opposed to allowing error to vary across time in the model (EU), it was found that power for the difference in slopes parameter was not substantially influenced. Differences in power across these two conditions (EC vs. EU) ranged from only .002 to .008.

In considering error variances that are not equal across time, two magnitudes of increase were investigated: Error variance was set to increase from an initial variance of 50 by increments of either 5 or 15 at each successive wave. When the analysis model was correctly specified to allow error variance to vary across time, the powers were not substantially different when population error variances were generated to increase by 5 or 15 with successive waves. Although the model with reduced error variance (due to smaller incremental increases) did have slightly higher power, this boost in power ranged from .002 to .03, with a greater difference in power apparent with a larger slope difference.

Table 7

Type I Error and Power across Level-1 Error Variance Conditions, with Level-1 Error Variance at Wave One of 50, 6 Repeated Measures, and Slope Variance of 20

Level-1 error variance condition	Slope difference	Sample size		
		200	500	800
Equal error variance in population, unconstrained in model (EU) df = 20	0	.054	.047	.040
	.5278	.112	.220	.343
	.7917	.205	.437	.631
	1.0556	.346	.676	.853
Equal error variance in population, constrained to be equal in model (EC) df = 25	0	.051	.047	.040
	.5278	.110	.223	.343
	.7917	.213	.441	.634
	1.0556	.354	.676	.853
Unequal error variances in population (increase in increments of 5 across time), unconstrained in model (UU) df = 20	0	.052	.047	.040
	.5278	.109	.214	.335
	.7917	.200	.429	.620
	1.0556	.337	.668	.841
Unequal error variances in population (increase in increments of 15 across time), unconstrained in model (UU) df = 20	0	.050	.045	.036
	.5278	.107	.212	.318
	.7917	.194	.409	.600
	1.0556	.320	.641	.819
Unequal error variances in population (increase in increments of 5 across time), constrained to be equal in model (UC) – <i>Misspecified</i> df = 25	0	.050	.044	.039
	.5278	.115	.215	.331
	.7917	.199	.429	.620
	1.0556	.339	.663	.837
Unequal error variances in population (increase in increments of 15 across time), constrained to be equal in model (UC) – <i>Misspecified</i> df = 25	0	.052	.045	.036
	.5278	.107	.211	.311
	.7917	.187	.405	.603
	1.0556	.316	.644	.817

As with the base conditions, biases and efficiencies of the slope difference parameter (see Table 8), RMSEA values (see Table 9), and chi-square values were collected. Again, bias was minimal, with absolute bias ranging from -0.03 to -0.02 and relative bias ranging from -1.12% to -4.16%. In terms of the chi-square statistic, when the population error variances were equal across time and the analysis model was constrained such that error variances were required to be equal (rather than leaving them unconstrained), the difference between chi-square and degrees of freedom was still minimal, although slightly higher (up to .65) than for models that did not constrain the model error variances to be equal (up to .49). RMSEA values ranged from .01 to .02 for correctly specified models (EU, EC, and UU conditions), with RMSEA decreasing slightly as sample size increased.

Table 8

Absolute Bias (Population value Minus Model Estimate) and Efficiency (in parentheses) across Level-1 Error Variance Conditions, with Level-1 Error Variance at Wave One of 50, 6 Repeated Measures, and Slope Variance of 20

Level-1 error variance condition	Slope difference	Sample size		
		200	500	800
Equal error variance in population, unconstrained in model (EU) df = 20	0	-.025 (.72)	-.016 (.45)	-.017 (.36)
	.5278	-.025 (.72)	-.016 (.45)	-.017 (.36)
	.7917	-.025 (.72)	-.016 (.45)	-.017 (.36)
	1.0556	-.025 (.72)	-.016 (.45)	-.017 (.36)
Equal error variance in population, constrained to be equal in model (EC) df = 25	0	-.020 (.67)	-.014 (.42)	-.015 (.34)
	.5278	-.020 (.67)	-.014 (.42)	-.015 (.34)
	.7917	-.020 (.67)	-.014 (.42)	-.015 (.34)
	1.0556	-.020 (.67)	-.014 (.42)	-.015 (.34)
Unequal error variances in population (increase in increments of 5 across time), unconstrained in model (UU) df = 20	0	-.021 (.68)	-.014 (.43)	-.015 (.34)
	.5278	-.021 (.68)	-.014 (.43)	-.015 (.34)
	.7917	-.021 (.68)	-.014 (.43)	-.015 (.34)
	1.0556	-.021 (.68)	-.014 (.43)	-.015 (.34)
Unequal error variances in population (increase in increments of 15 across time), unconstrained in model (UU) df = 20	0	-.022 (.70)	-.015 (.44)	-.016 (.35)
	.5278	-.022 (.70)	-.015 (.44)	-.016 (.35)
	.7917	-.022 (.70)	-.015 (.44)	-.016 (.35)
	1.0556	-.022 (.70)	-.015 (.44)	-.016 (.35)
Unequal error variances in population (increase in increments of 5 across time), constrained to be equal in model (UC) – <i>Misspecified</i> df = 25	0	-.022 (.68)	-.015 (.43)	-.016 (.34)
	.5278	-.022 (.68)	-.015 (.43)	-.016 (.34)
	.7917	-.022 (.68)	-.015 (.43)	-.016 (.34)
	1.0556	-.022 (.68)	-.015 (.43)	-.016 (.34)
Unequal error variances in population (increase in increments of 15 across time), constrained to be equal in model (UC) – <i>Misspecified</i> df = 25	0	-.026 (.70)	-.017 (.44)	-.018 (.35)
	.5278	-.026 (.70)	-.017 (.44)	-.018 (.35)
	.7917	-.026 (.70)	-.017 (.44)	-.018 (.35)
	1.0556	-.026 (.70)	-.017 (.44)	-.018 (.35)

Table 9

Mean RMSEA across Level-1 Error Variance Conditions, with Level-1 Error Variance at Wave One of 50, 6 Repeated Measures, and Slope Variance of 20

Level-1 error variance condition	Slope difference	Sample size		
		200	500	800
Equal error variance in population, unconstrained in model (EU) df = 20	0	.017 (.02)	.010 (.01)	.008 (.01)
	.5278	.017 (.02)	.010 (.01)	.008 (.01)
	.7917	.017 (.02)	.010 (.01)	.008 (.01)
	1.0556	.017 (.02)	.010 (.01)	.008 (.01)
Equal error variance in population, constrained to be equal in model (EC) df = 25	0	.016 (.02)	.010 (.01)	.007 (.01)
	.5278	.016 (.02)	.010 (.01)	.007 (.01)
	.7917	.016 (.02)	.010 (.01)	.007 (.01)
	1.0556	.016 (.02)	.010 (.01)	.007 (.01)
Unequal error variances in population (increase in increments of 5 across time), unconstrained in model (UU) df = 20	0	.017 (.02)	.010 (.01)	.008 (.01)
	.5278	.017 (.02)	.010 (.01)	.008 (.01)
	.7917	.017 (.02)	.010 (.01)	.008 (.01)
	1.0556	.017 (.02)	.010 (.01)	.008 (.01)
Unequal error variances in population (increase in increments of 15 across time), unconstrained in model (UU) df = 20	0	.017 (.02)	.010 (.01)	.008 (.01)
	.5278	.017 (.02)	.010 (.01)	.008 (.01)
	.7917	.017 (.02)	.010 (.01)	.008 (.01)
	1.0556	.017 (.02)	.010 (.01)	.008 (.01)
Unequal error variances in population (increase in increments of 5 across time), constrained to be equal in model (UC) – <i>Misspecified</i> df = 25	0	.032 (.02)	.032 (.01)	.033 (.01)
	.5278	.032 (.02)	.032 (.01)	.033 (.01)
	.7917	.032 (.02)	.032 (.01)	.033 (.01)
	1.0556	.032 (.02)	.032 (.01)	.033 (.01)
Unequal error variances in population (increase in increments of 15 across time), constrained to be equal in model (UC) – <i>Misspecified</i> df = 25	0	.075 (.02)	.076 (.01)	.075 (.01)
	.5278	.075 (.02)	.076 (.01)	.075 (.01)
	.7917	.075 (.02)	.076 (.01)	.075 (.01)
	1.0556	.075 (.02)	.076 (.01)	.075 (.01)

Misspecification in Error Variance and Form of the Growth Model

Two forms of misspecification were considered using the Monte Carlo technique: misspecification in the error variance structure and misspecification in the growth structure. In looking at misspecification in the error variance (in which the population error variances were set to increase in increments of 5 or 15 over time, but the analysis model constrained error variances to be equal over time), power to detect the difference in slopes was minimally influenced (see Table 7, UC conditions). Given error variance was increasing in the population, the maximum difference in power between conditions in which the error variances were not constrained to be equal and those that did constrain error variances to be equal was .007, holding constant all other factors. This difference was found under two conditions: (1) population error variance was increasing in increments of 15 with an effect size of approximately .3 and sample size of 200 and (2) error variance was increasing in increments of 15 with an effect size of approximately .2 and a sample size of 800.

As expected, chi-square values for models that are misspecified with respect to the error variance structure were larger than the degrees of freedom, with a larger increase in error variance (and thus increased misspecification) resulting in larger chi-square statistics relative to model degrees of freedom (see Table 10). Similarly, RMSEA values were greater for the conditions with misspecified error variance structures compared to the correctly specified models (see Table 9). In the misspecified models, when error variance increased in

increments of 5, the RMSEA was approximately .03, and when error variance increased in increments of 15, RMSEA was approximately .08.

Table 10

Average Chi-Square and Standard Deviation (in Parentheses) across Misspecified Level-1 Error Variance Conditions, with Level-1 Error Variance at Wave One of 50, 6 Repeated Measures, and Slope Variance of 20

Level-1 error variance condition	Slope difference	Sample size		
		200	500	800
Unequal error variances in population (increase in increments of 5 across time), constrained to be equal in model (UC) – <i>Misspecified</i> df = 25	0	31.63 (8.47)	40.02 (10.26)	47.90 (11.56)
	.5278	31.63 (8.47)	40.02 (10.26)	47.90 (11.56)
	.7917	31.63 (8.47)	40.02 (10.26)	47.90 (11.56)
	1.0556	31.63 (8.47)	40.02 (10.26)	47.90 (11.56)
Unequal error variances in population (increase in increments of 15 across time), constrained to be equal in model (UC) – <i>Misspecified</i> df = 25	0	54.78 (12.32)	97.49 (17.72)	139.36 (21.33)
	.5278	54.78 (12.32)	97.49 (17.72)	139.36 (21.33)
	.7917	54.78 (12.32)	97.49 (17.72)	139.36 (21.33)
	1.0556	54.78 (12.32)	97.49 (17.72)	139.36 (21.33)

**Note:* Slope difference calculated as average slope difference across level-1 error variance values of 50 and 100 for effect sizes of 0, 0.2 (slope difference of .5278), 0.3 (slope difference of .7917), and 0.4 (slope difference of 1.0556).

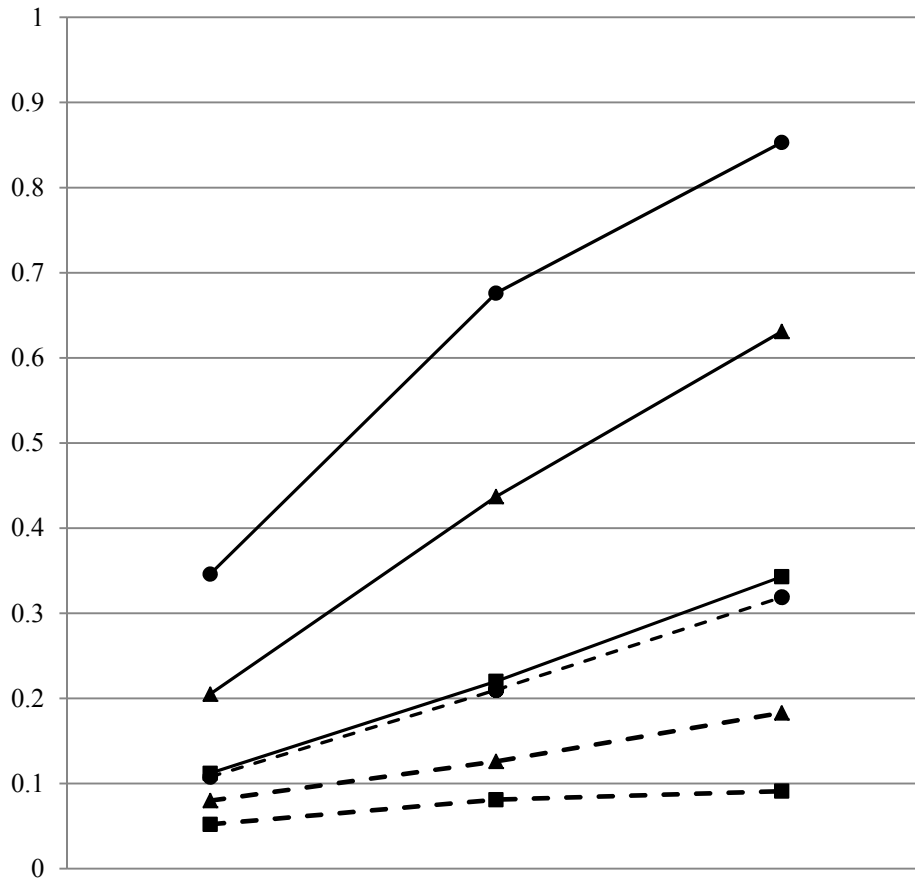
The investigation of the effect of misspecification of the form of the growth curve suggested a substantially larger impact on power than for the error variance misspecification, given the parameter values used in this study; this generalization is made with caution because the degree of misspecification is not equated between these conditions. In these conditions, the population growth model was quadratic in form, but was specified as linear in the analysis model. See Table 11 and Figure 4 for empirical powers as they relate to misspecification in the form of the growth model.

Table 11

Type I Error and Power under Correctly or Incorrectly Specified Form of the Growth Model for with 6 Repeated Measures, Slope Variance of 20, and Error Variance of 50 across Time

Growth form in population vs. model	Slope difference	Sample size		
		200	500	800
Quadratic – Linear (misspecified)	0	.056	.051	.048
	.5278	.052	.081	.091
	.7917	.080	.126	.183
	1.0556	.108	.210	.319
Linear – Linear	0	.054	.047	.040
	.5278	.112	.220	.343
	.7917	.205	.437	.631
	1.0556	.346	.676	.853

**Note:* Slope difference calculated as average slope difference across level-1 error variance values of 50 and 10, yielding effect sizes of approximating 0, 0.2 (slope difference of .5278), 0.3 (slope difference of .7917), and 0.4 (slope difference of 1.0556).



--- Quadratic growth in population-Linear growth in model (misspecified)
 — Linear growth in population-Linear growth in model

- Linear growth difference = .5278 (effect size approximately .2)
- ▲ Linear growth difference = .7917 (effect size approximately .3)
- Linear growth difference = 1.0556 (effect size approximately .4)

Figure 4. Comparison of power with and without growth form misspecification with 6 repeated measures, slope variance of 20, and error variance of 50.

As seen in Figure 4, incorrectly modeling quadratic growth and thereby analyzing a linear model under these conditions produces substantially lower power than analyzing a correctly specified linear model. Whereas a moderate slope difference produces a maximum power of approximately .85 when the linear model is correctly specified, a quadratic model that is analyzed as a linear

model reaches a maximum power of slightly more than .3. Therefore, if one is investigating power for a model they assume to be linear, but is, in actuality, quadratic, resulting power estimates will be deceptively high.

As expected, bias and efficiency (see Table 12), RMSEA values (see Table 13), and chi-square values (see Table 14) are higher in the presence of this form of misspecification. The substantial size of the chi-square values relative to the degrees of freedom suggests a large degree of misfit between the population and the specified model.

Table 12

Absolute Bias (Population value Minus Model Estimate) and Efficiency (in parentheses) under Correctly or Incorrectly Specified Form of the Growth Model for with 6 Repeated Measures, Slope Variance of 20, and Error Variance of 50 across Time

Growth form in population vs. model	Slope difference	Sample size		
		200	500	800
Quadratic – Linear (misspecified)	0	-.146 (.123)	-.139 (.78)	-.135 (.61)
	.5278	-.146 (.123)	-.139 (.78)	-.135 (.61)
	.7917	-.146 (.123)	-.139 (.78)	-.135 (.61)
	1.0556	-.146 (.123)	-.139 (.78)	-.135 (.61)
Linear – Linear	0	-.021 (.67)	-.014 (.42)	-.015 (.34)
	.5278	-.021 (.67)	-.014 (.42)	-.015 (.34)
	.7917	-.021 (.67)	-.014 (.42)	-.015 (.34)
	1.0556	-.021 (.67)	-.014 (.42)	-.015 (.34)

Table 13

Mean RMSEA and Standard Deviation (in Parentheses) under Correctly or Incorrectly Specified Form of the Growth Model for with 6 Repeated Measures, Slope Variance of 20, and Error Variance of 50 across Time

Growth form in population vs. model	Slope difference	Sample size		
		200	500	800
Quadratic – Linear (misspecified)	0	.219 (.02)	.219 (.01)	.219 (.01)
	.5278	.219 (.02)	.219 (.01)	.219 (.01)
	.7917	.219 (.02)	.219 (.01)	.219 (.01)
	1.0556	.219 (.02)	.219 (.01)	.219 (.01)
Linear – Linear	0	.017 (.02)	.010 (.01)	.008 (.01)
	.5278	.017 (.02)	.010 (.01)	.008 (.01)
	.7917	.017 (.02)	.010 (.01)	.008 (.01)
	1.0556	.017 (.02)	.010 (.01)	.008 (.01)

Table 14

Average Chi-Square and Standard Deviation (in Parentheses) under Correctly or Incorrectly Specified Form of the Growth Model for with 6 Repeated Measures, Slope Variance of 20, and Error Variance of 50 across Time

Growth form in population vs. model	Slope difference	Sample size		
		200	500	800
Quadratic – Linear (misspecified)	0	223.75 (31.60)	525.00 (50.96)	824.56 (62.99)
	.5278	223.75 (31.60)	525.00 (50.96)	824.56 (62.99)
	.7917	223.75 (31.60)	525.00 (50.96)	824.56 (62.99)
	1.0556	223.75 (31.60)	525.00 (50.96)	824.56 (62.99)

Comparison of Results across Power Analysis Approaches

Three different power analysis techniques (Monte Carlo, Satorra-Saris, and MacCallum-Browne-Cai) were used to estimate power to detect a significant difference in slopes. Table 15 reports the resulting power estimates across these three techniques and data conditions (slope variance of 20; level-1 error variance of 50 or 100; 3 or 6 measurement occasions; sample size of 200, 500, or 800; and slope difference of .5278, .7917, or 1.0556). Figures 5, 6, and 7 graphically depict power across these techniques and conditions.

In examining the power estimates for the Monte Carlo and Satorra-Saris technique, it is apparent that, across conditions for level-1 error variance, slope variance, number of measurement occasions, slope difference, and sample size, these two techniques provide comparable results. In fact, under no condition did power estimates from the Monte Carlo and Satorra-Saris technique differ by more than 0.03; this condition was for error variance of 100, sample size of 500, and six repeated measures, with the majority of estimates differing by 0.01 or less across these two techniques. The overlap of the Monte Carlo and Satorra-Saris power plots in Figures 5, 6, and 7 emphasize the similarities between these techniques across conditions. Even with a sample size of 200, Satorra-Saris power estimates, across all conditions, were quite close to the Monte Carlo estimates.

A limited investigation of smaller sample sizes smaller than 200 suggests that sample size can be quite small and these techniques can still yield comparable results. With an intercept variance of 100, slope variance of 20, error variance of 50, and slope difference of 0.7917, powers differed by no more than 0.01 between

the Monte Carlo and Satorra-Saris techniques for samples sizes from 50 to 200. For sample sizes of 200 and 150, the Satorra-Saris technique yielded power estimates that were approximately 0.01 greater than the Monte Carlo technique. With sample sizes of 100 and 50, however, the Satorra-Saris technique yielded power estimates that were approximately 0.01 less than the Monte Carlo technique.

However, any similarities between MBC technique estimates and the Monte Carlo (or Satorra-Saris) technique appear more arbitrary, as different model conditions result in different MBC power estimates more closely approximating the Monte Carlo power values. When using the MBC technique with six repeated measures, the lower RMSEA values that assume perfect fit for the full model (0 and .01; 0 and .02) tend to result in power estimates closest to the Monte Carlo and Satorra-Saris estimates, with the 0-.01 criterion being the most accurate with a small slope difference and the 0-.02 criterion being the most accurate with a moderate slope difference (with an effect size of approximately .3). For an effect size of approximately .4, the Monte Carlo and Satorra-Saris power estimates fall between the MBC power estimates with RMSEA pairs of 0-.02 and .04-.05. In all conditions with six repeated measures, the .04-.05 RMSEA criterion consistently overestimated power in comparison to the Monte Carlo and Satorra-Saris techniques. This relationship between the MBC technique and the Monte Carlo and Satorra-Saris techniques, in terms of which RMSEA corresponds most closely estimated Monte Carlo and Satorra-Saris powers, changes when the model includes only three repeated measures.

The number of repeated measures plays a large role in the MBC power estimates, as greater degrees of freedom result in increased power (see Figure 7). Therefore, whereas power estimates for a model with six repeated measures includes many conditions that overestimate power relative to the Monte Carlo and Satorra-Saris approaches, power estimates for a model with three repeated measures results in power often being underestimated in comparison with the Monte Carlo and Satorra-Saris techniques. Because of the substantial decrease in degrees of freedom between models with six vs. three repeated measures, when the model includes three repeated measures, only the .04-.05 RMSEA power estimates are remotely close to those of the Monte Carlo and Satorra-Saris estimates (with the increase in RMSEA values increasing power, bringing MBC estimates closer to the Monte Carlo approach).

With three repeated measures, however, the influence of effect size on power estimates using the Monte Carlo and Satorra-Saris techniques is quite apparent, emphasizing differences between these techniques and the MBC technique. Whereas conditions with six repeated measures included MBC estimates that were relatively close to the Monte Carlo and Satorra-Saris estimates (with larger RMSEA values better approximating power for conditions with larger effect sizes), the greatly underestimated power for the MBC technique with three repeated measures makes it obvious that as the slope difference increases, the Monte Carlo and Satorra-Saris power estimates become increasingly divergent from all MBC estimates. Whereas the Monte Carlo and Satorra-Saris techniques respond to the increased effect size with substantial increases in power, the MBC

power remains underestimated due to low degrees of freedom and perhaps inappropriate RMSEA values under the given conditions.

Finally, although the influence of level-1 error variance on power estimates in the Monte Carlo and Satorra-Saris techniques is small, the change in power seen across error variance conditions for the Monte Carlo and Satorra-Saris techniques is not apparent in the MBC technique because of the limited information required to estimate power using the MBC technique. This emphasizes the fact that data conditions and parameter values that are known to influence power (such as error and slope variance) do not influence MBC power estimates, which could result in more divergence in power estimates between techniques.

Table 15

Power across Power Analysis Techniques with Slope Variance of 20

Level-1 error variance	Number of measurement occasions	Slope difference	Monte Carlo			Satorra-Saris			MacCallum-Browne-Cai								
			Sample size			Sample size			Sample size								
			200	500	800	200	500	800	200	500	800	200	500	800	200	500	800
50	3 (unconstrained model df = 2; constrained model df = 3)	.5278	.143	.225	.337	.119	.226	.331	.057	.067	.078	.078	.121	.165	.152	.311	.468
		.7917	.218	.436	.645	.208	.438	.627	--	--	--	--	--	--	--	--	--
		1.0556	.328	.675	.861	.330	.671	.860	--	--	--	--	--	--	--	--	--
	6 (unconstrained model df = 20; constrained model df = 21)	.5278	.112	.220	.343	.122	.235	.345	.099	.176	.254	.253	.535	.736	.524	.892	.982
		.7917	.205	.437	.631	.215	.456	.647	--	--	--	--	--	--	--	--	--
		1.0556	.346	.676	.853	.344	.692	.876	--	--	--	--	--	--	--	--	--
100	3 (unconstrained model df = 2; constrained model df = 3)	.5278	.115	.197	.291	.109	.200	.292	--	--	--	--	--	--	--	--	--
		.7917	.194	.387	.564	.185	.386	.561	--	--	--	--	--	--	--	--	--
		1.0556	.289	.609	.791	.291	.605	.804	--	--	--	--	--	--	--	--	--
	6 (unconstrained model df = 20; constrained model df = 21)	.5278	.104	.199	.307	.114	.214	.313	--	--	--	--	--	--	--	--	--
		.7917	.188	.382	.583	.197	.414	.597	--	--	--	--	--	--	--	--	--
		1.0556	.305	.630	.809	.312	.641	.836	--	--	--	--	--	--	--	--	--

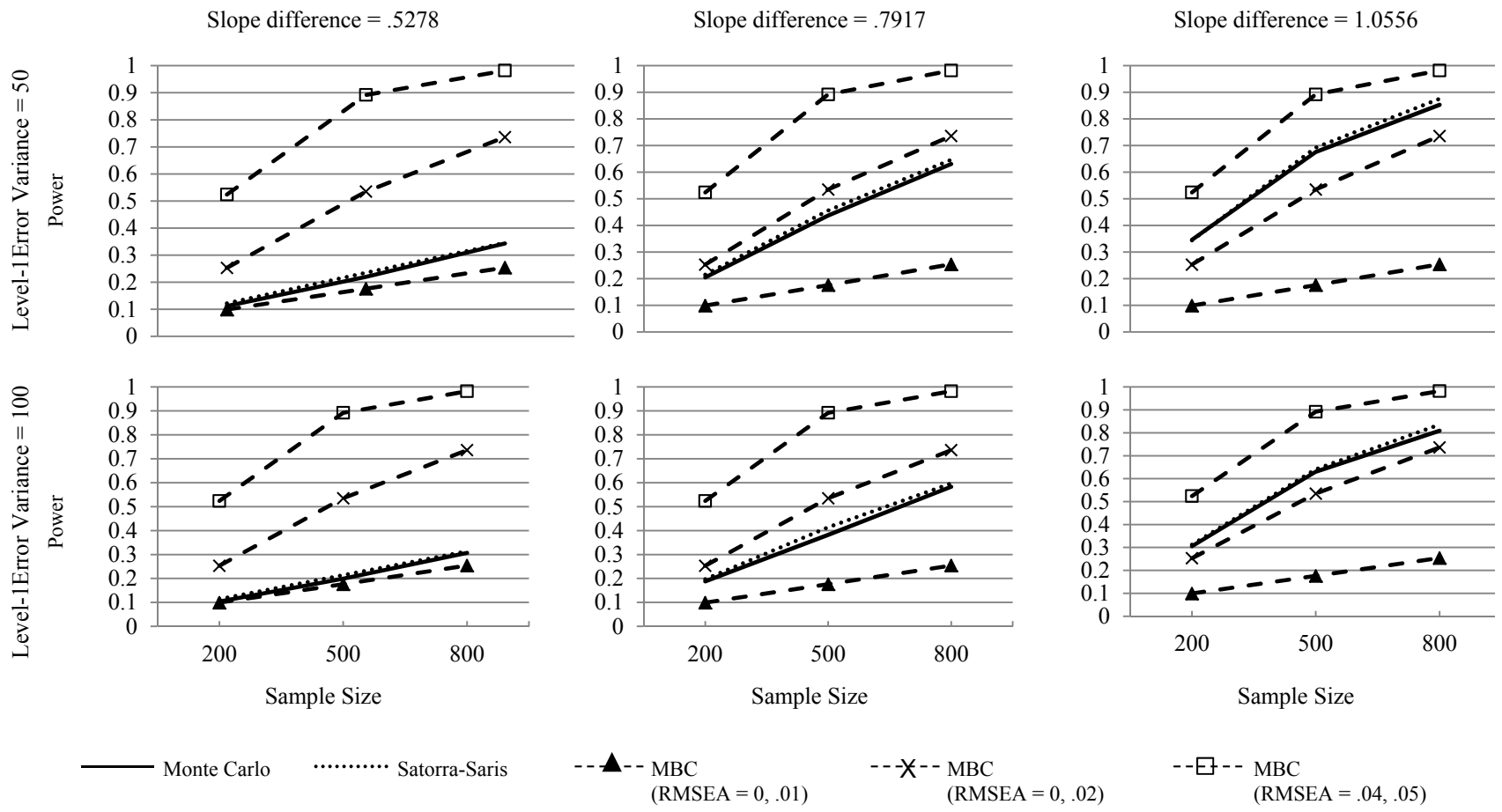


Figure 5. Comparison of power using different power analysis techniques with 6 repeated measures and slope variance of 20.

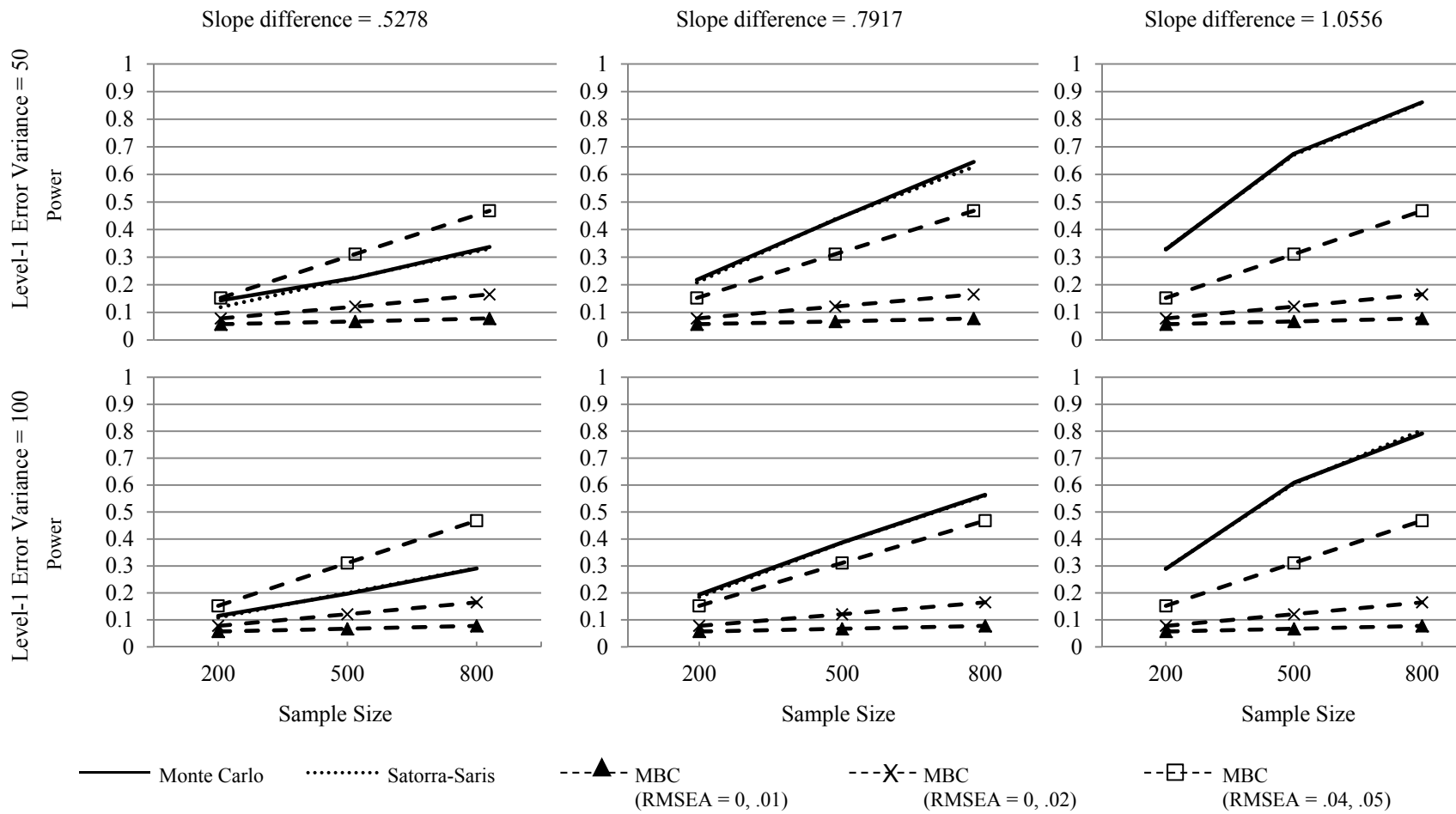


Figure 6. Comparison of power using different power analysis techniques with 3 repeated measures and slope variance of 20.

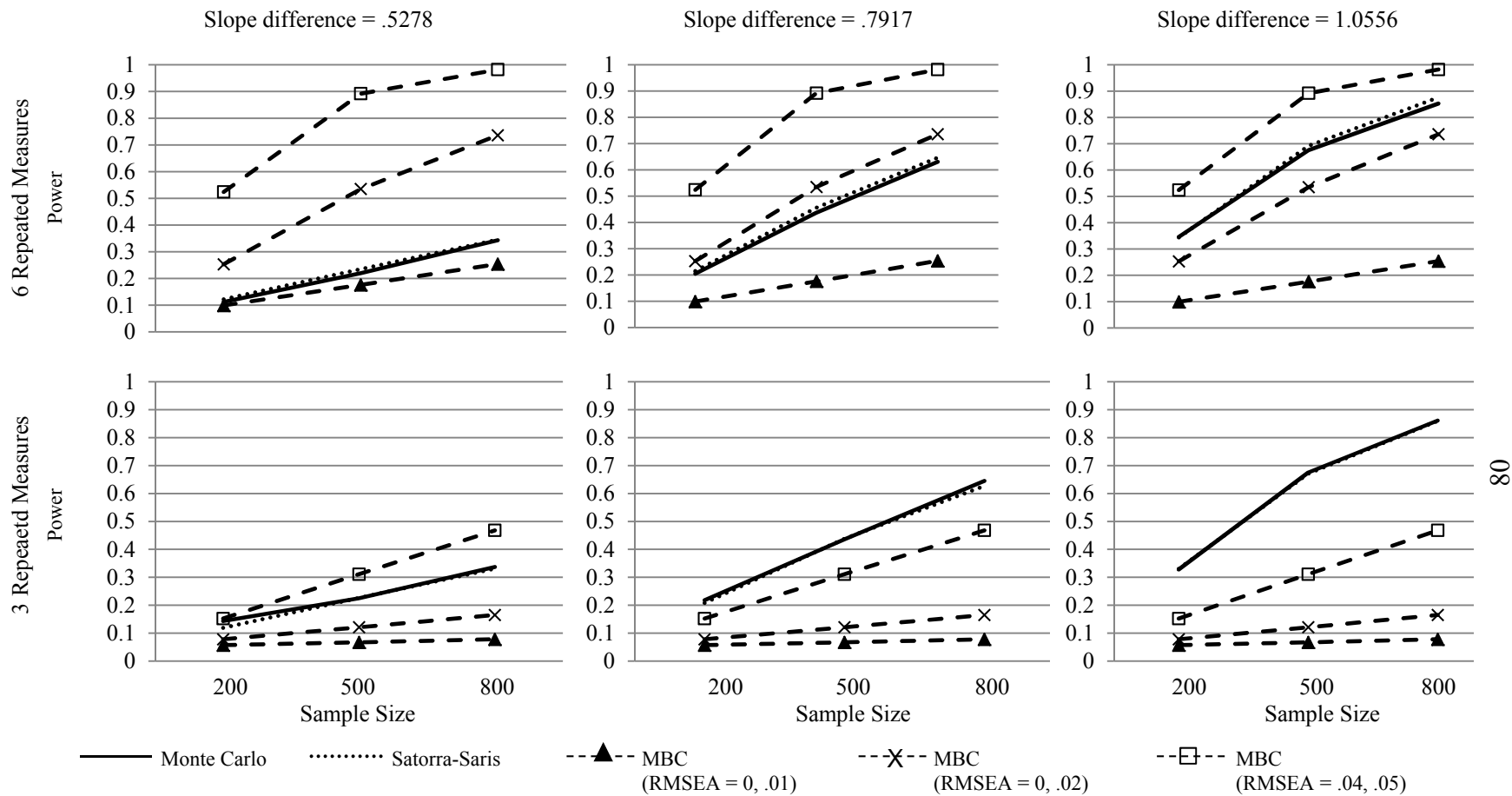


Figure 7. Comparison of power using different power analysis techniques with 3 and 6 repeated measures, level-1 error variance of 50, and slope variance of 20.

CHAPTER 4

DISCUSSION

The present investigation was designed to assist investigators in understanding conditions that affect power to detect the influence of a treatment variable on growth in latent growth modeling. By using the Monte Carlo approach to power analysis to investigate how slope variance, level-1 error variance, sample size, slope difference, and number of measurement occasions influenced power, it was possible to elucidate which parameter specifications require careful attention in conducting a priori power analyses. Additionally, by estimating power using three power analysis techniques commonly employed for SEM contexts, investigators can be informed about conditions in which these methods are likely to yield inconsistent results.

In the base power conditions using the Monte Carlo technique, number of repeated measures, slope variance, level-1 error variance, sample size, and slope difference had the greatest impact on power estimates. Additionally, although misspecification in the level-1 error variance structure slightly influenced power, misspecification in the growth structure had a substantial influence on power. Further, power estimates varied somewhat according to the choice of technique used to estimate power, with the MBC approach differing markedly from the Satorra-Sarris and Monte Carlo approaches for some conditions.

Type I error

Overall, Type I error rates were quite close to the prescribed alpha in all conditions. The conservative Type I error rates that occurred as sample size and

number of repeated measures increased could be potentially concerning; however, these Type I error rates were only slightly conservative. In fact, the most conservative Type I error estimate was .037 (for sample size of 800, slope variance of 20, error variance of 100, and 6 repeated measures), which still fell within Bradley's (1978) liberal criterion for an alpha of .05, requiring that Type I error values fall between .025 and .075.

Number of repeated measures

This investigation did not show that number of repeated measures, across a fixed study length, substantially influenced power to detect between-group differences in slopes. Although various studies have investigated how the number of repeated measures influence power in latent growth modeling (e.g., Hertzog et al., 2006; Hertzog et al., 2008; Muthén & Curran, 1997; Zhang & Wang, 2009), these studies typically focused on a different model parameter or focused on length of study rather than on number of repeated measures holding length of study constant.

Hertzog et al. (2006), for example, focused on the power to detect correlated change between two variables measured over time, whereas Hertzog et al. (2008) focused on the power to detect slope variance. In both studies, power was investigated across 1 to 19 repeated measures and an increase in repeated measures corresponded to an increase in study length, while varying conditions such as sample size, effect size, and growth curve reliability. Growth curve reliability was measured at wave one as intercept variance divided by total variance (i.e., intercept variance plus error variance). Thus, growth curve

reliability was manipulated by varying level-1 error variance values. The influence of number of measurement occasions on power to detect the covariance between slopes (2006) and slope variances (2008) was more substantial than what was found in the present study. For example, the 2006 study found that, under conditions of growth curve reliability between .5 and .7 (comparable to this study's range of growth curve reliability) with a correlation of .50 between variable growth and a sample size of 500, the maximum difference in power to detect correlated growth between three and six repeated measures was approximately .40 (with a power of .03 for three repeated measures and a power of .43 for six repeated measures). In the 2008 study, when growth curve reliability ranged from .5 to .7, with no correlation between slope and intercept variance, with a sample size of 500, the maximum difference in power to detect slope variance was .87 (.03 for three repeated measures, .90 for six repeated measures). While a number of other conditions also influenced the magnitude of the difference in power, the influence of number of measurement occasions combined with length of study on power for these particular parameters appears to have had a more substantial influence on power than the present study, which investigated the influence of number of repeated measures only (i.e., holding length of study constant).

Zhang and Wang (2009) took a similar approach to investigating number of repeated measures, with measurement occasions ranging from three to six and an increase in measurement occasions corresponding to an increased study length. Again, increased study length and number of measures was related to increased

power to detect growth. For example, in order to obtain a power of .80 with 3 repeated measures, a sample size of 300 was required, whereas the same power could be obtained with 6 repeated measures with a sample size of 210.

Unlike Hertzog et al. (2006, 2008) and Zhang and Wang (2009), Muthén and Curran (1997) considered the number of repeated measures for a model in a more detailed manner and for a parameter similar to the present investigation—the influence of a treatment variable on growth. Therefore, Muthén and Curran’s study yielded results more comparable to the present study. Rather than focusing solely on length of study, Muthén and Curran considered length of study, number of measurement occasions for a given study length (as was investigated in this study), and study length for given number of measurement occasions. Overall, similar to this investigation, the findings by Muthén and Curran suggest a small increase in power related to increased number of measurement occasions for a constant study length. For example, in comparing three versus five repeated measures with a constant study length, power differed by no more than .06 between numbers of repeated measures, which is similar to the maximum difference of .04 observed in this investigation. However, as with the other studies that considered length of study along with measurement occasions, Muthén and Curran report a more dramatic influence on power when study length increased from three time points to seven time points, where time between intervals is equal, making the length of a study with seven time points longer than a study with three time points. For example, with sample size of 500 and 3 time

points, power was approximately .52, whereas 7 time points and a correspondingly longer study period resulted in a power of .87.

Ultimately, it appears as though the influence of number of measurement occasions, given a constant study length, does influence power to detect the influence of a treatment condition on growth, with more measurement occasions increasing power; however, this difference was small within the limited range investigated. Clearly, the influence of study length and number of measurement occasions on power in latent growth modeling is an important consideration in study design. A more complete understanding of how number of measurement occasions influence power to detect various parameters in latent growth models requires an approach similar to Muthén and Curran, who considered a greater range of how length of study and number of measurement occasions can jointly impact power to detect differences in growth rates. Therefore, a limitation of the present study is that only one of the three measurement occasion conditions from Muthén and Curran's study was considered in the present study. Researchers seeking to optimize power to detect particular effects while minimizing costs are advised to consider the potential impact on power of various feasible combinations of length of study and number of measurement occasions, along with other conditions (e.g., error variance, sample size, effect size).

Additionally, although Hertzog et al. (2006, 2008) focused on different parameters and investigated the number of measurement occasions differently (as a combination of study length and number of repeated measures), their findings do suggest that a greater growth curve reliability increases the influence that

number of repeated measures have on power. Perhaps this relationship between growth curve reliability and number of repeated measures also occurs when the number of repeated measures is considered with a fixed study length. The present study considered a narrow range of growth curve reliabilities that did not exceed .70; however, Hertzog et al. findings indicated that the greatest difference in power, relating to measurement occasions, occurred between growth curve reliabilities of .70 to .99. Therefore, in addition to considering the number of measurement occasions more systematically, as was done by Muthén and Curran, future investigations should also consider a wider range of growth curve reliabilities, which can also be conceptualized as a wider range of error variances.

Finally, Venter, Maxwell, and Bolig (2002) report that three repeated measures in an ANCOVA setting, as compared to a pre-test/post-test study design, provide the advantage of additional information relating to growth form. Although they report that the biggest gains in power typically occur with five repeated measures (vs. two repeated measures), Venter et al. point out that even with minimal increases in power to detect growth with three versus two time points, the gains in understanding of growth and growth form resulting from three repeated measures can be reason enough to include an additional measurement occasion. Therefore, it is important to note that although power is an important consideration, small increases in power coupled with gaining additional information relating to growth can provide support for using additional measurement occasions. Therefore, investigators must consider their hypothesized form of growth, available resources, method of including additional

measurement occasions (i.e., holding study length constant or increasing study length), and implications for power in determining number of measurement occasions.

Slope variance

Hertzog et al. also manipulated slope variance in their investigations of power to detect the covariance between slopes (2006) and slope variance (2008) in latent growth modeling, considering slope variances of 25 or 50. However, these studies provide little insight into the influence of slope variance on power. In their 2006 investigation, Hertzog et al. investigated the influence of slope variance on power to detect growth covariance between two variables; however, the results of the slope variance conditions were not discussed. In their 2008 investigation, Hertzog et al. focused on power to detect slope variance, meaning an increase in slope variance increased power, making their scenario quite different from the present investigation.

As expected, this study found that increased variance decreased power to detect the treatment effect on growth. The power differences between slope variances of 10 and 20 were not inconsequential, with differences of up to .24 in power in the base conditions, holding all else constant. This suggests that slope variance estimates in conducting power analyses require consideration, and power should be investigated with a range of slope variance values in order to ensure an accurate range of power estimates and to determine ideal sample size for a given study. Additionally, researchers conducting power analyses might consider plotting their data based on parameter estimates, as done in Figure 3, in order to

get a visual sense of their slope variance estimates and to assist them in making more informed choices for slope variance.

In terms of the supplemental exploration of the slope difference to slope standard deviation ratio as it relates to power to detect the influence of a treatment effect on growth, it is evident that a larger ratio is related to increased power, as expected. However, power did increase within each ratio as slope difference increased. The relatively narrow range in power estimates, even when effect size ranged from .22 to .86, suggests that slope difference and slope variance should be considered together. The relationship between these two parameters could be helpful in providing an additional guide in parameter value selection as investigators consider their expectations relating to the magnitude of slope variance and effect size. Overall, this ratio could potentially serve as an alternative effect size for a difference in slopes; however, the difference in power estimates, even when this ratio was held constant, indicates that this ratio does not capture all individual effects of slope difference and slope variance. Because of the brief attention given to the ratio of slope difference to slope standard deviation in the present investigation, it is difficult to make strong conclusions at this point. Rather, additional consideration should be given to this matter in order to better understand this ratio's relationship to power and potentially yield useful guidelines in parameter value estimates.

Level-1 Error Variance

Unlike their brief investigation of slope variance, Hertzog et al. (2006, 2008) considered the influence of level-1 error variance more extensively. They

conceptualized the manipulation of error variance as a manipulation of growth curve reliability, measured (at wave one) as intercept variance divided by total variance (i.e., intercept variance plus error variance). Their investigation considered growth curve reliability values ranging from .50 to .99, resulting in error variance values of 100 to 1, respectively. Ultimately, as in their 2006 investigation of power to detect slope covariances, Hertzog et al. (2008) found that growth curve reliability (and thus error variance) had a “profound effect” on power to detect slope variance (p. 557). For example, with a large effect size (i.e., large slope variance), a sample size of 500, and five measurement occasions, power remained low when growth curve reliability was less than .91.

Although this same pattern in which low error variance results in higher power was seen in the present study, the influence of error variance (50 vs. 100) was not substantial. However, the range of growth curve reliability values in this investigation (.50 or .67) was substantially more restricted than that of Hertzog et al. In Hertzog et al. (2006), changes in power to detect a slope covariance across growth curve reliabilities of .50 to .70 (for both three and six measurement occasions) were evident; however, changes in power with lower growth curve reliabilities compared to changes for growth curve reliability values of .70 to .99 were much less substantial. Thus, the narrow range of the growth curve reliabilities in the present investigation and the fact that these values fell on the low end of the growth curve reliability scale (compared to Hertzog et al.) resulted in our finding that error variance did have an impact on power, albeit a less

pronounced effect than Hertzog et al. found across larger growth curve reliabilities.

This, however, begs the question of whether or not such high growth curve reliabilities are plausible and worth investigating in a priori power analyses. In determining population values for the base models, model-implied correlation values were considered in an effort to select parameter values that were realistic. Larger residual variance values played a key role in finding model-implied correlation values that were not unrealistically high. Employing error variances of 1 or 10 (corresponding to growth curve reliabilities of .99 or .91) resulted in very high model-implied correlations between waves. For example, with an intercept variance of 100, slope variance of 20, and error variance of 10, model-implied correlations ranged from .65 to .98, (with a number of correlations exceeding .90).

Although growth curve reliability, and, accordingly, error variance, can have a substantial impact on power, it appears that this impact occurs with error variance values that may be unrealistically low or would require special considerations in regard to study planning and execution to achieve. Perhaps, in most cases, level-1 error variance will play a more minor role in a priori power analyses as greater error variances that yield lower growth curve reliabilities are more likely to be investigated in an effort to use realistic parameter values. A review of applied study results and their level-1 error variance values and growth curve reliabilities may be helpful in determining a realistic range of growth curve reliability estimates.

As discussed further below, results similarly suggest misspecification of the error structure only slightly influences power, suggesting that, in many cases, error variance is not a parameter that requires a great deal of concern in power analyses. However, for those investigators who believe their study can minimize error variance (e.g., by decreasing measurement error through the use of multiple-indicators for the outcome variable; Hertzog et al., 2008), it could be useful to consider the increase in power that could result. Of course, investigating higher error variance values allows for more conservative power estimates and sample size planning. It must also be noted that these results apply only to the power to detect differences in slopes. Investigators who are interested in differences in means across groups at given time points may find that level-1 error variance is quite influential for the power to detect these differences in means.

Misspecification in Level-1 Error Structure and Form of the Growth Model

It is informative to consider how particular types of model misspecification can influence power to detect focal effects. In reviewing the influence of error variance structure on power, it appears that constraining error variances to be equal when they are increasing over time only slightly influences power. However, this could be related to the narrow range of error variance values considered. The largest incremental increase in the residual error variance was 15 at each wave, meaning growth curve reliability was 50 at wave one (yielding a growth curve reliability of .67) and was 125 at wave six (yielding a growth curve reliability of .44). As seen in Hertzog et al. (2006, 2008), the most substantial influence of error variance on power to detect slope covariances or

slope variances appeared when growth curve reliability approaches .7 or .8. Therefore, a greater influence of power due to misspecification could result with larger increments of increase or smaller error variance values. Overall, it appears that constraining error variances to be equal across time when error variances increase over time, as manipulated in this study (by 5 vs. 15), only slightly influences power to detect slope differences and may not be a large concern for power analysts.

Misspecifying the form of a growth model by modeling a quadratic growth model as linear, however, substantially influences the probability of detecting differences in slopes. This finding supports the tactic of looking at form of growth sequentially in order to avoid misspecification of the form of growth. That is, researchers should first specify a linear growth model and then examine whether a quadratic model improves upon this model. Although this tactic may have Type I error considerations, it helps ensure that the proper form of growth is not overlooked by testing a single omnibus hypothesis. In considering misspecification in the form of model growth in this investigation, however, only one form of misspecification was considered, making this portion of the investigation limited and only a first step in further studies. Ultimately, it would be best to consider past research and theory so that the most likely growth form can be considered. If past research is unclear regarding the likely shape of the growth trajectory for the focal outcome, power might consider investigating power across multiple forms of growth.

Power Analyses across Techniques

Past studies have found that, for the most part, the Monte Carlo and Satorra-Saris power analyses techniques are quite comparable with latent growth models in detecting the significance of a number of parameters (e.g., slope covariances, differences in slopes; Muthén & Curran, 1997; Hertzog et al., 2006). However, in some settings, the Satorra-Saris approach has been found lacking in its accuracy, such as when comparing two slightly misspecified models (Hertzog et al., 2008), with the accuracy of the Satorra-Saris approach defined by its consistency with Monte Carlo power estimates. In the present study conditions, however, the Satorra-Saris technique was found to approximate the Monte Carlo power estimates quite well, with power to detect a difference in slopes never differing by more than .03 between these two techniques. In fact, in explorations of sample sizes as low as 50, consistent results were still found between the methods. In choosing between the Satorra-Saris approach or the Monte Carlo approach under conditions in which both result in similar power estimates, researchers should therefore consider the feasibility of manipulating factors relevant to the study. For example, investigations of the impact of non-normality, model misspecification, and missing data on power can only be examined using the Monte Carlo approach. However, investigations that might focus on the influence of sample size for a particular model could use the Satorra-Saris approach, which allows for the calculation of noncentrality parameters across a range of sample sizes using the power estimate from just one sample size using a simple equation (multiplying the chi-square for a given sample size by the ratio of

the new sample size to the sample size used to determine the initial chi-square value). Although the Satorra-Saris approach requires more steps (3 steps using Mplus), this technique could save time in estimating power for a parameter across multiple sample sizes.

The MBC power analysis technique was most discrepant of the three analysis approaches. Because of the limited factors considered in the calculation of power using this technique, only the targeted RMSEA values can be manipulated in order to improve power estimates. Although MacCallum, Browne, and Cai (2006) recommend against using actual RMSEA values derived from samples in order to determine post hoc (i.e., observed) power, they do indicate that the selection of RMSEA values for a priori power analyses “be made with as much care as possible” (p. 30). Therefore, understanding how RMSEA values behave under certain conditions and exploring RMSEA values from past research may aid users of the MBC technique.

Ultimately, investigating power for a range of RMSEA values, in combination with considering model conditions and past research, may result in the most informative use of the MBC technique. For example, a larger expected effect size could be factored into the MBC approach by selecting RMSEA values that are more discrepant (e.g., .01 and .05 rather than .01 and .02). For example, in using the MBC SAS program that calculates power for a range of RMSEA values (Program F; MacCallum, Browne, & Cai, 2006), and focusing on a nested model with 3 degrees of freedom, a full model with 2 degrees of freedom, and sample size of 500, it was easy to see that by specifying the full model RMSEA to

be 0 (for compatibility with the Satorra-Saris technique), power estimates range from .07 to .97 as the nested model RMSEA increased from .01 to .10. This information alone tells an investigator very little about the likely power for detecting the focal parameter in their model. However, if one expects a small effect size for a parameter of interest, it might be useful to consider only the smaller discrepancies in RMSEA values (perhaps a nested model RMSEA value of .01 to .03, which yields a range in power estimates of .07 to .21; the Monte Carlo estimate was .23). If a larger effect size is expected, perhaps the nested model RMSEA values should range from .04 to .07 (which yields a range in power estimates of .34 to .77; the Monte Carlo estimate was .44 and .67 for effect sizes of approximately .20 and .30). Currently, these RMSEA values are selected arbitrarily; however, the general concept of selecting RMSEA pairings that reflect the expected effect size could help investigators determine a more specific range of power estimates that would more accurately reflect the Monte Carlo approach.

Additionally, knowing that the MBC technique results in low power estimates with low degrees of freedom, one might consider using RMSEA values with slightly larger discrepancies in order to compensate for this. Further, with fewer degrees of freedom, constraining a single parameter to be 0 could have a greater influence on model RMSEA compared to constraining a parameter to be 0 in a larger model; therefore, it would make sense to use more discrepant RMSEA values.

Following the completion of this investigation, Li and Bentler (2011) suggested an alternative method of using the MBC (2006) technique to determine

power to detect differences in model fit. Rather than selecting a pair of RMSEA values (i.e., a value for the full model and the nested model), Li and Bentler outline a technique that allows investigators to select only one RMSEA value. Whereas the selection of two RMSEA values can result in noncentrality parameters that are not necessarily comparable across models, even when the same RMSEA pairs are used, Li and Bentler's method of defining a single RMSEA value means the RMSEA value can retain its meaning and be compared across models that differ in degrees of freedom. This alternative method of using the MBC technique could be useful to practitioners in selecting RMSEA values that remain meaningful across models. Although Li and Bentler's approach appears to yield results comparable to the MBC approach, it may be useful to consider this alternative MBC approach in future investigations because of the benefits this approach has over the traditional MBC technique.

Overall, it seems that although the MBC approach at first appears to yield a wide range of power estimates that can deviate from the Monte Carlo estimates substantially, it is possible that the MBC approach could be used to approximate the power to detect a single parameter if an investigator takes specific model conditions and RMSEA behaviors into account. However, additional investigations of MBC power estimates across RMSEA values and across model conditions is required before any specific suggestions could be made.

Additionally, in considering the MBC versus the Monte Carlo and Satorra-Saris techniques, an important question then becomes whether or not the Monte Carlo and Satorra-Sarris power estimates are actually providing accurate power

estimates. Of course, power estimates are only accurate if parameters are properly specified; therefore, although the MBC approach may result in estimates substantially different from the Monte Carlo and Satorra-Saris technique, if the latter techniques involved the use of inaccurate parameter estimates--perhaps due to a lack of pilot data or past research--then it may not be possible to say which power estimate is best. Incorrect estimates for slope difference, for example, can substantially influence power and result in quite inaccurate estimates using the Monte Carlo and Satorra-Saris techniques. As the number of parameters that require estimation increase, the chance of misestimating parameters that can influence power also increases. Estimating the unknown, however, is an inherent concern in a priori power analyses, with any power analysis technique being riddled with estimates and guesses.

Limitations and Future Directions

Because of the nature of power analysis and simulation studies, this investigation considered only one set of manipulated conditions that influence power to detect a treatment effect in latent growth modeling. This study aimed to determine factors that may impact power to detect slope differences with the goal of providing power analysts with some guidance in terms of factors that may be more or less important to estimate carefully. Within any one factor, only a small number of possible parameter values were considered; more in-depth study of factors such as slope difference or slope variance, for example, could be examined systematically in future simulation work.

First, only two error variance values and two slope variance values were considered. Investigating a wider range of variances could help better identify trends in how they influence power across other model conditions. Ultimately, for power analysts working within a particular domain, a careful review of the literature and previous parameter estimates within particular construct domains (e.g., standardized achievement measures or personality measures) may be required to properly estimate model parameters for power analysis via either the Monte Carlo or Saris-Satorra approach. For example, the present study considered only homogenous level-1 error variance over time or consistently increasing variances over time; however, a thorough review of the literature in certain academic fields could suggest other observed patterns that should be investigated in terms of their influence on power. Additionally, this investigation considered only number of repeated measures within a fixed study length. Future research should also consider study length across fixed numbers of repeated measures in order to more fully understand when additional measurement occasions may or may not be fruitful in applied research.

In comparing Monte Carlo power estimates to Satorra-Saris power estimates, the present investigation found these estimates to be very comparable; however, the model conditions investigated were quite limited. In order to better understand how the Satorra-Saris approach compares to the Monte Carlo approach in power to detect differences in slope, it would be useful to consider more complex models and model conditions (e.g., additional covariates, quadratic

growth, missing data, nonnormal data) and models with varied slope and intercept relationships (as in Fan, 2003).

Finally, the method of determining significance (e.g., Wald test vs. likelihood ratio test; Hertzog et al., 2008) should be considered and varied. The design of this investigation involved conducting each power analysis approach as it would most likely be applied by a researcher. A possible limitation of this approach is different statistical indices were used in estimating power across these techniques. For example, power estimates for the Monte Carlo approach were based on the Wald tests for the parameter of regressing the binary treatment covariate on the slope factor. The Satorra-Saris approach, on the other hand, uses the likelihood-ratio chi-square test, with the resulting chi-square value for a slightly misspecified model (i.e., the focal parameter set to 0) approximating a noncentrality parameter. Although investigating these power analysis approaches using the significance test that researchers would typically employ is useful, it is also important to consider these techniques with comparable methods of significance testing as well as with other methods of testing significance, as done by Hertzog et al. (2008). Ultimately, the difference arising from the use of these two approaches is likely small, especially as sample size increases, as these two approaches are asymptotically equivalent (Buse, 1982).

Another key consideration for future research involves defining effect size in studies focusing on longitudinal growth modeling. Currently, the methods used to calculate effect size, particularly for differences in growth, vary across studies and capture different components of the model. For example, Muthén and Curran

(1997) and Fan (2003) defined the effect size for differences in growth as the difference between groups at a particular time point divided by the standard deviation. Further, the standard deviation used in this calculation could also vary, with investigators using either the standard deviation across the control and treatment groups or using the standard deviation for the control group only. Feingold (2009), on the other hand, suggested calculating effect size based on the overall growth rather than focusing on a particular point in time, with effect size calculated as the difference in growth multiplied by time and divided by the initial standard deviation of raw scores. Muthén and Muthén (2002), however, took yet another approach, defining effect size as the difference in slope means divided by the slope standard deviation. Varying approaches to calculating effect size for a single parameter makes equating findings difficult and complicates the power analysis process.

Future studies should consider how effect size is (or should be) defined, as well as other techniques of conceptualizing parameter variances within the context of longitudinal models. For example, further investigations of the slope difference to slope standard deviation ratio, two values that were found to substantially influence power, could yield information useful for estimating power and selecting parameter values. In turn, a better understanding of this ratio could inform investigators as to whether or not they need to consider multiple parameter values that ultimately yield the same slope difference to standard deviation ratio. Similarly, one might consider the use of intraclass correlation coefficients to understand important parameters, such as slope variance, within the larger model

context. Using such a technique could allow investigators to calculate the total percentage of variance in the dependent variable at a particular time accounted for by slope variance and therefore allow power analysts to better define what makes a large or small slope variance. Overall, consistency in defining effect size and additional methods of operationalizing parameters within the larger model context (understanding how much particular parameters contribute to total variance) are important considerations for future investigations and may allow simplification of parameter selection and power analyses.

Summary and Suggestions

Overall, this investigation suggests that effect size, which was manipulated here by slope difference, slope variance, and sample size are important parameters to thoroughly investigate when conducting power analyses. However, repeated measures (3 vs. 6) when study length is held constant and level-1 error variance require less consideration. Error variance has a small influence on power to detect the influence of a treatment on growth even in instances of model misspecification relating to error variance structure, when error variance values yield lower, potentially more realistic growth curve reliabilities. Additionally, one's decision to use the Monte Carlo technique or Satorra-Saris technique, under the conditions investigated here, depends largely on preference and focal conditions (e.g., power across sample sizes, missing data, non-normal data). Further, although the MBC technique appears quite discrepant from the Monte Carlo and Satorra-Saris techniques, the assumption of superiority of the Monte Carlo and Satorra-Saris techniques presumes that reasonable

parameter estimates are employed; poor estimates of important parameters (e.g., effect size or slope variance) could result in misleading estimates of required sample size. However, it is similarly apparent that although the MBC technique requires substantially less information (df, alpha, RMSEA), the selection of the RMSEA values greatly influences power estimates and results in a quite large range of power estimates depending on RMSEA values chosen.

Additionally, model-implied correlation matrices and plots based on estimated parameters should be generated and examined to ensure model plausibility and assist in the selection of realistic model parameters. Power analysts would be wise to avoid relying solely on parameter values used in past power investigations, as some of these values may yield unrealistic model-implied correlations and plots. Investigations of model-implied correlation matrices conducted while selecting parameter values in this study suggest that correlations among observations become more realistic as slope variance decreases and level-1 error variance increases. Therefore, these parameter values should be considered if correlations among observations are unrealistically high for a given measure or context. In addition to observing model-implied correlations, however, it is also useful to consider plots of hypothetical data based on various parameter values. The combination of graphing trajectories and observing model-implied correlations can help ensure parameter values imply data that are reasonable, therefore yielding more accurate power analyses.

Ultimately, a priori power analysis in latent growth contexts is a necessarily complex task that requires estimation of many parameters. Multiple

decisions must be made in terms of growth model specification and the focal parameters may vary across investigations, making the a priori power analysis process unique for each investigation. Further, any one study likely involves hypotheses that involve multiple parameters. Even in using a simplified analysis approach that requires minimal model information, such as the MBC approach, power estimates can vary widely and confidence in resulting estimates may be low. However, by investigating a limited number of systematically varied focal parameters and models, researchers may be able to provide more useful guidance to power analysts as they begin their exploration of a priori power for particular models.

In short, this investigation suggests that power analysts interested in differences between slopes in longitudinal growth models carefully consider values for effect size, sample size, and slope variance, as these model conditions and parameter values can substantially influence power estimates. Exploring a range of each of these values can assist an investigator in selecting a sample size that is appropriate across a variety of potential model conditions. Further, the fact that level-1 error variance and number of repeated measures within a given study length of six intervals did not appear to play a substantial role in power to detect differences in slopes suggests less attention is required in estimating these parameters. Additionally, considering the plausibility of models by exploring model-implied covariance matrices and plotting data based on selected parameter values can further guide investigators in their selection of parameter estimates.

References

- Abraham, W. T., & Russell, D. W. (2008). Statistical power analysis in psychological research. *Social and Personality Psychology Compass*, 2(1), 283-301.
- Bauer, D. J. (2003). Estimating multilevel linear models as structural equation models. *Journal of Educational and Behavioral Statistics*, 28(2), 135-167.
- Bradley, J. V. (1978). Robustness? *British Journal of Mathematical and Statistical Psychology*, 31, 144-152.
- Brock, L. L., Nishida, T. K., Chiong, C., Grimm, K. J., & Rimm-Kaufman, S. E. (2008). Children's perceptions of the classroom environment and social and academic performance: A longitudinal analysis of the contribution of the Responsive Classroom approach. *Journal of School Psychology*, 46(2), 129-149.
- Bryk, A. S., & Raudenbush, S. W. (1987). Application of hierarchical linear models to assessing change. *Psychological Bulletin*, 101, 147-158.
- Buse, A. The likelihood ratio, Wald, and Lagrange multiplier tests: An Expository Note. *The American Statistician*, 36(3), 153-157.
- Chou, C., Bentler, P. M., & Pentz, M. A. (1998). Comparisons of two statistical approaches to study growth curves: The multilevel model and the latent curve analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, 5(3), 247-266. doi: 10.1080/10705519809540104
- Curran, P. J. (2003). Have multilevel models been structural equation models all along? *Multivariate Behavioral Research*, 38(4), 529-569.
- Duncan, T. E., & Duncan, S. C. (2004). An introduction to latent growth curve modeling. *Behavior Therapy*, 35, 333-363.
- Duncan, T. E., Duncan, S. C., & Li, F. (2003). Power analysis models and methods: A latent variable Framework for power estimation and analysis. In Z. Sloboda and W. J. Bukosk (Eds.), *Handbook of Drug Abuse Prevention (Handbooks of Sociology and Social Research)* (pp. 609-626). New York, NY: Kluwer Academic/Plenum Publishers.
- Duncan, T. E., Duncan, S. C., & Li, F. (2003). Power analysis models and methods: A latent variable Framework for power estimation and analysis. In Z. Sloboda and W. J. Bukosk (Eds.), *Handbook of Drug Abuse Prevention (Handbooks of Sociology and Social Research)* (pp. 609-626). New York, NY: Kluwer Academic/Plenum Publishers.

- Duncan, T. E., Duncan, S. C., Strycker, L. A., & Li, F. (2002). A latent variable framework for power estimation within intervention contexts. *Journal of Psychopathology and Behavioral Assessment*, 24(1), 1-12.
- Farkas, G. (2008). A review of Latent curve models: A structural equation perspective. *Social Forces*, 87(1), 619-621. doi: 10.1353/sof.0.0084
- Feingold, A. (2009). Effect sizes for growth-modeling analysis for controlled clinical trials in the same metric as for classical analysis. *Psychological Methods*, 14(1), 43-53.
- Hertzog, C., Lindenberger, U., Ghisletta, P., & von Oertzen, T. (2006). On the power of multivariate latent growth curve models to detect correlated change. *Psychological Methods*, 11(3), 244-252.
- Hertzog, C., von Oertzen, T., Ghisletta, P., & Lindenberger, U. (2008). Evaluating the power of latent growth curve models to detect individual differences in change. *Structural Equation Modeling*, 15, 541-563. doi: 10.1080/10705510802338983
- Jung, S., & Ahn, C. (2003). Sample size estimation for GEE method for comparing slopes in repeated measurements data. *Statistics in Medicine*, 22, 1305-1315.
- Kim, K. H. (2005). The relation among fit indexes, power, and sample size in structural equation modeling. *Structural Equation Modeling*, 12(3), 368-390.
- Li, L., & Bentler, P. M. (2011). Quantified choice of root-mean-square errors of approximation for evaluation and power analysis of small differences between structural equation models. *American Psychological Association*, 16(2), 116-126.
- MacCallum, R. C., Browne, M. W., & Cai, L. (2006). Testing differences between nested covariance structure models: Power analysis and null hypotheses. *Psychological Methods*, 11(1), 19-35. doi: 10.1037/1082-989X.11.1.19
- MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, 1(2), 130-149. RMSEA method of power analysis
- Maxwell, S. E., Kelley, K., & Rausch, J. R. (2008). Sample size planning for statistical power and accuracy in parameter estimation. *Annual Review of Psychology*, 59, 537-563.

- Mehta, P. D., & West, S. G. (2000). Putting the individual back into individual growth curves. *Psychological Methods*, 5(1), 23-43.
- Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika*, 55(1), 107-122.
- Miles, J. (2003). A framework for power analysis using a structural equation modeling procedure. *BMC Medical Research Methodology*, 3(27), 1-11. Retrieved from: <http://www.biomedcentral.com/1471-2288/3/27>
- Mroczek, D. K. (2007). The analysis of longitudinal data in personality research. In R. W. Robins, R. C. Fraley, & R. F. Krueger (Eds.). *Handbook of research methods in personality psychology* (pp. 543-556). New York, NY: Guilford Press.
- Murphy, K. R., & Myors, B. (2004). *Statistical power analysis: A simple and general model for traditional and modern hypothesis tests* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Muthén, B. (2002). Using Mplus Monte Carlo simulations in practice: A note on assessing estimation quality and power in latent variable models. *Mplus Web Notes*, 1(2). Retrieved from: <http://www.statmodel.com/download/webnotes/mc1.pdf>
- Muthén, B. O., & Curran, P. J. (1997). General longitudinal modeling of individual differences in experimental designs: A latent variable Framework for analysis and power estimation. *Psychological Methods*, 2(4), 371-402.
- Muthén, L. K., & Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling*, 9(4), 599-620.
- Möller, I., & Krahé, B. (2008). Exposure to violent video games and aggression in German adolescents: A longitudinal analysis. *Aggressive Behavior*, 35(1), 75-89.
- Öhman, L., Bergdahl, J., Nyberg, L., & Nillson, L. (2007). Longitudinal analysis of the relation between moderate long-term stress and health. *Stress and Health*, 23(2), 131-138.
- Partridge, T., & Lerner, J. V. (2007). A latent growth-curve approach to difficult temperament. *Infant and Child Development*, 16(3), 255-265.
- Peterson, C., & Dodsworth, P. (1991). A longitudinal analysis of young children's cohesion and noun specification in narratives. *Journal of Child Language*, 18, 397-415.

- Raudenbush, S. W. (2001). Toward a coherent framework for comparing trajectories of individual change. In C. M. Collins & A. G. Sayers (Eds.). *New methods for the analysis of change* (pp. 35-64). Washington, DC: American Psychological Association.
- Satorra, A., & Saris, W. E. (1985). Power of the likelihood ratio test in covariance structure analysis. *Psychometrika*, *50*(1), 83-90.
- Singer, J. D. & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York, NY: Oxford University Press.
- Tomarken, A. J. & Waller, N. G. (2005). Structural equation modeling: Strengths, limitations, and misconceptions. *Annual Review of Clinical Psychology*, *1*, 31-65. doi: 10.1146/annurev.clinpsy.1.102803.144239
- Venter, A., Maxwell, S. E., & Bolig, E. (2002). Power in randomized group comparisons: The value of adding a single intermediate time point to a traditional pretest-posttest design. *Psychological Methods*, *7* (2), 194-209.
- von Oertzen, T. (2010). Power equivalence in structural equation modeling. *British Journal of Mathematical and Statistical Psychology*, *63*, 257-272. doi: 10.1348/000711009X441021
- Wu, W. (2008). *Evaluating model fit for growth curve models in SEM and MLM frameworks* (Doctoral dissertation). Retrieved from ProQuest. (3339593).
- Yu, C. (2002). *Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes* (Doctoral dissertation: University of California). Retrieved from <http://statmodel2.com/download/Yudissertation.pdf>
- Zhang, Z., & Wang, L. (2009). Statistical power analysis for growth curve models using SAS. *Behavior Research Methods*, *41*(4), 1083-1094.

APPENDIX A

SAMPLE MPLUS SYNTAX FOR MONTE CARLO POWER ANALYSIS FOR
A LINEAR GROWTH CURVE MODEL WITH A DUMMY CODED
COVARIATE


```

TITLE:                i100s20r50es4si2ss200gs55rm6EU
MONTECARLO:          NAMES ARE y1-y6 x;
                    CUTPOINTS = x (0); !split group;
                    NOBSERVATIONS = 200; !sample size;
                    NREPS = 1000;
                    SEED = 0802;
                    SAVE =save_i100s20r50es4si2ss200gs55rm6EU.sav;
                    RESULTS = i100s20r50es4si2ss200gs55rm6EU.txt;

MODEL MONTECARLO:
    [x@0]; x@1;

    i BY y1-y6@1;
    s BY y1@0 y2@1 y3@2 y4@3 y5@4 y6@5;
    [y1-y6@0];
    [i*0 s*0];
    i*100; !intercept variance;
    s*20; !slope variance;
    i WITH s*8.944; !intercept-slope covariance;
    y1*50 y2*50 y3*50 y4*50 y5*50 y6*50; !levellresvar;
    i ON x*0;
    s ON x*1.0556; !slope difference;

MODEL:
    i BY y1-y6@1;
    s BY y1@0 y2@1 y3@2 y4@3 y5@4 y6@5;
    [y1-y6@0];
    [i*0 s*0];
    i*100;
    s*20; !slope variance;
    i WITH s*8.944; !intercept-slope covariance;
    y1*50 y2*50 y3*50 y4*50 y5*50 y6*50; !levellresvar;
    i ON x*0;
    s ON x*1.0556; !slope difference;

OUTPUT:             TECH9;

```

APPENDIX B

SAMPLE MPLUS SYNTAX FOR THE THREE STEPS IN THE SATORRA-
SARIS POWER ANALYSIS FOR A LINEAR GROWTH CURVE MODEL
WITH A DUMMY CODED COVARIATE

```

TITLE:          Satorra-Saris Step 1, Output yields model
                 implied mean and correlation matrices

DATA:          FILE IS artific.dat;
                !data in artific.dat file is as follows,
                !first row is means followed by covariances
                !0 0 0 0 0 0 0
                !1
                !0 1
                !0 0 1
                !0 0 0 1
                !0 0 0 0 1
                !0 0 0 0 0 1
                !0 0 0 0 0 0 1
                TYPE IS MEANS COVARIANCE;
                NOBSERVATIONS = 1000;

VARIABLE:     NAMES ARE y1-y6 x;

ANALYSIS:     TYPE=MEANSTRUCTURE;

MODEL:        [x@.5]; x@.25;
                i BY y1-y6@1;
                s BY y1@0 y2@1 y3@2 y4@3 y5@4 y6@5;
                [y1-y6@0];
                [i@0 s@0];
                i@100;
                s@20; !slope variance;
                i WITH s@8.944; !intercept-slope covariance
                y1@50 y2@50 y3@50 y4@50 y5@50 y6@50; !levelresvar;
                i ON x@0;
                s ON x@1.0556; !slope difference;

OUTPUT:       STANDARDIZED RESIDUAL;

```

```

TITLE:          Satorra-Saris Step 2, Enter means and
                 covariances derived from Step 1 into pop.dat
                 file, use Step 2 to ensure parameter values
                 in Step 1 are retrieved in the output for
                 Step 2

DATA:          FILE IS pop.dat;
                TYPE IS MEANS COVARIANCE;
                NOBSERVATIONS = 1000;

VARIABLE:     NAMES ARE y1-y6 x;

ANALYSIS:     TYPE=MEANSTRUCTURE;

MODEL:        i BY y1-y6@1;
                s BY y1@0 y2@1 y3@2 y4@3 y5@4 y6@5;
                [y1-y6@0];
                s on x;
                i on x;
                i with s;
                [i s];

OUTPUT:       STANDARDIZED RESIDUAL;

```

```
TITLE:          Satorra-Saris Step 3,
                Constrain parameter of interest to 0

DATA:          FILE IS pop.dat;
                TYPE IS MEANS COVARIANCE;
                NOBSERVATIONS = 200;

VARIABLE:      NAMES ARE y1-y6 x;

ANALYSIS:      TYPE=MEANSTRUCTURE;

MODEL:         i BY y1-y6@1;
                s BY y1@0 y2@1 y3@2 y4@3 y5@4 y6@5;
                [y1-y6@0];
                s on x@0; !parameter fixed to 0
                i on x;
                i with s;
                [i s];

OUTPUT:        STANDARDIZED RESIDUAL;
```

APPENDIX C

SAMPLE SAS SYNTAX TO DETERMINE POWER GIVEN A
NONCENTRALITY PARAMETER DETERMINED USING THE SATORRA-
SARIS OR MBC TECHNIQUE

```
DATA POWER;
DF=1; CRIT=3.841459;
  *DF = difference in degrees of freedom
  between full and nested model
LAMBDA=1.026;
  *LAMBDA = noncentrality parameter derived
  from Satorra-Saris or MBC
POWER=(1-(PROBCHI(CRIT,DF,LAMBDA)));
RUN;
```

APPENDIX D

SAMPLE SAS SYNTAX TO DETERMINE POWER USING THE MBC
TECHNIQUE ACROSS A VARIETY OF RMSEA VALUES FOR A NULL
HYPOTHESIS OF NO DIFFERENCE

```

Title 'Power for different increments of RMSEA values';
data one;

** Start of user input;
alpha = 0.05; *alpha level;

** Enter the lower and upper limits of RMSEAs;
rmseaAL = 0.01; *model A RMSEA lower limit;
rmseaAU = 0.10; *model A RMSEA upper limit;
rmseaBL = 0.0; *model B RMSEA lower limit;

** Enter the step size for grid points to cover RMSEA range;
stepsize = 0.01; *step size;

** Enter model information;
da = 3; *df for model A;
db = 2; *df for model B;
n = 500; *sample size;
G = 1; *number of groups;
** End of user input;

** Power computation begins here;
nA = int((rmseaAU-rmseaAL)/stepsize)+1;
rmseaA = rmseaAL-stepsize;
do i=1 to nA;
  rmseaA = rmseaA+stepsize;
  nB = int((rmseaA-stepsize-rmseaBL)/stepsize)+1;
  rmseaB = rmseaBL-stepsize;
  do j=1 to nB;
    rmseaB = rmseaB+stepsize;
    ddiff = da-db; *df difference;
    fa = (da*rmseaA**2)/sqrt(G); *discrepancy fn value for model A;
    fb = (db*rmseaB**2)/sqrt(G); *discrepancy fn value for model B;
    ncp = (n-1)*(fa-fb); *non-centrality parameter;
    cval = cinv(1-alpha,ddiff); *critical value from central chi^2;
    Power = 1-probchi(cval,ddiff,ncp); *power;
    output;
  end;
end;
run;
proc print data=one; var rmseaA rmseaB Power; run;

```