12-1-2009

Functioning of global fit statistics in latent growth curve modeling

Kathryn K. DeRoche

UNIVERSITY OF NORTHERN COLORADO

Greeley, Colorado

The Graduate School

THE FUNCTIONING OF GLOBAL FIT STATISTICS IN LATENT GROWTH CURVE MODELING

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

Kathryn K. DeRoche

College of Education and Behavioral Sciences
School of Educational Research, Leadership, and Technology
Applied Statistics and Research Methods

December, 2009

This Dissertation by: Kathryn K. DeRoche

Entitled: *The Functioning of Global Fit Statistics in Latent Growth Curve Modeling*

has been approved as meeting the requirements for the Degree of Doctor of Philosophy in College of Education and Behavioral Sciences in School of Educational Research, Leadership, and Technology, Program of Applied Statistics and Research Methods

Accepted by the Doctoral Committee

____

Susan R. Hutchinson Ph.D., Chair

____

David Gilliam Ph.D., Committee Member

____

Lisa Rue Ph.D., Committee Member

____

Marilyn Welsh Ph.D., Faculty Representative

Date of Dissertation Defense ____

Accepted by the Graduate School

____

Robbyn R. Wacker, Ph.D.
Assistant Vice President of Research and Extended Studies
Dean of the Graduate School

ABSTRACT

DeRoche, Kathryn K. *The Functioning of Global Fit Statistics in Latent Growth Curve Modeling*. Published Doctor of Philosophy dissertation, University of Northern Colorado, 2009.

Latent growth curve (LGC) modeling is emerging as a preferred method of longitudinal analysis, which uses the structural equation modeling (SEM) framework to demonstrate growth or change (Meredith & Tisak, 1990). The purpose of this dissertation was to examine the performance of commonly utilized measures of model fit in LGC modeling data environments. A Monte Carlo simulation was conducted to examine the influence of LGC modeling design characteristics (i.e., sample size, waves of data, and model complexity) on selected fit indexes (i.e., χ^2, *NNFI*, *CFI*, and *RMSEA*) estimated in correct LGC models. The *CFI* performed the best, followed by the *NNFI*, χ^2, and finally, the *RMSEA* showed the least desirable characteristics. The *RMSEA* was found to over-reject correct models (i.e., suggest poor model fit) in conditions of small to moderate sample size ($N \leq 1,000$) and few waves of data. The χ^2 over-rejected correct multivariate models with more waves of data and small sample sizes ($N = 100$). The *NNFI* over-rejected unvariate and multivariate models with small sample size ($N = 100$) and three waves of data. Six guidelines were proposed for LGC modeling researchers, including: maximizing the chance of obtaining a plausible solutions, cautioning the use of the χ^2, adopting the novel LGC modeling cutoff values, using multiple fit indexes, and assessing the within-person fit. As LGC modeling applications escalate in the social and behavioral sciences,

there is a critical need for additional research regarding LGC model fit, specifically, the sensitivity of fit indexes to relevant types of LGC model misspecification.

ACKNOWLEDGEMENTS

I am extremely grateful to the numerous individuals that have helped to guide my dissertation journey. First, I would like to share my deepest appreciation to my boyfriend/husband, whom has encouraged me throughout our lives, allowing me to achieve my educational and personal goals. I would like to express gratitude to my family, specifically my mother for being a wonderful role model and to my father for your continuous love and support. To my brother, I could not have completed my educational goals without you. Thank you for being a wonderful brother and a great friend. Furthermore, I would like to thank my dissertation chair, Dr. Susan Hutchinson, and committee members, Dr. David Gilliam, Dr. Lisa Rue, and Dr. Marilyn Welsh, for their guidance and advice, as well as Dr. Antonio Olmos who served as mentor throughout my education career. To my friends, which have spent countless hours encouraging me to complete my dissertation, thank you for your support over the years. Last but not least, to my pets who have provided me unconditional love throughout my life.

TABLE OF CONTENTS

LISTS OF TABLES

TABLES Page

LIST OF FIGURES

CHAPTER I

INTRODUCTION

In contemporary science, many researchers, practitioners, and policy makers, across a variety of disciplines, have encountered the question of how to measure change. Measurements of change in the educational field are apparent by teachers', school administrators', and federal policy makers' concerns regarding change or growth in student achievement. Researchers in the pharmaceutical industry are interested in the discovery of changes in symptom relief between members of control and experimental groups who receive a novel drug treatment. Psychologists, therapists, early education specialists, and parents are focused on the growth of developmental characteristics, including predictors or moderators of cognitive development. In addition, the field of program evaluation is committed to determining if change has occurred as a result of a program or intervention.

Along with a general interest in change, accountability of change has become apparent at the federal, state, and local levels. For example, the No Child Left Behind Act (NCLB; U.S. Department of Education, 2001) requires state governments to monitor change in students' achievement. A potential consequence of NCLB is that schools can lose federal funding and even risk termination if they fail to display adequate progress. Furthermore, most state and foundation funding agencies in the health and human services require quantitative evaluations to assess change in desired program outcomes (e.g., U.S. Department of Health and Human Services, 2005). While the implications of quantitative evaluations of change vary by state, department, discipline, and agency, in general, those programs that fail to display change, risk loss of all or part of their funding. Therefore, agencies funded through federal, state, and local venues have a fiscal responsibility to demonstrate change. The federal mandates have driven statistical and methodological researchers to develop and examine models to adequately measure change for complex traits (i.e., achievement, cognition, etc.). Typically, advances in research guide policy but, in the last decade, policy has driven the research on the methods (i.e., statistical models) to estimate change.

Traditional Models of Change

Appropriate, quantitative techniques for the measurement of change have been debated in the methodological literature for the last century, and a novel perspective of change has provided the framework for modern statistical models (Rogosa, Brandt, & Zimowaski, 1982). Historically, change has been predominantly conceptualized as a raw change score, the difference between pre- and posttest scores. However, raw change scores may produce low reliability estimates and high correlations among initial status and change scores (Cronbach & Furby, 1970).

To address the complications associated with raw change scores, several researchers advocated that the conceptualization of change should be altered from an *increment* of time (i.e., pre- to posttest) to a continuous *process* of development (Rogosa et al., 1982; Rogosa & Willett, 1983**,** 1985; Willett, 1989). Specifically, Rogosa et al. proposed that change can be measured with precision and adequate psychometric properties when more than two waves of data, or data collection points, are collected.

Furthermore, Willett demonstrated that the reliability of change can substantially increase as additional waves of data are included, with an approximate 250% increase in reliability when three waves of data are adopted as opposed to two waves of data. When change is measured as a process with more than two time points, complications of low reliability estimates, typically found with raw change scores, are no longer of concern.

In the contemporary study of change, residual change scores, repeated measures analysis of variance (ANOVA), and regression techniques can be used to statistically assess change over more than two waves of data (Field, 2005). However, these procedures have been criticized due to their limitations. Residual change scores, the difference between the residuals at two time points, were developed to avoid the high correlation between raw change scores and initial status as discussed by Cronbach and Furby (1970). Considerable debate has occurred in regard to the corresponding interpretability of residual change, and theorists (Rogosa & Willett, 1985; Willett, 1989) have advocated avoiding the use of residual change scores

Typically, behavioral and social sciences researchers apply repeated measures ANOVA and regression techniques to measure change across three or more time points (Voelkle, 2007). Based on variance decomposition, ANOVA and regression models determine group differences by partitioning variance into between-person (i.e., inter-individual differences) and within-person (i.e., intra-individual differences) variations. The between-person variance represents variations accounted for in the model (e.g., variations due to a manipulated variable between an experimental and control group) and the within-person variance represents variation not accounted for in the model (i.e., variation due to individual differences) sometimes labeled as error variance. Inherent in its name**,** error variance, the individual variations in a trait are not of interest because attention is placed on the between group differences (i.e., between-person variance). However, questions about the accountability of change may be enhanced by an examination of variation within individuals (e.g., Is the rate of change in academic growth for one student different from the rate of change for another student?). As a novel extension of the ANOVA/regression family of techniques, growth curve modeling procedures include the analysis of within and between-person variations to measure change. Therefore, subsequently, the modern movement toward the analysis of change has adopted Rogosa and Willett's (1983) notion of growth models, because their application increases the accuracy of measuring, accounting for, and interpreting individual variations across examinees.

Growth Curve Models

The legislative mandates (e.g., NCLB) of the last few decades have guided the advances in a collection of statistical models, referred to as *growth curve models*, which includes an assortment of models to investigate growth at the within and between participant levels. All growth modeling procedures can be conceptualized as two distinct steps: (a) the within-person model symbolizing individual change over time, and (b) the between-person model which characterizes inter-individual change across time (Willett & Sayer, 1994). In all growth models, the within-person changes are accounted for in the statistical analysis of between-persons changes; however, the specific statistical methods used to achieve the result vary among the different types of growth models.

The list below contains potential research questions that are commonly encountered when an analysis of change is conducted with traditional statistical approaches:

1. Has there been a change or growth in the trait?
2. Do individuals differ in their growth in the trait? Or, for multiple groups, does the control group differ in the amount of change from the experimental group?

While growth models can answer the two previous questions, also, they can answer the questions listed below (Voelkle & Wittmann, 2007):

1. What is the trajectory of change for the group? Is change related to time?
2. What is the variability of individual trajectories for change?
3. Can the individual growth trajectories be predicted?
4. Are the growth trajectories the same for multiple groups?
5. Does a covariate explain the same amount of growth in multiple groups?

This list provides a glimpse into potential questions that can be answered when growth models are applied to longitudinal research designs. Consequently, state and federal mandates frequently apply growth curve modeling techniques to demonstrate change.

Growth curve models have been referred to as: (a) *developmental models* (Rindskopf, 1987); (b) *hierarchical linear models* (HLM) *or multilevel models* (Byrk & Raudenbush, 1992); (c) *random-effects ANOVA models* (Vangel & Rukhin, 1999); (d) *random coefficients models* (Rovine & Molenaar, 2000); and (d) *latent growth curve models* (LGC model; Duncan, Duncan, & Strycker, 2007). Estimation of LGC modeling techniques includes the analysis of variance-covariance structures and latent means to determine growth or change with use of a structural equation modeling (SEM) framework (Meredith & Tisak, 1990). Duncan et al. and Schulenberg and Maggs (2001) have mathematically discussed or methodologically investigated the similarities and differences between LGC models and other growth curve models, specifically, HLM-based growth models. The preference for LGC models, compared to other growth curve models, is rooted in the methodological advantages of SEM (i.e., also known as latent variable modeling). Accordingly, LGC models represent growth in latent traits (i.e., unobservable traits), whereas other growth model designs do not distinguish between latent and observed traits. Other advantages of LGC modeling include: (a) flexibility in modeling complex phenomena, (b) ability to account for measurement error, and (c) capability of testing model fit (Byrne & Crombie, 2003; Duncan & Duncan, 2004; Meredith & Tisak, 1990; Muthén & Curran, 1997; Voelkle, 2007; Willett & Sayer, 1994). Thus, LGC modeling techniques are emerging as a preferred method of researchers to assess change among complex latent traits.

Latent Growth Curve Models

Meredith and Tisak (1990) presented an SEM model that accounted for individual changes in a trait across time, referred to as a LGC model. As a subcategory of SEM, LGC models have similar benefits to SEM in general. However, LGC models have two dominant characteristics not found in standard SEM models. First, the goal of LGC modeling is to examine growth over time within longitudinal research designs, while typical SEM applications model cross-sectional designs (Meredith & Tisak). As a result of differences in the underlying conceptualization, LGC model researchers fix the relationships between the latent and observed variables (i.e., factor loadings), whereas in

SEM methods, these relationships are estimated. While LGC modeling includes the analysis of variance-covariance structures found in general SEM applications, transformations convert this data structure to means and variances in order to interpret overall growth parameters. Thus, a second variation of LGC modeling includes the estimation of latent means as well as the variance-covariance matrix.

Hypothesized theories of change in the social and behavioral sciences typically include additional study characteristics (e.g., variations in sample size, waves of data, potential covariates, quadratic growth, and multivariate growth) to properly represent the complex changes in the latent traits of interest. Due to the flexibility in LGC modeling, researchers typically include additional constructs in the LGC model to correspond to the hypothesized theory of change. For example, a social science researcher may include additional participants to achieve adequate statistical power or may include additional waves of data to ensure that the latent trait is measured adequately during the hypothesized period of change. Furthermore, several theories in the social and behavioral sciences assume multifaceted growth, requiring complex models to properly represent traits of interest. For example, longitudinal researchers may hypothesize that quadratic growth, or initial growth that levels off, is a more appropriate representation for the trait of interest than linear growth. A LGC modeling researcher may also be interested in growth in more than one trait; therefore, requiring a multivariate LGC to properly represent the hypothesized trait. Typically, theories in the social and behavioral sciences include covariates that may influence the process of change. In conclusion, LGC models applied to traits found in the social and behavioral science will include variations in the number of participants, waves of data, and model complexity. Therefore, extensions of

LGC models are of concern to longitudinal behavioral and social science researchers. In Chapter II**,** a comprehensive description of LGC models and the similarities and differences compared to general SEM are discussed.

As in all SEM models, LGC models are used to examine the hypothesized model to determine how well it *matches* the data. As a critical component of LGC modeling, model fit determines whether the statistical model matches or fits the data collected from the participants over time. Fit indexes are a collection of descriptive and inferential statistics that represent indicators of how well the data fit the hypothesized model. Due to the debatable advantages and disadvantages of the different indexes, SEM researchers frequently report multiple fit indexes (Hu & Bentler, 1999). Assessment of model fit determines how well the hypothesized model is supported by the data. Therefore, a fundamental concept of LGC modeling is an evaluation of the model fit to properly represent change.

A review of the literature produced a few simulation studies which investigated the functioning of LGC models under various conditions (Fan, 2003; Hertzog, Lindenberger, Ghisletta, & von Oertzen, 2006; Leite, 2007; Muthén & Curran, 1997; Muthén & Muthén, 2002). Simulation studies have expanded the knowledge of optimal conditions for the application of LGC models; however, the focus of these studies has been on statistical power and assumptions, with one exception. Coffman and Millsap (2006) conducted the only other known study in which the concept of model fit was investigated in regard to LGC modeling. By examination of the two conditions of linear and quadratic growth, the authors concluded that fit indexes may not accurately represent shape in individual growth trajectories. While Coffman and Millsap's simulation provided support for their conclusions, the authors examined their hypothesis under the limited conditions of a single sample size (*N* = 500) and waves of data (three waves of data) for only two fit indices. Consequently, the authors' research design lacked representation of many common conditions found in LGC modeling applications (e.g., in terms of model complexity, waves of data, sample size, and multiple fit indexes).

The use of LGC modeling has progressed in the areas of theoretical development and simulation studies of statistical power, with a need to examine the critical procedures of assessing model fit in the LGC model. Consequently, the examination of fit indexes, under various data environments encountered in LGC modeling applications, is a novel area of exploration.

Statement of the Problem

The gap in the methodological literature in regard to fit indexes for LGC models under various conditions has inhibited applied researchers from being able to fully understand and interpret change. Longitudinal research environments include varying conditions of sample size, waves of data collected, and model complexity which have not yet been examined in terms of their corresponding effect on fit indexes. Coffman and Millsap (2006) provided a critical hypothesis in regard to procedures for the assessment of model fit in LGC modeling that needs to be examined under additional conditions. Currently, applied longitudinal researchers do not know if values of a particular fit index suggest adequate fit for LGC models.

In addition, simulation studies (Bentler, 1990; Hu & Bentler, 1999; Yadama & Panday, 1995) of general SEM applications have demonstrated that fit indexes fluctuate with sample sizes and model complexity. Currently, LGC modeling researchers reference methodological studies of fit statistics under general SEM conditions to justify their interpretation of model fit (e.g., Hu & Bentler). However, given the lack of research on fit in the LGC modeling context, it is unknown if guidelines for assessing fit in standard SEM are applicable to LGC models. Assistance could be provided to applied researchers by increasing the understanding of fit indexes for LGC modeling data environments. For example, applied LGC modeling researchers could benefit from a better understanding of how fit indexes are affected under conditions of: (a) sample size, (b) waves of data collected, and (c) model complexity. In a recent review, Voelkle (2007) noted an apparent lack of methodological guidance for applied LGC researchers specifically related to interpretation of fit indexes and stated:

> Clearly, there is a need for future research to shed light on the complex interactions between these factors (sample size, underlying assumptions, and model complexity) in order to determine the optimal procedures for the analysis of change for a given set of data. Similar arguments can be made for most fit indexes employed in LGC modeling, which are greatly affected by sample size. This topic has been deliberately ignored because it is no different from standard structural equation modeling and a more detailed discussion would go far beyond the scope of this article. (p. 411)

Rationale for the Study

Even though the LGC modeling simulation literature is scarce, applications of the procedure have escalated in the last decade, especially in the behavioral and social sciences. Even a cursory review of the recent published literature reveals numerous LGC modeling application studies across a variety of disciplines, which varied considerably in sample size, number of waves of data, and model complexity, as well as in the type of fit indexes reported. The application studies cluster around conditions of three to five waves of data and fewer than 500 participants. In the majority of studies, fit indexes are reported to examine model fit, typically including the: (a) chi-squared likelihood ratio test (χ^2); (b) non-normed fit index (*NNFI*; Bentler & Bonett, 1980); (c) comparative fit index (*CFI*; Bentler, 1990); and (d) the root mean square error of approximation (*RMSEA*; Steiger & Lind, 1980). Unfortunately, many researchers who use LGC models seem to be unaware of the influence of study characteristics on fit indexes and follow guidelines developed upon more general SEM models. Consequently, there is a need for methodological research on the performance of various fit indexes under conditions commonly encountered in LGC modeling applications.

Purpose of the Study

The purpose of this dissertation was to examine varying LGC modeling conditions on the functioning of selected model fit indexes. By application of Monte Carlo simulation techniques, I examined how fit indexes function under simulated conditions that are commonly encountered in applied LGC environments. Data were generated by replication of the conditions of LGC modeling data environments found in the social and behavioral sciences, including varying levels of: (a) overall sample size, (b) waves of data collected, and (c) model complexity. The latter represent various design characteristics of LGC models including: (a) shape of growth, (b) number of dependent variables, and (c) inclusion of a covariate. Subsequently, the simulated data conditions were examined with use of LGC modeling techniques to estimate the following fit indexes: (a) χ^2, (b) *NNFI*, (c) *CFI*, and (d) the *RMSEA*. All other parameters required for LGC modeling estimation were held constant across simulation conditions, with the explanation and discussion of these parameters included in Chapters II and III. The results of the simulation study can be used to develop more informative guidelines for applied researchers to assess the fit of their LGC models.

Research Questions and Hypotheses

The overall hypothesis suggests that the selected fit indexes vary by sample size, waves of data collected, and model complexity. The explicit research questions and hypotheses are presented below:

Q1 Do model convergence rates vary under conditions of sample size, waves of data, and model complexity?

H1 Large models (including an increase in both waves of data and model complexity) with small sample sizes will have lower convergence rates compared to parsimonious models with large sample sizes. For example, the condition of a multivariate LGC model with 3 waves of data and N=100 will have the lowest frequency of model convergences. On the contrary, a univariate linear LGC model with 6 waves of data and N = 2,500 will have all samples converge, resulting in 100% convergence rate.

Q2 Do fit indexes (i.e., χ^2, *NNFI*, *CFI*, and *RMSEA*) differ under varying conditions of sample size?

H2 Regarding the influence of sample size, it is hypothesized that all fit indexes will display a difference among sample size conditions; with fit indexes in small sample size conditions (N = 100) deteriorating and implying a lack of model fit, while fit indexes under large sample size conditions ($N \geq$ 1,000) will suggest excellent model fit. However, the magnitude of variation will fluctuate among fit indexes. The χ^2 will display a large effect size and the *NNFI*, *CFI,* and *RMSEA* will display a small effect size among sample sizes conditions.

Q3 Do fit indexes (i.e., χ^2, *NNFI*, *CFI,* and *RMSEA*) differ under varying conditions of waves of data?

H3 Due to the increase in waves of data requiring additional observed measures in the LGC modeling, all fit indexes will display a difference among the waves of data conditions with fit indexes deteriorating, suggesting inadequate model fit, with increasing waves of data. According to previous simulations, researchers suggested that the *CFI* will have a medium effect, suggesting worse fit with more waves of data, with the χ^2, *NNFI*, and

RMSEA having small effect sizes that may be negligible in the context of practical changes in fit index values.

Q4 Do fit indexes (i.e., χ^2, *NNFI*, *CFI*, and *RMSEA*) differ under varying conditions of model complexity, defined in the current dissertation as a univariate linear LGC model, quadratic LGC model, multivariate linear LGC model, and a linear LGC model with a covariate?

H4 Extensions to the parsimonious linear LGC model require additional parameters to be estimated, increasing in model complexity with the addition of a covariate, representation of quadratic growth, and the most complex multivariate linear LGC model. Based on previous studies of model fit in general SEM, it is expected that as the model complexity increases, the fit indexes will depreciate. According to previous simulations, researchers suggested that the *CFI* will have a medium effect, with the χ^2, *NNFI*, and *RMSEA* having small effect sizes that may be negligible in the context of practical changes in fit index values.

The four fit indexes were investigated separately for research questions two through four. Support for the research questions and corresponding hypotheses is provided in Chapter II.

Limitations

In this dissertation, I replicated common application scenarios of LGC models, but did not attempt to simulate data specific to every possible scenario. For example, in many applications of LGC models, there may be: (a) missing data, (b) varying degrees of measurement error, (c) non-normality, (d) different variances within each measurement point, and (e) assumption violations; all of which were held constant in the current study. This dissertation was designed to simulate conditions of typical data environments in the social and behavioral sciences and, therefore, may not reflect all fields and applications of study. Furthermore, only selected fit indexes were examined despite the myriad fit indexes that have been developed for use in SEM with rationale and justification discussed in Chapter II. As a result, the external validity of the study is limited to the conditions which are explored in the current investigation and the results may not apply universally to all LGC modeling applications.

Conclusion

This study has built on the methodological research in regard to LGC models and enhanced the understanding of the functioning of fit indexes used to assess plausibility of tested models. The results from the simulation study have illuminated how fit indexes function under: (a) various sample sizes, (b) waves of data, and (c) model complexity. By an increased understanding of the influence of longitudinal design characteristics on model fit indexes, guidelines are provided for the applied LGC modeling researcher to assist her or him in interpretation of model fit. Fit indexes are utilized in SEM applications to support hypothesis testing (e.g., how well the data match the hypothesized model); therefore; guidelines specific to LGC modeling model fit may increase the rigor of hypothesis testing in applied longitudinal research.

The dissertation is organized to convey the study of fit indexes in LGC modeling through five chapters. Chapter I established the background for the current political and scientific study of change, the advantages of LGC modeling, and the need for additional studies investigating model fit in LGC models. Chapter II further explains the concepts of LGC modeling, continuing to review LGC model simulation studies, as well as simulation studies of fit indexes in general SEM. The methodological procedures are presented in Chapter III, including a discussion of the independent and dependent variables, in addition to methods for data generation and analysis. Furthermore, Chapter IV explains the results of varying LGC modeling design characteristics on the four measures of fit. In Chapter V, I discuss the implications for the findings providing six suggestions to improve the validity of LGC assessments of model fit. Finally, the appendix contains supplemental information for the study, including tables to present the results, as well as program syntax for data analysis procedures. The outline of the dissertation provides a clear description of the need, methodology, results, and interpretation of an investigation into model fit in LGC models.

CHAPTER II

REVIEW OF LITERATURE

The purpose of this dissertation is to assist applied LGC modeling researchers in their assessment of model fit by examining the functioning of selected model fit indexes under simulated LGC modeling conditions commonly found in application. Provided in Chapter II is a summary of the current literature related to the evaluation of model fit for latent growth curve (LGC) models and is divided into two sections, including: (a) a description of LGC models and relevant literature, and (b) a description of selected fit indexes and relevant literature. The chapter begins with a discussion of structural equation models (SEM) to describe the general family of statistical procedures in which LGC models are a subcategory. Next, a detailed description of LGC models is presented to demonstrate the procedures undertaken to estimate necessary parameters, along with a brief review of conditions found in LGC modeling applications and simulation studies. Then, I focus on fit indexes utilized in LGC model applications, and the computation of fit indexes used in this dissertation. To establish the foundation for procedures and hypotheses, simulation studies of fit indexes in general SEM are discussed. Finally, Chapter II is concluded by a discussion of the only prior LGC modeling simulation study to investigate fit indexes and the chapter summary. The information presented in Chapter II provides the framework for the methods proposed in Chapter III.

Structural Equation Modeling

Structural equation modeling (SEM) is a collection of statistical techniques commonly referred to as latent variable modeling. Latent variables are defined as unobservable variables that we, as human beings, have constructed. Accordingly, LGC models represent growth in latent traits, whereas other growth model designs do not distinguish between latent and observed traits. The roots of SEM are founded in the social and behavioral sciences, with applications in the majority of journals related to human behavior (Kline, 2005). Due to the predominant presence of latent traits within social and behavioral science (e.g., depression, self-esteem, and substance abuse recovery), it is understandable that, frequently, researchers in these fields apply SEM to capture the abstract phenomena they commonly encounter.

Extended from regression procedures, SEM is a family of statistical techniques that allows for various applications, which include: (a) path modeling, (b) confirmatory factor analysis (CFA), (c) structural covariance analysis, and (d) LGC modeling among others. Applications can be exploratory or confirmatory, applicable to all types of designs, and include predictors and covariate effects. Therefore, SEM applications have two general goals: (a) to understand the relationships among a collection of variables, and (b) to explain variability in a model based on a theoretical rationale (Kline, 2005). While these goals are achievable through other statistical venues, SEM is a flexible procedure that incorporates measurement error in the model, whereas all other procedures assume perfect measurement (e.g., regression, ANOVA, and time-series analysis). It is unrealistic to assume that the types of variables often studied in the social and behavioral sciences will include perfect measurement because of the abstract nature of the latent traits of interest.

Applications of SEM are partitioned into two components: (a) a measurement model (i.e., same as a CFA) and (b) a structural model. The measurement model identifies the relationships between the observed variables and latent variables (i.e., factor loadings), whereas the structural model displays the directional relationships among the latent variables. Unique to SEM application is the specification of relationships in the measurement and structural models, based on theory and previous research to define the relationships. By specification of the model, matrix equations are derived, which represent the theoretical relations among and between the observed and latent variables (Jöreskog & Sörbom, 2001). In addition to matrix equations, SEM applications can jointly display the measurement and structural model in diagram form. Both venues of the communication of SEM models (i.e., matrix notation and diagram) will be discussed specific to LGC models in the latter portion of this chapter.

The foundations of SEM procedures are focused on the analysis of variance-covariance matrices (e.g., also referred to as the unstandardized correlation matrix), which assess the strength of the relationships among two or more variables. Two types of variance-covariance matrices are at the forefront of SEM parameter estimation and model fit, specifically the observed variance-covariance matrix (Σ) and the model implied variance-covariance matrix ($\Sigma(\theta)$). The Σ represents the relationships (covariance) between all observed variables and the variance on each observed variable, whereas $\Sigma(\theta)$ is a variance-covariance matrix that is explained by the specified model (Jöreskog & Sörbom, 2001). Unique to SEM applications is the $\Sigma(\theta)$, which is an a priori hypothesized model specified by the researcher, of the relationships among and between the observed and latent variables. The framework of SEM is rooted in an examination of how well the $\Sigma(\theta)$ *matches* or accounts for the relationships in the Σ. Essentially, a SEM analysis determines how well the researchers' hypothesized model ($\Sigma(\theta)$), accounts for the relationships found in the data (Σ). The comparison of the $\Sigma(\theta)$ to the Σ can be conducted by use of a variety of estimation procedures (e.g., unweighted least squares, generalized least squares, generally weighted least squares, etc.); however, the focus of this dissertation is on maximum likelihood (ML) estimation techniques because of their frequent use in application studies. A ML fitting function is used to estimate the $\Sigma(\theta)$ by minimization of the discrepancies between the $\Sigma(\theta)$ and the Σ (Jöreskog & Sörbom). A detailed discussion of parameter estimation procedures is provided specific to LGC modeling techniques in the following section.

In summary, SEM applications have the benefit of flexible modeling and estimation of measurement error, through the use of measurement and structural models where the observed variance-covariance matrix is compared to the model implied variance-covariance matrix to examine the plausibility of the model. As a type of SEM, LGC models include identical components of estimation of the variance-covariance structure; however, LGC procedures include the estimation of latent means and are exclusively focused on longitudinal change in traits across time.

Latent Growth Curve Models

A basic LGC model can be considered a special case of a CFA. The most parsimonious LGC model is a two factor CFA with three indictor variables, meaning there are two latent factors measured by three observed variables total, across both latent factors. In a typical SEM application, a parsimonious CFA model with three observed variables could only be estimated for one latent trait, not two latent traits as represented in the LGC model. The underlying differences are due to issues of identification where LGC models fix or free (i.e., estimate) different variables than in a standard SEM application. The topic of LGC model identification is described in detail in the latter portion of this chapter.

Standard SEM procedures for estimation of the variance-covariance parameters can be described in CFA matrix notation, as displayed in Equation 2.1.

$$y = \Lambda \eta + \varepsilon \tag{2.1}$$

Where, y is a vector of the observed measures for each time point, Λ is matrix of fixed factor loadings to represent time, η is vector of latent factors, and ε is a vector of the residuals (Bollen & Curran, 2006). For each observed trait, a LGC model includes two latent factors to represent the trajectory of growth, including: (a) a latent intercept (η_{i1})**,** which represents the initial level in the trait at baseline; and (b) a latent slope (η_{i2}), which indicates growth in the trait over the specified time period (Duncan et al., 2007). The individual growth trajectories for each *i*th participant are estimated based on the vector of a latent factor (η).

A basic LGC model with three time points is represented in expanded matrix notation in Equation 2.2.

$$\begin{pmatrix} y_{i1} \\ y_{i2} \\ y_{i3} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} \eta_{i1} \\ \eta_{i2} \end{pmatrix} + \begin{pmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \\ \varepsilon_{i3} \end{pmatrix} \tag{2.2}$$

The observed measures (i.e., y_{i1} is time point 1, y_{i2} is time point 2, y_{i3} is time point 3 for each *i*th individual) are treated as indicators of the latent intercept and latent slope (i.e., characteristics of the growth trajectory). To represent the anticipated change in the trait of interest, LGC modeling researchers fix the factor loadings depending on the time intervals of data collection, which differs from the standard SEM procedures where the majority of factor loadings are estimated (Duncan et al., 2007). The factor loadings (λ_{11}, λ_{12}, λ_{13}) from the latent intercept (η_{i1}) to the observed variables (y_{it}) are fixed to a value of 1.0, which represents the equal influence of all observed measurement points to the latent intercept. The factor loadings from the latent slope (η_{i2}) to the observed variables (y_{it}) are typically fixed with linear trend contrasts to represent the coding of time (λ_{21}, λ_{22}, λ_{23} in general SEM; λ_t, in LGC modeling, where *t* represents the time point; Duncan et al.). A typical LGC model application, with equal intervals of data collection, would fix $\lambda_t = 0$, 1, 2. . . $n - 1$ to represent baseline, time point two, and time point three, respectively. Time can be coded with the use of alternative procedures; however, in this study standard polynomial coding was applied because of its frequent use in applications. Equations 2.1 and 2.2 are occasionally referred to as the *level one* model, similar to concepts found in other growth curve models (e.g., HLM).

To assist in the explanation of the underpinning of LGC modeling, I applied the example of the measurement of student achievement with a basic linear LGC model over three time points (i.e., Grades 9, 10, and 11). For example, a researcher interested in the trajectory of growth in student achievement between Grades 9 to 11 could apply a univariate linear LGC model to estimate the linear growth trajectory (including the initial level and rate of growth) for each student and for the entire sample. Equation 2.2 can be explained in terms of the student achievement example where the achievement scores from each *i*th participant for the three time points of Grade 9 (y_{i1}), Grade 10 (y_{i2}), and Grade 11 (y_{i3}) are a function of: (a) the fixed factor loadings (e.g., 1 in the first column of the factor loading matrix and the λ_t = 0, 1, and 2 represent the equal interval time periods from Grades 9-11 in the second column of the factor loading matrix); (b) the latent estimate of the initial level of achievement in Grade 9 (η_{i1}) and latent growth from Grades 9-11 grade (η_{i2}); and (c) residuals associated with the achievement scores in Grades 9 (ε_{i1}), 10 (ε_{i2}), and 11 (ε_{i3}) for each *i*th participant. Henceforth, the *i*th notation, which represents each participant's individual growth trajectory**,** will not be included until the discussion of latent means. In other words, an applied researcher enters the known information of three observed measure of academic achievement from Grades 9 to 11 and the coding of time into the LGC model equation. From this information, the errors associated with each observed measure estimated, along with the initial level of academic achievement in 9th grade and rate of growth in achievement over Grades 9 to 11. In summary, the univariate linear LGC model of student achievement assumes that the observed measures of student achievement are a function of the coding of time, the growth trajectory, and errors in the observed measures of achievement.

Commonly, both SEMs and LGC models are described in diagram format due to its ease in interpretation compared to matrix notation. The student achievement example is displayed in Figure 2.1. For audiences not familiar with SEM diagram notation, rectangles represent observed variables (e.g., test scores), circles imply latent variables, the triangle is a constant term, the single headed arrow suggests the direction of the relationship from one variable to another, and the double headed arrow implies the covariance between two variables (Kline, 2005). The diagram notations correspond to the description of LGC models presented in the following three pages.

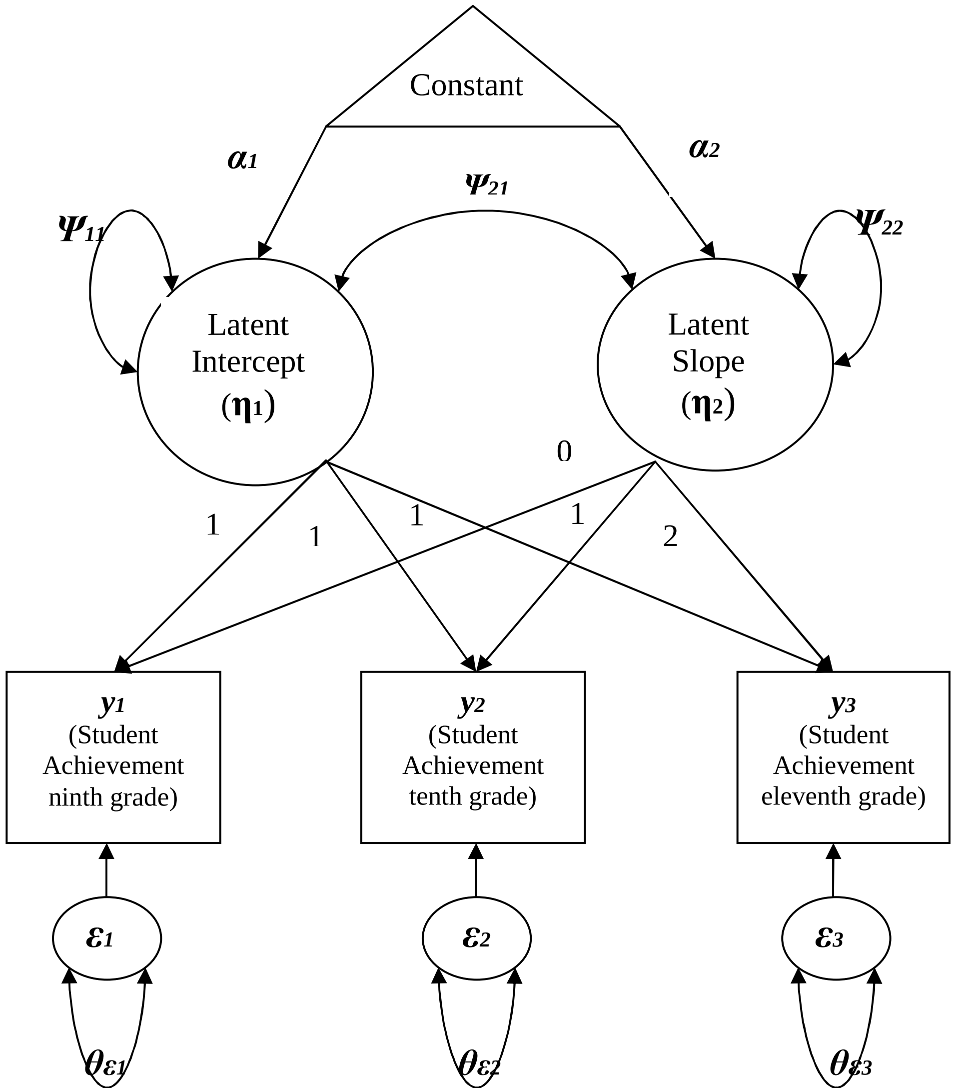

Figure 2.1. Univariate linear LGC model of student achievement

As displayed in Figure 2.1, η_1, the latent intercept, represents initial level of student achievement in Grade 9, and η_2, the latent slope, indicates the rate of growth in student achievement from Grades 9-11. The observed measures of student achievement in Grades 9, 10, and 11 are denoted as y_1 , y_2, and y_3, respectively. Observed measures of student achievement for Grades 9, 10, and 11 are indicators of the latent constructs of initial level of achievement in Grade 9, and the rate of growth in achievement from Grades 9-11. The λ_t are fixed with linear trend contrasts of 0, 1, and 2 to represent the coding of time for Grades 9, 10, and 11, with the factor loadings from the latent intercept to the latent slope fixed to the value of 1.0.

As in all SEM models, the symbols ε_1, ε_2, and ε_3 represent the random errors of measurement associated with each observed variable, and $\theta_{\varepsilon 1}$, $\theta_{\varepsilon 2}$, and $\theta_{\varepsilon 3}$ signify the residual variances for each of the random errors (Kline, 2005). Also, the variance and co-variation among the latent variables are estimated with use of traditional SEM parameters; Ψ_{11} symbolizes the variances of the latent intercept, η_1, Ψ_{22} stands for the variance of the latent slope, η_2, and Ψ_{12} represents the covariance between the η_1 and η_2 factors (Bollen & Curran, 2006). In the achievement example, the ε_1, ε_2, and ε_3 are the random errors of the observed achievement scores, the $\theta_{\varepsilon 1}$, $\theta_{\varepsilon 1}$, and $\theta_{\varepsilon 1}$ are the variance estimates of the errors of observed achievement scores in Grades 9, 10, and 11; Ψ_{11} is the variance estimate of the initial level of achievement in Grade 9, Ψ_{22} is the variance estimate of rate of change in achievement from Grades 9-11, and Ψ_{12} is the covariance between the initial level of achievement in Grade 9, and the rate of change from Grades 9-11. In the student achievement example, variances on the latent intercept would indicate that students differed on their initial achievement scores in Grade 9 and variances on the latent slopes would indicate that they grew at different rates from Grades 9 - 11; therefore, some students might improve on achievement at a faster rate across time than others.

In addition to the standard parameters of SEM described above, LGC models include estimation of a vector of latent means (α) to represent the trajectory of growth. The means are estimated by "fixing the intercept of the repeated measures to zero" (Bollen & Curran, 2006, p. 36). By inclusion of the constant of zero in Figure 2.1, the η_1 and η_2 factors can be expressed as a function of the α (latent means) and individual deviations away from the latent mean (ζ, also known as the disturbance), as displayed in Equation 2.3 (Bollen & Curran).

$$\eta = \alpha + \zeta \tag{2.3}$$

In LGC modeling, ζ have a mean of zero, variance and covariance corresponding to the variance and covariance of the latent trait, and are assumed to be uncorrelated with the residuals of the observed variables (e.g., covariance of ζ and ε_{it} is equal to zero) (Bollen & Curran). Drawing from HLM concepts, Equation 2.3 is referred to as the *level two* model. In the achievement example, Equation 2.3 can be written in expanded form to represent the growth trajectory, including the estimation of the initial level of achievement in Grade 9 (α_{i1}) and rate of growth in achievement from Grades 9-11 (α_{i2}) for each *i*th individual in the sample. The equations specific to each latent variable in the achievement example are displayed in Equations 2.4 and 2.5.

$$\eta_{i1} = \alpha_{i1} + \zeta_{i1} \tag{2.4}$$
$$\eta_{i2} = \alpha_{i2} + \zeta_{i2} \tag{2.5}$$

Estimation of the initial level of achievement in ninth grade (η_{i1}) is a function of: (a) the latent mean of the initial level of achievement in ninth grade (α_{i1}) and (b) some deviations away from the average achievement level in Grade 9 (ζ_{i1}). Similarly, estimation of the growth in achievement from Grades 9-11 (η_{i2}) is a function of: (a) the latent mean of the growth in achievement (α_{i2}) and (b) some deviations away from the average growth in achievement from Grades 9-11 (ζ_{i2}). The ζ_{i1} and ζ_{i2} have a mean of zero, variance of Ψ_{11} and Ψ_{22}, a covariance of Ψ_{12}, and are assumed to be uncorrelated with ε_{it}.

The estimation of the latent means and deviations can be expressed within the standard matrix notation of a CFA by the substitution of Equation 2.3 into Equation 2.1, as displayed in Equation 2.6, where α_η is the estimated latent means for each η (Bollen & Curran, 2006).

$$y = \Lambda(\alpha_\eta + \zeta) + \varepsilon \tag{2.6}$$

Consequently, the diagram notation presented in Figure 2.1 corresponds to Equation 2.6, with joint representation of the variance-covariance matrices and the latent means (e.g., includes both level one and two equations).

Parameter Estimation

The additional estimation of latent means in LGC modeling requires a fundamental addition to standard CFA estimation procedures. As discussed in the previous section on SEM, the central concept of parameter estimation and model fit is to examine how well the model implied covariance matrix ($\Sigma(\theta)$) can reproduce the observed covariance matrix (Σ). The LGC models are based on testing the same null hypothesis, shown in Equation 2.7, that is tested in standard SEM.

$$\Sigma = \Sigma(\theta) \tag{2.7}$$

For the example of student achievement with three time points, Equation 2.7 can be written in expanded matrix notation, as displayed below in Equation 2.8 (Bollen & Curran, 2006).

$$\begin{pmatrix} VAR(y_1) & COV(y_1,y_2) & COV(y_1,y_3) \\ COV(y_2,y_1) & VAR(y_2) & COV(y_2,y_3) \\ COV(y_3,y_1) & COV(y_3,y_2) & VAR(y_3) \end{pmatrix} = \tag{2.8}$$

$$\begin{pmatrix} \psi_{11}+VAR(\varepsilon_1) & \psi_{11}+\psi_{12} & \psi_{11}+2\psi_{12} \\ \psi_{11}+\psi_{12} & \psi_{11}+\psi_{22}+2\psi_{12}+VAR(\varepsilon_2) & \psi_{11}+2\psi_{22}+3\psi_{12} \\ \psi_{11}+2\psi_{12} & \psi_{11}+2\psi_{22}+3\psi_{12} & \psi_{11}+4\psi_{22}+4\psi_{12}+VAR(\varepsilon_3) \end{pmatrix}$$

The first matrix is the observed covariance matrix where *VAR* is the variance of each time point (i.e., variance in achievement scores for Grades 9, 10, and 11), the *COV* is the covariance between two time points (e.g., *COV* (*y1, y2*), the covariance of achievement scores between Grades 9 and 10 grade). The latter is the model implied matrix that matches the variances and covariances described in Figure 2.1 for the sample. Thus, the objective is to specify a model so that a model implied variance-covariance matrix reproduces, or is similar to, the observed covariance matrix. According to the student achievement example, the goal is to specify a model of growth in student achievement so that the hypothesized relationships in student achievement (model-implied variance-covariance matrix) reproduce the relationships found in the data (observed variance-covariance matrix).

In addition to estimation of the variance-covariance matrices, LGC models are used to examine the latent means of the growth trajectory, where the model-implied mean vector, $\mu(\theta)$, is estimated to determine how closely it reproduces the observed mean vector, μ, as displayed in Equation 2.9.

$$\mu = \mu(\theta) \tag{2.9}$$

According to the student achievement example with three time points, Equation 2.9 would be expressed as two mean vectors, shown in Equation 2.10.

$$\begin{pmatrix} \mu_{y1} \\ \mu_{y2} \\ \mu_{y3} \end{pmatrix} = \begin{pmatrix} \mu_{\alpha 1} \\ \mu_{\alpha 1} + \mu_{\alpha 2} \\ \mu_{\alpha 1} + 2\mu_{\alpha 2} \end{pmatrix} \tag{2.10}$$

The first vector consists of the estimated mean values on achievement for each time point (i.e., Grades 9 (μ_{y1}), 10 (μ_{y2}), and 11 (μ_{y3})). The vector on the right side of the equals sign is the model-implied mean vector which includes the means of growth trajectory, where $\mu_{\alpha 1}$ is the population mean of the latent intercept, and $\mu_{\alpha 2}$ is the population mean of the latent slope (Bollen & Curran, 2006). According to the hypothesized achievement example: (a) the estimated mean achievement level in Grade 9 (μ_{y1}) is equal to the population mean of the initial level of achievement in Grade 9 ($\mu_{\alpha 1}$); (b) the estimated mean achievement level in Grade 10 (μ_{y2}) is equal to the population means of the initial level of achievement in Grade 9 ($\mu_{\alpha 1}$) plus the rate of growth in achievement from Grades 9 to 11 ($\mu_{\alpha 2}$); and (c) the estimated achievement level in Grade 11(μ_{y3}) is equal to the population mean of the initial level of achievement in Grade 9 ($\mu_{\alpha 1}$) plus two times the rate of growth in achievement from Grades 9 to 11($\mu_{\alpha 2}$). In Equations 2.8 and 2.10, the parameters estimates are substituted in place of the population parameters to estimate the model implied covariance and mean vectors (e.g., $\mu_{\alpha 1} = \alpha_1$). For example, the variances of achievement in Grades 9, 10, and 11, the covariance in achievement among Grades 9 to 11, and the latent means in Grades 9 to 11 are substituted in place of the population parameter of student achievement to allow for estimation of the model-implied variance, covariance, and latent mean parameters.

The implied variance-covariance matrices and mean vectors, found in Equations 2.7 and 2.9, are critical to answering longitudinal research questions. The ML estimation procedures are used to estimate the values of the model-implied variance-covariance matrices and mean vectors to determine how well they reproduce the observed sample values. As previously discussed, the population means (μ) and variance-covariance matrix (Σ) are unobtainable; therefore, the sample values of the means ($\bar{y}$) and covariances (S) are used in the ML fitting function. For ML estimation, the latent means (α_1 and α_2), the variances of the latent means (Ψ_{11} and Ψ_{22}), and the covariance between the latent means (Ψ_{12}) are jointly denoted by θ, and "the goal is to choose values of θ to make $\mu(\theta)$ close to $\bar{y}$ and $\Sigma(\theta)$ close to S" (Bollen & Curran, 2006, p. 41). In Equation 2.11, the ML fitting function for LGC model parameter estimation is described.

$$F_{ML} = \ln|\Sigma(\theta)| - \ln|S| + tr\left[\Sigma^{-1}(\theta)S\right] - p - \left[\bar{y} - \mu(\theta)\right]' \Sigma^{-1}(\theta)\left[\bar{y} - \mu(\theta)\right] \tag{2.11}$$

A F_{ML} value can range from zero to infinity, with a value of zero indicating that the model implied variance-covariance matrix $\Sigma(\theta)$ is the same as the sample covariance matrix (S), and the model-implied mean vector ($\mu(\theta)$ is identical to the sample values of the means ($\bar{y}$) (Bollen & Curran). When a model has zero degrees of freedom (df), the F_{ML} value will equal zero because the model implied parameters are identical to the sample parameters. In such case, the model would fit the data perfectly, which is discussed in detail in the model identification section. However, when a model has $df \geq 1$, a F_{ML} value equal to zero occurs only theoretically because an exact match between the model implied and observed variance-covariance matrix and mean vectors does not occur in applications. As discrepancies between the model-implied parameters ($\Sigma(\theta)$ and $\mu(\theta)$) and the observed parameters (S and $\bar{y}$) increase, the F_{ML} value simultaneously increases;

therefore, large F_{ML} values are not desirable as they suggest the researcher's model does not fit the data.

In LGC modeling, the F_{ML} value is a measure of the discrepancy between the observed and model-implied variance-covariance matrices and latent mean vectors. Therefore, the F_{ML} values are utilized in the computation of fit indexes. Compared to standard model fit estimation procedures in SEM that usually evaluate only the variance-covariance matrices, evaluation of LGC model fit includes the evaluation of discrepancies in latent means as well. The additional estimation of the latent mean vectors in the F_{ML} may influence the interpretation and standard cutoff values for fit indexes that are used to assess poor vs. adequate model fit in general SEM applications. A comprehensive discussion of fit indexes and the influences of the estimation of latent means is presented in the second portion of Chapter II.

The estimation of the $\Sigma(\theta)$ and $\mu(\theta)$ parameters is used to answer a variety of research questions related to change or growth in a latent trait. The $\mu(\theta)$ values describe the overall expression of the growth trajectories, which answers two questions in regard to the achievement example: (a) What is the average initial level of achievement in Grade 9? and (b) What is the average rate of growth in achievement between Grades 9 and 11? (similar to the standardized betas in multiple regression). The $\Sigma(\theta)$ matrix includes the variances and covariances of the latent factors and observed variables. The variances reflect the amount of intra-individual variation in the initial status and rate of growth. The covariance between the latent means measures the relationship between initial status and the rate of growth for the selected time period. Confidence intervals are computed for the variance and covariance estimations, and can be computed for the latent means**,** to answer three research questions in terms of the achievement example: (a) Do students significantly differ in their initial level of achievement in Grade 9? (b) Do students significantly differ in their rate of growth in achievement from Grades 9-11? and, (c) Is there a significant relationship between the achievement in Grade 9 and the rate of growth in achievement from Grades 9-11? In addition, confidence intervals are computed for the residual variance estimates to answer the question, Is there significant variability that is unexplained in the repeated measure of achievement from Grades 9 to 11? If significant variability is found in the residual variances, then covariates and/or predictors variables should be examined. In the student achievement example, parental involvement could be considered a predictor and account for a significant amount of variability in the trajectory of student achievement, which includes both the initial status and rate of growth.

Statistical Assumptions

When the two families of statistical techniques are merged, LGC models display the benefits of both general SEM and growth curve models. Accompanying these benefits are assumptions or required conditions that allow for proper interpretation of longitudinal change within the LGC modeling framework. Since LGC modeling is a novel field, methodologists are rapidly discovering new advances to address what were previously defined as assumptions or restrictions (Preacher, Wichman, MaCallum, & Briggs, 2008). Thus, the specific assumptions associated with LGC models are highly dependent upon the date of publication of the source, and in some cases, the specific author(s) of the source. The following section describes the assumptions imposed in the majority of LGC modeling applications and simulations, and these were the assumptions for the current research study.

First and foremost, LGC modeling researchers must assume that the latent trait of interest is theoretically assumed to change over the time period of measurement. For example, theoretically, an intelligence quotient is not related to the passage of time and would not be hypothesized to change from Grades 9-11; thus, a LGC model would not be appropriate to measure intelligence during those years. Secondly, a minimum of three waves of data are required to estimate model parameters and model fit due to issues of model identification, which are discussed in a later section of Chapter II.

Furthermore, a third collection of assumptions is related to the residuals and disturbance terms, which differs from standard SEM assumptions. First, the means of the residuals and disturbance terms are fixed to zero within each participant, and for the residuals, the means are also fixed to zero at each time point ($\varepsilon_{it} = 0$, for $i = 1, 2, \ldots N$, $t = 1, 2, \ldots t$; $\zeta_{i1} = 0$ for $i = 1, 2, \ldots N$; and $\zeta_{i2} = 0$ for $i = 1, 2, \ldots N$; Bollen & Curran, 2006; Preacher et al., 2008). According to the student achievement example, if it were possible to observe multiple measures of the latent trait of student achievement in Grade 9 (i.e., or Grade 10, or Grade 11), the researcher would assume that the average disturbances (i.e., random measurement error) of the achievement scores in Grade 9 would be equal to zero across the multiple measurement. In other words, the deviations away from the latent intercept and latent slope factors in Grade 9 will average to zero if multiple measurements were taken at the same time point. Accordingly, the covariances between the residuals and disturbance terms are fixed to zero within and between waves of data for each participant, assuming there is no relationship among and between the disturbance terms and the residual errors ($COV(\varepsilon_{it}, \zeta_{i1}) = 0$; $COV(\varepsilon_{it}, \zeta_{i2}) = 0$; $COV(\zeta_{i1}, \zeta_{i2}) = 0$; $COV(\varepsilon_{it}, \varepsilon_{it}) = 0$ (Bollen & Curran; Preacher et al.). For instance, the errors in the student achievement scores are assumed to be unrelated within each participant and between the measurement points in Grades 9, 10, and 11. Likewise, all co-variations between the residual terms and the disturbances of the latent factors (e.g., the latent intercepts and slope factors) are fixed to zero for each individual, which implies no relationship between the errors in the observed measurements of student achievement and the deviations away from the initial status of student achievement in Grade 9, and the rate of growth in achievement from Grades 9-11. The assumptions placed on the residual and disturbance terms are related to model identification, or the ability to estimate growth parameters and model fit, as discussed later in Chapter II. However, applied researchers may modify the residual and disturbance assumptions in application depending on the specific LGC model examined.

Finally, the computation of the F_{ML} fitting function, used to estimate the desired model parameters, requires the necessary assumption of multivariate normality (Bollen & Curran, 2006). To summarize, the customary assumptions associated with LGC modeling include: (a) theoretical support, (b) a minimum of three waves of data, (c) restriction imposed on the residual terms, and (d) multivariate normality. However, depending on the author and year of publication, LGC modeling assumptions may vary.

Characteristics of Latent Growth Curve Models

Thus far, a parsimonious linear LGC model has been discussed, specifically, a CFA model with two latent factors and three observed variables. However, applications utilize more complex models in comparison to the basic model of student achievement discussed earlier. Typically, social and behavioral sciences applications include variations in sample sizes, and may require additional waves of data to capture the hypothesized time period of change. Although sample size is not a model characteristic per say, variations in sample sizes effect model estimation; therefore, decisions regarding sample sizes are critical to applied longitudinal researcher and are considered modeling characteristics in this dissertation. In addition, theories of change commonly anticipate non-linear growth, multivariate growth, and inclusion of covariate(s). In this dissertation, three design features of LGC models are explored as independent variables, which include variations in**:** (a) sample size, (b) waves of data, and (c) model complexity. The following sections describe the three characteristics of LGC models examined, as well as descriptions of previous LGC modeling simulation studies and data conditions found in LGC modeling applications. To understand the methodological conditions utilized in applied studies, a synthesis of published LGC modeling articles (i.e., from 2006-2008) was conducted based on a total of 29 application studies[1].

Sample Size

Depending on resources, a LGC modeling researcher may encounter large variations in sample sizes. For example, state-wide educational assessments include large sample sizes, while a pilot study of growth in adolescents' psychological development may include a small sample size. Therefore, depending on the latent trait of interest and associated resources, LGC modeling researchers may investigate designs varying in sample size.

As in all SEM models, varying the sample size changes the participants (i.e., *i*) at each time point, but does not directly alter the structure of the LGC model. For example, the diagram presented in Figure 2.1, corresponding to Equation 2.6, does not change with the exception of more participants (i.e., *i*) included in the model. The current study simulated LGC models with no missing data; therefore, all time points had the same number of observations. In the LGC modeling applications reviewed, sample sizes ranged from 65 participants (Hardy & Thiels, 2007) to 3,602 participants (Grimm, 2007), with a median of 356 participants (M = 690.75, SD = 911.70). Thus far, LGC methodologists have focused on investigations of the statistical power of the LGC model, in order to provide methodological guidelines of sample sizes, which are necessary to obtain adequate statistical power. To briefly summarize, LGC models under ideal conditions can produce adequate power with a relatively small sample size (i.e., 100-200; Muthén & Curran, 1997; Muthén & Muthén, 2002). However, conditions found in application studies include attrition, reduced reliability, inclusion of covariates, and missing data, which require larger sample sizes to adequately detect group differences and test parameter estimates (i.e., N > 500 or N > 1,000; Fan, 2003; Hertzog et al., 2006; Muthén & Muthén). Concisely, LGC modeling applications include a large range of sample sizes, with methodological findings to suggest adequate statistical power with as few as 100 participants for a parsimonious model and as large as 1,000 for more complex models.

Waves of Data

As discussed in Chapter I, additional waves of data increase the reliability of the LGC model; however, additional waves require additional resources (e.g., time and money for data collection). Furthermore, waves of data define the time period of interest in which growth parameters are estimated and should correspond to the research question and/or theory being investigated. For example, a researcher may be interested in growth in student achievement from K-12 and could measure achievement with 13 waves of data corresponding to kindergarten and Grades 1 through 12. However, another researcher may only be interested in growth in student achievement in high school; thus, the corresponding LGC model would only include Grades 9 – 12. Therefore, the waves of data represented in a LGC model vary to be consistent with the research question(s) as well as available resources. As a result, specification of the number of waves of data in a LGC model is a critical decision for applied researchers.

As the waves of data, or time points, are added to a LGC model, an additional observed variable (y_{ti}) is included for each time point (t = 0, 1, 2,… n) with additional fixed paths from the latent factors to the observed variables (λ_1 and λ_t). In the achievement example, the research question(s) could be adapted to learn about the growth in achievement from Grade 9 to freshman year of college. The LGC model would then include five waves of data, as displayed in Figure 2.2.

Compared to the more parsimonious univariate linear LGC model with three waves of data presented in Figure 2.1, the model with five waves of data includes the two additional observed variables (y_4 and y_5), which represent achievement in Grade 12 and freshman year of college, respectively. Also, the model estimates four additional parameters, including: two residuals of the observed measurement (ε_4, residual of achievement in Grade 12 and ε_5, residual of achievement freshman year of colleges), and two variances of the residual (θ_4, variance of residual of achievement in Grade 12 and θ_5, the variance of residual of achievement freshman year of college). Therefore, additional waves of data in a LGC model increase the model complexity by including additional observed variables.

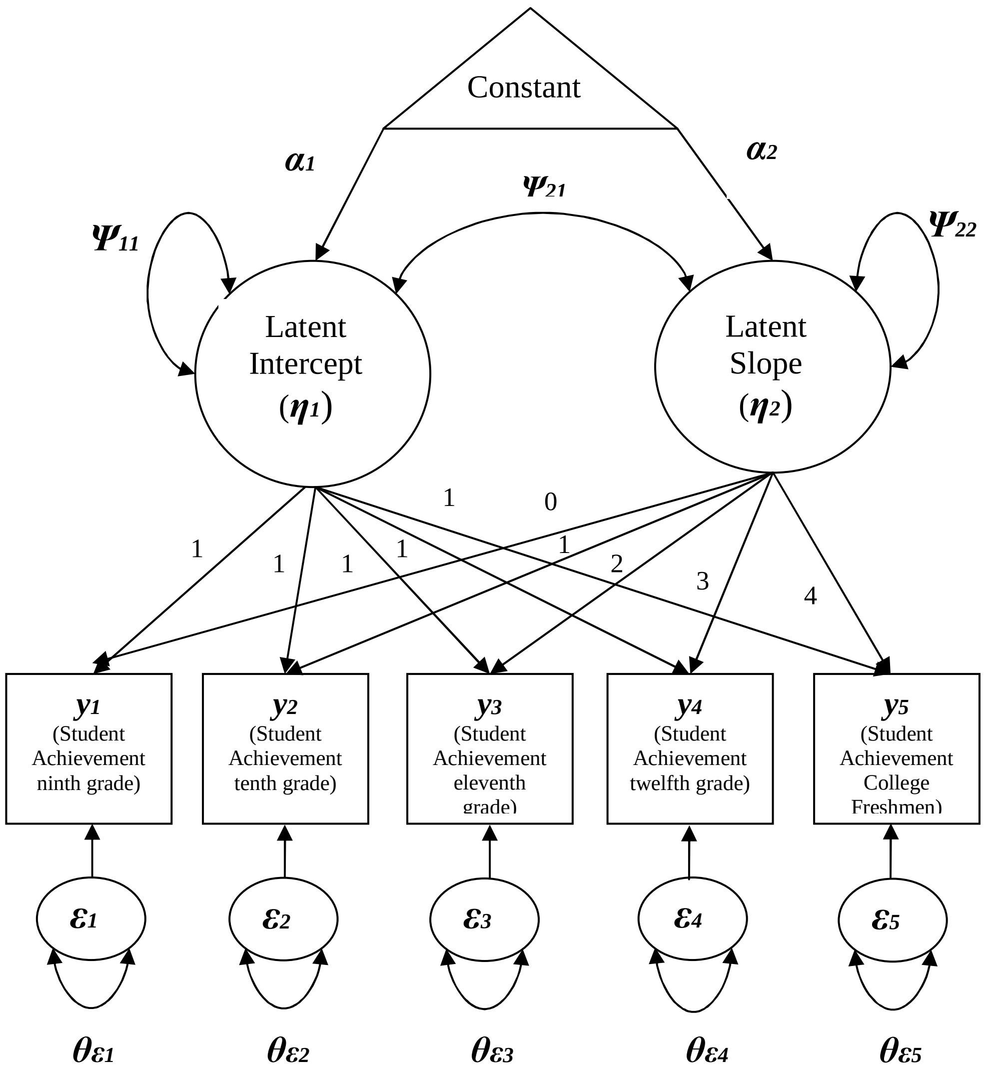

Figure 2.2. Univariate linear LGC model with five waves of data

Previous LGC modeling simulation studies have investigated waves of data in relation to statistical power that have ranged from three to eight waves of data. However, the majority of investigations included only the conditions of three, four, and five waves of data, which suggests that minor improvements are obtained beyond five waves when adequate sample size is achieved (Fan, 2003; Hertzog et al., 2006; Muthén & Curran, 1997; Muthén & Muthén, 2002). Similar patterns were found in the LGC modeling applications reviewed, as displayed in Figure 2.3, with three to five waves of data being frequently utilized. In conclusion, while LGC modeling applications and simulation studies have investigated a range of waves of data (e.g., three to eight time points), typically, researchers have focused on three to five waves of data.

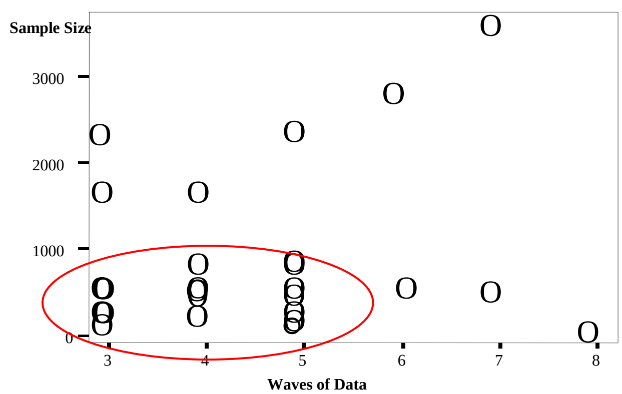

Figure 2.3. Waves of data and sample size of 29 application studies

Model Complexity

In LGC modeling, additional waves of data increase the model complexity by the inclusion of additional observed variables and fixed paths. However, in this dissertation and in most of the LGC modeling literature, waves of data are defined as a separate construct from model complexity because of the difference in the LGC modeling researcher's decisions in regard to specification of the model. Decisions regarding the number of waves of data should be defined according to the hypothesized time period of change. Typically, conditions of model complexity include additional characteristics of the growth trajectory (e.g., the shape of the growth trajectory, univariate or multivariate growth trajectories, and potential covariates of the trajectory of change). The following three sections define the conditions of model complexity utilized in this dissertation and include: (a) reasons for inclusion of the condition of model complexity, (b) the representation of model complexity in LGC modeling, and (c) simulation studies and applications that used the condition of model complexity.

Nonlinear growth. Frequently encountered in human growth or human development is nonlinear growth, which requires additional latent factors to represent the curvilinear nature of the growth (Burchinal & Appelbaum, 1991). While an assortment of procedures can be applied to represent nonlinear growth in LGC modeling (Bollen & Curran, 2006), the majority of nonlinear applications specified quadratic growth in the trait of interest (i.e., 31% of the studies reviewed). Moreover, Coffman and Miller's (2006) LGC model study simulated quadratic growth and compared the results to linear growth models under limited conditions. Therefore, this dissertation included the condition of quadratic growth.

As opposed to linear growth, quadratic growth assumes that a latent trait of interest begins with a slight growth and moves into moderate and high growth, and then plateaus with a slight decrease at the end of growth. A primary example of quadratic growth in the social and behavioral sciences is cognitive functioning across a lifetime. For example, cognitive functioning is expected to have slight increase at birth, with a large increase in youth and young adults, a plateau in mid to late adulthood, and a decrease in a geriatric population. In term of academic achievement from Grades 9 to 11, researchers in the field of education do not assume that achievement has a slight increase in Grade 9, with a moderate to large increase in Grade 10, a plateau in the end of Grade 10 and beginning of Grade 11, and slight decrease at the end of Grade 11. However, by changing the trait of interest from academic achievement to knowledge retention, an educational researcher may apply a quadratic LGC model because retention of knowledge is assumed to increase, plateau, and then have a slight decrease. Therefore, application of quadratic LGC modeling is dependent on the expected pattern of growth. Figure 2.4 illustrates a quadratic LGC model for knowledge retention from Grades 9 to 11.

Notice, the latent quadratic slope factor (η_3), the variance of the latent quadratic slope factor (ψ_{33}), and the latent mean of the quadratic slope factor (α_3) are incorporated in the model to represent quadratic growth in the trait of interest. According to the achievement example, η_3 is the quadratic growth in achievement from Grades 9-11, ψ_{33} is the variance of the quadratic growth in achievement from Grades 9-11, and α_3 is the latent mean of the quadratic growth in achievement from Grades 9-11.

Along with the additional latent factor, variance and mean parameters are associated with unidirectional paths from the quadratic factor to the observed variables (λ_q = 0, 1, 4), which are fixed to the squares of the linear factor loadings (Preacher et al., 2008). Additional covariances are represented in the model, including the covariance between the intercept factor and quadratic slope factor (Ψ_{13}) and the covariance between the linear slope factor and the quadratic slope factor (Ψ_{23}). In terms of the hypothesized achievement example, Ψ_{13} represents the relationship between the initial level of achievement in Grade 9 and the quadratic growth in achievement from Grades 9-11, and Ψ_{23} indicates the relationship between the linear growth in achievement from Grades 9-11 and the quadratic growth in achievement from Grades 9-11. For example, a large value of the covariance between the intercept and quadratic slope factor (Ψ_{13}) suggests that students' rate and pattern of growth may be related to where they started, as indicated by their initial achievement level.

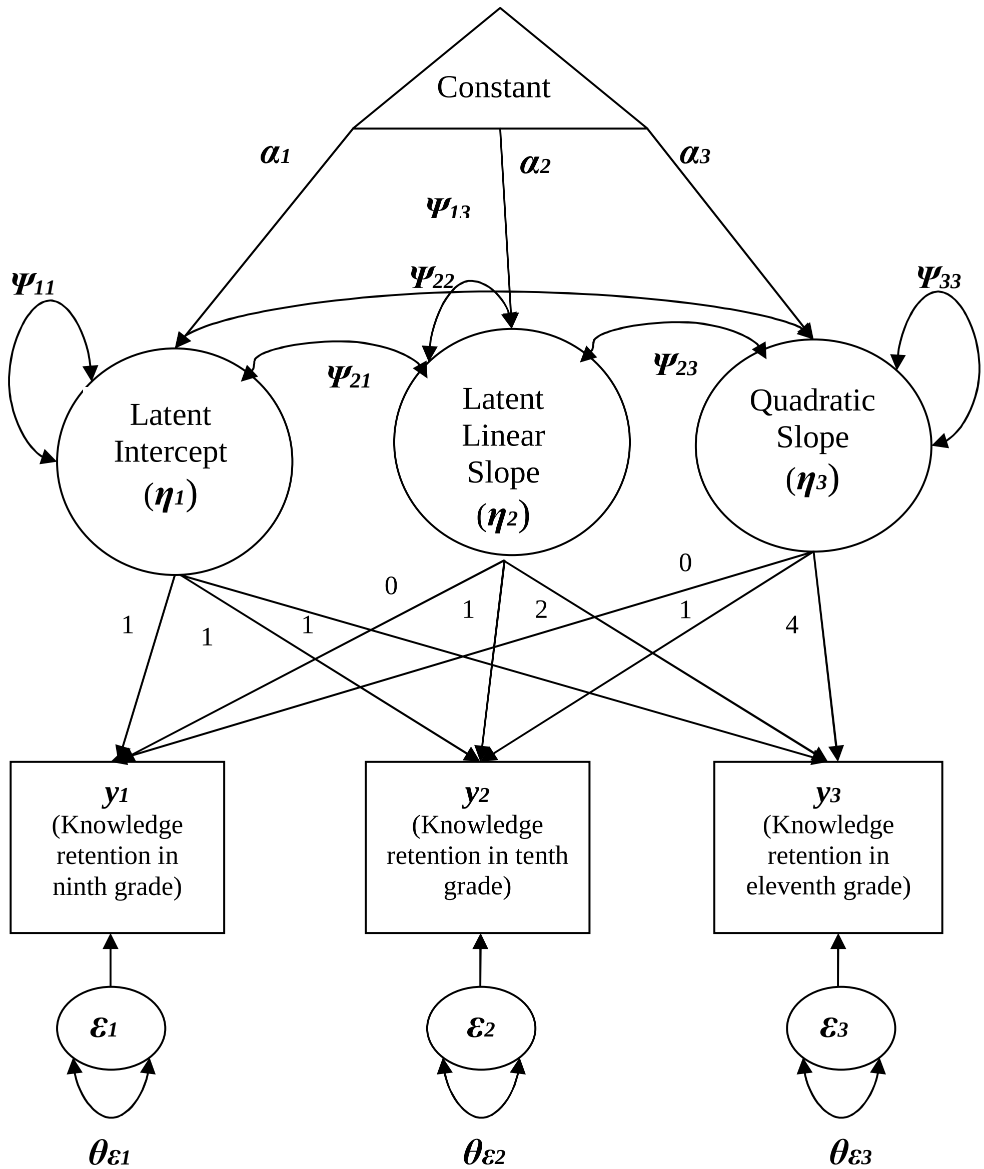

Figure 2.4. Univariate quadratic LGC model of knowledge retention

The η_3 is used to estimate the overall quadratic growth estimate (α_3) and deviation from the quadratic slope (ζ_3), through the procedures described in Equations 2.3 to 2.5. The latent quadratic growth factor in achievement from Grades 9-11, (η_3) is separated into the overall quadratic growth in achievement from Grades 9-11, (α_3), and some deviation from the quadratic slope of achievement from Grades 9-11, (ζ_3). Therefore, α_3 is the estimate of latent quadratic growth in student achievement and ζ_3 is a measure of external factors (e.g., students' concentration level). In summary, to represent quadratic growth an additional latent factor, variance of the latent factor, latent mean, deviation term, and two covariances are added to the model.

The validity of the representation of growth is commonly assessed by three different methods. First, confidence intervals can be computed for the estimated linear growth (α_2) and the estimated quadratic growth (α_3) to determine if the latent means of the linear and quadratic slope factors are significant across participants (Bollen & Curran, 2006). According to the achievement example, confidence intervals for the estimated mean linear growth from Grades 9-11, (α_2), and the estimated mean quadratic growth from Grades 9-11, (α_3), would be computed to determine if they are significant. If the quadratic growth factor includes a significant amount of variability, potential covariates and predictors should be explored. In the student achievement example, if the quadratic latent growth factor (α_3) was found to be significant, then the researcher should explore if parental involvement, or some other potential predictor or covariate, could explain some of the variability in growth in student achievement. Secondly, a χ^2 difference test is computed to determine if the quadratic slope factor improves the model fit, compared to linear growth, with significant results suggesting preference for the quadratic representation of growth (Willett & Sayer, 1994). Thirdly, fit indexes for a linear model can be compared to fit indexes for a quadratic model to assess which model displays the most desirable fit. While the procedures to assess the validity of the shape of growth are frequently applied, recent methodological sources have cautioned the sole use of these data-driven procedures.

Preacher et al. (2008) cautioned against the sole use of significance testing to determine the shape of growth, re-enforcing that theory must define the appropriate type of growth, and researchers should not “capitalize on possible idiosyncratic characteristics of the particular sample under scrutiny” (p. 51). Furthermore, Bollen and Curran (2006) emphasized that theories of change in latent traits found in the behavioral and social science are rarely hypothesized to represent quadratic growth, despite the frequent application of quadratic LGC models. For example, quadratic growth is expected to represent lifetime cognitive functioning; however, researchers rarely have the resources to measure cognitive function over a lifetime and are commonly examining segments of development that may require different representations of growth (e.g., cognitive functioning in adolescents). Finally, the application of a quadratic LGC model requires the estimation of additional parameters, in comparison to a linear LGC model. Therefore, researchers need to find the most appropriate balance among: (a) theoretical support; (b) effects of additional parameter estimates (i.e., increase in model complexity); and (c) the hypothesized bias of significance testing. Due to the complex debate that LGC modeling researchers encounter related to representation of growth, this dissertation included the condition of a univariate quadratic LGC model as a level of model complexity.

Multivariate LGC models. Within human development, growth simultaneously occurs among multiple traits and, frequently researchers are interested in the covariation among growth in two or more latent traits. The achievement example could be expanded to include: (a) growth in students' mathematical achievement, growth in students' verbal achievement, and of particular interest, the covariation between mathematical and verbal growth in achievement from Grades 9-11. Multivariate LGC modeling representation is separated into models to demonstrate first-order and second-order characteristics. The latter implies a higher order latent factor, which symbolizes the joint, global trait comprised of the two more specific traits (e.g., an additional global latent factor of achievement including paths to the mathematical and verbal latent factors; McArdle, 1988). The first-order representation, utilized in this dissertation, is referred to as an associative LGC model, which estimates the covariation among growth in two or more latent variables, but does not include a second order factor, which represents the combined traits of the first order factors (Duncan et al., 2007). The multivariate representation of achievement includes growth in both mathematical and verbal achievement from Grades 9-11, as displayed in Figure 2.5. Notice that estimation of a multivariate LGC model drastically increases the model complexity through inclusion of: (a) additional observed parameters to represent the second latent trait (y_4 is verbal achievement in Grade 9, y_5 is verbal achievement in Grade 10, and y_6 is verbal achievement in Grade 11); (b) associated estimated residuals ($\varepsilon_4, \varepsilon_5, \varepsilon_6$); (c) fixed factor loadings from the observed variables to the latent factors; (d) variance of the additional latent factors (Ψ_{33}, Ψ_{44}); (e) the covariation between the latent factors (Ψ_{24}, Ψ_{34}); and (f) the latent means of the additional latent factors (α_3 and α_4).

Despite the fact that previous simulation studies have not included multivariate LGC models, the practical appeal to human developmental research and the dramatic increase in model complexity requires methodological attention. Thus, this dissertation included a multivariate LGC model as a level of model complexity due to the expected increase in application in the fields of behavioral and social sciences.

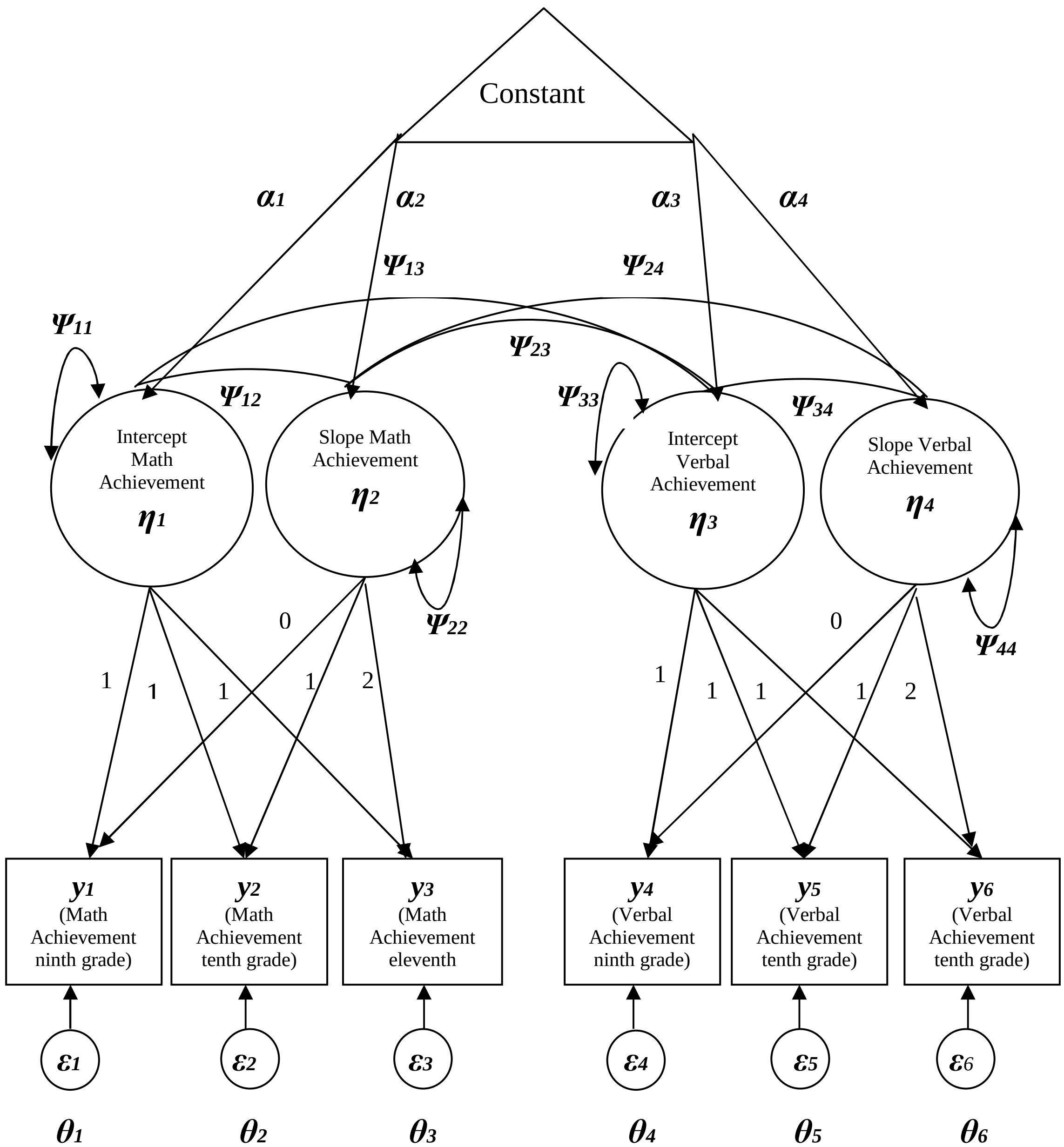

Figure 2.5. Multivariate linear LGC model of student achievement

Inclusion of a covariate. In most models of human behavior, change does not occur independently of other contextual factors, with covariates of growth necessary to provide a comprehensive model of change. A covariate in LGC modeling can be either continuous or categorical and is presented in the same manner as in general SEM applications where an additional observed variable is added to the model. The LGC models can include two categories of covariates: (a) time-varying covariates where an observed measure of the covariate is collected at each time point (e.g., classroom attendance in Grades 9, 10 and 11), and (b) time invariant covariates where one measure of the covariate is added to the model under the assumption that the variable will not change during the selected time period (e.g., participants' gender; Duncan et al., 2007). Frequently, time invariant covariate models are utilized in application studies and have been examined by Hertzog et al. (2006) and Muthén and Muthén (2002) in LGC modeling simulation studies. Therefore, the condition of a single time-invariant covariate added to a univariate LGC model is included as a level of model complexity.

Similar to all growth curve models, the addition of a covariate or predictor to a basic LGC model is called a *conditional LGC model*, whereas the models presented thus far are considered *unconditional LGC models* (Meredith & Tisak, 1990). Conditional LGC models are generally applied when an unconditional LGC model displays a significant amount of variance for the latent intercept and/or latent slope factors (i.e., a significant Ψ_{11} Ψ_{22}, and Ψ_{33} for a quadratic LGC model, and a Ψ_{44} for a multivariate LGC model). Significant variability of the latent factors (latent means) may be explained by another observed or latent trait; therefore, potential covariates of growth in the latent trait should be explored. For example, if the univariate LGC model of student achievement displayed significant variability for the latent intercept and latent slope factors (significant Ψ_{11} and Ψ_{22}), the corresponding interpretation is that students have significant variability in their trajectories of growth from Grades 9 to 11. The subsequent question arises as to what is influencing the variability in the initial level of achievement in Grade 9 and growth in achievement from Grades 9 to 11? Potential covariates could include gender, parental education levels, school attendance, IQ, and numerous others. Potential covariates can be latent or observed; however, this dissertation will only include observed variables. Conditional LGC models are typically applied to explain variability in the latent growth trajectories.

While not typically referenced as a rationale for applying conditional LGC models, it is important to highlight that most, if not all theories of change in the social and behavioral sciences include characteristics or traits hypothesized to modify or predict growth. Behavioral and social science theories tend to be complex, and I cannot identify a single theory of change in the social and behavioral sciences that assumes that growth in a latent trait is independent of any other trait or characteristic. For example, educational psychologists have numerous hypotheses about potential predictors (e.g., time spent in the classroom, school attendance, extracurricular activities) and covariates (e.g., parent’s education, teacher’s experience, school characteristics, socioeconomic status) of academic achievement. The conditional LGC model allows the researcher to examine the potential covariate of change in a latent trait. Therefore, the complex nature of human development supports the use of conditional LGC models in behavioral and social sciences to adequately represent the multiple influences on a hypothesized theory of change.

Conditional LGC models require alteration to the estimation of the latent means presented in Equations 2.4 and 2.5 (i.e., the level two growth curve model) to include a time-invariant covariate, as displayed in Equations 2.12 and 2.13.

$$\eta_{i1} = \alpha_1 + \beta_{\eta_{i1}} x_i + \zeta_{\eta_{i1}} \tag{2.12}$$

$$\eta_{i2} = \alpha_2 + \beta_{\eta_{i2}} x_i + \zeta_{\eta_{i2}} \tag{2.13}$$

The x_i symbol represents a single time-invariant covariate, $\beta_{\eta i1}$ is the random intercept parameter, and $\beta_{\eta i2}$ is the random slope parameter, which can both be interpreted similar to beta coefficients in a regression equation. However, the ζ_{ni1} and ζ_{ni2} have an alternative interpretation and are disturbances (i.e., conditional variances) with: (a) mean of zero; (b) variance of ψ_{11} and ψ_{22}, respectively; and (c) a covariance of ψ_{12} as opposed to the variances of the α_1 and α_2 discussed in Equations 2.4 and 2.5 (Bollen & Curran, 2006). Displayed in Figure 2.6 is the alternative structure of the level two LGC model with the addition of a single time-invariant covariate of gender in the student achievement example.

Notice the addition of the covariate effect on the random latent intercept ($\beta_{\eta i1}$) and the random slope parameter ($\beta_{\eta i2}$). In the conditional LGC model, the observed time points of student achievement are a function of the following: (a) the covariate of gender, (b) the disturbance terms, (c) the vector of factor loadings, (d) the latent intercept and latent slope, and (e) the residual terms. Therefore, an applied researcher can examine and test if the trajectory of growth in achievement, specifically the initial level and rate of growth, is moderated by gender. For example, male and female students may begin at the same level of achievement in Grade 9 (non-significant $\beta_{\eta i1}$); however, males and females may differ in their linear slope or growth from Grades 9 to 11 (significant $\beta_{\eta i2}$).

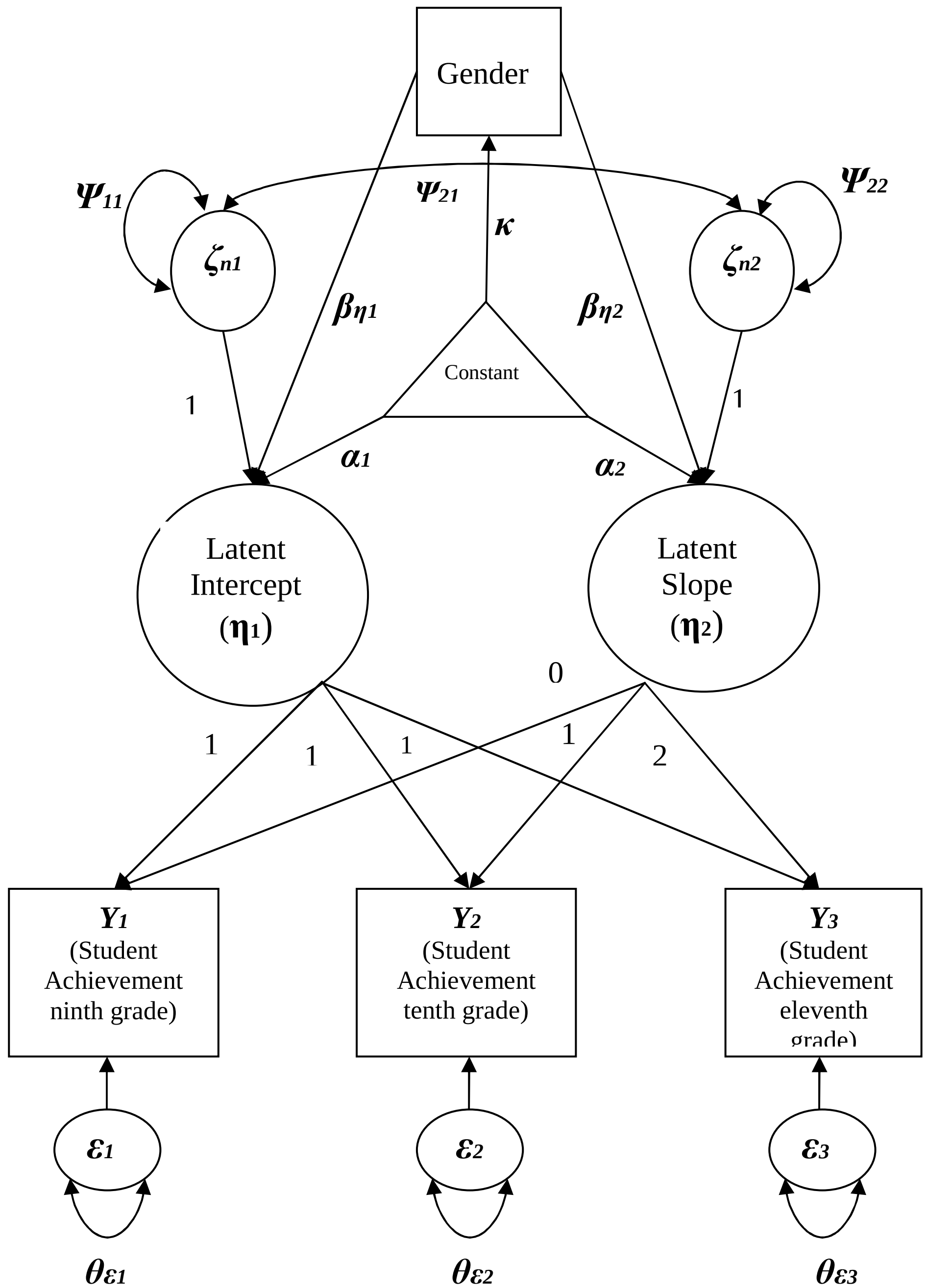

Figure 2.6. Univariate linear LGC model of student achievement with a covariate

Typically, the covariate effects are displayed in graphical format, as displayed in Figure 2.7 for the hypothesized gender effect of growth in student achievement from Grades 9 to 11. Notice, both males and females have the same initial status in achievement; however, the rate of growth for females is high than males from Grades 9 to 11. The conditional LGC model with a single time invariant covariate is included in this dissertation to examine whether the alternative structure, and representation of a conditional LGC model, influences the interpretation of fit indexes.

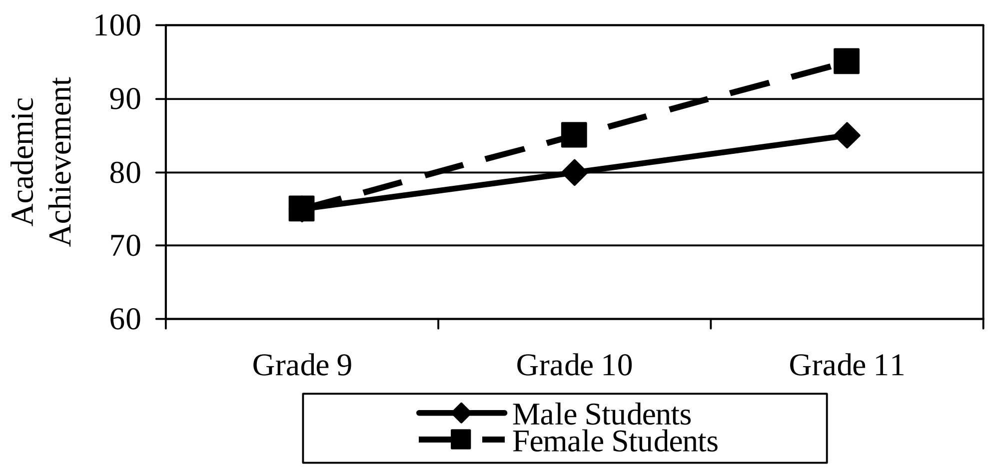

Figure 2.7. Academic achievement between grades 9 to 11 by gender

Summary of LGC Model Characteristics

This dissertation includes three characteristics of LGC models (e.g., sample size, waves of data, and model complexity), which are the independent variables in the simulation of model fit indexes. To briefly review, variations in sample size do not alter the LGC model structure; however, the sample size is of critical importance to statistical power, with methodological literature providing general guidelines for LGC models. An increase in waves of data will increase the number of observed variables and fixed factor loadings, as well as increase the reliability of the analysis. Conversely, increasing waves of data require additional resources and should have theoretical or contextual support. Model complexity conditions are associated with methodological decisions in regard to: (a) the trait(s) of interest represented in the growth trajectory (i.e., univariate or multivariate LGC model); (b) the shape of the growth trajectory (i.e., univariate or quadratic growth); and (c) need to represent traits which may influence the growth trajectory (i.e., inclusion of a covariate). Quadratic growth requires: (a) an additional latent factor, (b) variance of the latent factor, (c) covariances among the latent factors, and (d) estimation of a latent mean for the quadratic growth factor. Multivariate growth dramatically increases model complexity with: (a) additional observed variables and latent factors, (b) variance of the latent factors, (c) covariances among the latent factors, and (d) estimation of additional latent means. The inclusion of a single time invariant covariate creates a conditional LGC model and alters the structure and interpretation of the variances of the latent factors. Due to the differences in model specification among the conditions of model complexity, interpretation of fit indexes under varying conditions of LGC model complexity should be of critical concern to applied researchers. In summary, the three independent variables and levels of model complexity provide a range of LGC modeling conditions that directly relate to decisions frequently encountered in applications.

Model Identification

Inherent to all applications of SEM are issues of model identification or the ability “to derive a unique estimate of each parameter” (Kline, 2005, p. 105). The LGC model parameters, and all SEM applications, can be divided into two categories: (a) parameters that are *known to be identified*, commonly referred to as *known* parameters, that include the observed means, variances, and covariances; and (b) *unknown parameters* which are estimated in the model including the model implied latent means, error variance of observed measures, and the variance of latent factor (Bollen & Curran, 2006). In LGC modeling, the number of known parameters, minus the unknown parameters, equals the *df*.

To be able to estimate the parameters of the growth trajectory and model fit, the number of known parameters must be greater than the number of unknown parameters to be estimated (*df* > 0), which is termed an *identified model*. Equality of the number of known and unknown parameters results in a *just identified* model (*df* = 0), which permits estimation of the desired parameters of change; however, just identified models assume that the specified model fits the data perfectly. As a result, researchers who use a just identified model must assume perfect measurement, cannot estimate the residual errors of the latent means, and lack the ability to produce tests of model fit. Finally, desired parameters of change and model fit cannot be estimated when there are a greater number of unknown parameters to be estimated than known parameters (*df* < 0), referred to as an *unidentified* model.

As previously discussed, LGC models differ from standard SEM applications by fixing the paths from the observed variables to the latent factors; therefore, factor loadings are not unknown parameters as typically labeled in standard SEM applications. Furthermore, LGC models fix the means of the error variances and disturbance terms at each time point, which suggests that error variances of the observed time points are the same for all participants; however, the error variances can differ across time points (Bollen & Curran, 2006). In addition, two of the three independent variables for this study require the inclusion of additional parameters (i.e., waves of data and model complexity), which influence the model identification.

Bollen and Curran (2006) developed general equations to compute the number of known and unknown parameters for each LGC model utilized as independent variables in this dissertation. Described in Table 2.1 are: (a) the equations used; (b) the number of known parameters; and (c) the number of unknown parameters for the four levels of model complexity by waves of data (i.e., range from three to six). In Table 2.1, T represents the waves of data (i.e., or time points), and K symbolizes the number of covariates. Sample size is not included in Table 2.1 because it is not related to model identification.

Beginning with the univariate linear LGC model with three waves of data, presented in Figure 2.1 and Equation 2.8, there are nine parameters known, which correspond to the observed means, variances, and covariances. Specifically, the model includes: (a) three means of the observed achievement scores in Grades 9, 10, and 11 [($E(y_{i1})$, $E(y_{i2})$, $E(y_{i3})$]; (b) variances of Grade 9, 10, and 11 achievement scores [$VAR(y_{i1})$, $VAR(y_{i2})$, $VAR(y_{i3})$]; and (c) their covariances [($COV(y_{i1}, y_{i2})$, $COV(y_{i1}, y_{i3})$, $COV(y_{i2}, y_{i3})$], defined as known parameters (e.g., (½)(3)(3+3) = 9). There are eight unknown parameters corresponding to the model-implied variance-covariance matrix and latent means, including the estimation of the latent means ($\mu_{\alpha 1}$ and $\mu_{\alpha 2}$), variances of the latent factors (ψ_{11} and ψ_{22}), and covariance of the latent factors (ψ_{11}), in addition to the estimation of error variances for each time point [($VAR(\varepsilon_{i1})$, $VAR(\varepsilon_{i2})$, $VAR(\varepsilon_{i3})$] (e.g., 2 + 2 + 1 + 3= 8). Thus, the univariate linear model with three waves of data is identified with one degree of freedom.

Table 2.1

Model Identification for LGC Models

LGC Model (corresponding figure)	Waves of Data	Known Parameters	Unknown parameters	Degree of Freedom
Univariate Linear LGC model (2.1)		(½)(T)(T+3)	5 + T	
	3	9	8	1
	4	14	9	5
	5	20	10	10
	6	27	11	16
Univariate Quadratic LGC model (2.2)		(½)(T)(T+3)	9 + T	
	3	9	12	Unidentified
	4	14	13	1
	5	20	14	6
	6	27	15	12
Multivariate LGC model (2.3)		n.a.	14 + 2T	
	3	27	20	7
	4	44	22	22
	5	65	24	41
	6	90	26	64
Univariate Linear LGC model with time invariant covariate (2.4)		(½)(T+K)(T+K+3)	T + 2K + 5	
	3	14	10	4
	4	20	11	9
	5	27	12	15
	6	35	13	22

Note. T = the number of waves of data, K = the number of covariates; n.a. represents that estimation of multivariate condition cannot be captured in a simple equation.

The quadratic linear LGC model, presented in Figure 2.3, includes the same number of known parameters; however, four additional unknown parameters are added to the model, including the: (a) latent mean ($\mu_{\alpha 1}$); (b) variance (ψ_{33}) of the quadratic factor; and (c) two covariances between the quadratic factor, latent intercept, and latent slope factors (ψ_{13} and ψ_{23}). Consequently, a quadratic linear LGC model with three waves of data is an unidentified model; thus, fit indexes and parameters cannot be estimated.

A multivariate linear LGC model is the most intricate level of model complexity examined and substantially increases the known parameters (i.e., ranges between 27-90 parameters); therefore, all multivariate LGC models are over-identified. A univariate linear LGC model with a covariate increases the number of known parameters by including the covariate, variance, and co-variance of the covariate, as well as increasing the number of unknown parameters to include the covariate coefficients (i.e., $\beta_{\alpha 1}$ and $\beta_{\alpha 2}$). Therefore, inclusion of a covariate increases the *df* in LGC modeling with all model conditions being an identified model ($df \geq 1$). In conclusion, there is one condition where fit indexes cannot be estimated (i.e., univariate quadratic LGC with three waves of data).

Summary of Latent Growth Curve Modeling

The LGC models hold many similar characteristics to typical SEM procedures; however, LGC models estimate latent means as well as the variance-covariance matrices and are designed to answer longitudinal research questions. In the first portion of Chapter II, I explained parameter estimation in LGC modeling and detailed how F_{ML} reproduces the model implied variance-covariance matrix and means to minimize the discrepancies to the observed variance-covariance matrix and latent means. The discussion continued to describe three extensions of LGC models used as independent variables in this dissertation (i.e., sample size, waves of data, and model complexity), which are associated with critical design decisions made by LGC modeling researchers and methodologists. In the last section of Chapter II, the focus shifts to description of the four fit indexes, dependent variables in this dissertation, and relevant literature that was used to formulate the research hypotheses.

Description of Fit Indexes

Excluding Coffman and Millsap's (2006) study of LGC model fit, the methodological knowledge of LGC model fit is derived from SEM simulation studies. Conversely, SEM literature is immersed with methodological studies investigating fit indexes under various conditions of: (a) model type, (b) sample size, (c) estimation nmethod, (d) model misspecification, and (e) normality (Beauducel & Wittmann, 2005; Cheung & Rensvold, 2002; Curran, Bollen, Paxton, Kirby, & Chen, 2002; Davey, Savla, & Luo, 2005; Fan & Sivo, 2005; Fan & Wang, 1998; Hu & Bentler, 1998; 1999; Jackson, 2007; La Du & Tanaka, 1995; Sivo, Fan, Witta, & Willse, 2006; Tanguma, 2001; Widaman & Thompson, 2003; Yuan, Bentler, & Zhang, 2005). Historically, the χ^2 likelihood ratio test has been used to assess model fit, but due to its limitations, numerous alternative indexes have been proposed. The array of fit indexes available is problematic (Sivo et al.), and applied researchers and methodologists have varied preferences in regard to the optimal venue to assess model fit. Therefore, despite the abundance of investigations, procedures to establish model fit lack congruency and are still highly debated (Marsh, Hau, & Wen, 2004).

The authors of LGC modeling books have selectively endorsed a collection of fit indexes for LGC modeling techniques; however, rationales for the indexes are extracted from general SEM simulation studies and lack discussion of how LGC modeling variations may influence model fit interpretation (Bollen & Curran, 2006; Duncan et al., 2007; Preacher et al., 2008). Described in Table 2.2 are the fit indexes suggested for use in the three LGC modeling books that are the primary resources for training and teaching of LGC modeling procedures. In addition, Table 2.2 displays the percentage of authors, among the 29 applications studies reviewed, which reported the indexes in their published manuscript. Notice, the standardized root mean square residual (*SRMR*) and incremental fit index (*IFI*) are rarely utilized and lack endorsement in the majority of LGC modeling books. Even though references to *SRMR* are found in LGC modeling literature (e.g., the square root of the squared absolute difference between the *S* and $\Sigma(\theta)$ matrices), the *SRMR* lacks assessment of the mean vectors (i.e., comparison of $\bar{y}$ and $\mu(\theta)$; for a review of notation, see section on parameter estimation). Similarly, the *IFI* is seldom found in LGC modeling literature, and researchers have shown similar performance of the *IFI* and *CFI* in simulated CFA models (Bentler, 1990, La Du & Tanaka, 1995; Yadama & Pandey, 1995).

Table 2.2

Recommended Fit Indexes in LGC Modeling

Source	χ^2	*NNFI*	*CFI*	*RMSEA*	*SRMR*	*IFI*
Bollen & Curran (2006)	✓	✓		✓		✓
Duncan et al. (2007)	✓	✓	✓	✓		
Preacher at al. (2008)	✓	✓	✓	✓	✓	
29 Application Studies	93.1%	51.7%	65.5%	65.5%	17.2%	3.4%

Note. *NNFI* = non-normed fit index, *CFI* = comparative fit index, *RMSEA* = root mean squared error of approximation, *SRMR* = standardized mean square residual, and *IFI* = incremental fit index.

The χ^2, *NNFI*, *CFI*, and *RMSEA* are used in this dissertation to represent fit indexes utilized in LGC modeling applications. The four fit indexes are trifurcated into three categories based on the manner in which they assess model fit, including: (a) χ^2 likelihood ratio test; (b) incremental fit indexes (*NNFI* and *CFI*); and (c) absolute fit indexes (*RMSEA*; Bollen, 1989; Hu & Bentler, 1999). The following discussion describes the: (a) computation of the fit indexes, (b) proposed cutoff values to determine adequate model fit, and (c) summaries of selected CFA model fit simulations. Due to the countless simulation studies on model fit, the following review is focused on authors who utilized conditions relevant to the independent and dependent variables used in this dissertation (e.g., similar to LGC modeling data environments, continuous outcomes, ML estimation procedures, etc.).

Chi-Squared Likelihood Ratio Test

In the behavioral and social sciences, the χ^2 likelihood ratio test is a frequently applied procedure, where a dichotomous decision is made to retain or reject the null hypothesis described below. In terms of LGC model fit, the χ^2 likelihood ratio test determines whether a significant difference simultaneously occurs between the observed variance-covariance matrix and mean vector, and model implied variance-covariance matrix and mean vector (Bollen & Curran, 2006). The null hypothesis implies that $\Sigma(\theta)$ exactly reproduces S, and $\mu(\theta)$ exactly reproduces $\bar{y}$ as displayed in Equation 2.14.

$$H_0: S = \Sigma(\theta) \text{ and } \bar{y} = \mu(\theta) \quad (2.14)$$

Unlike typical applications of the χ^2 test, a nonsignificant result is desirable, as it conceptually assesses the *badness of model fit*. Significant discrepancies between S and $\Sigma(\theta)$ and between $\bar{y}$ and $\mu(\theta)$ are interpreted as insufficient model fit and imply that the hypothesized model does not adequately account for relationships in the observed data. The LGC models lack adequate model fit when the estimated parameters, $\Sigma(\theta)$ and $\mu(\theta)$, are significantly different from observed parameters. In other words, if the model specified by the researcher does not adequately match the observed data, the hypothesized model might lack support. The χ^2 test statistic is equal to the F_{ML}, described in Equation 2.11, multiplied by the sample size (N) minus one, as shown in Equation 2.15.

$$\chi^2 = (N - 1)\, F_{ML} \tag{2.15}$$

When the assumptions of LGC modeling are satisfied, Equation 2.15 follows a χ^2 distribution with degrees of freedom equal to the difference between the number of unique elements of the observed variance-covariance matrix and the number of parameters estimated. The degrees of freedom for the LGC models utilized in this dissertation are presented in Table 2.1 in the discussion of model identification.

Intuitively, application of the χ^2 likelihood ratio test provides an adequate measure of model fit; however, numerous researchers have discussed limitations to this procedure (Beauducel & Wittmann, 2005; Bentler, 1990; Bollen & Curran, 2006; Duncan et al., 2007; Fan & Wang, 1998; Hu & Bentler, 1999; Kline, 2005). By conceptualizing model fit through significance testing, sample size becomes a confounding factor. Large sample sizes excessively increase statistical power to detect minor discrepancies, which result in a rejection of the null hypothesis (i.e., implies a lack of model fit) even when the model may adequately fit the data (Beauducel & Wittmann). Consequently, numerous SEM methodologists have demonstrated that, when sample sizes are large, the χ^2 likelihood ratio test will be too restrictive, resulting in excessive rejection of correct models (Beauducel & Wittmann; Hu & Bentler).

The use of LGC modeling may increase detection of minor discrepancies between observed and model-implied matrices for the χ^2 assessment in conditions of large sample sizes because of additional estimation of latent means. An increase in the number of parameters ultimately increases the model complexity (i.e., the size of the variance-covariance matrices and mean vectors), which in turn, increases the statistical power of the significance testing (Bollen & Curran, 2006). Thus, there is a higher chance of having a minor discrepancy between S and $\Sigma(\theta)$ and $\bar{y}$ and $\mu(\theta)$ result in a significant result, implying a lack of model fit. To demonstrate with the student achievement example, a trivial discrepancy may occur between the observed and model-implied mean of Grade 9 student achievement, with all remaining model implied parameters being exactly reproduced (e.g., the variances-covariances and means of Grade 10 and 11 student achievement). In this situation, a significant χ^2 test would be obtained, which would imply inadequate model fit; however, minor discrepancies in the single Grade 9 mean achievement scores may be negligible in terms of practical significance. A comparable cross-sectional CFA model would include only the variance-covariance matrices resulting in a non-significant χ^2, which would imply acceptable model fit (Bollen & Curran). As a consequence, the LGC model structure may increase the bias of the χ^2 with large sample sizes due to the additional estimation of the latent means.

Furthermore, when sample sizes are small, the χ^2 test may not contain enough statistical power to adequately reject a misspecified model (Field, 2005). Therefore, the χ^2 likelihood ratio test is commonly understood as being too relaxed for small sample sizes and too conservative for large samples. In recent years, the popularity of assessing the χ^2/df ratio is preferred to correct for the χ^2 tendency to penalize more complex models. Numerous SEM methodologists and researchers have found the χ^2/df to be appropriately sensitive to model misspecficiation, with complex models and extreme sample sizes (Jackson, 2007), compared to the standard χ^2 test. However, SEM methodologists generally report the standard χ^2 test, despite its shortcomings; therefore, both the χ^2 test the χ^2/df were reported in this dissertation in order to compare to other typical CFA model fit simulations. Along with the χ^2 test, and the χ^2/df assessment, SEM researchers frequently use alternative descriptive methods to assess model fit. These alternative fit indexes are descriptive and assume a continuum of fit, as opposed to the χ^2 likelihood ratio test which is an inferential test with an *all or nothing* interpretation. Continuous debate occurs in regard to what values constitute acceptable model fit, and methodologists discuss the falsified nature of any single cutoff values among the vast SEM application conditions (Hu & Bentler, 1999; March et al., 2004). However, applied researchers, textbooks, and journals continue to present and support standard cutoff values for individual fit indexes. The following section describes the computation of fit indexes, as well as suggested cutoff values for assessing model fit.

Incremental Fit Indexes

Incremental fit indexes compare the hypothesized model, specified by the researcher, to a more restrained baseline model to determine the proportion of improvement in model fit (Hu & Bentler, 1999). Conceptually, a continuum is created**,** which ranges from the largest chi-squared value of a baseline model to a saturated model (df = 0; Bentler, 1990). In typical SEM applications, LGC modeling applications, and the current dissertation, the baseline model is the independence null model, specified to estimate (i.e., free) the variances of the observed variables and fix the covariances among the observed variables to zero to imply no underlying common or latent factors (Bentler & Bonett, 1980). Although not commonly found in SEM applications, alternative null models may provide a more appropriate measure of model fit because the assumptions of no common variance among the observed data points (i.e., zero covariance) is rarely, if ever, found in social and behavioral science applications (Widaman & Thompson, 2003). While use of an alternative null baseline model may provide a more suitable assessment of model fit specific to the growth or change in a hypothesized latent trait, the widespread application would affect the standardization of fit statistics across studies, resulting in overall lower fit index values. Therefore, this dissertation specified the independence null model for variance-covariance structure. Even less agreement is found for the specification of appropriate values for the mean vectors of the baseline LGC model. Following guidelines provided by Bollen and Curran (2006), in this dissertation, I freed (estimated) the mean parameters in the baseline model. In summary, incremental fit indexes in this dissertation compared the hypothesized LGC model to a baseline model with: (a) estimated variances of the observed variables, (b) estimated latent means, and (c) fixed covariances at zero.

The nonnormed fit index (*NNFI*; also referred to as the Tucker Lewis Index) is an incremental fit index that requires model comparison to determine the value of the fit index (Bentler & Bonett, 1980). The *NNFI* utilizes the likelihood ratio χ^2 test statistic and degrees of freedom, displayed as

$$NNFI = \frac{{\chi_b}^2 / df_b - {\chi_h}^2 / df_h}{{\chi_b}^2 / df_b - 1} \tag{2.16}$$

where, ${\chi_b}^2$ is the likelihood ratio χ^2 test statistic for the baseline model, df_b are the degrees of freedom for the χ^2 test statistic for the baseline model, ${\chi_h}^2$ is the likelihood ratio χ^2 of the hypothesized model, and df_h are the degrees of freedom of the χ^2 for the hypothesized model. The equation compares the relative difference between the baseline model denoted with a *b* subscript, and the hypothesized (i.e., or specified) model represented by the subscript *h*. By dividing the χ^2 value by the *df* prior to computing the relative difference, the *NNFI* is known to compare model fit per *df* (Bentler, 1990). The *NNFI* is a nonnormed fit index with the majority of values ranging from 0-1.0; however, values can fall outside of this range. Values of 1.0 indicate perfect model fit; values greater than 1.0 occur when the $\chi_h^{\ 2}$ is greater than df_h (Bentler). The value produced by the *NNFI* would rarely be negative, because the baseline model ($\chi_b^{\ 2} / df_b$) is expected to be larger than the corresponding hypothesized model ($\chi_h^{\ 2} / df_h$) due to the hypothesized model imposing more restrictions than the baseline model.

A more recent incremental fit index is the *CFI*, designed to have the benefits of the *NNFI,* while reducing the undesirable characteristics (e.g., large variance, over-parameterization), as discussed in the following section. The *CFI* is a normed fit index, proposed by Bentler (1990) that ranges from 0-1.0, where 1.0 indicates a perfect fit. The *CFI* is computed by

$$CFI = 1 - \left(\tilde{d}_h / \tilde{d}_b\right) \tag{2.17}$$

where, $\tilde{d}_h$ is the $\max(d_h, 0)$, $\tilde{d}_b$ is the $\max(d_b, d_h, 0)$, d_h is $\left(\chi_h^{\ 2} - df_h\right)/n$, d_b is $\left(\chi_b^{\ 2} - df_b\right)/n$, with all other variables defined according to Equation 2.16. In application, the baseline model χ_b^2 should be greater than the χ_h^2 for the less restrictive hypothesized model, which should both be greater than zero (i.e., unless the model is just-identified); therefore $\tilde{d}_b$ > $\tilde{d}_h$ > 0. $\tilde{d}_h$ is the maximum value of the range produced by the value for $\left(\chi_h^{\ 2} - df_h\right)/n$, which adjusts for χ^2 values *df* and sample sizes (Bollen & Curran, 2006).

The $\tilde{d}_b$ is the maximum value of the continuum created by the baseline model, hypothesized model, and zero, also corrected for sample size and *df*.

Within the last decade, methodologists have reported evidence for the limited sensitivity of incremental fit indexes in the detection of model misspecification (Fan, Thompson, & Wang, 1999; Jackson 2007). For example, Fan et al. found the *CFI*, *NNFI*, and χ^2 to have approximately half of the sensitivity of the *RMSEA* to detect appropriate models. Although the results of Fan et al. study lack support for the *CFI* and *NNFI*, the authors defined model misspecification as additional unspecified latent paths (i.e., factor loadings), according to standard CFA model fit simulation procedures. Misspecification of latent paths lack relevance within the LGC modeling framework, which fixes the factor loadings to represent time. However, Hu and Bentler (1998) manipulated both measurement model and structural model misspecification, defined as misspecification between latent factors, and found that the *CFI* was highly sensitive to structural misspecification and only moderately sensitive to measurement model misspecification. Therefore, the limited sensitivity of incremental fit indexes discussed in the model fit literature may lack relevance for LGC modeling data environments.

Originally, Bentler and Bonett (1980) proposed cutoff values of .90 or greater to constitute sufficient model fit for the *NNFI*. To date, Hu and Bentler's investigations (1998, 1999) into model fit are the most well accepted methodological references and provide a foundation upon which subsequent application studies and methodological investigations base their results (Beauducel & Wittmann, 2005; Sivo et al., 2006). Even though Hu and Bentler (1999) discussed the erroneous nature of any single cutoff values among the vast SEM application conditions, they provided cutoff values for individual fit indexes that have become the *golden rules* of SEM (March et al., 2004). Hu and Bentler (1999) proposed more restrictive cutoffs than originally proposed for the incremental fit indexes and suggested that values of .95 (*NNFI*) and .96 (*CFI*) or greater imply sufficient model fit.

To complicate the debate of adequate cutoffs, researchers have provided evidence to support that *NNFI* and *CFI* are influenced by sample size. In conditions of small sample size (i.e., $N < 200$), researchers have found the *NNFI* to produce low values (i.e., which suggests inadequate model fit) and extremely high values that lack interpretability (Jackson, 2007; Sharma, Mukherjee, Kurmer, & Dillion, 2005; Tanguma, 2001). Understandably, authors who have utilized sample sizes greater than 150 report no substantial influence of sample size for the *NNFI* and *CFI* (Bentler, 1990; Cheung & Rensvold, 2002; Fan et al., 1999; Sivo et al., 2006). Moreover, the *NNFI* has been found to produce standard deviations that are substantially larger than other fit indexes (Bentler; Jackson; Sharma et al.; Yadama & Panday, 1995). The large range of the *NNFI* reflects contradictory interpretations that for some simulated samples would suggest a lack of model fit, whereas for others would over-estimate model fit.

To further the understanding of cutoff for incremental fit indexes, Sivo et al. (2006) examined the optimal cutoff values for *CFI* and *NNFI* in two situations: (a) the minimum value without rejection of any correctly specified models (i.e., Type I error) and (b) the maximum value to reject all misspecified models (i.e., Type II error). The range between the two optimal values creates guidelines of acceptable values dependent on sample size. Sivo et al. found the ranges to be identical for both incremental fit indexes, including: (a) $N = 150$ (*NNFI* and *CFI* = .95 – 1.0); (b) $N = 250$ (*NNFI* and *CFI*

= .97 – 1.0); (c) N = 500 (*NNFI* and *CFI* =.98 - .99); (c) N = 1,000 (*NNFI* and *CFI* = .98 - .99); (d) N = 2,500 (*NNFI* and *CFI* = .98 -.99); and (e) N = 5,000 (*NNFI* and *CFI* = .98 - .99). Notice, the range decreases as sample size increases with all conditions where $N \geq$ 1,000 indicating an identical range. Collectively, the simulated results support that the *NNFI* and *CFI* are influenced by variability in sample size.

Contrary to sample size, which is investigated in the majority of model fit simulations, model complexity is less frequently included as a condition of interest. Sharma et al. (2005) simulated CFA model complexity by an increase of the number of latent factors (e.g., 2, 4, 6, and 8) and corresponding observed variables (e.g., 8, 16, 24, and 32); they found *NNFI* values to vary among conditions of model complexity, with the magnitude of the effect of greater model complexity increasing with low sample sizes; however, they did not investigate the *CFI*. Although the results support the influence of model complexity on *NNFI* assessment of fit, conditions examined by Sharma et al. may lack applicability to LGC modeling environments (e.g., 16 or more waves of data). Comparable to LGC modeling conditions, Cheung and Rensvold (2002) simulated CFA models as they varied the number of: (a) latent factors (e.g., 2 and 3); (b) observed variables (e.g., 3, 4, and 5); and sample size (e.g., 150 and 300). They found that the *CFI* and *NNFI* values were higher, suggesting adequate fit, with a lower number of latent factors and observed variables. Values decreased as model complexity increased (i.e., as additional latent factors and observed variables were added to the model, the values suggested worse fit). In regard to the magnitude of effects, the *CFI* was more influenced than the *NNFI* for both latent factors and observed variables.

Based on prior research on model fit literature for incremental fit indexes, in the current dissertation the *NNFI* and *CFI* are hypothesized to be afflicted by variation in sample size, when N < 200 resulting in biased estimates of model fit. Specially, the *NNFI* and *CFI* values will vary among the different conditions of sample size; however, the effect sizes and mean values will suggest negligible difference in terms of practical implications for sample sizes greater than 200. For waves of data and model complexity conditions, the *NNFI* and *CFI* were hypothesized to vary under conditions of waves of data (i.e., observed variables) and model complexity (i.e., additional factor loadings). The *CFI* was expected to suggest worse model fit with increased model complexity, in comparison to the *NNFI*. Finally, the *NNFI* was expected to produce large variations, in comparison to all other fit indexes.

Absolute Fit Indexes

Unlike incremental fit indexes, absolute fit indexes do not use a baseline model to assess model fit, but examine to what degree a hypothesized variance-covariance matrix and mean vectors can be reproduced (Bollen & Curran, 2006). As opposed to the three previously described fit indexes that are based on exact model fit, the *RMSEA* is based on close approximation to the correct model. The *RMSEA* is an absolute fit index, computed as

$$RMSEA = \sqrt{\max\left[\left(\chi_h{}^2 - df_h\right)/\left(n \times df_h\right), 0\right]} \qquad (2.18)$$

where all values are defined as in Equation 2.16, including: df_b are the degrees of freedom for the χ^2 test statistic for the baseline model, $\chi_h{}^2$ is the likelihood ratio χ^2 of the hypothesized model, and df_h are the degrees of freedom of the χ^2 for the hypothesized model. The numerator of the first term is the unbiased estimate of “the noncentrality parameter for the noncentral chi-square distribution underlying hypothesized model” $\left(\chi_h{}^2\right)$ (Bollen & Curran, 2006, p. 47). The expression in the denominator of the first term corrects for the sample size effect and penalizes for increasing *df*, commonly found in complex models (Bollen & Curran). The values of the *RMSEA* range from zero to infinity where values of zero indicate a perfect fit.

Preference for the *RMSEA* is related to its ability to be highly sensitive to model misspecification (Fan et al., 1999; Fan & Wang, 1998; Hu & Bentler, 1998; Jackson, 2007; Sivo et al., 2006). Another reported advantage of the *RMSEA* index is that confidence intervals can be computed, based on upper and lower limits of the non-central chi-squared distribution (Curran et al., 2002). In contrast, Chen, Curran, Bollen, Kirby, and Paxton (2008) found a lack of support for the value added by the construction of *RMSEA* confidence intervals with an upper bound of 0.1 and lower bound of .05, corresponding to standard cutoff values for acceptable fit. Chen et al. concluded that confidence intervals are afflicted by the use of universal cutoff values as upper and lower bound limits for the abundance of conditions found in research environments. Furthermore, Curran et al. found *RMSEA* confidence intervals to be biased when $N < 200$, due to deviations from the non-central chi-square distribution with small sample sizes. Other research has been less favorable toward the *RMSEA* on the basis of its relative lack of sensitivity to model misspecification. For example, Sharma et al. (2005) endorsed the *NNFI* over the *RMSEA* and suggested that the *NNFI* is more sensitive to model misspecification than the *RMSEA* in a CFA model with varying sample sizes and conditions of model complexity similar to LGC modeling environments (e.g., additional latent factors and observed variables; see section on incremental fit indexes for description of conditions). Regardless of the simulated evidence in favor of or against the *RMSEA* as a measure of model fit, researchers frequently utilize the *RMSEA,* and its functioning under LGC modeling conditions are of critical concern.

Steiger's (1989) original guidelines, in conjunction with support from other methodologists, endorsed *RMSEA* values of: (a) less than .05 to suggest good fit, (b) .08 for reasonable fit, and (c) values beyond .10 to indicate model misfit (MacCallum, Browne, & Sugawara, 1996). Hu and Bentler (1999) supported values less than .05 to assume adequate model fit; however, the authors cautioned the use of an *RMSEA* universal cutoff of .05 cutoff with small sample sizes because of the tendency to over-reject correct models. Also, the range of optimal values between the reduction of Type I and Type II errors for the *RMSEA* were computed by Sivo et al. (2006) for the sample sizes of: (a) N = 150 (*RMSEA* = .06 - .01); (b) N = 250 (*RMSEA* = .05 - .01); (c) N = 500 (*RMSEA* = .03 - .01); (d) N = 1,000 (*RMSEA* = .03 - .005); (e) N = 2,500 (*RMSEA* = .02 - .003); and (f) N = 5,000 (*RMSEA* = .01 - .002). Notice that the *RMSEA* range is large for small sample sizes and reduces as the sample size increases, where a global cutoff of greater than or equal to .05 over-rejects the correct model with N <150 and under-rejects the incorrect model with N > 500. Subsequent researchers have provided supportive evidence of the *RMSEA's* tendency to over-reject correct models with small sample sizes (N < 200; Chen et al., 2008; Fan & Wang, 1998; Sharma et al., 2005). Moreover, Sivo et al. reported that global cutoff values of .05 will tend to under-reject incorrect models with large sample sizes (N > 500). The collective evidence indicated that the *RMSEA* is influenced by variations in sample size, leading to the tendency of the *RMSEA* to over-reject models with small sample sizes and under-reject models with large sample sizes.

In regard to model complexity, Chen et al. (2008) applied a *RMSEA* cutoff value of < .05 for a simulated, correct, three factor CFA and found a tenfold decrease in the percentage of models rejected when the same model with six additional observed variables was examined. Although, the Chen et al. findings were based on six additional observed variables added to the CFA model, the trend may extrapolate to LGC models where the addition of a single observed variable (i.e., wave of data) may result in worse model fit according to the *RMSEA* value (i.e., higher RMSEA values). However, Sharma et al. (2005) simulated CFA model complexity when they increased the number of factors and indictors (i.e., see description under incremental fit indexes), and reported a negligible effect for sample size and model complexity on the *RMSEA* value. Furthermore, Cheung and Rensvold (2002) simulated conditions similar to LGC modeling and found the *RMSEA* was not affected by variations in the number of observed variables or latent factors.

In summary, *RMSEA* was hypothesized in the current study to be influenced by sample size with inappropriate estimates that occur in small sample sizes (N < 200). Due to the common variations in LGC model complexity, which include few latent factors (e.g., 2- 4) and observed variables (e.g., 3-6), it wass hypothesized that significant differences would occur for the conditions of waves of data and model complexity. However, the effect sizes and mean values of the *RMSEA* were expected to suggest that the significant differences lack practical importance in terms of assessing model fit due to the anticipated small effect size.

Latent Growth Curve Model Fit Investigation

Coffman and Millsap (2006) initiated the model fit research specific to LGC analysis and investigated model misspecification related to the shape of growth. One condition of model misspecification was constructed to represent a small quadratic term examined with the use of two fit indexes. The χ^2 and *RMSEA* displayed poor fit for the linear model and adequate fit for the quadratic model; therefore, the fit indexes suggested a preference for quadratic growth, even when the majority of individual trajectories exhibited linear growth. Interestingly, estimates of a covariate effect on a univariate LGC model over five time points suggested similar parameter estimates for both the linear and quadratic models, despite the lack of fit for the linear model. As a result, Coffman and Millsap concluded that fit indexes for LGC models may be influenced by shape misspecification and suggested estimation of log likelihood values ($-2P_{LL}$) for each individual subject as a measure of within person fit, as well as the fit indexes to assess global fit of the overall model.

While the novel study conducted by Coffman and Millsap (2006) provided interesting insight into the consequences of LGC model misspecification, application studies have not utilized investigations into individual-level fit statistics. In addition, in the Coffman and Millsap preliminary study, they reviewed a single condition of model misspecification, reducing the external validity. Due to the practical focus of this dissertation to examine characteristics found in application studies in the social and behavioral sciences, global fit indexes are of fundamental interest; however, the avenue of model misspecification and individual level fit requires additional attention in future research endeavors.

Summary of Model Fit Indexes

The model fit literature consists of a gap specific to LGC modeling environments; nevertheless, LGC modeling educators and applied researchers endorse and frequently apply the χ^2, *NNFI*, *CFI*, and *RMSEA*. Excluding the Coffman and Millsap (2006) investigation, this dissertation is the first investigation to examine the influence of typical LGC modeling environments on selected global fit indexes. Drawing on the relevant SEM literature, the hypothesis for Research Question 1 suggests that model complexity would affect convergence rates, with complex models based on low sample sizes displaying the lowest convergence rates. The hypothesis for Research Question 2 suggests that sample sizes would influence all fit indexes, especially when $N < 200$. It is expected that the χ^2 will be most affected with a general trend to over-reject correct models in small and large sample size conditions. The *NNFI*, *CFI*, and *RMSEA* were expected to be less influenced with a general trend to over-reject correct models in small sample size conditions and to under-reject incorrect models with large sample sizes. Regarding Research Questions 2 and 3, it was expected that varying waves of data and LGC model complexity would influence all four fit indexes following the pattern to over-reject correct complex models (i.e., and more waves of data), as well as to under-reject correct, parsimonious models (i.e., fewer waves of data). The χ^2 and *CFI* were hypothesized to be most influenced by model complexity and increasing waves of data, followed by the *NNFI*. The influence of model complexity and waves of data on the *RMSEA* values was expected to have limited practical importance as displayed by the effect sizes and mean values.

Chapter Two Summary

In Chapter II, I conveyed the procedures and relevant literature of LGC models and assessment of model fit with the χ^2, *NNFI*, *CFI*, and *RMSEA*. As discussed, LGC modeling is a flexible tool that can model various types of longitudinal research environments by the estimation of variance-covariance matrices and mean vectors. Of particular interest to applied LGC modeling researchers is how variations in sample size, waves of data, and model complexity (i.e., defined as linear univariate LGC model, quadratic univariate LGC model, multivariate linear LGC model, and linear univariate LGC model with a covariate) may affect estimation of model fit. Due to the lack of literature that pertains to LGC model fit, inferences were drawn from SEM simulation studies with similar model structures. Based on the review of SEM literature, it was hypothesized that all four fit indexes would be afflicted by variations in sample size, waves of data, and model complexity when $N < 200$; however, the true questions lie in the magnitude of difference and practical relevance to applied researchers. In Chapter II, the background literature was established to allow for a discussion of methods utilized to examine the functioning of model fit indexes in LGC modeling environments, which are presented in Chapter III.

CHAPTER III

METHODOLOGY

In this chapter, I address the methods applied to investigate the functioning of fit indexes in latent growth curve (LGC) models under conditions of: (a) sample sizes, (b) waves of data, and (c) model complexity. In this dissertation, LGC modeling simulation techniques were applied to answer the following four questions:

Q1 Do model convergence rates vary under conditions of sample size, waves of data, and model complexity?

Q2 Do fit indexes (χ^2, *NNFI*, *CFI*, and *RMSEA*) differ under varying conditions of sample size?

Q3 Do fit indexes (χ^2, *NNFI*, *CFI*, and *RMSEA*) differ under varying conditions of waves of data?

Q4 Do fit indexes (χ^2, *NNFI*, *CFI*, and *RMSEA*) differ under varying conditions of model complexity, defined in the current dissertation as a: (a) univariate linear LGC model, (b) quadratic LGC model, (c) multivariate linear LGC model, and (d) a linear LGC model with a covariate?

Models to Be Tested

The four research questions were investigated with the utilization of two types of LGC models. As established in Chapter II, LGC models are divided into: (a) unconditional LGC models (i.e., measurement models); and (b) conditional models, which include additional structural components (i.e., inclusion of a covariate). In this dissertation, unconditional LGC models were examined in three conditions of model complexity including: (a) univariate linear LGC models, (b) quadratic univariate LGC models, and (c) linear multivariate LGC models. Conditional LGC models were examined in a single condition of model complexity based on a univariate linear LGC model with a time-invariant covariate. The two population models were described in Chapter II; however, they are briefly restated in the following section. The unconditional LGC model, which jointly represents both Level 1 and 2 models, is displayed in Equation 3.1.

$$y_{it} = \Lambda\left(\alpha_{\eta_i} + \zeta_{it}\right) + \varepsilon_{it} \text{、} \tag{3.1}$$

In this equation, y_{it} is a vector of the observed measures for each ith participant (i = 1, 2…N) at each t time point (t = 0, 1… t-1), Λ is a matrix of fixed factor loadings to represent time, $\alpha_{\eta i}$ is a vector of latent means for each latent factor (i.e., η_i, represents the growth trajectory), ζ_{it} is a vector of the individual deviations away from the latent means, and ε_{it} is a vector of the residuals (Bollen & Curran, 2006). The conditional LGC model, which jointly represents both Level 1 and 2 models, is presented in Equation 3.2.

$$y_{it} = \Lambda\left(\alpha_{\eta_i} + \beta_{\eta_i} x_i + \zeta_{it}\right) + \varepsilon_{it} \tag{3.2}$$

All parameters were described in Equation 3.1 except for $\beta_{\eta i}$ which is a vector of the random parameters for each latent factor and x_i represents a single time invariant covariate. All parameters were estimated with use of maximum likelihood techniques (F_{ML}), the predominant estimation method found in LGC modeling applications and simulations. The population means, variance, and covariance parameters are discussed according to each independent variable.

Coding of time in LGC modeling is represented by Λ which describes the paths from the observed variables (e.g., waves of data) to each latent factor (e.g., intercept, slope, and additional factors). In all LGC models, the paths from the observed variables to the latent intercept are set to one, with paths from the observed variables to the latent slope, and additional factors, fixed to represent the coding of time (λ_{it}). Frequently, linear trend contrasts are utilized to represent time in LGC modeling and were applied in this dissertation (λ_{it} = 0, 1, 2,… t-1).

Independent Variables

The current analysis is among the first simulation studies to focus on model fit indexes within LGC modeling data environments. As in all novel areas of research, the most fundamental variations need to be considered prior to examination of more complex conditions. Consequently, an assortment of independent variables and potential levels need to be inspected in regard to their corresponding functioning of fit indexes. The current investigation examined only conditions deemed to be essential to most LGC modeling applications. The rationale for the selected independent variables and associated levels within each condition were based on three predominant considerations: (a) the review of conditions employed in current LGC modeling applications in behavioral and social sciences journal articles; (b) the conditions examined in previous simulation studies of LGC models (Fan & Sivo, 2005; Leite, 2007; Muthén & Curran, 1997, Muthén & Muthén, 2002) and studies of fit indexes in SEM (Hu & Bentler, 1999; Sivo et al., 2006); and (c) a combination of a reasonable number of conditions to allow for proper interpretation of the influence of each variable being examined. The estimated values for the parameters in the models were based on previous LGC modeling simulations, which have typically followed the procedures suggested by Muthén and Muthén (2002), who are known as experts in the field and are the developers of the software being used in the current study. The levels and justification for independent variables investigated in the proposed dissertation are described below.

Sample Size

The independent variable of sample size included five levels (*N* = 100, 250, 500, 1,000, and 2,500). This range encompasses the majority of LGC modeling applications reviewed (96.5%). Previous LGC modeling simulation studies have included a similar range, excluding the most extreme level of 2,500, for example: (a) 200-500, Hertzog et al., 2006; and (b) 100-1,000, Leite, 2007; Muthén & Curran, 1997). However, extremely large sample sizes (e.g., > 2,000) are common in general SEM simulation studies of model fit to examine the tendency of chi-square tests (χ^2) to produce biased estimates of fit with large sample sizes (Hu & Bentler, 1999; Sivo et al., 2006). Therefore, it is critical to examine an extreme level of sample size in the evaluation of model fit, although this has not been typical in LGC modeling simulations.

Waves of Data

The waves of data included four levels of 3, 4, 5, and 6, creating a range which includes 89.6% of the applied studies reviewed. The four levels mimic conditions found in previous LGC modeling simulations (Hertzog et al., 2006; Muthén & Muthén, 2004; Leite, 2007; Sivo et al., 2006), with the exclusion of Muthén and Curran's (1997) simulation study which investigated seven waves of data. Six or more waves of data were rarely found in LGC modeling applications; thus, three and six waves of data were used as the extreme condition of waves of data.

Model Complexity

Based on the complexity of traits examined in LGC modeling applications in the social and behavioral sciences, it is reasonable to believe that most applied studies examine complex traits. In addition, SEM methodologists have debated between preference for model parsimony vs. proper representation of change through more complex representations (Raykov & Marcoulides, 1999). Consequently, conditions of model complexity typically require decisions by LGC modeling researchers. The four conditions of model complexity examined were chosen based on common decisions required in LGC modeling application: (a) What shape of growth occurs in this trait? (e.g., linear or quadratic); (b) What type of model represents the trait? (e.g., univariate or multivariate); and (c) Does a covariate account for variations in growth on the trait(s)? (e.g., inclusion or exclusion of a covariate). The condition of univariate linear LGC modeling is the most parsimonious LGC model examined, with model complexity increasing to a quadratic LGC model, a LGC model with inclusion of a time invariant covariate, and to the most complex multivariate LGC model. Figure 3.1 is the expanded matrix notation of Equations 3.1 and 3.2, with an emphasis on the portions included in each condition of model complexity.

Covariate Univariate

$$\begin{matrix} y_{i1} \\ y_{i2} \\ y_{i3} \\ \cdots \\ y_{i(t-1)} \end{matrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ \cdots & \cdots \\ 1 & t-1 \end{bmatrix} \begin{bmatrix} \alpha_1 + \beta_1 x + \varsigma_1 \\ \alpha_2 + \beta_2 x + \varsigma_2 \\ \alpha_3 + \varsigma_3 \\ \alpha_4 + \varsigma_4 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \cdots \\ \varepsilon_{t-1} \end{bmatrix}$$

Quadratic Multivariate

Figure 3.1. LGC model highlighted for model complexity conditions

For example, the linear LGC model parameters (e.g., parameters included in the red univariate linear LGC model, excluding the green covariate term) include the population means and variances for two latent factors that represent the growth trajectories. In this dissertation, population means of the latent intercept (α_1) were set to zero, the linear latent slope (α_2) was fixed at 0.2, with the variance of the linear slope factor (ψ_{22}) at 0.1, and the variance of the latent intercept (ψ_{11}) at 0.5, which have been suggested to mimic common LGC application conditions (Leite, 2007; Muthén & Muthén, 2002). Therefore, the latent growth curve models estimated have a baseline value of zero and a slight, positive linear growth over the selected time period. Furthermore, the proportion of the variance of linear slope factor to the variance of the intercept factor represents a 1:5 ratio, as suggested by Muthén and Muthén, to replicate commonly encountered variances in applied longitudinal research environments. The covariance between the intercept and linear slope factor (ψ_{12}) will be set to 0.2, to represent a small relationship. While the population means and variance vary depending on model complexity, the error variances of the observed variables (ε_{it}) were set at 0.5 for each wave of data (e.g., observed variable) in all models and were assumed henceforth at this value. By setting the error variances to 0.5, the corresponding R^2 values of the observed variables mimic commonly found conditions in applied longitudinal environments (i.e., $R^2(y_1) = .50$, $R^2(y_2) = .55$, $R^2(y_3) = .64$, $R^2(y_4) = .74$; Muthén & Muthén). This paragraph described the model parameters to generate data for a linear LGC model, and the following subsections describe the rationale and population values for the three other conditions of model complexity.

Quadratic growth. The dominant form of nonlinear growth discovered in the review of application studies was a quadratic growth trajectory. Although linear growth is more frequently applied, utilization of linear growth models may be attributed to convenience, as opposed to theory or strong contextual evidence (Burchinal & Appelbaum, 1991). Coffman and Millsap's (2006) LGC modeling simulation study of model misfit was focused on the representation of growth and utilized a linear and quadratic trajectory of growth. For models which examine quadratic growth, the additional paths from the observed variables to the latent quadratic slope were fixed with a quadratic polynomial representation (λ_t = 0, 1, 4, 16,…*t* -1). The quadratic model included identical variances, covariance, and population mean of the latent intercept described for the linear LGC model, with additional parameters to represent the quadratic growth as displayed in blue in Figure 3.1 (i.e., excluding the green covariate representation). Following a similar rationale to that of the univariate linear LGC model, the latent intercept was set at zero (α_3), the latent linear slope factor (α_2) was fixed at 0.1, the quadratic slope factor (α_3) was fixed at 0.2, representing a slight positive quadratic increase in the latent trait over the selected time period. The variance of the quadratic factor (ψ_{33}) was fixed at 0.1 to represent a 1:5 ratio with the variance of latent intercept factor, to replicate typical longitudinal research environments. The covariance between the latent intercept and the linear slope factor (ψ_{12}) was set at 0.1. The covariance between the latent intercept and the quadratic slope factor (ψ_{13}) was set a 0.2, which is the identical relationships set in the linear LGC model between the intercept and slope factor. Finally, the covariance between the linear slope factor and the quadratic slope factor (ψ_{23}) was set at .05, representing the smallest relationship among the latent variables.

Multivariate growth. In regard to LGC modeling simulation studies, although Leite (2007) investigated a multivariate factor-of-curve model, neither his study nor any of the other LGC modeling simulations reviewed included an associative multivariate LGC model such as the one applied in this dissertation. The majority of LGC applications reviewed used a univariate model of growth with limited use of a multivariate representation of growth. While application studies have infrequently applied multivariate LGC models, it may be understandable to assume that multivariate applications will increase in the coming years for two reasons. First, within the social and behavioral sciences, the complexity of research questions is escalating in an attempt to represent the complex phenomena of human behavior in which multivariate growth may have greater theoretical support. Secondly, LGC modeling software programs allow for easy programming of multivariate growth in LGC models, as compared to software programs used to estimate other types of growth curve models (e.g., HLM-6). Therefore, an associative multivariate LGC model was chosen due to an expected increase in application among the social and behavioral sciences and it has yet to be investigated.

For the associative multivariate LGC model, the parameters described for the univariate linear LGC model are assumed in addition to parameters, which represent the second growth trajectory and the relationship among the latent factors of the growth trajectories. In Figure 3.1, the multivariate model includes parameters in the orange area, excluding the green area representing the covariate parameters. The latent intercept factor of the second trait (α_3) was fixed at 0.5, and the latent slope factor of the second trait (α_4) was fixed at 0.1. In contrast to the growth in the first latent trait, the trajectory of the second latent trait begins slightly higher, representing a lower rate of growth over time.

The variance of the linear slope factor for the second trait (ψ_{44}) was fixed at 0.1, and the variance of the latent intercept of the second trait (ψ_{33}) at 0.5, suggested to mimic longitudinal data environments with a slope to intercept variance ratio of 1:5 (Muthén & Muthén, 2002). In the first trait, the covariance between the intercept and slope factors (ψ_{12}) was set a 0.2, whereas in the second latent trait the same relationship (ψ_{34}) was set a 0.1. All other covariances among latent variables were set to zero. Therefore, the relationship between the initial level and the rate of growth was stronger in the first trait than in the second trait.

Time invariant covariate. The simulated models included representation of a univariate linear LGC model with a single time invariant covariate, similar to the LGC model used in Hertzog et al.'s (2006) simulation study of statistical power. While some application studies reviewed included representation of multiple covariates, the majority of studies included only a single covariate in the LGC model. In Figure 3.1, the univariate linear LGC model with a single time invariant covariate is represented by the parameters in the green and red areas. The single time-invariant covariate, x_i, was set with a mean of .5 and a variance of .25, representing a dichotomous covariate (e.g., gender). The random regression coefficients to both the latent intercept and latent slope factor, β_{it}, were set at 0.2, with a variance of 0.09 (Muthén & Muthén, 2002). The values were selected due to their correspondence to a representation of medium effect size (d = .63), which was also used in a previous LGC modeling simulation with a covariate (Hertgoz et al.; Muthén & Muthén).

Summary of Independent Variables

In summary, the LGC model simulation included the independent variables of: (a) sample size (100, 250, 500, 1,000, and 2,500); (b) waves of data (3, 4, 5 and 6); and (c) model complexity (i.e., univariate linear, univariate quadratic, inclusion of a covariate, and multivariate linear LGC model). Due to the requirements of identification, one cell could not be computed; thus, a completely crossed design was not applied. The condition of a univariate quadratic LGC model with three waves of data results in an under identified model, which does not allow for estimation of model fit indexes.

Dependent Variables

The dependent variables included four fit indexes, χ^2, *NNFI*, *CFI*, and *RMSEA*, which were described in detail in Chapter II. To briefly reiterate, the χ^2 is the most historic measure and uses significance testing; however, the *NNFI*, *CFI*, and *RMSEA* are all descriptive indexes based on a continuum of fit. The *NNFI* and *CFI* are incremental fit indexes that assess a ratio between the hypothesized model and baseline model and the *RMSEA* is an absolute fit index that assesses approximate fit. The equations used to compute the fit indexes are presented in Table 3.1. Moreover, LGC modeling procedural guides endorse the use of the four selected fit indexes in application (Bollen & Curran, 2006; Duncan et al., 2007; Preacher et al., 2008). In addition, the four selected indexes are reported as defaults in most LGC modeling software (i.e., M*plus*, LISREL, EQS), which results in frequent reporting in applied studies.

Table 3.1

Fit Indexes and Recommended Cutoff Values

Fit Index	Range	Perfect fit	Hu and Bentler's (1999) cutoff values
$\chi^2{}_b = (N - 1)\, F_{ML}$	χ^2 distribution	Non-significant result (adequate fit)	n.a.
$NNFI = \dfrac{\chi_b{}^2 / df_b - \chi_h{}^2 / df_h}{\chi_b{}^2 / df_b - 1}$	0 – infinity but generally between 0 -1.0	1.0	.95 – 1.00
$CFI = 1 - \left(\tilde{d}_h / \tilde{d}_b\right)$	0 – 1	1.0	.96 – 1.00
$RMSEA = \sqrt{\max\left[\left(\chi_h{}^2 - df_h\right) / \left(n \times df_h\right), 0\right]}$	0 – infinity	0.0	.05 – < 0.1

Note. *FML* = full information maximum likelihood estimation, *N* = sample size, $\chi_b{}^2 = \chi^2$ test statistic for the baseline model, df_b = degree of freedom for the χ^2 test statistic for the baseline model, $\chi_h{}^2$ = likelihood ratio χ^2 of the hypothesized model, df_h = degrees of freedom of the χ^2 hypothesized model, $\tilde{d}_h = \max(d_h, 0)$, $\tilde{d}_b = \max(d_b, d_h, 0)$, $d_h = \left(\chi_h{}^2 - df_h\right)/n$, and $d_b = \left(\chi_b{}^2 - df_b\right)/n$.

In addition to the four fit indexes examined in this dissertation, there is an assortment of additional measures of model fit, which include the: (a) goodness of fit index (*GFI*; Jöreskog & Sörbom, 2001); (b) adjusted goodness of fit index (*AGFI*; Jöreskog & Sörbom); (c) normed fit index (*NFI*; Bentler & Bonett, 1980); (d) root mean square residual (*RMR*; Jöreskog & Sörbom); and (e) McDonald's centrality index (*Mc*;

McDonald, 1989), among others. Previous general SEM simulation studies of fit indexes suggested that the four selected indexes have potential benefit over the excluded alternatives (Hu & Bentler, 1998, 1999). Also, the χ^2/df ratio was computed in the current study based on methodological preference over the χ^2 value alone. In summary, the χ^2 (χ^2/df), *NNFI*, *CFI*, and *RMSEA* were examined in the conditions of sample size, waves of data, and model complexity in this dissertation.

Generating the Data

The data were generated by use of Monte Carlo procedures available in M*plus* (Muthén & Muthén, 2004). The parameters were fixed or varied as previously discussed according to model conditions. A random seed was created based on the random numbers generator in Microsoft Excel, then input into the M*plus* syntax for data generation. One thousand replications were generated per design condition. The M*plus* syntax for this dissertation can be found in Appendix A.

Also, convergence and inadmissible solutions for each data condition were reported. To account for the large amount of inadmissible solutions (non-plausible values), numerous additional datasets were estimated, from which only the first 1,000 datasets that converged and had admissible solutions (plausible values) were used in the analysis. One thousand replications is the average number of replications found in other LGC modeling simulations (e.g., Liete, 2007). The data were checked for validation by examination of the number of replications reported in the analysis and the means and variances among the fit indexes. Based on the three research questions (excluding Q1 regarding convergence rates), the simulation included 75 cells, with 1,000 replications in each cell, resulting in a total 75,000 datasets generated and analyzed regarding model fit.

Data Analysis

Four fit indexes (i.e., x^2, *NNFI, CFI,* and *RMSEA*) were produced for each dataset resulting in a total of 300,000 indexes to be interpreted. To collect the indexes, an SPSS (Version 15.0) program selected the desired statistics from the *Mplus* output and placed them in a format which could be analyzed. Once the data were represented in an interpretable SPSS data set, descriptive and inferential analyses were conducted. First, a 5 x 4 x 4 incomplete factorial ANOVA was conducted for each of the four dependent variables (χ^2, *NNFI, CFI,* and *RMSEA*), as well as the χ^2/df ratio, to determine if LGC modeling design characteristics *influence* the fit index values. Partial eta-squared effect sizes were examined for all main effects and interactions to descriptively quantify the magnitude of the effect and were interpreted as: (a) .09 as a small effect, (b) .14 as a medium effect, and (c) .22 as a large effect (Gamst, Meyers, & Guarino, 2008). The mean fit indexes were reported descriptively in two formats. Line graphs are presented in the text for χ^2 and *RMSEA* and tables can be found in Appendix B. Due to the large sample size, the statistical power in this analysis was high which increased the chance to obtain significant outcomes. Therefore, an alpha level of .01 was applied to determine significance and more weight was placed on the effect sizes and mean values when interpreting the results.

Due to the practical focus of this dissertation, planned contrasts were conducted to answer common questions as encountered by applied researchers. For example, an applied researcher, who designs a longitudinal study, may debate between using N = 100 or using additional resources to increase sample size to N = 250, but would not typically contemplate between N = 100 and N = 1,000. Therefore, planned comparisons assessed whether a significant difference was obtained when small sample sizes are compared to conditions with larger sample sizes (N = 100 vs. N = 250), while also examining decisions among moderate sample sizes (N = 250 vs. N = 500), and large sample sizes (N = 500 vs. N = 1,000 and N = 1,000 vs. N = 2,500). Similarly, researchers may utilize resources to add an additional wave of data; therefore, planned comparisons assessed significant differences with one additional wave of data (t = 3 vs. t = 4; t = 4 vs. t = 5; t = 5 vs. t = 6). Furthermore, the test of main effects for model complexity lacks applicability for applied researchers because the conditions of model complexity are utilized for different purposes. For example, a LGC modeling researcher may debate between a univariate linear or multivariate linear LGC model; however, it would be rare for a researcher to change from a quadratic univariate model to a linear multivariate model. Therefore, a unvariate linear LGC model were compared to all other conditions of model complexity through planned comparisons (e.g., unvariatate vs. quadratic, univariate vs. multivariate, univariate vs. inclusion of covariate). An alpha level of .01 was applied to examine the planned comparisons. Cohen's *d* effect sizes were computed for all planned comparisons and were interpreted as .2 for a small effect, .5 for a moderate effect, and .8 for a large effect (Field, 2005).

In addition, the Type I error rate were computed in all 75 conditions to investigate the frequently applied *methods* for determining model fit. In this simulation, correctly specified models were estimated; therefore, theoretically, all fit statistics should imply adequate model fit. The Type I error rate is defined, in this dissertation, as the proportion of models that are rejected based on Hu and Bentler's (1999) proposed cutoff values.

Summary of Method

The simulation analysis provided insight into the performance of common fit indexes in LGC model environments by the generation and analysis of 75 data conditions (i.e., with 1,000 replications) among variations in sample size, waves of data, and model complexity. The levels of independent and dependent variables were designed, based on conditions found in application and simulation LGC modeling studies, in order to aid in the understanding of fit indexes for LGC modeling methodologists and applied researchers. A 5 x 4 x 4 incomplete factorial ANOVA was conducted for each of the four dependent variables, along with partial eta-squared effect sizes for the main effects and interactions. Planned comparisons were computed for significant main effects, along with Cohen's *d* effect sizes. Finally, Type I errors were computed to determine the proportion of models rejected based on frequently used methods for determining acceptable model fit. The examination of four common fit indexes in LGC models provided novel information in regard to the functioning of fit indexes, which can be utilized to provide applied longitudinal researchers with valid methodological information.

CHAPTER IV

RESULTS

The purpose of this dissertation was to provide guidance for applied longitudinal researchers regarding the evaluation of latent growth curve (LGC) model fit. In Chapter IV, I conveyed the results of four commonly used measures of global fit to understand their performance in correctly estimated LGC models under various conditions based on three variables: (a) sample size, (b) waves of data, and (c) model complexity (e.g., defined in this dissertation as a univariate LGC model, quadratic LGC model, multivariate LGC model, and a univariate LGC model with the inclusion of a time-invariant covariate). The specific fit indexes investigated include: (a) the likelihood ratio chi-squared (χ^2), (b) nonnormed fit index (*NNFI*), (c) comparative fit index (*CFI*), and (d) the root mean squared error of approximation (*RMSEA*).

The chapter begins with a discussion of model convergence and inadmissible solution rates for the 75 LGC modeling conditions estimated. Subsequently, five incomplete factorial analysis of variance (ANOVA) designs are discussed for each of the measures of global fit, according to the interactions, main effects, and planned comparisons among the three independent variables. To interpret the magnitude of the effect, partial eta-squared effect sizes (η^2) are presented for main effect and interaction findings, while Cohen's *d* effect sizes are presented for the planned comparison results. Furthermore, the mean values of the five measures of fit are presented, as well as the

Type I error rates, defined as the proportion of models rejected based on frequently utilized criteria for determining acceptable model fit, such as Hu and Bentler's (1999) proposed cutoff values. Chapter IV summarizes detailed evidence examining the influence of design characteristics (i.e., sample size, waves of data, and model complexity) on commonly utilized measures of LGC model fit, with supplemental information presented in the Appendices.

Model Convergence and Inadmissible Solutions Rates

Interestingly, all conditions achieved 100% convergence; however, the inadmissible solutions rate was problematic. Inadmissible solutions occur when maximum likelihood estimation results in an implausible value (i.e., also known as a *Heywood case*), including one or more of the following conditions: (a) a negative latent intercept variance, (b) a negative latent slope variance, or (c) a correlation between the latent intercept and slope factors beyond the acceptable range (i.e., -1 to 1; Kline, 2005; Leite, 2007). Consequently, inadmissible solutions should not be interpreted by longitudinal researchers, as noted in the error message that occurs in structural equation modeling (SEM) software (i.e., non-positive definite variance-covariance matrix).

Table 4.1 displays the percent of inadmissible solutions for the first 1,200 replications generated in each LGC modeling condition. Conditions with small sample sizes and few waves of data encountered the largest inadmissible solution rates, particularly in quadratic and multivariate LGC modeling conditions. The lowest rates were observed for all models with six waves of data and N = 2,500, ranging up to 49.9% to 74.5% for models with three waves of data and N = 100. Notice, if a researcher were to begin with the design conditions of a univariate LGC model with N = 100 and four waves of data, the inadmissible solutions rate would be 43.1%. To include a single covariate the rate would decrease to 42.6%, to change to a multivariate model the rate would increase to 59.4%, and to change to a quadratic LGC model the inadmissible rate would increase to 90.3%. Therefore, the highest rates were observed in quadratic models, followed by multivariate models, then univariate models, and finally, the lowest rates were observed for the covariate model. In conclusion, the percentage of inadmissible solutions increased in conditions with small sample sizes and few waves of data, especially in quadratic and multivariate model conditions.

Table 4.1

Percent of Inadmissible Solutions

Waves of Data	Sample Size				
	100	250	500	1,000	2,500
Univariate Linear LGC Model					
3	52.0%	46.9%	43.6%	39.3%	30.3%
4	43.1%	34.5%	27.3%	19.3%	7.7%
5	34.8%	23.5%	11.6%	4.9%	<.1%
6	26.4%	13.0%	5.6%	1.1%	0%
Quadratic LGC Model					
4	90.3%	84.3%	76.7%	64.2%	51.1%
5	69.3%	46.8%	33.1%	17.2%	5.3%
6	48.6%	26.2%	12.7%	4.3%	.3%
Multivariate Linear LGC Model					
3	74.5%	64.2%	55.5%	45.7%	33.3%
4	59.4%	40.8%	31.9%	22.0%	7.3%
5	46.1%	27.9%	16.8%	5.8%	.4%
6	36.5%	17.4%	6.4%	.7%	0%
Covariate Linear LGC Model					
3	49.4%	44.4%	44.3%	38.1%	31.3%
4	42.6%	36.1%	28.2%	17.6%	6.2%
5	37.3%	22.8%	15.1%	4.2%	.9%
6	29.3%	12.5%	5.7%	1.2%	0%

Note. Values based on the first 1,200 datasets generated.

Due to the applied nature of this dissertation, inadmissible solutions were removed and replaced with admissible solutions (i.e., plausible values). Although an additional 200 replications were estimated, in conditions with small sample sizes and few waves of data additional simulations were conducted to achieve 1,000 admissible datasets in each condition. In summary, even though all models achieved convergence, the high rate of inadmissible solutions was problematic, particularly in conditions with small sample sizes, few waves of data, and in quadratic and multivariate models.

Assessment of Differences in LGC Model Fit

The *influence* of LGC modeling design characteristics on global measures of model fit was investigated with five incomplete factorial ANOVAs and effect sizes for the χ^2, χ^2/df, *NNFI, CFI,* and *RMSEA.* The ANOVAs were used to examine the statistical significance of the design conditions (alpha level = .01), while the effect sizes exhibited the magnitude of the effects. Table 4.2 displays the effect sizes and significant differences for the five measures of model fit. Detailed ANOVA tables for the measures of fit are presented in Appendix B.

As expected, the majority of main effects and interactions were significant, most likely due to the large amount of statistical power typically encountered in simulation studies such as the current study. Therefore, the effect size results, which varied in magnitude among the fit indexes and model conditions, were weighted more heavily than significance testing in interpreting the results. Clearly, the χ^2 assessment of fit was highly, negatively affected by additional waves of data and increasing model complexity (excluding the quadratic model), while the *RMSEA* and *CFI* displayed a moderate tendency to suggest worse fit in small sample sizes, and the χ^2/df and *NNFI* displayed no notable effects. The detailed findings are discussed below, partitioned by each measure of model fit.

Table 4.2

Effect Sizes and Significant Differences for Model Fit Indexes

Condition	*df*	*Effect Sizes*				
		χ^2	χ^2/df	*NNFI*	*CFI*	*RMSEA*
		Partial Eta Squared (η^2) Effect Sizes[a]				
SS	4	<.01*	<.01*	<.01*	.11*	.13*
W	3	.75*	.00	.00*	<.01*	<.01*
C	3	.81*	.00	<.01*	.01*	<.01*
SS x W	12	<.01*	.00	<.01*	<.01*	<.01*
SS x C	12	<.01*	.00	<.01*	.01*	<.01*
W x C	8	.58*	.00	.00	<.01*	<.01*
SS x W x C	32	<.01*	.00	<.01*	<.01*	<.01*
		Cohen's *d* Effect Sizes[b]				
SS: 100 vs. 250	1	.02*	.03*	<.07*	.25*	.56*
SS: 250 vs. 500	1	.01	.02	<.03	.36*	.25*
SS: 500 vs. 1000	1	<.01*	.00	<.01	.35*	.22*
SS: 1000 vs. 2500	1	.01*	<.01	<.01	.43*	.46*
W: 3 vs. 4	1	.76*	-	.03*	.12*	.07*
W: 4 vs. 5	1	.72*	-	.02	.01	.02*
W: 5 vs. 6	1	.56*	-	.01	.05*	.05*
C: uni. vs. quadratic	1	-.26*	-	<.01	.29*	<.01*
C: uni. vs. multivariate	1	1.83*	-	.05*	.04*	.13*
C: uni. vs. covariate	1	.81*	-	<.01	.02*	.02*

Note. *NNFI* = Nonnormed Fit Index; *CFI* = Comparative Fit Index; *RMSEA* = Root Mean Squared Error of Approximation; SS = sample size; W = waves of data; C = model complexity.
[a]η^2 effect sizes of .09 or lower were interpreted as a small effect, .14 as a moderate effect, and .22 as a large effect. [b]*d* effect sizes were interpreted as .2 for a small effect, .5 for a moderate effect, and .8 for a large effect.
* $p < .01$.

Chi-Squared Ratio Test

Despite the significant difference observed for the χ^2 assessment of fit among conditions of sample size, the effect size displayed a negligible effect (i.e., <.01% of the variance explained). As previously discussed, the significant difference is most likely related to the high statistical power in this analysis, increasing the ability to detect minor differences. Consequently, the significant difference for χ^2 among sample size conditions lacks practical merit and will not be interpreted as a genuine effect. Similar trends were observed for most fit index comparisons (i.e., significant differences with negligible effect sizes), and will be interpreted in the same manner (i.e., concluding that the significant difference lacks practical value). For example, the interactions including sample size were also found to display negligible effects, despite the significant differences.

The χ^2 values displayed a large amount of disparity among the conditions of waves of data and model complexity, explaining 75% and 81% of the variation, respectively. As additional waves of data were added to the LGC model, the χ^2 values suggested a decrease in model fit. Similarly, as model complexity increased, the χ^2 values implied a decrease in model fit, excluding the quadratic LGC model. The planned comparisons for waves of data displayed moderate effect sizes (d = .56 - .76).

Among the model complexity conditions, the comparison of univariate to multivariate models displayed the largest effect (d = 1.83), followed by the comparison between univariate and covariate models that also displayed a large effect (d = .81). Therefore, univariate models displayed better model fit than covariate models, and much better fit than multivariate models. The comparison between univariate and quadratic models displayed a small effect (d = -.26); however, the direction of effect was negative suggesting that χ^2 values tend to be lower in quadratic models. Therefore, quadratic LGC models will display slightly better model fit than univariate linear LGC models.

The interaction between waves of data and model complexity accounted for 58% of the variation in χ^2 values. To understand the nature of the interaction, a test of simple main effects was conducted as presented in Table 4.3. Notice, within all four model complexity conditions, varying the waves of data was found to significantly affect the χ^2 values. More specifically, in all model complexity conditions as waves of data were added to LGC models, the χ^2 values suggested worse model fit.

Table 4.3

ANOVA Table of Simple Main Effects for the χ^2

Source	*SS*	*df*	*MSE*	*F*-value	*p*-value
Waves within univarite models	643,436.08	3	214,478.69	1,628.40	<.01*
Waves within quadratic models	2,161,568.08	3	720,522.87	5,470.48	<.01*
Waves within multivariate models	9,444,146.30	3	3,148,048.80	23,901.16	<.01*
Waves within covariate models	834,688.43	3	278,229.48	2,112.42	<.01*
Error	9,876,620.44	74987	131.71		
Total	22,960,459.88	74999			

Note. *SS* = sums of square; *df* = degrees of freedom; *MSE* = mean squared error; SS = sample size; W = waves of data; C = model complexity.
*$p < .01$.

To further investigate the nature of the interaction, Figure 4.1 graphically represents the mean values for the χ^2 assessments of fit, which can also be found in table format in Appendix C. Notice, the interaction between waves and data and model complexity is evident, where the effect of number of waves of data is considerably stronger for the multivariate model than for the other models. Therefore, multivariate models with more waves of data (i.e., five and six waves) displayed the worst model fit; whereas, univariate models with three waves of data and quadratic models with four waves of data displayed the best model fit. In summary, the χ^2 assessment of LGC model fit was found to be influenced by variations in waves of data and model complexity, but was not affected by differences in sample size.

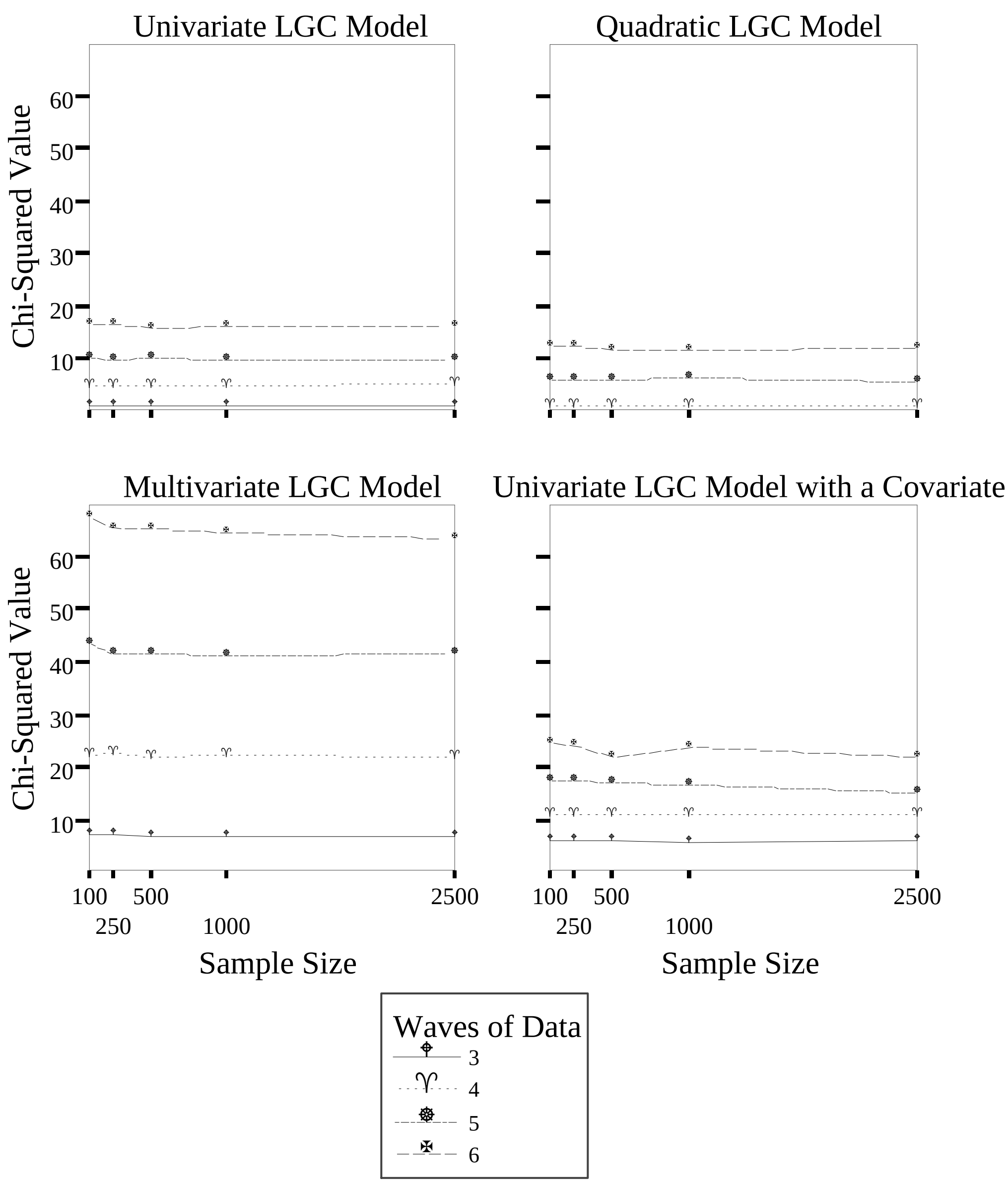

Figure 4.1. Mean chi-square values for LGC modeling conditions

By correcting for *df*, the χ^2/df ratio has been suggested to be superior to the traditional χ^2 assessment. Despite the significant difference observed for the χ^2/df assessment of fit among conditions of sample size, the effect size displayed a negligible effect (i.e., <.01% of the variance explained), as seen in Table 4.2. As previously discussed in regard to the χ^2, the significant difference is related to the high statistical power but the effect lacks practical merit and was not interpreted. No notable effects for χ^2/df were observed among the conditions of waves of data and model complexity. Furthermore, all interaction effects lacked evidence of discrepancies among the design conditions as well. Therefore, the χ^2/df was not influenced by the LGC modeling design conditions examined in this dissertation. Due to the lack of variation among conditions, mean values for the χ^2/df are not displayed in the text, but can be found in Appendix C.

In summary, the conditions of waves of data and model complexity produced substantial variation in the χ^2 assessment of LGC model fit. Increasing waves of data resulted in worse model fit (i.e., larger χ^2 values). Multivariate models produced the largest amount of variation, followed by covariate models, univariate model, and finally, the smallest differences were observed for the quadratic model. As expected, the χ^2/df was superior to the traditional χ^2 assessment of model fit, finding no variations based on the LGC modeling conditions examined in this dissertation. Sample size had only a negligible effect on either the χ^2 or χ^2/df.

Nonnormed Fit Index

The significant differences observed for the *NNFI* were found to lack practical relevance among the LGC modeling conditions examined in this dissertation, as displayed by the minimal effect sizes in Table 4.2. Due to the lack of effects for sample size, waves of data, and model complexity, the mean values are not displayed in the text, but can be found in Appendix C. In summary, *NNFI* values did not vary under LGC modeling conditions and functioned similarly to the χ^2/df, in performing in a superior manner to the traditional χ^2 assessment of model fit.

Comparative Fit Index

Similar to the *NNFI*, the *CFI* values displayed no notable main effects or interactions for the conditions of waves of data and model complexity; however, a moderate variation was observed for the main effect of sample size (η^2 = .11) and small to moderate effects were found for the planned comparisons of sample size (*d* = .25 - .43). Upon further examination, identical mean values were observed in all conditions (.99) with the only difference being standard deviations in the tenth (for model with N = 100 and 3 waves of data) or hundredth decimal place (N > 100), as presented in Appendix C. Therefore, the small to moderate differences in the *CFI* may lack practical relevance based on the limited change in mean values.

Root Mean Squared Error of Approximation

The *RMSEA* values displayed a moderate effect for sample size (η^2 = .13), with no notable differences among the conditions of waves of data and the majority of model complexity conditions. The planned comparisons revealed a moderate effect (*d* = .46 and *d* = .58) for the extreme comparisons of sample size (i.e., 100 vs. 250 and 1,000 vs. 2,500), respectively, and small effect for the remaining sample size comparisons (*d* = .25 - .22). As displayed in Figure 4.2, smaller sample sizes ($N \leq 250$) were found to produce higher *RMSEA* values (suggesting decrements in model fit), which began to stabilize at N = 1,000. The comparisons between univariate and multivariate models displayed a small effect size (d = .13), which can be seen in Figure 4.2 with an average of a .01 mean difference in *RMSEA* values among the two conditions of model complexity. In summary, the *RMSEA* was generally stable under varying conditions of waves of data and model complexity, but displayed a moderate effect for sample size suggesting worse model fit with smaller sample sizes.

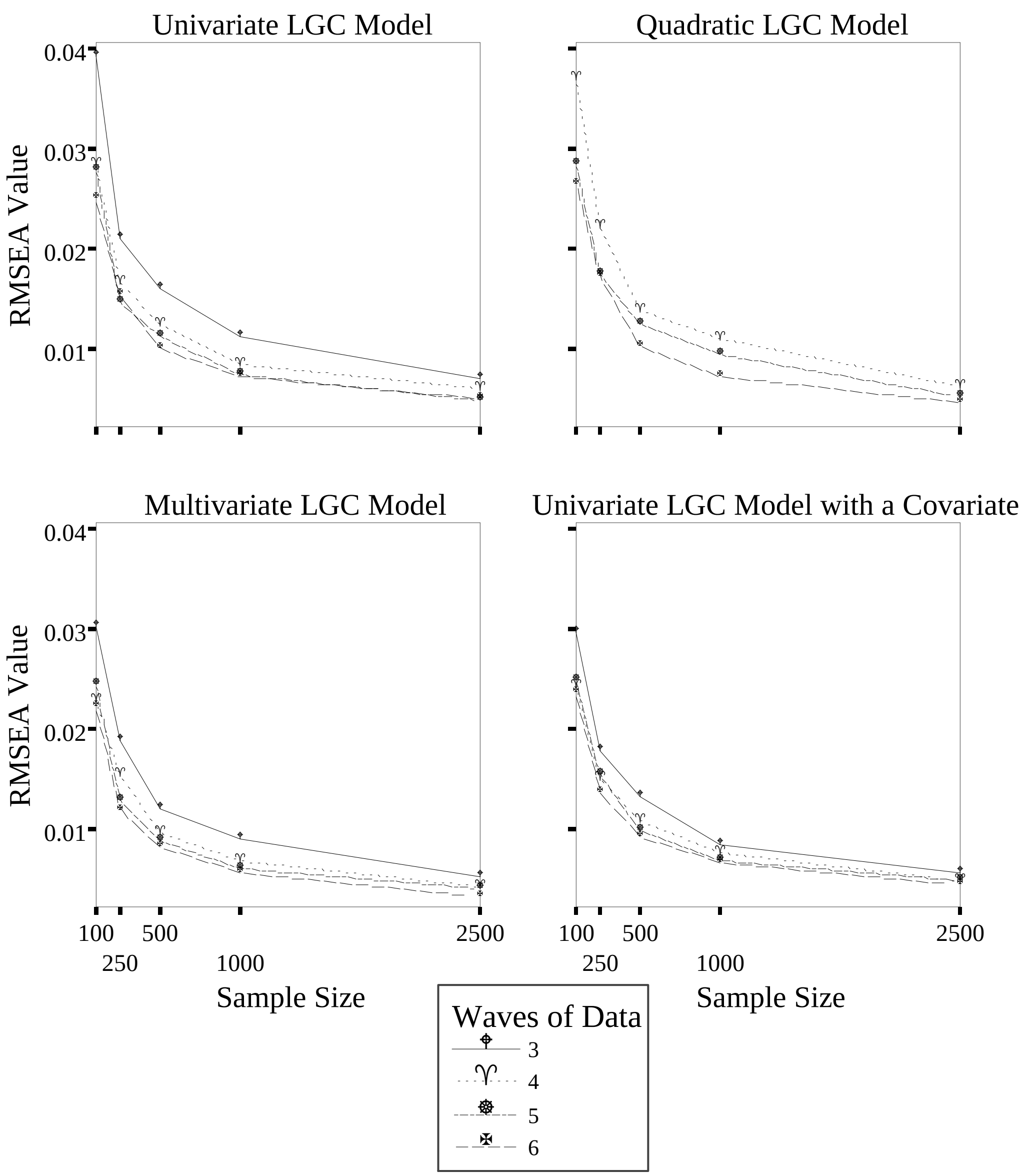

Figure 4.2. Mean values for the *RMSEA*

Summary of the Assessment of Mean Differences in LGC Model Fit

In summary, the ANOVAs, effect sizes, and mean fit index values highlighted the *influence* of LGC modeling conditions on selected global fit indexes. The χ^2/df and *NNFI* exhibited superior performance, lacking any differences due to varying LGC modeling data conditions examined in this dissertation. The *CFI* and *RMSEA* were found to have a moderate effect based on sample size; however, upon examining the mean values, the disparities in *CFI* values were found to lack practical relevance. The *RMSEA* mean values were found to vary tending toward worse model fit (i.e., larger values) in conditions with small sample sizes. The χ^2 assessment was the only fit index that varied among waves of data and model complexity, with higher values (suggesting worse model fit) in conditions with few waves of data and in multivariate and covariate LGC models. Unexpectedly, quadratic LGC model conditions displayed better model fit (i.e., lower χ^2 values) than the parsimonious univariate LGC model. In conclusion, the χ^2 assessment of fit suggested worse model fit with more waves of data and an increasing model complexity (excluding quadratic models), whereas the *RMSEA* displayed poorer fit in the presence of smaller sample size.

Type I Error Rates

To investigate the practical application, or *methods*, of determining adequate model fit, the Type I error rates were computed as the proportion of correct models that were falsely rejected using frequently utilized cutoffs. All models were correctly estimated (i.e., no model misspecifications) so, theoretically, all fit indexes should have displayed acceptable fit. Thus, any models classified as having inadequate fit would reflect a Type I error. For the χ^2 assessment of fit, the Type I error rate is the percentage of correct models that are rejected based on significance testing (i.e., *p*-value $<.05$). For the three descriptive fit indexes, Hu and Bentler's (1999) frequently utilized guidelines for acceptable model fit were applied to estimate the Type I error rates (i.e., $NNFI \leq .95$, $CFI \leq .96$, and $RMSEA \geq .05$.) Type I error rates of concern were defined as conditions that rejected more than 5% of the samples, comparable to the concept of an alpha level of .05.

Type I Error Rates: Chi-Square

Figure 4.3 graphically displays the Type I error rate, which ranged between 3% and 10%, for the χ^2 assessment of model fit, with the majority of conditions exhibiting Type 1 error rates between 4% and 6%. Small to moderate sample sizes resulted in variation in the χ^2 Type I errors, which stabilized at $N \geq 1{,}000$ among the different levels of model complexity. Type I error rates of concern were identified in multivariate models with $N = 100$ and more waves of data (i.e., 10%). In conclusion, the χ^2 is at moderate risk of displaying poor model fit for the multivariate model with more waves of data and small sample sizes.

Type I Error Rates: NNFI

Even though the *NNFI* was not found to vary among the LGC model data environments examined, excessive Type I errors were found in selected conditions, as shown in Figure 4.4. Conditions with small sample sizes (i.e., $N = 100$) and few waves of data (i.e., three waves, and four waves for quadratic models) displayed the worst model fit (i.e., encountered the highest Type I error rates). The Type I errors varied in model complexity conditions with the highest tendency to suggest poor fit in the univarite LGC model, followed by the multivariate LGC model, the quadratic LGC model, and finally, the covariate LGC model suggested the best fit. Therefore, applied researchers using the *NNFI* to assess model fit may falsely reject a correct univariate and multivariate LGC model with three waves of data and N = 100.

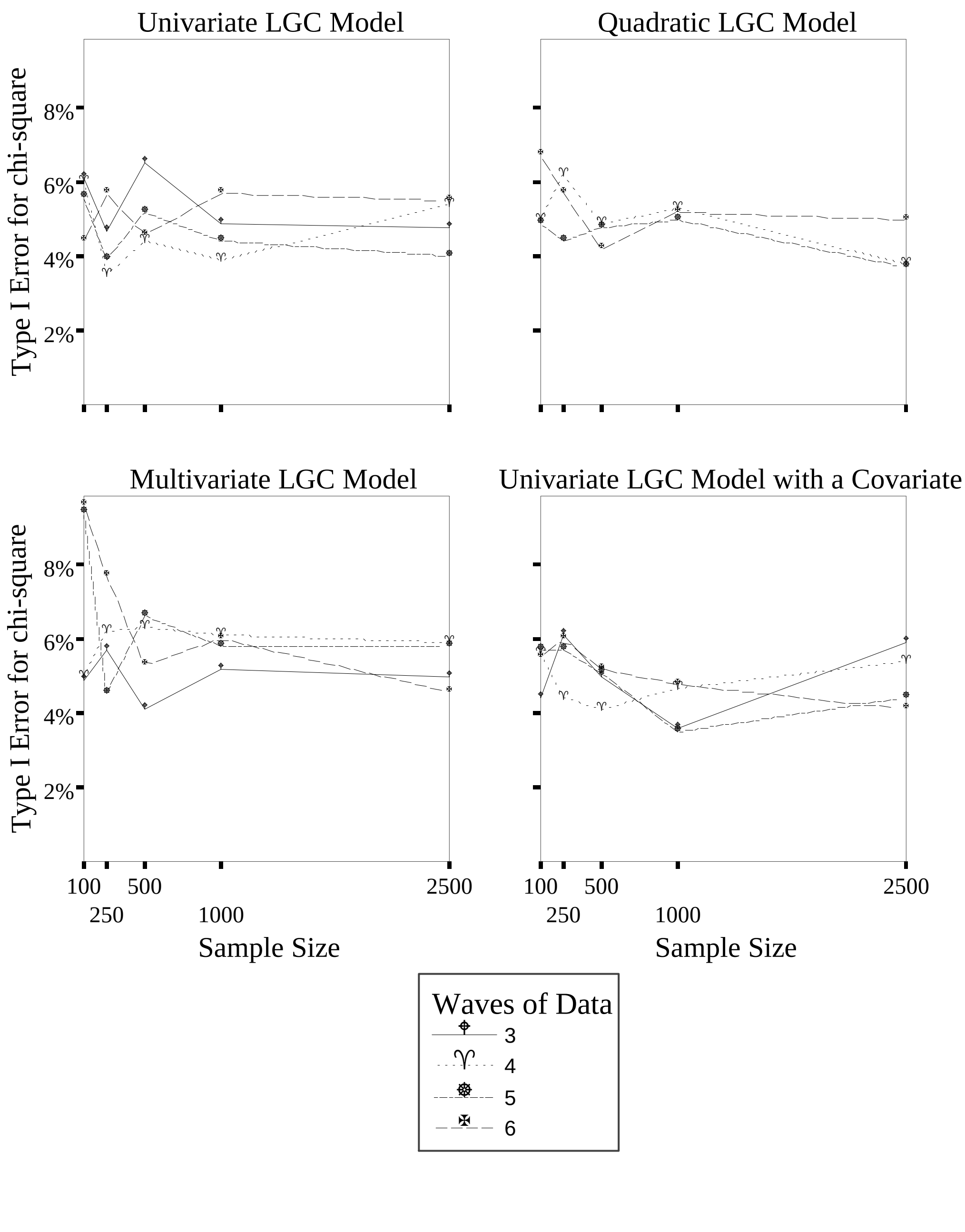

Figure 4.3. Type I error rates for the χ^2 assessment of fit

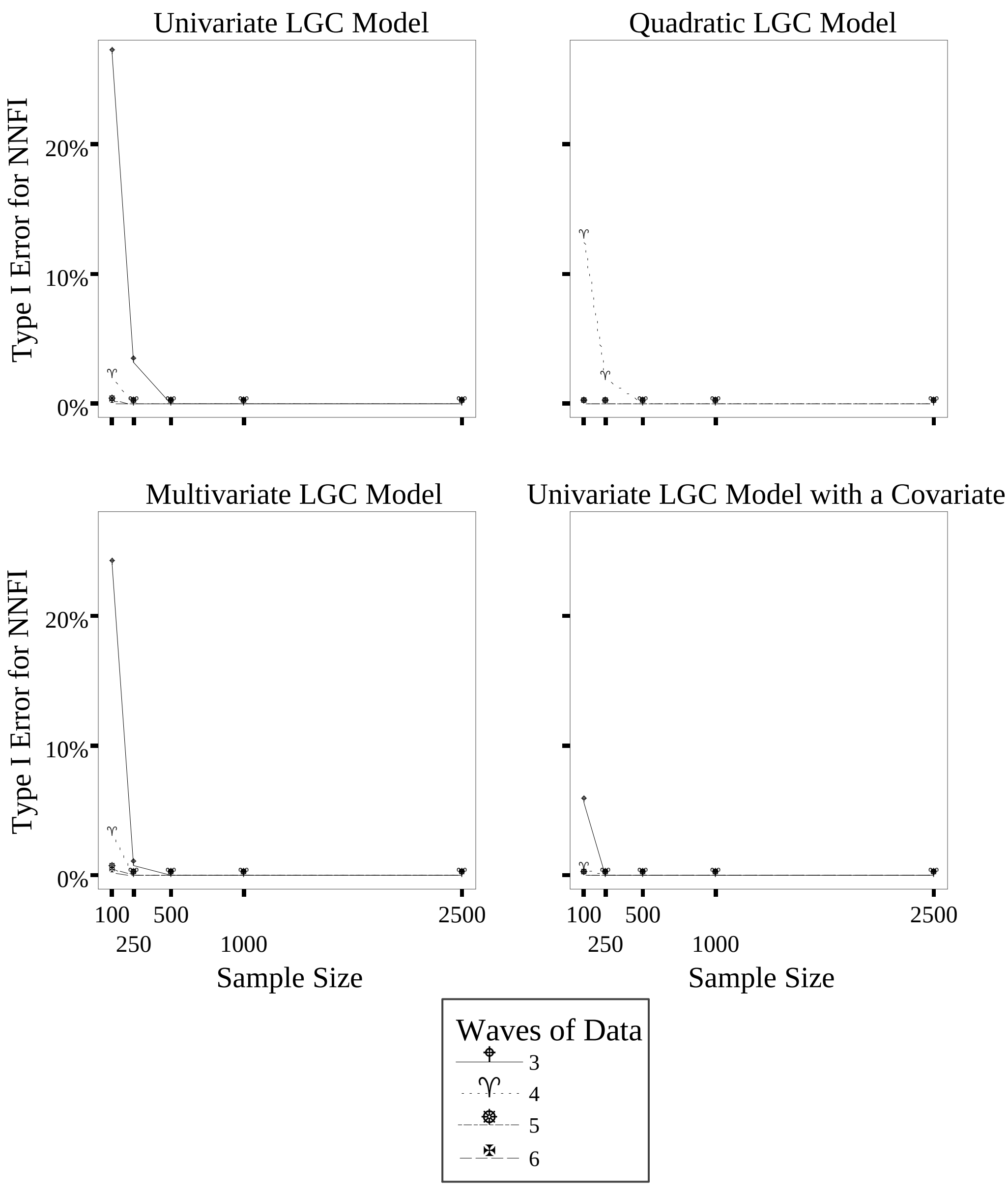

Figure 4.4. Type I error rates for the *NNFI*

Type I Error Rates: CFI

The *CFI* was found to perform the best, with the lowest Type I error rates among the four fit indexes, ranging from 0% to 4% (see Figure 4.5), although probability of obtaining poor model fit displayed a slight increase as more waves of data were added to models with $N = 100$. As for model complexity, the covariate model displayed the highest Type 1 error rates, followed by the multivariate models, univariate models, and finally, quadratic models which did not encounter any Type I errors. In summary, the *CFI* did not encounter any troubling Type I error among the LGC modeling conditions examined.

Type I Error Rates: RMSEA

The *RMSEA* did not perform well, with considerable Type I error rates ranging up to 30% in selected conditions with small sample sizes and few waves of data. In other words, the *RMSEA* incorrectly suggested poor model fit in conditions with small to moderate sample size and few waves of data. The Type I error rates were negligible when $N \geq 500$ for multivariate and covariate models, and $N \geq 1{,}000$ in univariate and quadratic models. In summary, the Type I error rates, based on the .05 *RMSEA* cutoff proposed by Hu and Bentler (1999), over-rejected correct models with small sample sizes ($N \leq 500$ to $N \leq 1{,}000$) and fewer waves of data.

Summary of Type I Error Rates

Regarding the *methods* for determining model fit (Type I error rates), the *CFI* performed the best, followed by the *NNFI*, χ^2, and *RMSEA* displayed the least desirable characteristics. The *NNFI* was found to over-reject (i.e., imply poor model fit) in correct univariate and multivariate models with three waves of data and $N = 100$. The χ^2 suggested inadequate model fit (i.e., excessive Type I errors) in multivariate models with $N = 100$ and five or six waves of data. The *RMSEA* displayed potentially problematic characteristics, tending toward poor fit, in LGC models with small to moderate sample sizes ($N \leq 1{,}000$) and few waves of data.

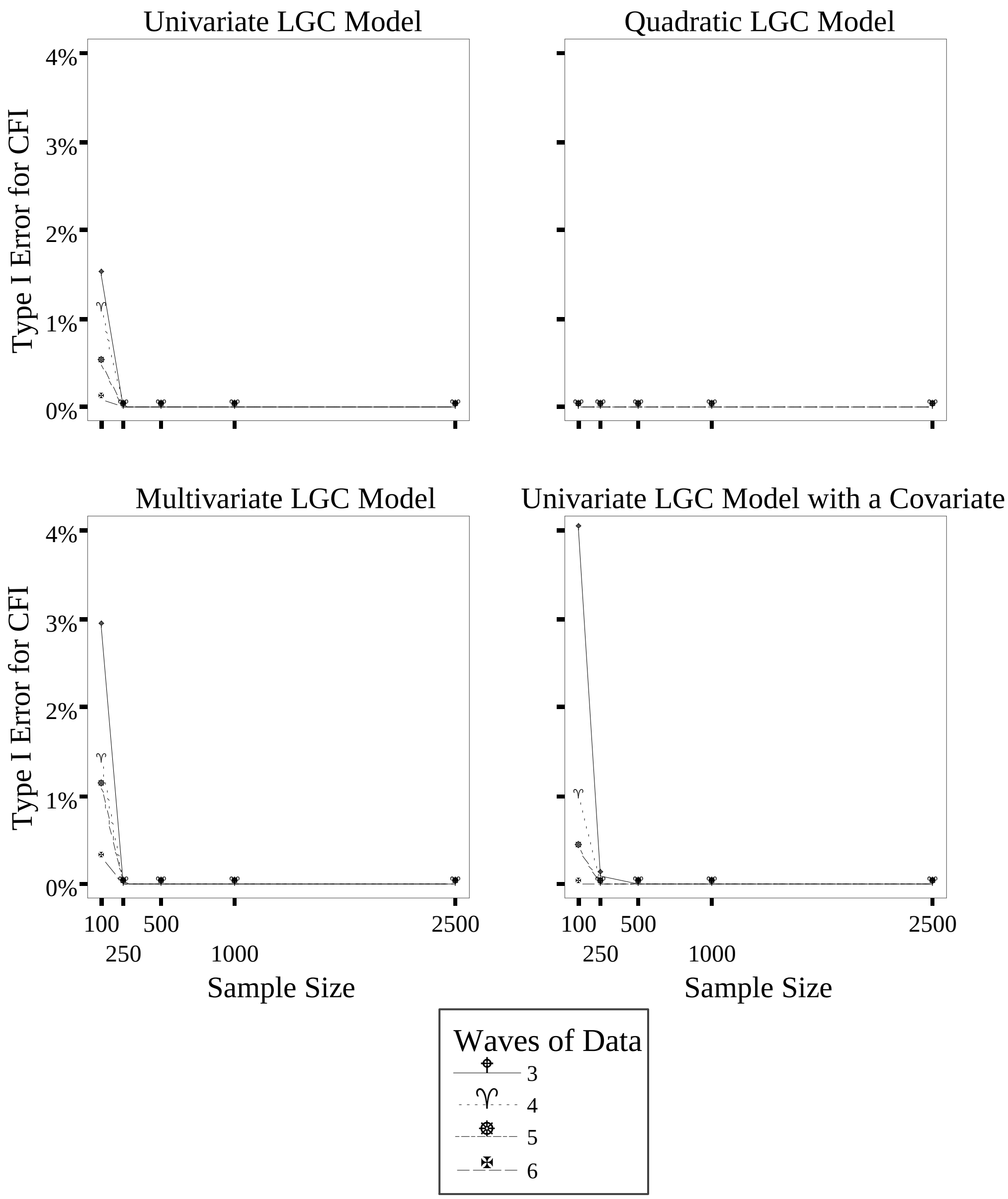

Figure 4.5. Type I Error Rates for the *CFI*

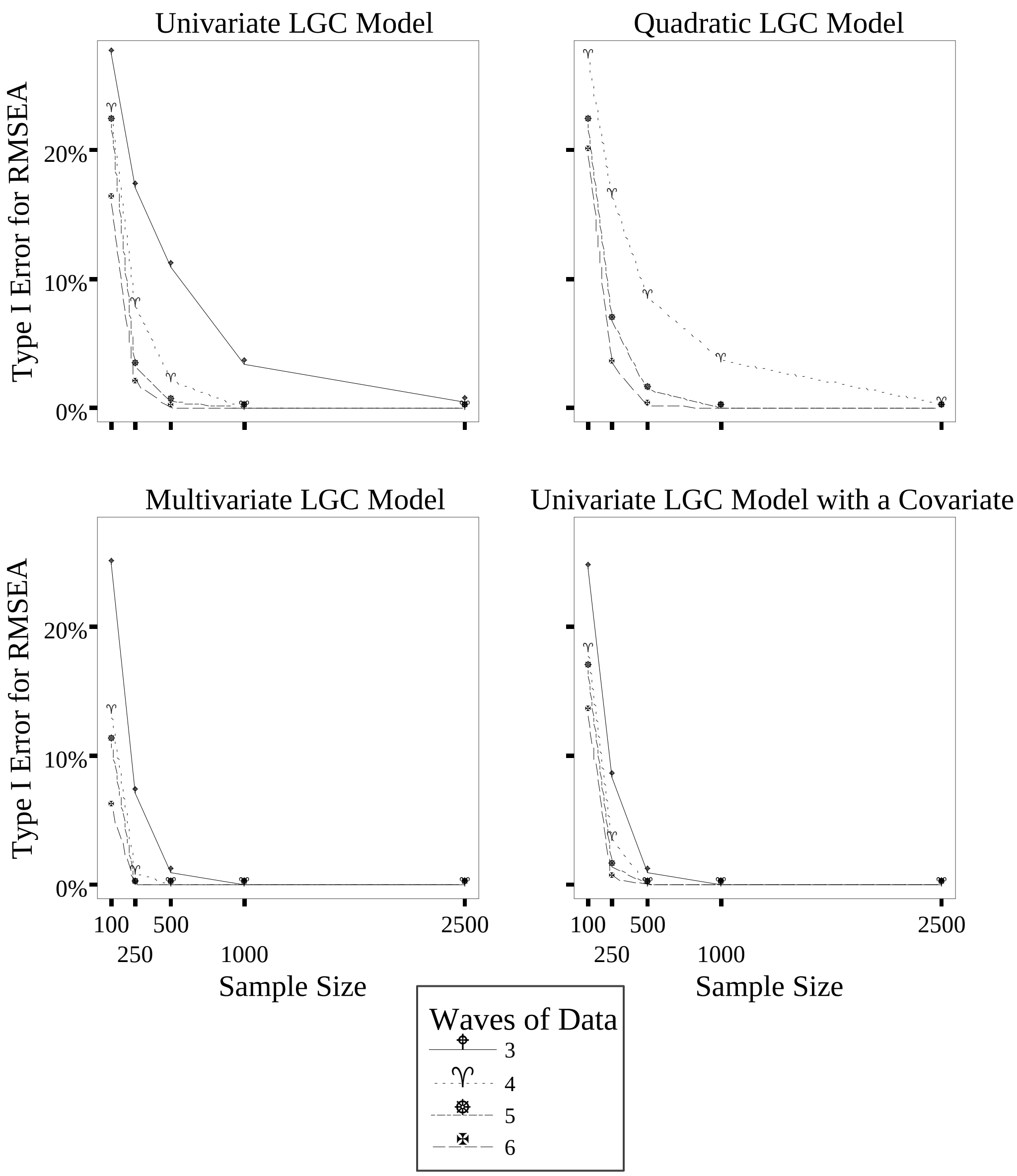

Figure 4.6. Type I Error Rates for *RMSEA*

Summary of Results

The investigation into the influence of LGC modeling design characteristics on commonly utilized fit indexes revealed several interesting findings. Four research questions were proposed and the results are summarized accordingly.

Research Question One: Model Convergence

The first research question inquired as to the model convergence rates among the data conditions, asking:

Q1 Do model convergence rates vary under conditions of sample size, waves of data, and model complexity?

It was hypothesized that complex models with small sample sizes and few waves of data would produce lower convergence rates. While all models converged, the rate of inadmissible solutions followed the anticipated convergence trends for sample size, waves of data, and selected conditions of model complexity. As expected, the multivariate models had a higher rate of inadmissible solutions than the covariate models, which in turn had higher rates than the univariate model. However, the quadratic model produced higher inadmissible solution rates than the more complex multivariate model conditions, particularly with three waves of data. In summary, all models converged, but the rate of inadmissible solutions was problematic, specifically in conditions of small sample size and few waves of data, particularly in quadratic and multivariate models.

Research Question Two: Sample Size

The second research question examined the influence of varying sample size conditions on four measures of model fit in LGC models, asking:

Q2 Do fit indexes (χ^2, *NNFI, CFI,* and *RMSEA*) differ under varying conditions of sample size?

It was hypothesized that the *CFI*, *NNFI,* and *RMSEA* will imply poor model fit in conditions of small sample size; whereas the χ^2 will imply good model fit in small sample sizes. Regarding the *influence* of LGC model conditions, the χ^2, χ^2/df, and *NNFI* performed well suggesting acceptable model fit in all sample size conditions. The *CFI* displayed a small to moderate effect for sample size; however, examination of mean values revealed the extremely minor discrepancies which lacked practical relevance. The *RMSEA* had the least desirable characteristics, displaying worse model fit in conditions with small to moderate sample sizes ($N \leq 1{,}000$) sample sizes.

Concerning the *methods* for determining model fit, the *CFI* performed the best, followed by the *NNFI* and χ^2, and finally, the *RMSEA* was the worst index examined. The latter three indexes were found to suggest poor model fit in conditions with smaller sample sizes. More specifically, the *NNFI* and χ^2 only suggested poor model fit when N = 100. However, the *RMSEA* was found to suggest inadequate fit in sample sizes as large as N = 1,000. To summarize, the *CFI* performed the best and the *NNFI* and χ^2 displayed moderate deficiencies, suggesting poor model fit when N = 100. However, the use of *RMSEA* in LGC models is of great concern, displaying unacceptable fit in conditions with small and moderate sample sizes (i.e., $N \leq 1{,}000$).

Research Question Three: Waves of Data

The third research question examined the effects of differing waves of data, asking:

Q3 Do fit indexes (χ^2, *NNFI, CFI,* and *RMSEA*) differ under varying conditions of waves of data?

It was hypothesized that as waves of data (observed variables) are added to a LGC model, the fit indexes will suggest worse model fit. However, the hypothesis was only supported for the χ^2 assessment of fit, suggesting worse model fit in models with more waves of data (five and six waves). The *CFI* performed in a superior manner compared with all other indexes, lacking any variations among waves of data conditions. While the *NNFI* and *RMSEA* were not found to be *influenced* by changing the waves of data in a LGC model, the *methods* for evaluating model fit displayed poor model fit with fewer waves of data. More specifically, the *NNFI* displayed unacceptable fit in univariate and multivariate models with three waves of data. Similarly, the *RMSEA* displayed worse model fit in all models with three waves of data and quadratic model with four waves of data. As waves of data increased, the χ^2 values increased suggesting increasingly poorer model fit. In summary, variations in waves of data suggested poor model fit for: (a) the χ^2 in conditions with five and six waves of data, in multivariate models, (b) the *NNFI* with three waves of data in univariate and multivariate models, and (c) the *RMSEA* with three waves of data and four waves of data in quadratic LGC model.

Research Question Four: Model Complexity

The fourth research question investigated the model complexity in LGC models asking:

> Q4 Do fit indexes (χ^2, *NNFI*, *CFI*, and *RMSEA*) differ under varying conditions of model complexity, defined in the current dissertation as a: (a) univariate linear LGC model, (b) quadratic LGC model, (c) multivariate linear LGC model, and (d) a linear LGC model with a covariate?

It was hypothesized that as model complexity increases (i.e., beginning with the most parsimonious univariate linear model to the quadratic model, the covariate model, and finally the most complex, multivariate models) the fit indexes will suggest worse model fit. The hypothesis was not fully supported in the findings, whereas the ordering of model complexity conditions differed from the expected direction of influence. For example, χ^2 was found to vary among conditions of model complexity; however, the quadratic model displayed better fit than the univariate model, even though the quadratic model is generally considered more complex (includes an additional latent variable). The superior model fit for the quadratic model is most likely related to the quadratic models having fewer *df* than univariate models. As expected, the χ^2 displayed the worse model fit in multivariate models compared to all other model complexity conditions. Multivariate models include more *df* than all other modeling conditions; thus, higher χ^2 values were found in these models. Furthermore, the *influence* of LGC modeling conditions did not affect the *NNFI*, but the *methods* of assessing model fit were found to over reject correct univariate and multivariate models, which were hypothesized to be the most extreme conditions examined. The *RMSEA* displayed a minor tendency to show better model fit in multivariate and covariate models, which were hypothesized to be the two most complex conditions examined. In summary, varying model complexity conditions did affect the fit indexes' assessment of model; however, not in the expected directions.

In conclusion, the results support the overall hypotheses that the fit indexes' values and the methods for determining model fit are affected by LGC modeling design conditions. The *CFI* performed the best, suggesting good model fit in the majority of conditions examined. The χ^2 suggested worse model fit in models with more waves of data (i.e., five or six waves), small sample sizes (i.e., $N = 100$), and multivariate models. The *NNFI* tended toward poor model fit in conditions with small sample sizes ($N = 100$), three waves of data, as well as univariate and multivariate models. The *RMSEA* tended toward poor model fit in conditions up to $N \leq 1{,}000$, with few waves of data (three and four), and in univariate and quadratic model. Presented in the following chapter is the discussion of the implications of these findings for applied and methodological researchers.

CHAPTER V

DISCUSSION

To date, there has been a lack of methodological guidance for applied longitudinal researchers in regard to the evaluation of latent growth curve (LGC) model fit, and they have had to follow the suggestions proposed for general structural equation modeling (SEM). In this dissertation, I examined the functioning of four commonly utilized fit indexes in LGC modeling data environments, including variations in sample size, waves of data, and model complexity. In Chapter IV, two venues for investigating LGC model fit were explored: (a) the *influence* of LGC modeling design characteristics on measures of model fit (e.g., effect sizes and the mean fit values), and (b) the *methods* for determining adequate model fit (e.g., Type I error rates based on significance testing and Hu and Bentler's (1999) cutoff values).

In regard to the *influence* of LGC modeling data environments, the chi-square divided by the degrees of freedom (χ^2/df), nonnormed fit index (*NNFI*), and comparative fit index (*CFI*) performed well, lacking any differences in the assessment of fit among LGC modeling conditions. Conversely, the chi-square likelihood ratio assessment (χ^2) and the root mean square error of approximation (*RMSEA*) performed poorly in selected LGC modeling conditions. Concerning the *methods* for determining acceptable fit, the Type I error findings suggest lack justification for the common practice of using universal cutoff values among the various LGC modeling environments, as found in numerous general SEM model fit simulations (Fan & Wang, 1998; Hu & Bentler, 1999; March, Hau, & Wen, 2004; Sivo et al., 2006). The *CFI* performed in a superior manner compared with all other indexes examined, using the standard cutoff of .96. Among the three remaining fit indexes examined, the *NNFI* performed the best (i.e., use of a cutoff of .95), followed by the χ^2, and finally, the *RMSEA* displayed the least desirable characteristics (i.e., application of a cutoff of .05).

Although authors of general SEM model fit simulations have drawn similar conclusions (Fan & Wang, 1998; Hu & Bentler, 1999; March et al., 2004; Sivo et al., 2006) it is critical to readdress these trends in terms of LGC modeling conceptualization. In Chapter V, I translate the findings into six general guidelines or suggestions to assist applied researchers in the assessment of LGC model fit, including: (a) the design of longitudinal studies to maximize the chance of obtaining a plausible solution, (b) cautious use of the χ^2 in selected modeling conditions, (c) relaxation of Hu and Bentler's (1999) cutoff for the *NNFI* in selected conditions, (d) adoption of novel LGC modeling cutoff values for the *RMSEA*, (e) use of multiple fit indexes in combinations to assess overall model fit, and (f) assessment of the within person fit as well as global model fit. In the following sections, I discuss evidence for the guidelines/suggestions, as well as specific design decisions that may allow applied researchers to increase the validity of the assessment of model fit.

Obtaining a Plausible Solution

Although all LGC modeling conditions converged, a high rate of inadmissible solutions was observed, specifically in conditions with small sample size, fewer waves of data, as well as in quadratic and multivariate models. As discussed in Chapter IV, inadmissible solutions occur when maximum likelihood estimation results in an implausible value (Bollen, 1989; Kline, 2005). Larger sampling distributions frequently occur in conditions with smaller sample sizes, which may extend beyond the range of plausible solutions and result in a larger percent of inadmissible solutions observed (Fan & Wang, 1998). In addition, models with three waves of data have a single *df* resulting in a limited amount of known information (Leite, 2007). For a review of model identification refer to pages 52-57. An identical rationale is attributed to the extremely high rates found in quadratic models with four waves of data, where there is only a single *df* in the model. Therefore, higher rates of inadmissible solutions for small sample sizes and fewer waves of data are rooted in the mathematical limitations of using parsimonious LGC models.

Interestingly, the topic of inadmissible solutions specific to LGC models is rarely addressed in the literature. Leite (2007), the only other known methodologist to discuss LGC modeling inadmissible solutions, observed rates which ranged up to 30.7% for univariate models and 32.4% for multivariate curve of factor models in conditions with three waves of data and $N = 100$. In this dissertation, higher rates were obtained which may have resulted from the parameter specification discrepancies between the two simulations (e.g., covariance between the latent intercept and latent slope was specified as .4 in Leite and .2 in this dissertation). Compared to general SEM simulations (Fan & Wang, 1998; Siemsen & Bollen, 2007), higher inadmissible solutions rates were observed in LGC modeling simulations (i.e., Leite and this dissertation). For example, Fan and Wang reported a 12.5% inadmissible solutions rate for general SEM models with $N =$ 100. Therefore, applied LGC modeling researchers should anticipate a higher probability of encountering an inadmissible solution, compared to what they might encounter when conducting standard SEM, and recognize that variations in parameter estimates may affect the probability of obtaining an inadmissible solution.

The corresponding practical interpretation of the high inadmissible solution rates is critical to the design of longitudinal studies. Specifically, applied researchers should attempt to design longitudinal studies that minimize the possibility of finding an inadmissible solution, with the use of larger sample sizes and more waves of data. For example, a researcher should avoid implementing a quadratic LGC model with four waves of data and N = 100 because of the limited chance in obtaining a proper solution (e.g., 9.3%). If it is not feasible to alter design characteristics, due to limited resources, applied researchers should be prepared to conduct a more traditional longitudinal analysis of change (e.g., repeated measures ANOVA) if an inadmissible solution is obtained. Therefore, Suggestion #1 is stated below.

Suggestion #1: Applied longitudinal researchers should be proactive by minimizing the chance of obtaining an implausible solution using design conditions with more waves of data and larger sample sizes. If design conditions cannot be altered, due to limited resources, applied researchers should prepare an alternative analysis in the chance that an inadmissible solution is obtained.

Cautions about the Use of the Chi-Square Likelihood Ratio Test

The tendency for χ^2 assessment of model fit to differ among sample size conditions is by far the most commonly referenced model fit limitation (Beauducel & Wittmann, 2005; Bollen & Curran, 2006; Bentler, 1990; Duncan et al., 2007; Fan & Wang, 1998; Hu & Bentler, 1999; Kline, 2005). Interestingly, in this LGC modeling simulation, the χ^2 was robust to variation in sample size, meaning the assessment of LGC model fit did not change by increasing or decreasing sample sizes. General SEM simulations have demonstrated acceptable performance of the χ^2 in ideal modeling conditions (e.g., normally distributed data, no missing data, continuous data using maximum likelihood estimation) and correctly specified models, similar to the conditions simulated in this dissertation (March, Hau, & Wen, 2004). Therefore, in application, when real data conditions are introduced into a LGC model, the good performance of the χ^2 among sample size conditions may diminish.

Despite the ideal modeling conditions simulated, the χ^2 suggested worse model fit as more waves of data were added to a LGC model (e.g., additional observed variables). As discussed in Chapter II, *df* for the χ^2 are equal to the difference between the number of unique elements of the observed variance-covariance matrix (i.e., known parameters) and the number of parameters estimated (i.e., unknown parameters), which were presented in Table 2.1 on page 54. As waves of data are added to a LGC model, additional observed variables (i.e., known parameters) are added, ultimately increasing the *df*. Although not directly examined in this dissertation, evidence supports the tendency for the χ^2 to vary with model *df*. By collectively interpreting the LGC simulation studies, including the findings from this dissertation, the use of increasing waves of data has benefits of increased reliability (Willett, 1989) and statistical power (Muthén & Curran, 1997; Muthén & Muthén, 2002). However, these design benefits are accompanied by worse model fit (i.e., increase in χ^2 values).

Also, the χ^2 assessment of model fit was found to vary among model complexity conditions, where quadratic models suggested better model fit than univariate models. The tendency for the χ^2 assessment of fit to imply better fit in quadratic models was not anticipated because these models are generally considered more complex. However, quadratic models require estimation of an additional latent parameter without an increase in observed variables leading to decreased *df*. For example, a quadratic model with three waves of data is just identified (i.e., model fit cannot be estimated)**,** even though an equivalent linear model is an over-identified model (i.e., model fit can be estimated). Applied LGC modeling researchers, who apply exploratory techniques, may compare χ^2 values between two or more competing model (e.g., linear univariate vs. quadratic models) to determine the most appropriate shape of the growth trajectory. However, the findings from the current investigation demonstrated that the design conditions of quadratic LGC models simply resulted in better model fit due to differences in *df*. Understandably, applied LGC modeling researchers may have drawn erroneous conclusions, by interpreting minor improvements in model fit (i.e., decrease in χ^2 values), as representing better model fit in a quadratic growth model, when the apparent improvement in fit may ban artifact of having fewer *df*. Applied longitudinal researchers should follow three general guidelines, based on the findings from this dissertation and previous methodological suggestions, to assess the shape of quadratic growth trajectories: (a) ensure that their underlying theory assumes quadratic growth in the hypothesized trait (Bollen & Curran, 2006; Preacher et al., 2008); (b) collectively interpret results from multiple indexes to interpret the shape of the growth trajectory, and (c) observe more than a 5.00 increase in χ^2 values per *df* to assume that the variations are more than what is expected from design condition alone. In conclusion, the χ^2 tendency to suggest a minor improvement in model fit for quadratic models (i.e., < 5.00), compared to univariate models, is an artifact of LGC model design characteristics and does *not* imply that the hypothesized shape represents quadratic growth.

In respect to the χ^2 assessment of fit, multivariate LGC models displayed the worst model fit compared to all other modeling conditions, when more than three waves of data were incorporated in the model. The χ^2 tendency toward poor model fit in complex models has been well documented in the general SEM literature (Beauducel & Wittmann, 2005; Cheung & Rensvold, 2002; Hutchinson & Olmos, 1998). Again, multivariate LGC models have more *df* than more parsimonious univariate models, therefore, complex models will produce higher χ^2 values, tending toward worse model fit. It is unknown if applied longitudinal researchers extrapolate this finding to LGC modeling by cautious comparisons of χ^2 statistics between multivariate and more parsimonious LGC models (e.g., univariate, quadratic models). To clearly reiterate in terms of LGC modeling terminology, larger χ^2 will occur in more complex models, which implies poor model fit (i.e., multivariate models), and applied researchers should avoid the sole use of χ^2 statistics to compare complex models to parsimonious models.

Despite the large variations in χ^2 model fit assessment regarding the *influence* of varying waves of data and model complexity conditions, the *methods* for determining model fit were less severely affected by LGC design conditions. The χ^2 displayed a moderate tendency (i.e., 10% Type I error) to suggest poor model fit (i.e., over-reject correct models) for multivariate models with N = 100 and more waves of data (i.e., five and six waves of data), which were the most complex conditions examined. Consequently, the current LGC modeling results differed slightly from the well documented tendency for the χ^2 assessment to over-reject complex models with small sample sizes and under-reject parsimonious models with large sample sizes (Beauducel & Wittmann, 2005; Fan & Wang, 1998), excluding the two most complex models examined.

As a whole, LGC models include a smaller range of model complexity conditions compared to general CFA simulations. For example, LGC models require two latent factors to represent a growth trajectory, and by nature, are more complex models than parsimonious CFA models, which can have one latent factor. Similarly, the LGC modeling data environments examined in this dissertation are relatively parsimonious compared to the number of latent and observed variables included in complex CFA designs. For example, Sharma et al. (2005) used more observed variables in their parsimonious conditions for CFA models than examined in the most complex conditions of this dissertation (e.g., 8-32 observed variables). Therefore, the LGC models examined in this dissertation represented less extreme conditions; thus, the well documented trend of over-rejecting complex models was observed only in the two most complex conditions (e.g., multivariate models with five and six waves of data when $N = 100$).

As discussed above, the use of the χ^2 assessment of fit should be cautioned in decisions made by applied LGC modeling researchers, which leads to Suggestion #2.

Suggestion #2: Applied researchers should use of the χ^2 with caution in the four following research design decisions.

1. Decisions to increase waves of data will result in a higher χ^2 values and tend toward poor model fit.
2. In comparison of quadratic models to univariate models, χ^2 values will be lower in quadratic models suggesting better model fit.
3. In comparison of multivariate models to a univariate model, the χ^2 values will be higher in multivariate models suggesting poor fit.
4. In complex models (i.e., multivariate models with five and six waves of data) with small sample sizes ($N = 100$), the χ^2 has a moderate tendency to suggest poor model fit (i.e., over-reject correct models).

LGC Modeling Cutoff Values for the Nonnormed Fit Index

While the *NNFI* lacked practical variations concerning the *influence* of design conditions, application of a .95 universal cutoff was found to have a moderate tendency (i.e., ≈ 10% Type I error) to suggest poor model fit in conditions with small sample sizes (*N* = 100), three waves of data, and in univariate and multivariate LGC models. This finding was expected based on the general SEM literature which has demonstrated that the *NNFI* will produce standard deviations that are substantially larger than other fit indexes in conditions of small sample size (Bentler, 1990; Jackson, 2007; Sharma et al., 2005; Yadama & Pandey, 1995). Interestingly, inflated Type I error rates were found only for univariate and multivariate models, which included the same ratio of observed to latent variables. Both quadratic and covariate LGC models included an additional latent variable; thus, the ratio of latent to observed variables was increased. Therefore, applied LGC modeling researchers should attempt to include an additional latent variable in their LGC model (e.g., covariate), which will reduce the *NNFI* tendency to suggest worse model fit based on design characteristics. If the addition of a covariate or predictor is not possible, researchers should be cautious in their use of a .95 universal cutoff for the *NNFI* in univariate and multivariate models with small sample sizes and three waves of data.

Selected SEM methodologists have discouraged the proposal of novel cutoff values, and even the use of fit indexes at all, because any cutoff values will be flawed among the vast number of potential modeling conditions (Chen et al., 2008). However, the practical reality is that in applied LGC modeling and SEM, researchers use fit indexes and procedural guides continue to endorse the use of the fit index along with Hu and

Bentler’s (1999) cutoff values. Therefore, I propose the adoption of novel cutoff values specific to LGC models that vary by design characteristics. The proposed cutoffs are *not* universally applicable, but may provide a more accurate assessment of fit compared to traditional SEM cutoff values. Based on the results from this dissertation, I propose that applied researchers should relax the *NNFI* cutoff values proposed by Hu and Bentler in conditions with $N \leq 100$ and three waves of data in univariate and multivariate LGC models, which leads to Suggestion #3.

Suggestion #3: Applied researchers should relax the *NNFI* cutoff values to the originally proposed value of .90 (Bentler & Bonett, 1980) in conditions of small sample sizes (i.e., $N \leq 100$) and few waves of data (i.e., three waves of data) for univariate and multivariate models. Or, applied researchers could simply add a covariate to the univariate and multivariate LGC model, which would elevate the *NNFI* and reduce its problematic tendencies; then the standard .95 cutoff could be applied.

LGC Modeling Cutoff Values for the Root Mean Error of Approximation

In conditions of small sample, the *RMSEA* suggested poor model fit, which has been well documented in the general SEM literature (Chen et al., 2008; Fan & Wang, 1998; Sharma et al., 2005; Sivo et al., 2006). However, in LGC models, tendency to suggest unacceptable model fit did not ameliorate until $N = 1{,}000$; which is higher than what has been found in the general SEM literature that suggests, on average, biased *RMSEA* values are only found in conditions with $N \leq 250$. The undesirable tendency of the *RMSEA* held true in the evaluation of the *methods* to determine acceptable model fit. By applying the standard cutoff value of .05, the *RMSEA* was found to suggest poor model fit in conditions with small to moderate samples sizes and few waves of data. Regarding model complexity, the undesirable characteristics of the *RMSEA* did not dissipate until $N \geq 500$ for multivariate and covariate models and $N \geq 1{,}000$ and in univariate and quadratic models. Again, the influences of sample size on *RMSEA* Type I error rates was expected based on the general SEM literature (Hu & Bentler, 1999; Sivo et al., 2006). Even when Hu and Bentler originally proposed the cutoff value of .05, they cautioned that this standard value will tend to over-reject correct models in conditions of small sample size. However, longitudinal researchers should be aware that *RMSEA* values may continue to vary and may imply poor model fit in LGC models conditions up to *N* = 1,000, which extends the cautionary range of recommendations from the general SEM literature.

Therefore, *RMSEA* cutoff values should be adjusted to account for the tendency to suggest poor model fit in conditions with small sample sizes and few waves of data, particularly in univariate and quadratic models. As discussed in Chapter II, Steiger's (1989) original guidelines endorsed *RMSEA* values of: (a) less than .05 to suggest good fit, (b) .08 for reasonable fit, and (c) values beyond .10 to indicate model misfit. Therefore, I propose using the cutoff values of .05, .08, and .10 for most LGC modeling conditions due to their familiarity. Specifically, that *RMSEA* values less than or equal to .10 may constitute appropriate model fit in the following conditions:

1. univariate LGC models with *N* = 100 and three, four, or five waves of data;
2. quadratic LGC models with *N* = 100 and four or five waves of data; and
3. multivariate and covariate LGC models with *N* = 100 and three waves of data.

RMSEA values less than or equal to .08 may suggest acceptable model fit in the following conditions:

1. univarite LGC models with six waves of data and *N* = 100;
2. univariate LGC models with three and four waves of data when *N* = 250;
3. univariate LGC models with three waves of data when *N* = 500;
4. quadratic LGC models with four waves of data when *N* = 250 or *N* = 500;
5. quadratic LGC models with five waves of data when *N* = 250;

6. multivariate LGC models when $N = 100$ with four, five, or six waves of data;
7. multivariate LGC models with three waves of data when $N = 100$;
8. covariate LGC models when $N = 100$ with three, four, and five waves of data; and
9. covariate models with three waves of data and $N = 250$.

Finally, all other more complex LGC modeling conditions may be able to determine appropriate model fit by using the frequently applied *RMSEA* values less than or equal to .05. Althought the previous guidelines are detailed; applied researchers typically follow more general guidelines. Therefore, novel cutoff values for the RMSEA are consisly stated in Suggestion #4.

Suggestion #4: Applied researchers should adopt novel cutoff values for the *RMSEA* in selected LGC modeling conditions. Values less than or equal to .10 may constitute appropriate fit in models with $N = 100$ and fewer waves of data. Values less than or equal to .08 may suggest acceptable model fit in model with $N = 100$ and more waves of data OR with moderate sample sizes ($N = 250$ or 500) and fewer waves of data. All other more complex LGC modeling conditions may be able to determine appropriate model fit by using the frequently applied *RMSEA* values less than or equal to .05.

To reiterate, the novel cutoff values proposed for the *NNFI* and *RMSEA* in LGC models are *rules of thumb*; therefore, these values are clearly limited and will not be appropriate for all LGC modeling conditions. Furthermore, the new cutoff values should be adjusted, based on future research regarding LGC model misspecification, and discussed in detail in a later section.

Using Combinations of Fit Indexes

Methodologists have discussed the deceptive nature of deducing a dichotomous decision regarding model fit based on any *single* fit index, suggesting that two or more fit indexes should be collectively interpreted (Beauducel & Wittmann, 2005; Chen et al., 2008; Hu & Bentler, 1999; Hutchinson & Olmos, 1998; Kline, 2005; Preacher et al., 2008). For example, the seminal article by Hu and Bentler was predominantly devoted to recommending pairs of model fit indexes used conjunctively to evaluate fit. Even though researchers use multiple fit indexes to increase the information obtained regarding model fit, the fit indexes may imply various degrees of model fit because they were developed on a different rationale of model fit (Bollen, 1998; Bollen & Curran, 2006). When fit indexes provide conflicting model fit interpretations, applied researchers need to review methodological model fit guidelines to determine which fit index(es) should be weighted more than other based on the currently modeling conditions. For example, if a researcher was examining a LGC with small sample sizes, less weight should be placed on the *RMSEA* based on the finding of this study. Therefore, when multiple fit indexes are collectively interpreted, LGC modeling researchers may not arrive at a unanimous decision regarding model fit, which requires a critical analysis of the type of information provided by each index (Hutchinson & Olmos), and a review of the relevant literature to determine data conditions that may have biased the index(es). Therefore, Suggestion #5 addresses the use of a collection of fit indexes to determine model fit.

Suggestion #5: Similar to guidelines found in the general SEM model, I recommend that applied LGC modeling researchers collectively interpret all four fit indexes examined in this dissertation, while recognizing and adjusting for the limitations of the χ^2, *CFI*, *NNFI*, and *RMSEA* in selected modeling conditions discussed above. By using multiple fit indexes, applied researchers may have to engage in a critical analysis of the type of model fit information presented by each fit index.

Evaluating Within Person Fit

Coffman and Millsap (2006), the authors of the only other known LGC model simulation of model fit, demonstrated the poor performance of the χ^2 and *RMSEA* to detect model shape misspecification. The authors recommended exploring within person fit in LGC models by using negative loglikelihood values (-2LL) for each participant, in addition to assessing global fit. While the evaluation of within person fit may benefit the overall assessment of LGC model fit, the computation of -2LL parameters are not standard in the SEM software used in the social and behavioral sciences (e.g., EQS, AMOS, LISREL, M*plus*). However, ambitious applied longitudinal researchers are encouraged to use the Mx syntax provided by Coffman and Millsap to evaluate within person fit, as well as global model fit.

Suggestion #6: Applied longitudinal researchers should attempt to evaluate within person model fit using -2LL values; however, the computation of these values requires the use of alternative software program, not frequently utilized in the social and behavioral sciences, which may limit the widespread use.

Dissemination of Information for Assessing LGC Model Fit

In general, SEM methodologists tend to communicate model fit findings to their colleagues through highly technical and complex methodological journals and conference presentations, which limits the impact of these findings in applied research fields. Due to the applied nature of this dissertation, it is imperative to discuss strategies that may encourage applied longitudinal researchers in the social and behavior sciences fields to adopt the proposed guidelines for assessing LGC model fit.

First, I propose that novel findings regarding LGC model fit be presented in a user-friendly manner (e.g., clear suggestions focused on methodological decisions), as well as being published in applied journals frequently read by longitudinal researchers. Secondly, novel information regarding LGC model fit should be conveyed to a key group of influential individuals who have the ability to modify methodological practices in the behavioral and social sciences. Typically, trends in the standard reporting of statistical concepts, for the behavioral and social sciences, are driven by journal editors' requirements for publishing and standards proposed by national organizations (e.g., the Manual of the American Psychological Association [APA]). For example, the concepts of statistical power and effect sizes were addressed in the methodological literature for decades, but did not fully emerge in the applied literature until journal editors and APA staff endorsed their use. Consequently, by proactively seeking out a selected group of researchers in the behavioral and social science (e.g., journal editors and national organizations), the recommendation proposed for assessing model fit may be adopted more rapidly.

Limitations and Future Research

Even though the findings from this dissertation provided interesting insights into the assessment of LGC model fit, there are clear limitations to this study. First and foremost, this dissertation lacked evaluation of the sensitivity of the fit indexes to LGC model misspecification, which is the primary purpose of using fit indexes. For example, while the incremental fit indexes (i.e., *NNFI* and *CFI*) were quite robust among LGC modeling design conditions, general SEM researchers have reported their limited ability to detect model misspecification (Fan et al., 1999; Jackson, 2007). Furthermore, incremental fit indexes have been found to exhibit discrepancies based on estimation methods and non-normality (Fan & Wang, 1998; Hutchinson & Olmos, 1998); therefore, their advantages may dissipate when additional variations are considered. In summary, all results and conclusions are limited without a comprehensive investigation of fit indexes in model misspecification conditions relevant to LGC modeling and under various estimation methods.

In contrast to general SEM, LGC modeling researchers conceptualize and specify model parameters differently; therefore, specification of model misspecification will differ in LGC model. For example, general SEM model misspecification is typically separated according to two categories: (a) measurement misspecification including misspecified paths or factor loadings, and (b) structural misspecification including misspecified latent traits and paths among latent traits. Although measurement misspecification is more frequently investigated in general SEM simulations, within a LGC modeling framework, measurement misspecification lacks relevance because all factor loadings are fixed to represent the coding of time. Consequently, fit indexes should be re-evaluated based on types of structural misspecification relevant to the conceptualization of LGC modeling, potentially, to include: (a) shape misspecification, where an additional latent trait(s) is estimated in the model to represent the growth of the trajectory of change; (b) time period misspecification, which would include the addition or reduction of the number of observed variables to ensure that proper time period of growth is measured in the LGC model; and (c) predictors or covariate misspecification, which would include examining whether a critical predictor(s) or covariate(s) should be included or excluded from a model. To summarize, all fit indexes proposed for assessing SEM model fit, even indexes deemed to lack sensitivity, should be reexamined to determine if they are sensitive to structural misspecification relevant to LGC modeling.

As noted by methodologists, numerous additional LGC modeling simulations must be conducted to expand the knowledge regarding the assessment of model fit (Coffman & Millsap, 2006; Voelkle, 2007). Specially, a large scale simulation study should be conducted to examine the sensitivity of multiple fit indexes to types of LGC model misspecification. Based on the results from this dissertation, future LGC model fit simulations should specifically focus on the examination of two LGC design characteristics including: (a) *df* in the LGC model and (b) the ratio of latent to observed variables in the model.

Discussion Summary

This dissertation has provided the applied longitudinal researcher with a preliminary understanding of the functioning of fit indexes in LGC modeling environments. Consistent with standard SEM simulation findings, results from the current study supported that the fit indexes commonly utilized in LGC modeling applications are influenced by variations in: (a) sample size, (b) waves of data, and (c) model complexity. The *CFI* was found to be quite robust among the LGC modeling design conditions examined; however, the sensitivity of the *CFI* to LGC model misspecification needs to be assessed in future research. All other fit indexes were found to suggest poor model fit (i.e., over-reject correct models) in select LGC modeling conditions. Six guidelines were proposed for LGC modeling researchers, including: (a) design longitudinal studies to maximize the chance of obtaining a plausible solutions (i.e., more waves of data and larger sample sizes); (b) be cautious in the use of the χ^2 in selected modeling conditions; (c) relax Hu and Bentler’s (1999) cutoff values for the *NNFI* in univariate and multivariate models with $N = 100$ and three waves of data; (d) adopt novel LGC model cutoff values for the *RMSEA* in conditions of small to moderate samples sizes and few waves of data; (e) use multiple fit indexes in combinations to assess overall model fit; and (f) assess the within person fit as well as global model fit.

As the use of LGC modeling applications increases in the social and behavioral sciences, there is a critical need for additional research regarding LGC model fit, specifically, the sensitivity of fit indexes to relevant types of LGC model misspecification. In conclusion, this dissertation provides novel information regarding the interpretation of LGC model fit; however, additional methodological research is needed to increase the rigor of applied longitudinal studies in the social and behavior sciences.

REFERENCES

Beauducel, A., & Wittmann, W. W. (2005). Simulation study on fit indexes in CFA based on data with slightly distorted simple structure. *Structural Equation Modeling, 12,* 41-75.

Bentler, P. M. (1990). Comparative fit indexes in structural equation modeling. *Quantitative Methods in Psychology, 107,* 238-246.

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88,* 588-606.

Bollen, K. A. (1989). A new incremental fit index for general structural equation model. *Sociological Methods and Research, 17,* 303-316.

Bollen, K. A., & Curran, P. J. (2006). *Latent curve modeling: A structural equation perspective.* Hoboken, NJ: John Wiley & Sons.

Burchinal, M., & Appelbaum, M. I. (1991). Estimating individual developmental functions: Methods and their assumptions. *Child Development, 62,* 23-43.

Byrk, A., & Raudenbush. S. (1992). *Hierarchical linear modeling in social and behavioral research: Applications and data analysis methods*. Newbury Park, CA: Sage.

Byrne, B. M., & Crombie, G. (2003). Modeling and testing change: An introduction to the latent growth curve model. *Understanding Statistics, 2,* 177-203.

Chen, F., Curran, P. J., Bollen, K. A., Kirby, J., & Paxton, P. (2008). An empirical evaluation of the use of fixed cutoff points in RMSEA test statistic in structural equation models. *Sociological Methods and Research, 36*, 462 – 494.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*, 233 – 255.

Coffman, D. L., & Millsap, R. E. (2006). Evaluating latent growth curve models using individual fit statistics. *Structural Equation Modeling, 13,* 1-27.

Cronbach, L. J., & Furby, L. (1970) How we should measure “change”- Or should we? *Psychological Bulletin, 74,* 86-80.

Curran, P. J., Bollen, K. A., Paxton, P., Kirby, J., & Chen, F. (2002). The non-central chi-square distribution in misspecified structural equation models: Finite sample results from a Monte Carlo simulation. *Multivariate Behavioral Research, 37*, 1-36.

Davey, A., Savla, J., & Luo, Z. (2005). Issues in evaluating model fit with missing data. *Structural Equation Modeling, 12,* 578-597.

Duncan, T. E., & Duncan, S. C. (2004). An introduction to latent growth curve modeling. *Behavioral Therapy, 35,* 333-363.

Duncan, T. E., Duncan, S. C., & Strycker, L. A. (2007). *An introduction to latent variable growth modeling: Concepts, issues and application* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.

Fan, X. (2003). Power of latent growth modeling for detecting group differences in linear growth trajectory parameters. *Structural Equation Modeling, 10,* 380-400.

Fan, X., & Sivo, S. A. (2005). Sensitivity of fit indexes to misspecified structural or measurement model components: Rationale of two-index strategy revisited. *Structural Equation Modeling, 12,* 343-367.

Fan, X., Thompson, B., & Wang, L. (1999). The effects of sample size, estimation methods, and model specification on SEM fit indices. *Structural Equation Modeling, 6,* 56-83.

Fan, X., & Wang, L. (1998). Effects of potential confounding factors on fit indices and parameter estimates for true and misspecified SEM models. *Educational and Psychological Measurement, 58,* 701-735.

Field, A. (2005). *Discovering statistics using SPSS* (2nd ed.). Thousand Oaks, CA: Sage.

Gamst, G., Meyers, L. S., & Guarino, A. J. (2008). Analysis of variance designs: A conceptual and computational approach with SPSS and SAS. Cambridge, UK: Cambridge University Press.

Grimm, K. J. (2007). Multivariate longitudinal methods for studying developmental relationships between depression and academic achievement. In*ternational Journal of Behavioral Development, 31,* 328-339.

Hardy, S. A., & Thiels, C. (2007). Using latent growth curve modeling in clinical treatment research: Comparing guided self-change and cognitive behavioral therapy treatments for bulimia nervosa. *European Psychiatry, 22, Supplement 1,* S341- S342.

Hertzog, C., Lindenberger, U., Ghisletta, P., & von Oertzen, T. (2006). On power of multivariate latent growth curve models to detect correlated change. *Psychological Methods, 11,* 244-252.

Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods, 3*, 424 – 453.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structural analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1-55.

Hutchinson, S. R., & Olmos, A. (1998). The behavior of descriptive fit indices in confirmatory factor analyses using ordered categorical data. *Structural Equation Modeling: A Multidisciplinary Journal*, *5*, 344-364.

Jackson, D. L. (2007). The effect of the number of observations per parameter in misspecified confirmatory factor analytic models. *Structural Equation Modeling, 14,* 48 – 76.

Jöreskog, K., & Sörbom, D. (2001). LISREL 8: User's reference guide. Chicago, IL: Scientific Software International.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York: Guilford Press.

La Du, T. J., & Tanaka, J. S. (1995). Incremental fit index changes for nested structural equation models. *Multivariate Behavioral Research, 30*, 289-316.

Leite, W. L. (2007). A comparison of latent growth models for constructs measured by multiple items. *Structural Equation Modeling, 14,* 581-610.

Marsh, H. W., Hau, K., & Wen, Z. (2004). In search of golden rules: Comments on hypothesis-testing approaches to setting cutoff values for fit indexes and danger in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling, 11*, 320 – 341.

McArdle, J. J. (1988). Dynamic but structural equation modeling of repeated measures data. In R. B. Cattell & J. Nesselroade (Eds.), *Handbook of multivariate experimental psychology* (2nd ed., p. 561-614). New York: Plenum.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*, 130-149.

McDonald, R. P. (1989). An index of goodness-of-fit based on noncentrality. *Journal of Classification Bulletin, 6*, 97-103.

Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika*, *55,* 107-122.

Muthén, B. O., & Curran, P. J. (1997). General longitudinal modeling of individual difference in experimental designs: A latent variable framework for analysis and power estimation. *Psychological Methods, 2,* 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*, *4*, 599-620.

Muthén, L. K., & Muthén, B. O. (2004). *Mplus user's guide.* Los Angeles, CA**:** Muthén & Muthén.

Preacher, K. J., Wichman, A. L., MacCallum, R. C., & Briggs, N. E., (2008). *Latent growth curve modeling.* In T. F. Liao (Ed.), Quantitative applications in the social sciences (Vol. 157). Thousand Oaks, CA: Sage**.**

Raykov, T.**,** & Marcoulides, G. A. (1999). On desirablity of parsimony in structural equation model selection. *Structural Equation Modeling, 6,* 292-300.

Rindskopf, D. (1987). Using latent class analysis to test developmental models. *Developmental Review, 7*, 66-85.

Rogosa, D., Brandt, D., & Zimowaski, M. (1982) A growth curve approach to the measurement of change. *Psychological Bulletin, 92*, 726-748.

Rogosa, D., & Willett, J. B. (1985). Understanding correlates of change by modeling individual difference in growth. *Psychometrika, 50,* 203-228.

Rogosa, D. R., & Willett, J. B. (1983). Demonstrating the reliability of the difference score in the measurement of change. *Journal of Educational Measurement, 20*, 335-343.

Rovine, M. J., & Molenaar, P. C. M. (2000). A structural modeling approach to a multilevel random coefficients model. *Multivariate Behavioral Research, 35*, 51-88.

Schulenberg, J., & Maggs, J. L. (2001). Moving targets: Modeling developmental trajectories of adolescent alcohol misuse, individual, peer risk factors and intervention effects. *Applied Developmental Sciences, 5,* 237-253.

Sharma, S., Mukherjee, S., Kurmer, A., & Dillion, W. R. (2005). A simulation study to investigate the use of cutoff values for assessing model fit in covariance structure models. *Journal of Business Research, 58*, 935-943.

Siemsen, E., and K. Bollen. 2007. LAD: An Alternative Estimator for Structural Equation Models. *Sociological Methods and Research, 36*, 2267-265.

Sivo, S. A., Fan, X., Witta, E. L., & Willse, J. T. (2006). The search for “optimal” cutoff properties: Fit indexes criteria in structural equation modeling. *The Journal of Experimental Education, 74*, 267-288.

SPSS Inc. (2006). *SPSS (Version 15.0)*.Chicago, IL.

Steiger, J. H. (1989). *EZPATH: A supplementary module for SYSTAT and SYGRAPH*. Evanston, IL: SYSTAT.

Steiger, J. H., & Lind, J. C. (1980). *Statistically based tests for the number of factors.* Paper presented at the annual meeting of the Psychometric Society, Iowa City, IA.

Tanguma, J. (2001). Effects of sample size on the distribution of selected fit indices: A graphical approach. *Educational and Psychological Measurement, 6,* 759-776.

U.S. Department of Education (2001). *No Child Left Behind Act*. (Public Law 107-110). Washington, DC: U.S. Government Printing Office.

U.S. Department of Health and Human Services. (2005). *Congressionally mandated evaluations of the State Children's Health Insurance Program* (HHS-100-01-002). Rockville, MD: Mathematical Policy Research. Retrieved February 5, 2009 from http://www.mathematica_mpr.com/publications

Vangel, M. G., & Rukhin, A. L. (1999). Maximum likelihood analysis for heteroscedastic one-way random effects ANOVA in interlaboratory studies. *Biometrics, 55,* 129-136.

Voelkle, M. C. (2007). Latent growth curve modeling as an integrative approach to the analysis of change. *Psychological Science, 49,* 375-414.

Voelkle, M. C., & Wittman, W. W. (2007, November 4). *An introduction to longitudinal analysis: From the classics to modern structural equation modeling.* AEA Professional Development work shop; Portland, OR.

Widaman, K. F., & Thompson, J. S. (2003). On specifying the null model for incremental fit indices in structural equation modeling. *Psychological Methods, 8,* 16-37.

Willett, J. B. (1989). Some results on reliability for the longitudinal measurement of change: Implications for the design of studies for individual growth. *Educational and Psychological Measurement, 49*, 587-602.

Willett, J. B., & Sayer, A. G. (1994). Using covariate structures analysis to detect correlates and predictors of individual change over time. *Quantitative Methods in Psychology*, *116*, 363-381.

Yadama, G., N., & Panday, S. (1995). Effect of sample size on goodness-of-fit indices in structural equation models. *Journal of Social Service Research, 20,* 49-70.

Yuan, K., Bentler, P. M., & Zhang, W. (2005). The effect of skewness and kurtosis on mean and covariance structural analysis: The univariate case and its multivariate implications. *Sociological Methods and Research, 34*, 240-258.

Footnote

[1] The specific databases examined for the review of LGC modeling applications in 2006 and 2007 included the following databases: Academic Search Premier, Agricola, America: History and Life, Art Abstracts, Biological Abstrcacts, Business Source Premier, CINAHL, Communication & Mass Media Complete, EconLit, Family & Society Studies Worldwide, GeoRef, Humanities International Complete, Information Science and Technology Abstract, MedicLatina, MEDLINE, Pre-CINAHL, PsychARTICLES, PsycINFO, and Regional Business news. The search terms included "latent growth curve," "latent growth modeling," "latent growth curve model," "latent growth curve modeling," and "latent growth curve models."

APPENDIX A

MPLUS PROGRAM SYNTAX

A.1 Unconditional model: Univariate linear LGC model data generation

```
montecarlo:
        names = y1-y3;                    (note: varies according to waves of data)
        nobs = 100;                       (note: N = 250, 500, 1,000, 2,500)
        seed = 99228;
        nreps = 200;
  repsave=all;
  save = C:\program files\Mplus\2rep*_uni_100_4waves.dat;
  model population:
       i s | y1@0, y2@1, y3@2;      (note: waves of data = y4@3, y5@4, y6@5)
       [y1-y3@0];                   (note: varies according to waves of data)
       y1-y3*.5;                    (note: varies according to waves of data)
       [i*0 s*.2];
       i*.5; s*.1; i with s*.2;
output:
         tech9;
```

A.2 Unconditional model: Univariate linear LGC model data analysis

```
title: analysis of uni_100_3 waves
data: file = C:\Program Files\Mplus\replist_uni_100_3waves.dat;
  type=montecarlo;
variable:
        names = y1-y3;              (note: varies according to waves of data)
model:
       i s | y1@0, y2@1, y3@2;      (note: waves of data = y4@3, y5@4, y6@5)
       [y1-y3@0];                   (note: varies according to waves of data)
       y1-y3*.5;                    (note: varies according to waves of data)
       [i*0 s*.2];
       i*.5; s*.1; i with s*.2;
savedata: results are results.sav;
output: tech9;
```

A.3 Unconditional model: Univariate quadratic LGC model data generation

```
montecarlo:
        names = y1-y4;                    (note: varies according to waves of data)
        nobs = 100;                       (note: N = 250, 500, 1,000, 2,500)
        seed = 17385;
        nreps = 100;
repsave=all;
save = C:\program files\Mplus\2rep*_quad_100_4waves.dat;
model population:
        i by y1-y4@1;                     (note: waves of data = y5@4, y6@5)
```

```
        l by y1@0, y2@1, y3@2, y4@3;       (note: waves of data = y5@4, y6@5)
        q by y1@0, y2@1, y3@4, y4@9;       (note: waves of data = y5@4, y6@5)
        [y1-y4@0];                         (note: varies according to waves of data)
        y1-y4*.5;                          (note: varies according to waves of data)
        [i*0 l*.1 q*.2];
        i*.5; l*.1 q*.1;
         i with l*.1;
         i with q*.2;
         l with q*.05;
output:
         tech9;
```

A.4 Unconditional model: Univariate quadratic LGC model data analysis

```
title: analysis of quad_100_4 waves
data: file = C:\Program Files\Mplus\replist_quad_100_4waves.dat;
  type=montecarlo;
variable:
        names = y1-y4;                      (note: varies according to waves of data)
model:
        i by y1-y4@1;                      (note: waves of data = y5@4, y6@5)
        l by y1@0, y2@1, y3@2, y4@3;       (note: waves of data = y5@4, y6@5)
        q by y1@0, y2@1, y3@4, y4@9;       (note: waves of data = y5@4, y6@5)
        [y1-y4@0];                         (note: varies according to waves of data)
        y1-y4*.5;                          (note: varies according to waves of data)
        [i*0 l*.1 q*.2];
        i*.5; l*.1 q*.1;
         i with l*.1;
         i with q*.2;
         l with q*.05;
savedata: results are results.sav;
output: tech9;
```

A.5 Unconditional Model: Multivariate linear LGC model data generation

```
montecarlo:
        names = y1-y6;                     (note: varies according to waves of data)
        nobs = 100;                        (note: N = 250, 500, 1,000, 2,500)
        seed = 43152;
        nreps = 1000;
   repsave=all;
save = C:\program files\Mplus\2rep*_multi_100_4waves.dat;
model population:
    Int1 by y1-y3@1;                       (note: varies according to waves of data)
    slp1 by y1@0, y2@1, y3@2;              (note: waves of data = y4@3, y5@4, y6@5)
```

```
     Int2 by y4-y6@1;                    (note: varies according to waves of data)
     slp2 by y4@0, y5@1, y6@2;           (note: waves of data = yX@3, 4 and 5)
     [y1 – y6@0];                        (note: varies according to waves of data)
     y1-y6*.5;                           (note: varies according to waves of data)
     [int1*0 slp1*.2 int2*.5 slp2*.1];
     int1*.5; slp1*.1; int2*.5; slp2*.1;
     Int1 with int2*0 slp1*.2 slp2*0;
     Int2 with slp1*0 slp2*.1;
     Slp1 with slp2*0;
output:
           tech9;
```

A.6 Unconditional Model: Multivariate linear LGC model data analysis

```
title: analysis of multi_100_3 waves
data: file = C:\Program Files\Mplus\replist_multi_100_3waves.dat;
  type=montecarlo;
variable:
          names = y1-y6;                 (note: varies according to waves of data)
model:
     Int1 by y1-y3@1;                    (note: varies according to waves of data)
     slp1 by y1@0, y2@1, y3@2;           (note: waves of data = y4@3, y5@4, y6@5)
     Int2 by y4-y6@1;                    (note: varies according to waves of data)
     slp2 by y4@0, y5@1, y6@2;           (note: waves of data = yX@3, 4 and 5)
     [y1 – y6@0];                        (note: varies according to waves of data)
     y1-y6*.5;                           (note: varies according to waves of data)
     [int1*0 slp1*.2 int2*.5 slp2*.1];
     int1*.5; slp1*.1; int2*.5; slp2*.1;
     Int1 with int2*0 slp1*.2 slp2*0;
     Int2 with slp1*0 slp2*.1;
     Slp1 with slp2*0;
savedata: results are results.sav;
output: tech9;
```

A.7 Conditional Model: Univariate linear LGC model with a time-invariant covariate data generation

```
montecarlo:
          names = y1-y3 x;               (note: varies according to waves of data)
          nobs = 100;
          seed = 19574;
          nreps = 2200;
 repsave=all;
 save = C:\program files\Mplus\rep*_cov_100_3waves.dat;
model population:
```

```
   i s | y1@0, y2@1, y3@2;                 (note: waves of data = y4@3, y5@4, y6@5)
   [x@0]; x@1;
   [y1-y3@0];                              (note: varies according to waves of data)
   [i*0 s*.2];
   i*.5; s*.1;
   i with s*.2;
   y1-y3*.5;
   i ON x*.5;
   s ON x*.1;
   [i@0 s@.2]
output:
         tech9;
```

A.8 Conditional Model: Univariate linear LGC model with a time-invariant covariate data analysis

```
title: analysis of cov_100_3 waves
data: file = C:\Program Files\Mplus\replist_cov_100_3waves.dat;
  type=montecarlo;
variable:
        names = y1-y3 x;                   (note: varies according to waves of data)
model:
   i s | y1@0, y2@1, y3@2;                 (note: waves of data = y4@3, y5@4, y6@5)
   [x@0]; x@1;
   [y1-y3@0];                              (note: varies according to waves of data)
   [i*0 s*.2];
    i*.5; s*.1;
   i with s*.2;
   y1-y3*.5;
    i ON x*.5;
    s ON x*.1;
    [i@0 s@.2]
savedata: results are results.sav;
                                     output: tech9;
```

APPENDIX B

ANOVA AND EFFECT SIZE TABLES

B.1 *ANOVA* Table and Effect Size for the χ^2

Source	*SS*	*df*	*MSE*	*F*-value	*p*-value	Effect Size
SS	5516.40	4	1379.10	41.77	<.01*	<.01
W	7716479.68	3	2572159.89	77921.74	<.01*	.75
C	10676023.23	3	3558674.41	107807.50	<.01*	.81
SS x W	5209.32	12	434.11	13.151	<.01*	<.01
SS x C	5106.61	12	425.55	12.892	<.01*	<.01
W x C	3511902.95	8	438975.36	13298.44	<.01*	.58
SS x W x C	6669.98	32	208.43	6.31	<.01*	<.01
Error	2473238.71	74925	33.01			
Total	22960459.87	75000				
Corrected Total		74999				
	Contrast Estimate	Std. Error	Lower CI	Upper CI	*p*-value	Effect Size
SS: 100 vs. 250	.37	.06	.24	.50	<.01*	<.01
SS: 250 vs. 500	.29	.06	.16	.42	<.01*	<.01
SS: 500 vs. 1000	-.07	.06	-.20	.05	.28	<.01
SS: 1000 vs. 2500	.28	.06	.15	.41	<.01*	<.01
W: 3 vs. 4	-10.75	.06	-10.88	-10.62	<.01*	.76
W: 4 vs. 5	-8.81	.05	-8.92	-8.70	<.01*	.76
W: 5 vs. 6	-10.58	.05	-10.69	-10.46	<.01*	.76
Univ. vs. Quad.	-4.41	.06	-4.54	-4.28	<.01*	.79
Uni. vs. Multi.	26.02	.05	25.90	26.13	<.01*	.79
Uni. vs. Cov.	6.24	.05	6.13	6.35	<.01*	.79

Note. *SS* = sums of square, *df* = degrees of freedom, *MSE* = mean squared error, SS = sample size; W = waves of data; C = model complexity; * indicates a significant effect at the .01 alpha level; η^2 effect sizes were interpreted as .09 as a small effect, .14 as a moderate effect, and .22 as a large effect.

B.2 *ANOVA* Table for the χ^2/df

Source	*SS*	*df*	*MSE*	*F*-value	*p*-value	Effect Size
SS	17.50	4	4.37	9.99	<.01*	<.01
W	3.43	3	1.14	2.61	.04	.00
C	1.75	3	.58	1.33	.26	.00
SS x W	4.55	12	.37	.86	.58	.00
SS x C	10.75	12	.89	2.04	.02	.00
W x C	2.76	8	.34	.78	.61	.00
SS x W x C	18.04	32	.56	1.28	.12	.00
Error	32814.90	74925	.43			
Total	108879.91	75000				
Corrected Total	32873.45	74999				
	Contrast Estimate	Std. Error	Lower CI	Upper CI	*p*-value	Effect Size
SS: 100 vs. 250	.02	<.01	<.01	.03	<.01*	<.01
SS: 250 vs. 500	.01	<.01	<.01	.02	.07	<.01
SS: 500 vs. 1000	<.01	<.01	-.01	.01	.76	<.01
SS: 1000 vs. 2500	<.01	<.01	-.01	.02	.50	<.01
W: 3 vs. 4	.01	<.01	<.01	.03	.01	<.01
W: 4 vs. 5	<.01	<.01	-.01	<.01	.32	<.01
W: 5 vs. 6	<-.01	<.01	-.01	<.01	.39	<.01
Univ. vs. Quad.	<-.01	<.01	-.01	.01	.62	<.01
Uni. vs. Multi.	.01	<.01	<-.01	.02	.13	<.01
Uni. vs. Cov.	<-.01	<.01	-.01	.01	.90	<.01

Note. *SS* = sums of square, *df* = degrees of freedom, *MSE* = mean squared error, SS = sample size; W = waves of data; C = model complexity; * indicates a significant effect at the .01 alpha level; η^2 effect sizes were interpreted as .09 as a small effect, .14 as a moderate effect, and .22 as a large effect.

B.3 *ANOVA* Table and Effect Size for the *NNFI*

Source	*SS*	*df*	*MSE*	*F*-value	*p*-value	Effect Size
SS	.01	4	<.01	43.26	<.01*	.00
W	<.01	3	<.01	4.54	<.01*	.00
C	<.01	3	<.01	15.30	<.01*	.00
SS x W	<.01	12	<.01	3.64	<.01*	.00
SS x C	<.01	12	<.01	7.76	<.01*	.00
W x C	<.01	8	<.01	1.33	.22	.00
SS x W x C	<.01	32	<.01	1.82	<.01*	.00
Error	6.41	74925	855522.18			
Total	74967.63	75000				
Corrected Total	6.45	74999				
	Contrast Estimate	Std. Error	Lower CI	Upper CI	*p*-value	Effect Size
SS: 100 vs. 250	<-.01	.00	<-.01	<-.01	<.01*	.00
SS: 250 vs. 500	<.01	.00	<.01	<.01	.12	.00
SS: 500 vs. 1000	<-.01	.00	<.01	<.01	.98	.00
SS: 1000 vs. 2500	<-.01	.00	<.01	<.01	.97	.00
W: 3 vs. 4	<.01	.00	<-.01	<.01	<.01*	.00
W: 4 vs. 5	<.01	.00	<-.01	<.01	.05	.00
W: 5 vs. 6	<.01	.00	<.01	<.01	.77	.00
Univ. vs. Quad.	<-.01	.00	<.01	<.01	.77	.00
Uni. vs. Multi.	<-.01	.00	<-.01	<.01	<.01	.00
Uni. vs. Cov.	<-.01	.00	<.01	<.01	.72	.00

Note. *SS* = sums of square, *df* = degrees of freedom, *MSE* = mean squared error, SS = sample size; W = waves of data; C = model complexity; * indicates a significant effect at the .01 alpha level; η^2 effect sizes were interpreted as .09 as a small effect, .14 as a moderate effect, and .22 as a large effect.

B.4 *ANOVA* Table and Effect Size for the *CFI*

	Contrast Estimate	Std. Error	Lower CI	Upper CI	*p*-value	Effect Size
SS: 100 vs. 250	<-.01	.00	<.01	<.01	<.01*	.14
SS: 250 vs. 500	<-.01	.00	<.01	<.01	<.01*	.14
SS: 500 vs. 1000	<.01	.00	<.01	<.01	<.01*	.14
SS: 1000 vs. 2500	<.01	.00	<.01	<.01	<.01*	.14
W: 3 vs. 4	<.01	.00	<-.01	<.01	<.01*	<.01
W: 4 vs. 5	<-.01	.00	<.01	<.01	<.01*	<.01
W: 5 vs. 6	<.01	.00	<.01	<.01	<.01*	<.01
Univ. vs. Quad.	<.01	.00	<.01	<.01	<.01*	.01
Uni. vs. Multi.	<-.01	.00	<-.01	<-.10	<.01*	.01
Uni. vs. Cov.	<.01	.00	<-.10	<.10	<.01*	.01

Note. *SS* = sums of square, *df* = degrees of freedom, *MSE* = mean squared error, SS = sample size; W = waves of data; C = model complexity; * indicates a significant effect at the .01 alpha level; η^2 effect sizes were interpreted as .09 as a small effect, .14 as a moderate effect, and .22 as a large effect.

B.5 *ANOVA* Table and Effect Size for the *RMSEA*

Source	*SS*	*df*	*MSE*	*F*-value	*p*-value	Effect Size
SS	4.81	4	1.20	2414.79	<.01*	.11
W	.24	3	.08	165.51	<.01*	.00
C	.18	3	.06	123.69	<.01*	.00
SS x W	.07	12	<.01	11.96	<.01*	.00
SS x C	.05	12	<.01	9.82	<.01*	.00
W x C	.04	8	<.01	10.44	<.01*	.00
SS x W x C	.04	32	<.01	2.72	<.01*	.00
Error	37.37	74925	<.01			
Total	56.52	75000				
Corrected Total	42.69	74999				
	Contrast Estimate	Std. Error	Lower CI	Upper CI	*p*-value	Effect Size
SS: 100 vs. 250	.01	.00	.01	.01	<.01*	.11
SS: 250 vs. 500	<.01	.00	<.01	<.01	<.01*	.11
SS: 500 vs. 1000	<.01	.00	<.01	<.01	<.01*	.11
SS: 1000 vs. 2500	<.01	.00	<.01	<.01	<.01*	.11
W: 3 vs. 4	<.01	.00	<.01	<.01	<.01*	.00
W: 4 vs. 5	<.01	.00	<.01	<.01	<.01*	.00
W: 5 vs. 6	<.01	.00	<.01	<.01	<.01*	.00
Univ. vs. Quad.	<.01	.00	<.01	<.01	<.01*	.00
Uni. vs. Multi.	<-.01	.00	<-.01	<-.01	<.01*	.00
Uni. vs. Cov.	<-.01	.00	<-.01	<-.10	<.01*	.00

Note. *SS* = sums of square, *df* = degrees of freedom, *MSE* = mean squared error, SS = sample size; W = waves of data; C = model complexity; * indicates a significant effect at the .01 alpha level; η^2 effect sizes were interpreted as .09 as a small effect, .14 as a moderate effect, and .22 as a large effect.

APPENDIX C

MEAN FIT INDEX TABLES

C.1 χ^2 mean fit values, standard deviations, p-values, and χ^2/df mean values

Model & Waves of Data		Sample Size				
		100	250	500	1,000	2,500
Uni.	3	$\chi^2 = 1.09$	$\chi^2 = .97$	$\chi^2 = 1.04$	$\chi^2 = 1.01$	$\chi^2 = 1.01$
		$df = 1$	$df = 1$	$df = 1$	$df = 1$	$df = 1$
		$p = .48$	$p = .51$	$p = .48$	$p = .49$	$p = .50$
		$\chi^2/df = 1.09$	$\chi^2/df = .97$	$\chi^2/df = 1.04$	$\chi^2/df = 1.01$	$\chi^2/df = 1.01$
	4	$\chi^2 = 4.98$	$\chi^2 = 4.87$	$\chi^2 = 4.87$	$\chi^2 = 4.84$	$\chi^2 = 5.11$
		$df = 5$	$df = 5$	$df = 5$	$df = 5$	$df = 5$
		$p = .50$	$p = .50$	$p = .51$	$p = .50$	$p = .49$
		$\chi^2/df = .99$	$\chi^2/df = .97$	$\chi^2/df = .97$	$\chi^2/df = .96$	$\chi^2/df = 1.02$
	5	$\chi^2 = 10.32$	$\chi^2 = 9.67$	$\chi^2 = 9.98$	$\chi^2 = 9.89$	$\chi^2 = 9.79$
		$df = 10$	$df = 10$	$df = 10$	$df = 10$	$df = 10$
		$p = .48$	$p = .51$	$p = .50$	$p = .50$	$p = .51$
		$\chi^2/df = 1.03$	$\chi^2/df = .96$	$\chi^2/df = .99$	$\chi^2/df = .98$	$\chi^2/df = .97$
	6	$\chi^2 = 16.36$	$\chi^2 = 16.35$	$\chi^2 = 15.90$	$\chi^2 = 16.04$	$\chi^2 = 16.26$
		$df = 16$	$df = 16$	$df = 16$	$df = 16$	$df = 16$
		$p = .47$	$p = .48$	$p = .50$	$p = .49$	$p = .48$
		$\chi^2/df = 1.02$	$\chi^2/df = 1.02$	$\chi^2/df = .99$	$\chi^2/df = 1.00$	$\chi^2/df = 1.01$
Quad	4	$\chi^2 = 1.07$	$\chi^2 = 1.05$	$\chi^2 = .92$	$\chi^2 = 1.01$	$\chi^2 = .92$
		$df = 1$	$df = 1$	$df = 1$	$df = 1$	$df = 1$
		$p = .48$	$p = .49$	$p = .51$	$p = .50$	$p = .50$
		$\chi^2/df = 1.07$	$\chi^2/df = 1.05$	$\chi^2/df = .92$	$\chi^2/df = 1.01$	$\chi^2/df = .92$
	5	$\chi^2 = 6.01$	$\chi^2 = 5.93$	$\chi^2 = 5.99$	$\chi^2 = 6.20$	$\chi^2 = 5.77$
		$df = 6$	$df = 6$	$df = 6$	$df = 6$	$df = 6$
		$p = .49$	$p = .50$	$p = .50$	$p = .47$	$p = .51$
		$\chi^2/df = 1.00$	$\chi^2/df = .98$	$\chi^2/df = .99$	$\chi^2/df = 1.03$	$\chi^2/df = .96$
	6	$\chi^2 = 12.31$	$\chi^2 = 12.40$	$\chi^2 = 11.80$	$\chi^2 = 11.76$	$\chi^2 = 11.92$
		$df = 12$	$df = 12$	$df = 12$	$df = 12$	$df = 12$
		$p = .48$	$p = .47$	$p = .51$	$p = .51$	$p = .50$
		$\chi^2/df = 1.02$	$\chi^2/df = 1.03$	$\chi^2/df = .98$	$\chi^2/df = .98$	$\chi^2/df = .99$
Multi	3	$\chi^2 = 7.37$	$\chi^2 = 7.21$	$\chi^2 = 6.98$	$\chi^2 = 7.05$	$\chi^2 = 6.92$
		$df = 7$	$df = 7$	$df = 7$	$df = 7$	$df = 7$
		$p = .46$	$p = .48$	$p = .50$	$p = .49$	$p = .50$
		$\chi^2/df = 1.05$	$\chi^2/df = 1.03$	$\chi^2/df = .99$	$\chi^2/df = 1.00$	$\chi^2/df = .98$
	4	$\chi^2 = 22.26$	$\chi^2 = 22.79$	$\chi^2 = 21.98$	$\chi^2 = 22.13$	$\chi^2 = 21.99$
		$df = 22$	$df = 22$	$df = 22$	$df = 22$	$df = 22$
		$p = .48$	$p = .46$	$p = .50$	$p = .49$	$p = .50$
		$\chi^2/df = 1.01$	$\chi^2/df = 1.03$	$\chi^2/df = .99$	$\chi^2/df = 1.00$	$\chi^2/df = .99$
	5	$\chi^2 = 43.48$	$\chi^2 = 41.36$	$\chi^2 = 41.36$	$\chi^2 = 41.10$	$\chi^2 = 41.32$
		$df = 41$	$df = 41$	$df = 41$	$df = 41$	$df = 41$
		$p = .43$	$p = .48$	$p = .49$	$p = .49$	$p = .49$
		$\chi^2/df = 1.05$	$\chi^2/df = 1.00$	$\chi^2/df = 1.00$	$\chi^2/df = 1.00$	$\chi^2/df = 1.00$
	6	$\chi^2 = 67.31$	$\chi^2 = 65.07$	$\chi^2 = 65.22$	$\chi^2 = 64.48$	$\chi^2 = 63.46$

		df = 64	df = 64	df = 64	df = 64	df = 64
		p = .42	p = .47	p = .46	p = .48	p = .51
		χ^2/df = 1.05	χ^2/df = 1.01	χ^2/df = 1.01	χ^2/df = 1.00	χ^2/df = .99
Cov.	3	χ^2 = 6.13	χ^2 = 6.09	χ^2 = 6.12	χ^2 = 5.91	χ^2 = 6.05
		df = 6	df = 6	df = 6	df = 6	df = 6
		p = .48	p = .49	p = .48	p = .50	p = .50
		χ^2/df = 1.02	χ^2/df = 1.01	χ^2/df = 1.02	χ^2/df = .98	χ^2/df = 1.00
	4	χ^2 = 10.95	χ^2 = 10.94	χ^2 = 10.91	χ^2 = 10.94	χ^2 = 10.88
		df = 11	df = 11	df = 11	df = 11	df = 11
		p = .50	p = .49	p = .50	p = .50	p = .50
		χ^2/df = .99	χ^2/df = .99	χ^2/df = .99	χ^2/df = .99	χ^2/df = .98
	5	χ^2 = 17.52	χ^2 = 17.37	χ^2 = 16.85	χ^2 = 16.79	χ^2 = 15.11
		df = 17	df = 17	df = 17	df = 17	df = 17
		p = .47	p = .48	p = .51	p = .50	p = .49
		χ^2/df = 1.03	χ^2/df = 1.02	χ^2/df = .99	χ^2/df = .98	χ^2/df = 1.00
	6	χ^2 = 24.73	χ^2 = 24.16	χ^2 = 21.94	χ^2 = 23.80	χ^2 = 22.08
		df = 24	df = 24	df = 24	df = 24	df = 24
		p = .46	p = .49	p = .50	p = .50	p = .49
		χ^2/df = 1.03	χ^2/df = 1.00	χ^2/df = .99	χ^2/df = .99	χ^2/df = 1.00

C.2 NNFI mean values and standard deviations

Model	Waves of Data	Sample Size				
		100	250	500	1,000	2,500
Uni. LGC	3	.99 (.03)	1.00 (.01)	.99 (<.01)	.99 (<.01)	.99 (<.01)
	4	.99 (.01)	1.00 (.01)	1.00 (<.01)	1.00 (<.01)	.99 (<.01)
	5	.99 (.01)	1.00 (<.01)	1.00 (<.01)	1.00 (<.01)	1.00 (<.01)
	6	.99 (.01)	.99 (<.01)	1.00 (<.01)	.99 (<.01)	.99 (<.01)
Quad. LGC	4	.99(.02)	.99 (<.01)	1.00 (<.01)	.99 (<.01)	.99 (<.01)
	5	.99 (<.01)	1.00 (<.01)	1.00 (<.01)	.99 (<.01)	1.00 (<.01)
	6	.99 (<.01)	.99 (<.01)	1.00 (<.01)	1.00 (<.01)	1.00 (<.01)
Multi. LGC	3	.99 (.03)	.99 (.01)	1.00 (<.01)	.99 (<.01)	1.00 (<.01)
	4	.99 (.03)	.99 (<.01)	1.00 (<.01)	.99 (<.01)	1.00 (<.01)
	5	.99 (.01)	.99 (<.01)	.99 (<.01)	.99 (<.01)	.99 (<.01)
	6	.99 (.01)	.99 (<.01)	.99 (<.01)	.99 (<.01)	1.00 (<.01)
Cov. LGC	3	.99 (.01)	.99 (<.01)	.99 (<.01)	1.00 (<.01)	.99 (<.01)
	4	1.00 (.01)	1.00 (<.01)	1.00 (<.01)	1.00 (<.01)	1.00 (<.01)
	5	.99 (.01)	.99 (<.01)	1.00 (<.01)	1.00 (<.01)	.99 (<.01)
	6	.99 (<.01)	.99 (<.01)	1.00 (<.01)	1.00 (<.01)	.99 (<.01)
					1.00 (<.01)	

Note: *NNFI* = Non-normed fit index; Uni = univariate LGC Model; Quad = Quadratic LGC Model; Multi = Multivariate LGC model; Cov = Univariate Linear LGC model with a single time invariant covariate.

C.4 *CFI* mean values and standard deviations

Model	Waves of Data	Sample Size 100	250	500	1,000	2,500
Uni. LGC	3	.99 (.01)	.99 (<.01)	.99 (<.01)	.99 (<.01)	.99 (<.01)
	4	.99 (<.01)	.99 (<.01)	.99 (<.01)	.99 (<.01)	.99 (<.01)
	5	.99 (<.01)	.99 (<.01)	.99 (<.01)	.99 (<.01)	.99 (<.01)
	6	.99 (<.01)	.99 (<.01)	.99 (<.01)	.99 (<.01)	.99 (<.01)
Quad. LGC	4	.99 (.01)	.99 (<.01)	.99 (<.01)	.99 (<.01)	.99 (<.01)
	5	.99 (<.01)	.99 (<.01)	.99 (<.01)	.99 (<.01)	.99 (<.01)
	6	.99 (<.01)	.99 (<.01)	.99 (<.01)	.99 (<.01)	.99 (<.01)
Multi. LGC	3	.99 (.01)	.99 (<.01)	.99 (<.01)	.99 (<.01)	.99 (<.01)
	4	.99 (<.01)	.99 (<.01)	.99 (<.01)	.99 (<.01)	.99 (<.01)
	5	.99 (<.01)	.99 (<.01)	.99 (<.01)	.99 (<.01)	.99 (<.01)
	6	.99 (<.01)	.99 (<.01)	.99 (<.01)	.99 (<.01)	.99 (<.01)
Cov. LGC	3	.99 (.01)	.99 (<.01)	.99 (<.01)	.99 (<.01)	.99 (<.01)
	4	.99 (<.01)	.99 (<.01)	.99 (<.01)	.99 (<.01)	.99 (<.01)
	5	.99 (<.01)	.99 (<.01)	.99 (<.01)	.99 (<.01)	.99 (<.01)
	6	.99 (<.01)	.99 (<.01)	.99 (<.01)	.99 (<.01)	.99 (<.01)

Note: *CFI* = Comparative Fit Index; Uni = univariate LGC Model; Quad = Quadratic LGC Model; Multi = Multivariate LGC model; Cov = Univariate Linear LGC model with a single time invariant covariate

C.5 *RMSEA* mean values and standard deviations

Model	Waves of Data	Sample Size 100	250	500	1,000	2,500
Uni. LGC	3	.03 (.06)	.02 (.03)	.01 (.02)	.01 (.01)	<.01 (.01)
	4	.02 (.04)	.01 (.02)	.01 (.01)	<.01 (.01)	<.01 (<.01)
	5	.02 (.03)	.01 (.02)	.01 (.01)	<.01 (.01)	<.01 (<.01)
	6	.02 (.02)	.01 (.01)	<.01 (.01)	<.01 (.01)	<.01 (<.01)
Quad. LGC	4	.03 (.06)	.02 (.04)	.01 (.01)	.01 (.01)	<.01 (<.01)
	5	.02 (.03)	.01 (.02)	.01 (.01)	<.01 (.01)	<.01 (<.01)
	6	.02 (.03)	.01 (.02)	.01 (.01)	<.01 (<.01)	<.01 (<.01)
Multi. LGC	3	.03 (.03)	.01 (.02)	.01 (.01)	<.01 (.01)	<.01 (<.01)
	4	.02 (.02)	.01 (.01)	<.01 (.01)	<.01 (<.01)	<.01 (<.01)
	5	.02 (.02)	.01 (.01)	<.01 (.01)	<.01 (<.01)	<.01 (<.01)
	6	.02 (.02)	.01 (.01)	<.01 (.01)	<.01 (<.01)	<.01 (<.01)
Cov. LGC	3	.02 (.03)	.01 (.02)	.01 (.01)	<.01 (<.01)	<.01 (<.01)
	4	.02 (.02)	.01 (.01)	.01 (.01)	<.01 (<.01)	<.01 (<.01)
	5	.02 (.02)	.01 (.01)	<.01 (.01)	<.01 (<.01)	<.01 (<.01)
	6	.02 (.02)	.01 (.01)	<.01 (.01)	<.01 (<.01)	<.01 (<.01)

Note: *RMSEA* = Root Mean Squared Error of Approximation; Uni = Univariate LGC Model; Quad = Quadratic LGC Model; Multi = Multivariate LGC model; Cov = Univariate Linear LGC model with a single time invariant covariate.